

基于特征增强的多分支U-Net肺结节分割

温帆, 杨萍, 张鑫, 田吉, 王金华
北京联合大学智慧城市学院, 北京 100101

【摘要】针对肺结节尺度差异大、边界纹理不清晰、背景干扰严重导致的肺结节分割不精确的问题,以3D U-Net为基础,引入Transformer结构,设计一种基于特征增强的多分支U-Net肺结节分割算法。Transformer从全局角度提取肺结节及周围组织的结构特征,浅层3D U-Net提取图像纹理特征;利用上述结构特征及纹理特征进行特征增强;多尺度残差块和3D坐标注意力对3D U-Net进行改进,用于提取特征增强后的肺结节多尺度信息,并在3D U-Net解码器基础上,对深层语义信息进行复用,最终实现肺结节分割。在LIDC-IDRI数据集上对该模型进行验证,精确度、敏感度、Dice相似性系数分别达到90.04%、86.64%、88.80%,综合分割性能优于其他算法。

【关键词】肺结节;3D U-Net;Transformer;多尺度残差块;坐标注意力

【中图分类号】R318;TP391.41

【文献标志码】A

【文章编号】1005-202X(2023)11-1343-07

Pulmonary nodule segmentation using multi-branch U-Net based on feature enhancement

WEN Fan, YANG Ping, ZHANG Xin, TIAN Ji, WANG Jinhua
Smart City College, Beijing Union University, Beijing 100101, China

Abstract: To address the problem of the inaccurate segmentation of pulmonary nodules caused by large scale differences, unclear boundary texture and serious background interference, a multi-branch U-Net based on feature enhancement is designed for pulmonary nodules segmentation. The method uses Transformer to extract structural features of pulmonary nodules and surrounding tissues from a global perspective, and shallow 3D U-Net to extract the texture features. The extracted both structural and texture features are used for feature enhancement. In addition, a multi-scale residual block and 3D coordinate attention module are designed to modify 3D U-Net for obtaining multi-scale information of pulmonary nodules with enhanced features. Based on 3D U-Net decoder, the deep semantic information is reused for accomplishing the segmentation of pulmonary nodules. The verification on LIDC-IDRI dataset shows that the proposed model has a precision, sensitivity and Dice similarity coefficient of 90.04%, 86.64% and 88.80%, respectively, exhibiting superior comprehensive segmentation performance as compared with other algorithms.

Keywords: pulmonary nodule; 3D U-Net; Transformer; multi-scale residual block; coordinate attention

前言

肺癌是全球死亡率最高的恶性肿瘤,其早期诊断对于制定最佳治疗策略并提高预后效果至关重要^[1]。根据世界卫生组织/国际癌症研究署(WHO/IARC)发布的《2020全球癌症报告》,在世界范围内,肺癌的发病率在所有癌症中排名第二,死亡率排名

第一^[2]。肺癌的早期诊断对提高患者的生存率有很大帮助。肺癌早期在CT图像上表现为肺结节,医生通过审阅CT图像中肺结节的特征来判断肺结节的良恶性,大量的阅片工作不仅耗时,还会加重医生的阅片负担,将肺结节从肺实质中准确分割出来可以减轻医生的负担,帮助医生进行病变分析,提高诊断效率和准确率。

深度学习在近年来发展迅速,Ronneberger等^[3]提出U-Net用于医学图像分割,该网络是一个典型的编解码结构,编码器部分通过卷积与池化操作提取图像不同层次的特征,解码器部分通过上采样与卷积及与编码部分的跳跃连接融合各层次特征得到最后的分割结果。U-Net在医学图像分割领域有不俗的表现,引起研究者的关注并进行多方改进。Yu

【收稿日期】2023-10-26

【基金项目】国家自然科学基金(62172045,62272049)

