DOI:10.16016/j.2097-0927.202409145

基于小波变换和CNN-Transformer的超声甲状腺结节分割算法研究

郑水婧[1],杨 君[2],蔡瑜娇[2],文 静[1] [1]重庆大学计算机学院模式识别实验室,重庆;[2]陆军军医大学(第三军医大学)第二附属医院普通外科,重庆

[摘要] **目的** 融合小波变换和CNN-Transformer构建甲状腺结节自动分割网络,以提升甲状腺结节超声影像的智能分割效率和精准度。**方法** 收集2023年5月至2024年2月在陆军军医大学第二附属医院超声科获取的1 371套甲状腺结节超声影像。经过预处理和标准化后,数据按照8∶1∶1的比例划分为训练集、验证集和测试集。以UNet为基础,将CNN与Swin-Transformer并联作为编码器,并在编码器与解码器之间插入小波变换模块,完成甲状腺结节分割网络的构建。使用准确率、IoU和Dice系数指标在收集的内部数据集上评估分割模型性能。**结果** 本研究最终在收集的1 371套超声甲状腺结节上进行验证,平均Dice系数达到79.63%,IoU达到67.3%。相较于UNet,分割准确度提升1.02%。其甲状腺结节分割结果位置准确,边缘平滑。分割出的甲状腺结节相比于其他方法分割出的结节,在轮廓形态上与医师的手工分割结果吻合度更高。相较于UNet方法,本方法对结节的纹理学习更加充分,规避了结节容易错误分割为周围组织的情况。**结论** 构建的基于小波变换和CNN-Transformer的分割模型精度优于UNet、Attention-UNet、UNetv2等UNet变种网络以及SAM Med2D分割大模型等先进分割方法,能有效应用于超声甲状腺结节的精确分割,提升医师的工作效率。

[关键词] 甲状腺结节分割;小波变换;超声图像诊断;深度学习
[中图法分类号] R312;R445.1;R581.04 [文献标志码] A

Ultrasound thyroid nodule segmentation algorithm based on wavelet transform and CNN-Transformer

ZHENG Shuijing[1], YANG Jun[2], CAI Yujiao[2], WEN Jing[1] [1]Laboratory of Pattern Recognition, College of Computer Science, Chongqing University, Chongqing; [2]Department of General surgery, Second Affiliated Hospital, Army Medical University (Third Military Medical University), Chongqing, China

[**Abstract**] **Objective** To develop an automatic segmentation network for thyroid nodules by integrating wavelet transform and CNN-Transformer in order to improve the efficiency and accuracy of ultrasound image segmentation. **Methods** A total of 1 371 sets of ultrasound images of thyroid nodules were collected from Department of Ultrasonography of Second Affiliated Hospital of Army Medical University between May 2023 and February 2024. After preprocessing and normalization, the data were divided into training, validation, and testing sets in a ratio of 8∶1∶1. Based on UNet, CNN and Swin-Transformer were used in parallel as the encoder, with a wavelet transform module inserted between the encoder and decoder to construct a thyroid nodule segmentation network. The performance of the segmentation model was evaluated

[基金项目] 重庆市卫健委与科技局联合项目(CSTB2023TIAD-KPX0050,2024ZDXM007);中央高校基本科研基金(2022CDJYGRH0156)
[通信作者] 文 静,E-mail:wj@cqu.edu.cn

2097-0927/Copyright © 2025,陆军军医大学(第三军医大学)。这篇开放获取文章遵循CC BY许可协议(https://creativecommons.org/licenses/by/4.0/)
引用格式:郑水婧,杨君,蔡瑜娇,等.基于小波变换和CNN-Transformer的超声甲状腺结节分割算法研究[J].陆军军医大学学报,2025,47(14):1595-1601.DOI:10.16016/j.2097-0927.202409145.
Citation:ZHENG S J, YANG J, CAI Y J, *et al*. Ultrasound thyroid nodule segmentation algorithm based on wavelet transform and CNN-Transformer[J]. J Army Med Univ,2025,47(14):1595-1601.DOI:10.16016/j.2097-0927.202409145.