【作者简介】温帆,硕士研究生,研究方向:医学图像处理,E-mail: 905525170@qq.com

【通信作者】杨萍,博士,副教授,研究方向:信号与信息处理、医学图像处理,E-mail: xxyangping@buu.edu.cn

等^[4]利用残差结构捕获图像中更丰富的特征;Amara等^[5]提出一个先上采样后进行下采样的U型结构,与原U-Net组成了一个O型结构,旨在通过提高图像分辨率对目标实现精准分割;Wang等^[6]为更好地结合纹理特征与语义特征,使用双向长短期记忆网络代替U-Net中简单的跳跃连接;Keetha等^[7]则利用双向特征金字塔^[8]实现不同层级的特征融合;Li等^[9]将U-Net不同层级的特征融入到每一次上采样的结果中,并对编码器最后的输出进行多尺度处理,然后将多尺度处理后的结果送入解码器实现图像尺寸的复原。为更好地提取目标特征,Liu等^[10]提出金字塔空洞卷积,一方面扩大感受野,另一方面捕捉肺结节的多尺度特征,同时在一定程度上解决空洞卷积在反向传播时的梯度消失问题;文献[11-13]将形变卷积引入U-Net,提高了对目标轮廓特征的提取能力;Tang等^[14]通过对图像的频谱进行裁剪和填充后转换到空间域分别实现池化和上采样操作,并提出双核傅里叶卷积,即将图像的高低频分离后再转换到空间域进行卷积,以此利用肺结节的频域特征帮助分割肺结节。卷积神经网络的感受野受限于卷积核的大小,无法在浅层捕捉图像的全局特征,随着网络加深虽然可以扩大感受野,但此时图像分辨率已大为降低,难以保证全局特征的精度,导致对肺结节周围相似组织的误判,影响分割精度。

Transformer结构具有长程依赖性,在捕捉全局信息方面具有优势^[15]。Dosovitskiy等^[16]将Transformer应用在计算机视觉中,提出视觉Transformer模型。Wang等^[17]将Transformer模块加入到U-Net模型的跳跃连接部分,将U-Net各层的纹理信息通过Transformer模块进行全局特征提取。Chen等^[18]在U-Net编码器中使用卷积与Transformer模块串联的特征提取路径,首先使用卷积提取图像的纹理特征,然后使用Transformer提取图像的全局结构特征,但在使用Transformer前已丢失部分全局信息。王振华等^[19]利用Transformer结构设计一个多尺度通道交叉融合模块,增强U-Net不同层级特征图与全局信息之间的联系。Hatamizadeh等^[20]使用Transformer构建编码器,将不同层的图像序列经过重塑、上采样为不同的分辨率,获得不同尺度的特征图,通过跳跃连接与解码器相连,得到最终的分割结果,并将其应用于3DCT图像的分割任务。Transformer结构始终保持全局注意,有效地解决卷积在浅层无法捕提高精度肺结节图像全局特征的问题,但是由于使用Transformer结构的图像在输入时需要映射为序列,一定程度上破坏了图像的纹理特征。

为解决上述问题,本研究提出一种基于特征增强的多分支U-Net:通过Transformer充分获取图像的全局信息,捕捉肺结节及周边组织的结构特征;使用浅层卷积网络获取肺结节的纹理信息,捕捉肺结节的边界轮廓和内部纹理特征;利用上述结构特征和纹理特征对图像进行增强;并设计多尺度残差块(Muti-Scale Residual Block, MSResBlock)和3D坐标注意力(3D Coordinate Attention, 3D CoordAtt),捕捉肺结节不同尺度的特征,并对深层语义进行单独解码,增强网络对肺结节整体结构的把握,提高肺结节分割精度。

1 方法

1.1 模型总体结构

模型由全局结构特征提取模块、局部纹理特征提取模块、分割模块组成,采用双特征强化策略。利用Transformer结构的长程优势,构建图像的结构特征提取模块,充分提取肺结节的整体结构特征;利用浅层3D U-Net捕获局部纹理特征的优势,构建纹理提取模块,捕捉肺结节的纹理特征;并将提取的局部纹理特征和全局结构特征叠加到原始图像进行特征增强;特征增强后的图像导入分割模块,分割模块是融合了MSResBlock、3D CoordAtt和深层语义复用的改进3D U-Net,通过分割模块进行多尺度信息的充分提取和利用,得到最终的分割结果。模型整体结构如图1所示。