on the collected internal dataset using accuracy, IoU, and Dice coefficient metrics. **Results** The finally verified 1 371 sets of ultrasonic thyroid nodules had an average Dice coefficient of 79.63% and an IoU of 67.30%. Compared with UNet, the segmentation accuracy was improved by 1.02%. The segmentation model obtained accurate location and smooth edges of thyroid nodules, and the segmentation was more consistent in thyroid nodule edge and morphology with those marked by doctors manually when compared with other segmentations. Compared with UNet, this segmentation method can learn the texture of nodules more fully and avoid the situation that nodules had been incorrectly divided into surrounding tissues. **Conclusion** Our developed segmentation model based on wavelet transform and CNN-Transformer demonstrates better segmentation accuracy in comparison to conventional UNet variants, such as UNet, Attention-UNet, and UNetv2, and medical segment anything models like SAM Med2D. This segmentation method enables accurate segmentation of ultrasound thyroid nodules, thereby enhancing clinical workflow efficiency through automated precise delineation.

[**Key words**] thyroid nodule segmentation; wavelet transform; ultrasound image diagnosis; deep learning

Supported by the Joint Project of Chongqing Science and Technology Commission and Chongqing Health Commission (CSTB2023TIAD-KPX0050, 2024ZDXM007), and the Fundamental Research Fund for the Central Universities of China (2022CDJYGRH015). Corresponding author: WEN Jing, E-mail: wj@cqu.edu.cn

近年来,随着现代生活节奏的加快,人群中甲状腺结节发病率不断上升[1],如不及时治疗可能会发展为甲状腺癌。鉴于甲状腺结节的潜在风险,对甲状腺的定期检查显得愈加重要[2]。超声成像技术是诊断甲状腺疾病的常用方法,该技术具有实时、廉价、无创、无放射性等特点[3]。在临床检查中,医师需要对甲状腺超声图像进行精确分割,判断甲状腺结节大小,以区分结节的良性与恶性。此外,准确的超声图像分割还可以帮助医师建立结节的动态监测系统,通过对比不同时间点的结节变化,提供对结节生长或缩小的定量分析,这对于患者的随访和预后评估尤为重要。

目前临床仍多采用手工勾画甲状腺结节,这种方法依赖于医师的主观判断。在实际操作中,手工勾画容易导致分割精度不足,从而影响后续的诊断和量化过程,降低了诊断的效率和准确性。通过引入图像自动分割技术,为准确、高效地分割超声图像提供了思路。由于超声图像成像质量较差,普遍存在衰减、斑点、阴影、对比度低、信号丢失[4]、图像中的器官组织模糊且不均匀[5]等问题。特别是对于甲状腺超声图像,其甲状腺结节纹理不清晰,与周围组织的边界模糊,并且自身又具有较高的变化性,这些都给甲状腺结节的自动分割带来了巨大挑战。

目前甲状腺结节的自动分割方法主要有3种:基于轮廓模型、区域特征以及深度学习方法[3]。其中轮廓模型难以处理不规则边缘和恶性结节;区域特征方法在灰度不均匀的图像中存在一定的局限性;而深度学习通过学习大量数据提取特征,近年来得到广泛应用且效果显著。研究发现,频率信息能够很好地对纹理特征进行表征,而深度学习模型如卷积神经网络(convolutional neural network, CNN)和Transformer仅在空间域提取特征,无法模拟纹理分析的频率域分布[6]。小波变换因其独特的时频局部化特征和多尺度分解能力,成为纹理特征提取和分类的重要工具。因此,本研究在分割网络设计中添加小波变换模块获取频域特征。