1.2 全局结构特征提取模块

基于Transformer的全局结构特征提取模块整体为编解码结构,由12个Transformer模块串联构成的编码器和一个逐步复原分辨率的解码器构成。在编码器部分,对于一个单通道 $32 \times 32 \times 32$ 大小的输入图像切块后展平成为一系列向量并进行线性映射,接着进行位置信息嵌入,在一定程度上弥补切块时带来的位置信息丢失的问题,然后送入到串联的Transformer模块构成的编码器中进行结构特征提取。每个Transformer模块由两个层归一化(Layer Normalization, LN)、一个多头注意力(Multi-Head Attention, MHA)、一个多层感知机(Multi-Layer Perceptron, MLP)构成;LN对数据进行归一化处理,有利于模型的训练;MHA通过计算图像中各像素点之间的关联性获得权重图,将权重图与向量化图像相加后进行LN处理,然后送入MLP提取所需特征,将MLP的输出与加权后的向量化图像相加,得到最终的输出。Transformer模块结构如图2所示。

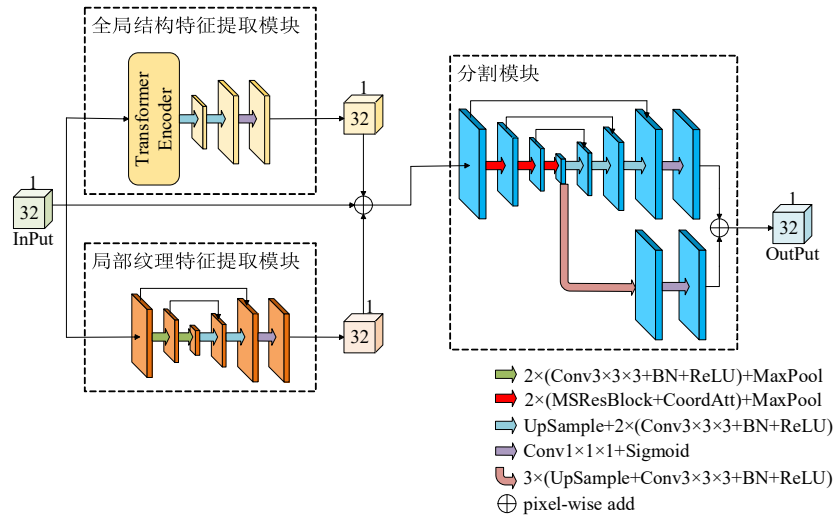


图 1 模型总体结构

Figure 1 Overall structure of the proposed model

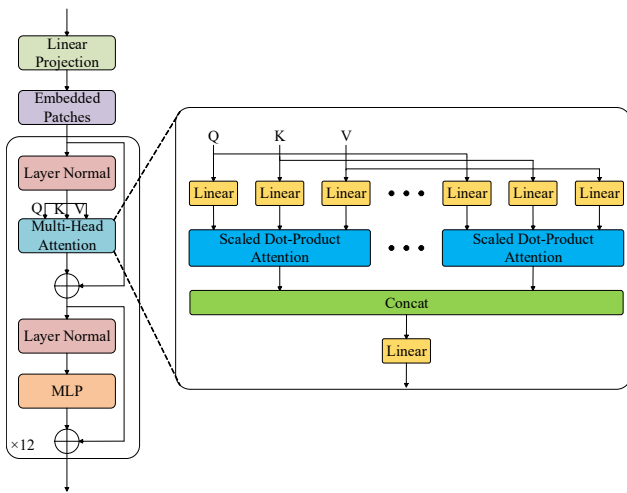


图 2 Transformer 模块

Figure 2 Transformer module

MHA 的 Q 、 K 、 V 输入均为经 LN 处理的向量化图像, 在进行自注意力前, 先将 Q 、 K 、 V 通过不同的线性运算映射到不同的线性空间 QW_i^Q 、 KW_i^K 、 VW_i^V , 以此在不同的子空间中学习相关信息, 接着在每个线性空间中进行自注意力计算得到 $head_i$, 为防止计算时内积过大, 在反向传播时易导致梯度消失问题, 对 QW_i^Q 、 KW_i^K 的内积进行尺度变换。 $head_i$ 通过捕获像素点之间的关联性, 为每个像素点根据重要性提供不同的权重值, 将不同线性空间计算出的权重图进行拼接后通过线性运算进行融合, 得到最终的权重图。整个过程可表示如下:

$$MHA(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (1)$$

$$\text{head}_i = \text{softmax}\left(\frac{QW_i^Q(KW_i^K)^T}{\sqrt{d_k}}\right)VW_i^V \quad (2)$$

其中, h 表示总线性空间的个数, Q 、 K 、 V 代表 3 个向

量化的图像, W_i^Q 、 W_i^K 、 W_i^V 分别代表 Q 、 K 、 V 在第 i 个线性空间的映射矩阵, d_k 代表向量的维度, $head_i$ 表示第 i 个线性空间的权重矩阵, W^O 表示融合输出矩阵。

在解码器部分, 由上采样和卷积交错进行, Transformer 编码器提取到的全局特征通过解码逐步还原图像中的结构信息, 最后通过一个核大小为 $1 \times 1 \times 1$ 的卷积构成的输出层得到结构特征掩膜。考虑分割任务的本质特征以及使用 Transformer 结构与 3D 卷积会带来大量参数冗余, 易导致过拟合问题, 本模型中所有卷积均使用空间可分离卷积。空间可分离卷积将普通的 $N \times N \times N$ 大小的 3D 卷积核分解为 $N \times 1 \times 1$ 、 $1 \times N \times 1$ 、 $1 \times 1 \times N$ 的一维卷积, 使参数量从 N^3 降为 $3N$, 同时不会影响卷积核最终的感受野。空间可分离卷积的具体结构如图 3 所示。

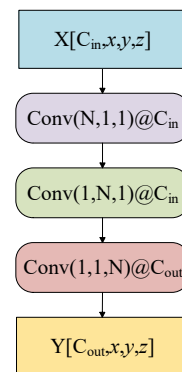


图 3 空间可分离卷积

Figure 3 Spatially separable convolution

1.3 局部纹理特征提取模块

利用浅层网络提取细节特征的优势, 局部纹理特征提取模块使用 3 层的 3D U-Net 结构, 通过卷积和

下采样操作,提取肺结节图像的细节纹理特征,如结节外延的条索状组织和结节的边界,并在上采样过程中,通过跳跃连接,不断融合编码器提取的特征信息,逐步构建与输入图像大小相同的纹理特征掩膜。局部纹理特征提取模块结构如图4所示。

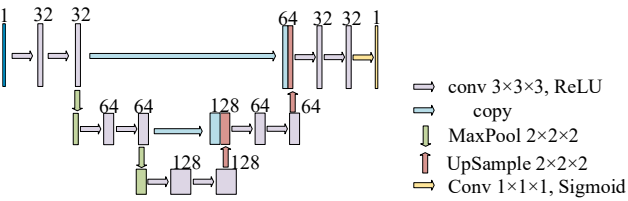


图4 3D U-Net模型
Figure 4 3D U-Net model

1.4 分割模块

1.4.1 改进的3D U-Net模块 分割模块以经典3D U-Net结构为基础,将3D U-Net各层中两次连续的卷积替换为 MSResBlock,并跟随一个 CoordAtt 模块。MSResBlock通过加深网络深度,捕捉肺结节不同尺度的特征,并在网络深层捕捉强化后的全局信息;3D CoordAtt 模块可以为肺结节所在位置赋予更大的权重,进而抑制背景的噪声干扰。3D U-Net在解码器部分使用跳跃连接,跳跃连接使得深层的语义信息在每次上采样后都可以补充浅层的细节纹理特征,但也因此稀释了深层语义信息在生成分割结果时的贡献,为解决这一问题,使用卷积和上采样串联构建并行解码路径,对深层语义信息进行复用,使网络充分利用编码器提取的深层语义特征,提高分割性能。

1.4.2 MSResBlock 本研究设计的MSResBlock主要由3条路径构成,一条路径由4个核大小为3的普通卷积构成,感受野为3;一条路径由4个核大小为3、空洞率为2的空洞卷积构成,感受野为5;最后一条路径由核大小为1的普通卷积核构成。核大小为3的普通卷积路径和核大小为3的空洞卷积路径分别从不同尺度捕捉肺结节的纹理特征,两条路径卷积得到的特征图通过逐像素相加进行融合,感受野小的卷积为感受野大的卷积补充细节,感受野大的卷积为感受野小的卷积补充视野外的信息,核大小为1的卷积保持卷积前后通道一致,补充卷积过程中丢失的原始图像信息,最后将3条路径提取的特征图进行逐像素相加,得到最终输出。使用空洞卷积可以在不增加模型参数的同时扩大图像的局部感受野。MSResBlock结构如图5所示。