为实现特征学习过程的自动化,已有学者将传统纹理分析方法以可学习的"层"形式融入神经网络中,并取得了更高的诊断精度[7]。针对小波变换,RODRIGUEZ等[8]结合提升小波提出深度自适应小波网络(deep adaptive wavelet network, DAWN),首次在CNN背景下提出可训练的小波滤波器。本研究借鉴DAWN的经验,将自适应小波变换与神经网络融合,基于小波变换和CNN-Transformer构建超声甲状腺结节分割网络,将CNN与Swin Transformer并联作为编码器,分别学习局部信息和全局信息,以增强模型的表征能力,补充长程依赖关系。同时,引入小波变换模块进一步强化关键局部细节,提升分割框架的性能。

1 资料与方法

1.1 数据集来源

本研究所使用的数据集来源于2023年5月至2024年2月陆军军医大学第二附属医院超声科收治的

甲状腺结节患者的超声图像资料,其中包含符合研究需求的1 371套甲状腺结节超声影像。纳入标准:临床资料完整,超声影像完整清晰,每张图片至少含有1个结节。排除标准:①图像质量差,或有明显影响分割的伪影;②图像信息不全;③除甲状腺结节以外的其他病理类型的甲状腺疾病。患者的影像资料包含了原始超声图像及对应的结节分割标签(掩膜)。每位患者的结节分割标签由经验丰富的超声医师手工标记,以确保训练数据的准确性。本研究经陆军军医大学第二附属医院医学伦理委员会伦理审查(2023-研第165-01)。

1.2 数据预处理

数据预处理主要包含下述步骤:①数据集划分。将1 371套数据集按照8∶1∶1的比例随机划分为训练集、验证集和测试集。为了提高模型的泛化能力和评估的可靠性,本研究对划分后的训练集进行了5折交叉验证。②数据进行离差标准化。对每张图像的像素值进行归一化处理,将像素值映射到[0,1]的范围内。其中,归一化的公式如下:

$$x_j = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{1}$$

1.3 融合小波变换的超声甲状腺结节分割网络

基于小波变换的超声甲状腺结节分割网络架构设计如图1所示,本结构是一个编码器解码器结构。提出的架构提供了一个端到端的训练策略,该策略集成了Swin Transformer的全局上下文表示和编码器中CNN模块的局部特征。然后使用双级特征融合获得更丰富的特征表示。引入一个可学习的小波变换分支,用以增强解码器对纹理特征的表达能力,最终解码器输出精确的结节分割图。

1.3.1 编码器模块

在网络的编码器部分,受RODRIGUEZ等[8]提出的深度自适应小波网络启发,本研究引入一个可学习的小波分支,用以替代原始UNet中的第1个跳跃连接,从而增强解码器的纹理特征表示。同时,通过并联CNN和Swin Transformer的结构,有效结合了局部和全局信息,提高了对医学超声图像复杂结构的理解能力。这种设计不仅保留空间信息和细节特征,还能生成与原始图像尺寸对齐的精确分割结果,为医师提供了可靠的分析工具。每个CNN和Swin Transformer都包括3个不同的层级,通过跳跃连接将每个CNN层的局部特征传输到相应的Swin Transformer层,以获得通用的特征表示。然后,将每个传输后的CNN层与其对应的Transformer层进行并联,并利用图像块合并模块生成分层的特征表示。本研究设计CNN与Transformer分层并联结构以利用多尺度的特征表示能力,并通过双级特征融合交换来自不同尺度的信息。

1.3.2 解码器模块

解码器通过逐步上采样和跳跃连接,将来自编码器的特征信息和低级特征有效结合,以重建高分辨率的输出。这种设计不仅能够保留空间信息并恢复细节特征,还能有效生成与原始输入尺寸对齐的分割图,为超声医学图像分割任务提供精确的预测能力。

1.4 联合损失函数

医学图像分割领域常使用交叉熵损失函数,可定义为:

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\widehat{y_i}\right) + \left(1 - y_i\right)\log\left(1 - \widehat{y_i}\right)\right] \tag{2}$$