1.4.3 3D CoordAtt 模块 在Hou等^[21]的启发下,本研究提出3D CoordAtt,相较Woo等^[22]提出的卷积块注

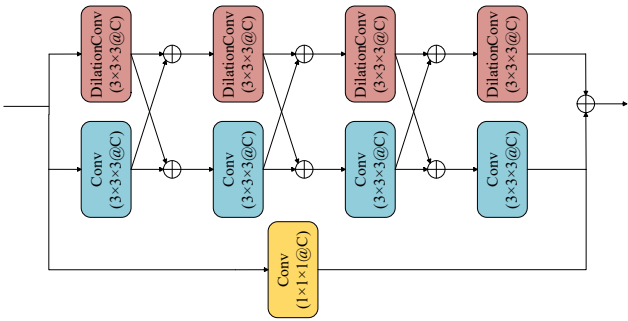


图5 MSResBlock
Figure 5 Multi-scale residual block

意力机制模块(Convolutional Block Attention Module, CBAM),3D CoordAtt模块的参数量更少,在数据匮乏的情况下更易于训练。图6展示其具体的实现方式。GAP与GMP分别代表全局平均池化与全局最大池化操作。首先对输入图像分别沿X、Y、Z轴进行全局平均池化与全局最大池化操作,在保留通道信息的同时捕捉空间位置信息,每个轴向的平均池化结果与最大池化结果融合后得到轴向信息,紧接着捕捉X、Y、Z轴之间的联系,最后分别生成3个轴向的注意力权重,并与输入图像进行逐像素相乘,得到最后的输出。通过轴向的全局最大池化可更好地定位肺结节的位置,轴向的全局平均池化帮助排除与肺结节灰度相近的异常点,从而给每个像素点赋予恰如其分的权重,加强模型对肺结节特征的关注。

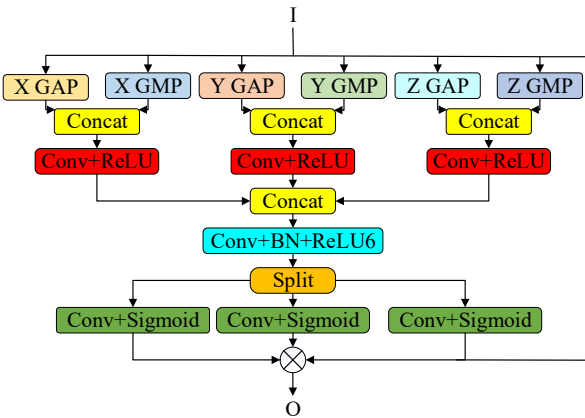


图6 3D CoordAtt
Figure 6 3D coordinate attention

1.5 损失函数

损失函数采用二值交叉熵(Binary Cross Entropy, BCE)损失函数,定义如下:

$$L_{BCE} = -\frac{1}{N} \sum_{i=1}^N (y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)) \quad (3)$$

其中, N 代表图像中像素点总数, y_i 代表像素点真实标签, p_i 代表模型预测出的像素点标签。BCE 损失函数在模型训练初始阶段可以计算出较大的损失值, 使模型参数更新幅度更大, 从而加快模型参数收敛。交叉熵损失计算依赖于所有像素点, 在模型训练后期不会因局部损失较大产生波动, 可以保证模型训练过程的稳定。

2 实验

2.1 数据集及数据预处理

实验数据采用 LIDC-IDRI 公开肺结节数据集^[23], 该数据集包含从 7 个学术中心和 8 家医学影像公司搜集到的 1018 个病例的肺部 CT 图像, 图像是一系列大小为 512×512 的切片, 并以 DICM 格式进行存储。每个病例都由 4 位专家进行两阶段的标注, 以确保标注的准确性, 所有的标注均以 XML 格式进行存储。

数据预处理包含原始图像处理和图像标签处理, 本研究采用 Pydicc 工具^[24]进行数据预处理。在原始图像处理部分, 剔除切片厚度大于 3 mm 的数据, 并剔除只有一到两名专家标注的肺结节, 最终保留 837 个结节。所有图像的像素值都转换为 0~255 的灰度值。从每个病例对应的 XML 文件中提取图像标签, 肺结节部分的像素点的像素值为 1, 非肺结节部分的像素点的像素值为 0, 图像标签同样保存为 .npy 格式。由于 3D 数据在训练时对显存要求很大, 所以以肺结节为中心裁剪出 32×32×32 大小的 3D 图像, 以降低模型训练时的显存需求。