其中y_i和$\widehat{y_i}$分别代表标签和第i个像素的像素值预测结果,N=C×W×H代表总的像素数量。

甲状腺结节分割往往存在数据分布不平衡的特性。相较于背景区域,结节区域在尺寸上小很多。因

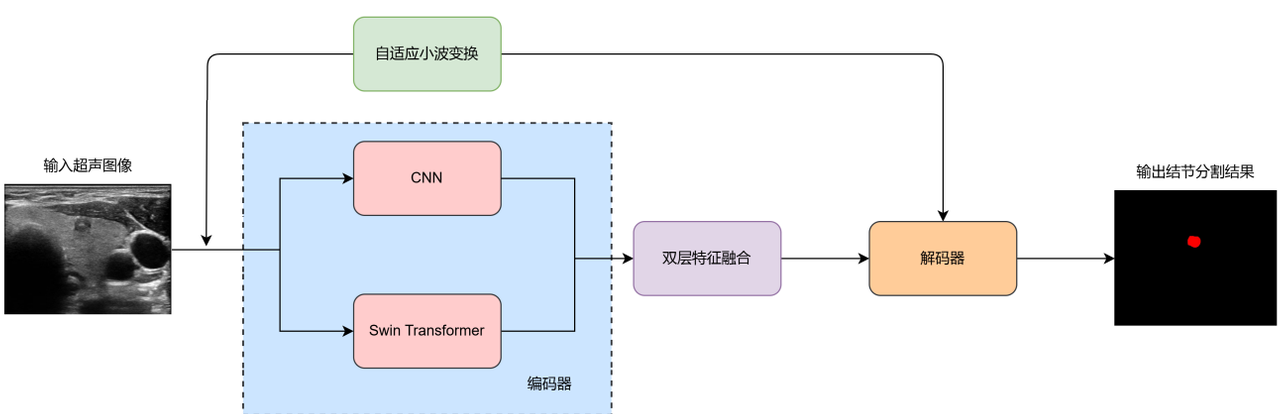


图1 甲状腺结节分割算法的网络架构

此,本研究引入Dice损失以解决此问题,可定义为:

$$L_{Dice} = 1 - \frac{2\sum_{i=1}^{N} y_i \widehat{y_i}}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \widehat{y_i}} \quad (3)$$

IoU也用于衡量分割模型的预测结果和真实标签的重叠程度。可以一起使用来进一步优化模型对分割任务的学习能力。IoU损失可定义为:

$$L_{IoU} = 1 - \frac{\sum_{i=1}^{N} y_i \widehat{y_i}}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \widehat{y_i}} \quad (4)$$

最终,构建了分割网络的联合损失函数,可表示为:

$$L_{Total} = L_{BCE} + L_{Dice} + L_{IoU} \quad (5)$$

1.5 模型训练方法

该框架在PyTorch 1.12.1中实现。所有模型均使用带有CUDA11.4的NVIDIA 3090 GPU进行训练。在训练中采用Adam优化器,初始学习速率baselr设为0.01,采用DeepLabV2[9]中的策略进行学习率更新,并按照公式6进行计算,其中iter是当前迭代的步数,max_iter是训练过程中总的迭代步数。每个网络训练200个轮次,训练批大小设置为8。学习率lr的计算公式如下:

$$lr = baselr \times (1 - \frac{iter}{\max_iter})^{0.9} \quad (6)$$

在甲状腺结节分割实验中,将所有网络的输入图像尺寸调整为512×512。

1.6 模型评价指标

本研究利用图像分割中常用的Accuracy、IoU和Dice系数来评价模型的分割效果,公式如下:

1)准确率:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (7)$$

2)IoU:

$$IoU = \frac{TP}{TP + FN + FP} \quad (8)$$

3)Dice系数:

$$Dice\ Coefficient = \frac{2TP}{2TP + FP + FN} \quad (9)$$

其中TP、FP、TN和FN分别表示真阳性、假阳性、真阴性和假阴性。

2 结果

2.1 验证模型在不同分割方法中的泛化能力

利用PyTorch深度学习框架对本研究所设计的分割网络进行编程实现。本研究与ResUNet[10]、UNet[11]、Attention U-Net[12]、UNetv2[13]、Deeplabv3[14]、SAM Med2D[15]等主流医学图像分割方法进行了比较,结果如表1所示。

表1 各种方法在甲状腺结节分割上的表现

方法	准确率(%)	IoU	Dice
ResUNet[10]	83.47	67.02	79.25
UNet[11]	84.85	62.04	73.57
Attention-UNet[12]	79.73	60.95	73.79
UNetv2[13]	82.67	65.42	74.12
Deeplabv3[14]	82.90	64.50	72.27
SAM Med2D[15]	73.46	63.38	71.15
本研究方法	85.87	67.30	79.63

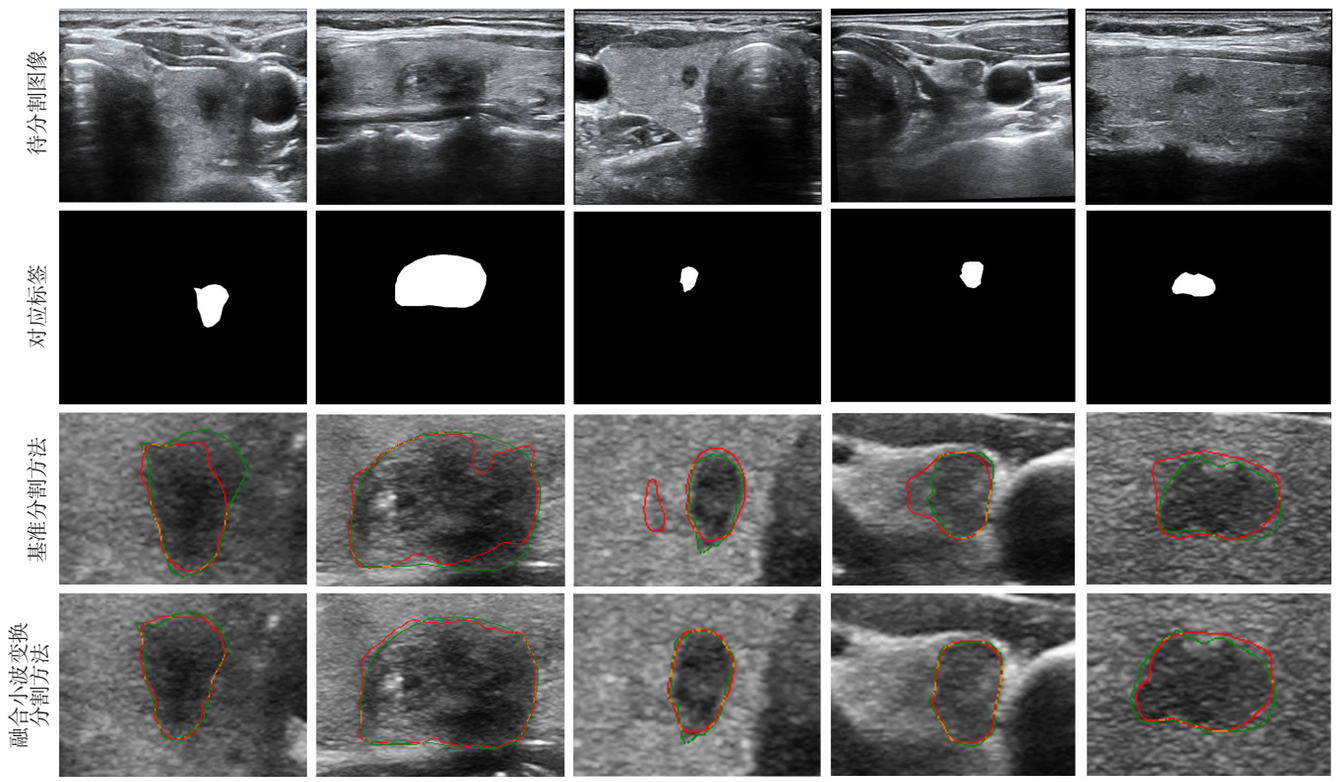
从实验结果可以看出,SAM Med2D等医学图像大模型虽然在分割方面取得了一定进展,但在处理甲状腺复杂结构和低对比度超声图像时,仍然存在精度不足的问题,效果不如传统CNN。本研究提出的方法在结节数据集上评估指标优于对比实验的其他方法,这证明了所提出的模型具有良好的学习能力和泛化能力。同时,本方法体现出了与UNet、SAM Med2D相当的性能,对于结节在准确率上取得了显著的提高,与对比实验中最好的UNet相比准确率提高了1.02%。因此,融合小波变换和CNN-Transformer的分割网络结构增强了特征提取能力,对结节分割的提升效果很明显。