2.2 实验环境

实验使用 Pytorch 深度学习框架进行实现, cuda 版本为 11.6, 模型训练在英伟达 RTX3090Ti 上进行, 24 G 显存, 优化器选择 Adam, 学习率设置为 0.001, BatchSize 为 8, 共训练 300 轮。

2.3 评价指标

模型的性能评估采用医学图像分割领域的公认指标: 精确度 (Precision, Prec)、敏感度 (Sensitivity, Sen)、Dice 相似性系数 (Dice Similarity Coefficient, DSC), 各指标定义如下:

$$\text{Prec} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4)$$

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (5)$$

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

其中, TP 代表真正例, 即预测为正例且真实标签为正例的像素点; FN 代表假反例, 即预测为反例且真实标签为正例的像素点; FP 代表假正例, 即预测为正例且

真实标签为反例的像素点。当假正例越少时, 精确度越高; 当假反例越少时, 敏感性越高。如果将有病变的区域预测为无病变区域将会导致严重的后果, 所以敏感性这一指标尤为重要; DSC 是用于评价轮廓拟合程度的指标, 当预测的病变区域轮廓与真实标签标记的病变区域轮廓越接近, 这一指标越高。

3 实验结果及分析

3.1 消融实验

首先验证在没有特征增强的情况下, MSResBlock、3D CoordAtt、深层语义复用对肺结节分割性能的影响。MSResBlock 提高了编码器对于肺结节特征的提取能力, 并通过残差连接补充卷积时丢失的原始图像信息; 3D CoordAtt 可以给予肺结节更高的权重, 使模型关注的重点集中于肺结节, 从而使编码器最终的编码结果包含更多的肺结节敏感信息, 抑制背景噪声。MSResBlock 与 3D CoordAtt 从不同角度提升分割性能, 所以将二者叠加使用时分割模块性能有进一步提升; 对深层语义的复用可以使模型在一定程度上把握肺结节整体结构, 但没能解决卷积在全局特征提取上的不足。接着验证使用全局结构特征和局部纹理特征进行特征强化的有效性, 通过使用分割模块直接分割、仅加入全局结构特征提取模块、仅加入局部纹理特征提取模块、同时加入全局结构特征提取模块和局部纹理特征提取模块分别进行实验。实验结果表明在改进分割模块的基础上, 引入全局结构特征提取模块或局部纹理特征提取模块进行特征增强的分割性能优于直接分割的分割性能。同时引入两种特征增强方式后, 模型性能有进一步提升, 证明特征增强方式的有效性。详见表 1。

model_1 至 model_7 的对比实验验证了 MSResBlock、3D CoordAtt、深层语义复利用有助于提高分割性能, model_8 至 model_10 是在使用 MSResBlock、3D CoordAtt、深层语义复利用改进的分割模块的基础上, 使用全局结构特征、局部纹理特征进行特征强化后的分割效果。实验结果证明上述特征强化策略的有效性。

3.2 不同模型的分割效果对比

为验证对 3D U-Net 的改进效果, 将本模型结果与近年来其他学者的研究结果进行对比。本文方法的精确度、敏感度、DSC 分别达到 90.04%、86.64%、88.80%, 相较基准的 3D U-Net 均有明显提升, 不仅如此, 与近年来基于 U-Net 改进的方法^[9, 25]和基于生成对抗网络的方法^[26-27]相比, 本文方法的精确度和 DSC 依然具有优势。详见表 2。