2.2 模型的实验结果可视化分析

本研究所有分割示例图中,绿线勾勒的区域为对应的医师手工标注图,红线勾勒的区域为当前方法的分割预测结果。图2是融合小波变换的方法与没有引入小波变换模块相比的部分可视化示例,红线对比突出显示了本研究提出的方法优于基准分割方法。本研究所提出的分割网络在复杂低质量的超声甲状腺图像中能更准确地找到结节的位置。图2第1、2列可以看出没有加入小波变换模块的基准方法会出现欠分割的情况,第3列可以看到基准方法会出现将非甲状腺结节区域误识别为甲状腺结节的情况,第4、5列可以看出基准方法会出现过分割的情况。本研究方法生成的分割结果最接近真实情况,同时能够识别不同形状和大小的结节。

3 讨论

精准的甲状腺结节超声图像分割对于提高临床诊断效率、减轻医师工作负担以及提升诊断准确性具有重要意义。本研究提出的基于深度学习和小波变换的分割模型,旨在解决超声图像中甲状腺周围组织



绿线为医师手工标注,红线为当前方法的分割预测结果

图2 超声甲状腺结节的分割结果

结构复杂、纹理不清晰以及图像分辨率低的问题,提高甲状腺结节的自动分割精度,从而帮助医师更快、更准确地识别结节,减少由于人工操作引入的主观性误差。另外,结节的准确分割结果有助于有效量化结节的大小和形状,帮助医师在制定相关手术方案时做出更加科学和精确的决策,进而提升手术成功率。

目前,在深度学习领域,医学图像分割网络主要基于卷积神经网络(CNN)。CNN模型在各种计算机视觉任务中取得了成功,最典型的如UNet[11]和V-Net[16]。向永嘉等[17]采用DenseUnet模型进行核磁共振图像下肛提肌的自动分割,吴哲等[18]利用Vgg16-Unet模型进行MRI图像下前列腺分区智能分割研究。此外,深度卷积神经网络在超声结节诊断中也得到广泛应用[19]。HU等[20]和MA等[21]率先采用卷积神经网络进行乳腺肿块分割和甲状腺结节分割。为了进一步提高CNN的分割性能,研究人员探索了对模型架构的修改。XIE等[22]改进了Mask R-CNN[23],提高了其分割性能。BYRA等[24]在模型架构中引入了注意力机制。YING等[25]提出了一种级联方法,迭代训练多个CNN模型,每个模型都利用前一个模型的输出,逐步提高分割结果的准确性。但这些方法由于其感受野有限并且具有固有的归纳偏差,限制了CNN在图像中构建全局上下文和长程依赖关系的能力,从而影响了它们在图像分割中的性能。

Transformer类方法完全基于自注意力机制,能对长程依赖关系进行较好的建模,最早应用于自然语言处理[26],近年来也应用到了图像分割,例如MISSFormer[27]、DSTransUNet[28]和TransCeption[29]等。但这些基于Transformer的方法缺乏局部特征,并且在进行训练时所需的数据量通常比CNN更大,而临床医学往往很难满足如此大的数据需求量。

此外,现有的大模型及CNN分割技术在处理甲状腺这类低对比度超声图像时,也仍然存在精度不足的问题。本研究利用小波变换的多分辨率分析特征,将深度学习方法与小波变换这一传统纹理分析方法进行有机结合,能充分利用超声图像的空间域和频域特征,克服传统模型在处理超声甲状腺结节时的局限性,从而帮助医师更准确地进行辅助诊断。通过专门的训练和数据集优化,本研究方法能够更好地捕捉到甲状腺结节的细微特征。与通用的医学分割大模型相比,该方法在分割精度和对噪声的适应性上具有明显优势。

本研究方法在超声甲状腺结节分割任务中表现良好,但仍存在改进空间。当前的数据集可能尚未涵

盖所有可能的变化情况,比如在临床诊断中,部分患者可能患有慢性炎症或甲状腺肿,导致甲状腺背景不均匀,这可能干扰结节的分割,限制模型的泛化能力。未来的工作可以通过增加数据集的多样性,以及采用预训练和微调的方法,进一步提升分割精度和泛化性能。本研究提出的融合小波变换与CNN-Transformer的分割模型为甲状腺结节的自动分割提供了有效的解决方案,并为后续研究和实际应用奠定了基础。

利益冲突 本文所有作者声明不存在利益冲突

参考文献:

[1] HAUGEN B R, ALEXANDER E K, BIBLE 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] DURANTE C, GRANI G, LAMARTINA L, *et al*. The diagnosis and management of thyroid nodules: a review[J]. JAMA, 2018, 319(9): 914-924. DOI: 10.1001/jama.2018.0898.

[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[C]//IEEE Transactions on Information Technology in Biomedicine. IEEE, 2009: 519-527. DOI: 10.1109/TITB.2008.2007192.

[4] JARDIM S V B, FIGUEIREDO M A T. Automatic contour estimation in fetal ultrasound images[C]//Proceedings 2003 International Conference on Image Processing. IEEE, 2003. DOI: 10.1109/ICIP.2003.1246869.

[5] WU S B, ZHU Q S, XIE Y Q. Evaluation of various speckle reduction filters on medical ultrasound images[C]//2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2013: 1148-1151. DOI: 10.1109/EMBC.2013.6609709.

[6] TUCERYAN M, JAIN A K. Texture analysis in handbook of pattern recognition and computer vision[J]. Singapore: World Scientific Publishing, 1993: 235-276. DOI: 10.1142/9789814343138_0010.

[7] PEEPLES J, XU W H, ZARE A. Histogram layers for texture analysis[J]. IEEE Trans Artif Intell, 2022, 3(4): 541-552. DOI: 10.1109/TAI.2021.3135804.

[8] RODRIGUEZ M X B, GRUSON A, POLANIA L F, *et al*. Deep adaptive wavelet network[C]//Workshop on Applications of Computer Vision. IEEE, 2020. DOI: 10.1109/WACV45572.2020.9093580.

[9] 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 Trans Pattern Anal Mach Intell, 2018, 40(4):834-848. DOI: 10.1109/TPAMI.2017.2699184.

[10] DIAKOGIANNIS F I, WALDNER F, CACCETTA P, *et al*. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data[J]. ISPRS J Photogramm Remote Sens, 2020, 162: 94-114. DOI: 10.1016/j.isprsjprs.2020.01.013.

[11] 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. DOI: 10.1007/978-3-319-24574-4_28.

[12] OKTAY O, SCHLEMPER J, FOLGOC L L, *et al*. Attention U-Net: learning where to look for the pancreas[EB/OL]. (2018-04-11). https://arxiv.org/abs/1804.03999.

[13] PENG Y P, SONKA M, CHEN D Z. U-Net v2: rethinking the skip connections of U-Net for medical image segmentation[EB/OL]. (2023-11-29). https://arxiv.org/abs/2311.17791v2.

[14] CHEN L C, PAPANDREOU G, SCHROFF F, *et al*. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-06-17). https://arxiv.org/abs/1706.05587.

[15] CHENG J, YE J, DENG Z, *et al*. SAM-Med2D[EB/OL]. (2023-08-30). https://arxiv.org/abs/2308.16184.