表 1 消融实验(%)
Table 1 Ablation experiment (%)

| 模型 | 分割模块 | | | 特征增强 | | 精确度 | 敏感度 | DSC |
|----------|--------|-------------|--------|--------|--------|-------|-------|-------|
| | 多尺度残差块 | 3D CoordAtt | 深层语义复用 | 结构提取模块 | 纹理提取模块 | | | |
| 3D U-Net | - | - | - | - | - | 84.17 | 81.29 | 81.10 |
| model_1 | √ | - | - | - | - | 86.03 | 83.33 | 84.81 |
| model_2 | - | √ | - | - | - | 84.40 | 85.62 | 84.25 |
| model_3 | - | - | √ | - | - | 87.41 | 80.00 | 83.97 |
| model_4 | √ | √ | - | - | - | 86.63 | 84.51 | 85.46 |
| model_5 | √ | - | √ | - | - | 86.37 | 83.59 | 85.06 |
| model_6 | - | √ | √ | - | - | 85.61 | 84.79 | 85.21 |
| model_7 | √ | √ | √ | - | - | 87.12 | 86.01 | 86.33 |
| model_8 | √ | √ | √ | √ | - | 87.83 | 86.57 | 87.20 |
| model_9 | √ | √ | √ | - | √ | 88.55 | 86.00 | 87.86 |
| model_10 | √ | √ | √ | √ | √ | 90.04 | 86.64 | 88.80 |

√表示模型中包含该模块,-代表未包含

表 2 不同分割方法的对比(%)
Table 2 Comparison among different segmentation methods (%)

| 模型 | 精确度 | 敏感度 | DSC |
|----------|-------|-------|-------|
| 3D U-Net | 84.17 | 81.29 | 81.10 |
| 3D U-Det | 84.85 | 80.94 | 82.54 |
| 文献[9] | 84.77 | 84.73 | 84.76 |
| 文献[25] | 85.35 | 83.81 | 84.48 |
| 文献[26] | 77.52 | 82.56 | 76.36 |
| 文献[27] | 78.98 | 88.51 | 83.16 |
| 本文 | 90.04 | 86.64 | 88.80 |

3.3 分割效果展示

图7展示了同一个结节的不同切片的分割效果,可以看出:3D U-Net在肺结节的边缘切片(1)和(8)上分割性能欠佳;本文方法在边缘切片的分割上虽然还未能与标注完全吻合,但与3D U-Net模型相比有了明显提升,不仅如此,在中心切片上的分割结果也优于3D U-Net模型。图8展示了不同肺结节中心切片的分割效果,不同的肺结节在形态、轮廓、位置上存在很大差异,一些存在清晰轮廓的结节(1)和(2),本文方法以及3D U-Net都有不错的分割效果,但在类似于图8中内部结构不清晰的结节(3)和(4)、与其他组织发生粘连的结节(5)和(6)、位于肺实质边缘的结节(7)和(8),本文方法的分割效果相较于原3D U-Net有了明显的提升。

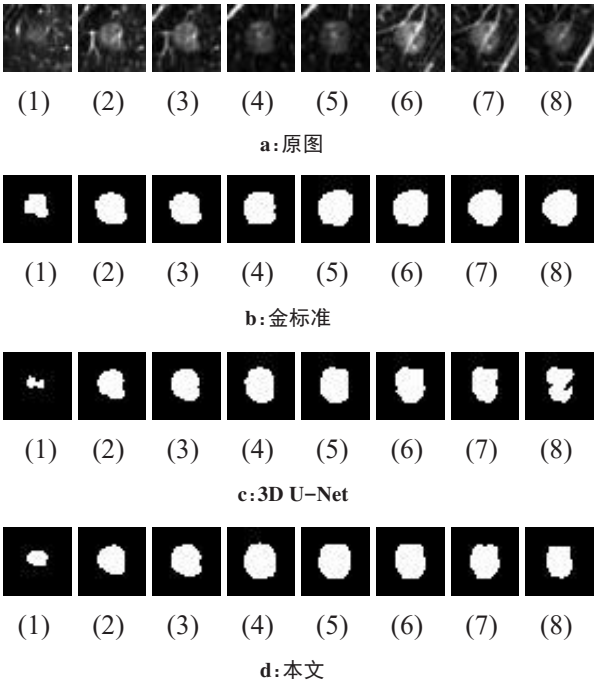


图 7 同一结节不同切片的分割结果对比
Figure 7 Comparison of segmentation results of different sections of the same nodule

4 结 语

本研究针对肺结节尺度差异大、纹理结构不清晰、背景干扰严重导致的肺结节分割不精确的问题,提出一种基于特征增强的多分支 U-Net 进行肺结节分割,通过 Transformer 充分捕捉图像中的全局信息,解决背景干扰的问题;使用卷积网络捕捉肺结节的