[16] 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. DOI: 10.1109/3DV.2016.79.

[17] 向永嘉, 吴毅, 张小勤, 等. 基于DenseUnet模型的核磁共振图像下肛提肌的自动分割[J]. 第三军医大学学报, 2021, 43(18): 1720-1728. DOI: 10.16016/j.1000-5404.202102089.
XIANG Y J, WU Y, ZHANG X Q, *et al*. Automatic segmentation of levator ani muscle in MRI images based on DenseUnet model[J]. J Third Mil Med Univ, 2021, 43(18): 1720-1728. DOI: 10.16016/j.1000-5404.202102089.

[18] 吴哲, 许杉杉, 崔慧林, 等. 基于Vgg16-Unet模型的MRI图像下前列腺分区智能分割研究[J]. 陆军军医大学学报, 2023, 45(13): 1441-1449. DOI: 10.16016/j.2097-0927.202211045.
WU Z, XU S S, CUI H L, *et al*. Intelligent segmentation of prostate partitioning in MRI images based on Vgg16-Unet modeling[J]. J Army Med Univ, 2023, 45(13): 1441-

1449. DOI:10.16016/j.2097-0927.202211045.

[19] AVOLA D, CINQUE L, FAGIOLI A, *et al*. Multimodal feature fusion and knowledge-driven learning *via* experts consult for thyroid nodule classification[J]. IEEE Trans Circuits Syst Video Technol, 2022, 32(5): 2527-2534. DOI: 10.1109/TCSVT.2021.3074414.

[20] HU Y Z, GUO Y, WANG Y Y, *et al*. Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model[J]. Med Phys, 2019, 46(1): 215-228. DOI: 10.1002/mp.13268.

[21] MA J L, WU F, JIANG T A, *et al*. Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks[J]. Int J Comput Assist Radiol Surg, 2017, 12(11): 1895-1910. DOI: 10.1007/s11548-017-1649-7.

[22] XIE X Z, SHI F Q, NIU J W, *et al*. Breast ultrasound image classification and segmentation using convolutional neural networks [C]//Advances in Multimedia Information Processing——PCM 2018. Cham: Springer International Publishing, 2018: 200-211. DOI: 10.1007/978-3-030-00764-5_19.

[23] HE K M, GKIOXARI G, DOLLAR P, *et al*. Mask R-CNN [J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(2): 386-397. DOI: 10.1109/tpami.2018.2844175.

[24] BYRA M, JAROSIK P, SZUBERT A, *et al*. Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network[J]. Biomed Signal Process Control, 2020, 61: 102027. DOI: 10.1016/j.bspc.2020.102027.

[25] YING X, YU Z H, YU R G, *et al*. Thyroid nodule segmentation in ultrasound images based on cascaded convolutional neural network [C]//Neural Information Processing. Cham: Springer International Publishing, 2018: 373-384. DOI: 10.1007/978-3-030-04224-0_32.

[26] VASWANI A, SHAZEER N M, PARMAR N, *et al*. Attention is all you need[EB/OL]. (2023-08-02). https://arxiv.org/abs/1706.03762.

[27] HUANG X H, DENG Z F, LI D D, *et al*. MISSFormer: an effective medical image segmentation transformer[EB/OL]. (2021-09-15). https://arxiv.org/abs/2109.07162v2.

[28] LIN A L, CHEN B Z, XU J Y, *et al*. DS-TransUNet: dual swin transformer U-Net for medical image segmentation[J]. IEEE Trans Instrum Meas, 2022, 71: 4005615. DOI: 10.1109/TIM.2022.3178991.

[29] AZAD R, JIA Y W, AGHDAM E K, *et al*. Enhancing medical image segmentation with TransCeption: a multi-scale feature fusion approach [EB/OL]. (2023-06-25). https://arxiv.org/abs/2301.10847v1.

(**收稿**:2024-09-26;**修回**:2024-11-02)

(编辑 张 维)