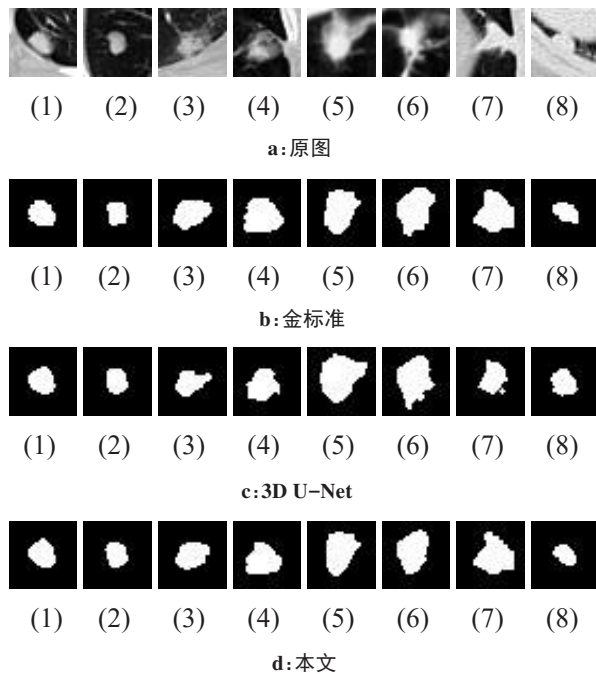


图8 不同结节中心切片分割结果对比

Figure 8 Comparison of the segmentation results of the central sections of different nodules

纹理信息,把握肺结节的内部结构;最后使用引入MSResBlock和3D CoordAtt捕捉肺结节不同尺度的特征,并对深层语义信息进行复用,在把握肺结节轮廓的基础上得到更精准的分割结果。最终在LIDC-IDRI数据集上进行测试,精确度、敏感度、DSC分别达到90.04%、86.64%、88.80%,测试结果表明本文方法相比其他方法具有一定的优势,有助于提高肺结节分割精度。

【参考文献】

- [1] 陈雅婧,徐肖攀. 机器学习在肺癌液体活检中应用的研究进展[J]. 空军军医大学学报, 2022, 43(2): 235-239.
Chen YJ, Xu XP. Research progress on application of machine learning in liquid biopsy of lung cancer[J]. Journal of Air Force Medical University, 2022, 43(2): 235-239.
- [2] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA-Cancer J Clin, 2021, 71(3): 209-249.
- [3] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]. Cham: Springer, 2015: 234-241.
- [4] Yu H, Li J, Zhang L, et al. Design of lung nodules segmentation and recognition algorithm based on deep learning [J]. BMC Bioinformatics, 2021, 22(5): 314.
- [5] Amara K, Aouf A, Kennouche H, et al. COVIR: a virtual rendering of a novel NN architecture O-Net for COVID-19 Ct-scan automatic lung lesions segmentation[J]. Comput Graph-UK, 2022, 104: 11-23.
- [6] Wang X, Li M, Yang J. Segmentation of pulmonary nodules based on BBCLstm Unet[J]. JPCS, 2021: 012037.
- [7] Keetha NV, Babu SA, Annavarapu CS. U-DET: a modified U-NET

- architecture with bidirectional feature network for lung nodule segmentation[J]. arXiv: 2003. 09293, 2020.
- [8] Tan M, Pang R, Le QV. Efficientdet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE/CVF, 2020: 10781-10790.
- [9] Li D, Yuan S, Yao G. Pulmonary nodule segmentation based on REMU-Net[J]. Physl Eng Sci Med, 2022, 45: 995-1004.
- [10] Liu W, Liu X, Li H, et al. Integrating lung parenchyma segmentation and nodule detection with deep multi-task learning[J]. IEEE J Biomed Health Inf, 2021, 25(8): 3073-3081.
- [11] Zhang M, Li X, Xu M, et al. Automated semantic segmentation of red blood cells for sickle cell disease[J]. IEEE J Biomed Health, 2020, 24(11): 3095-3102.
- [12] Hu X, Yang H. DRU-net: a novel U-net for biomedical image segmentation[J]. IET Image Processing, 2020, 14(1): 192-200.
- [13] Li H, Wang YK, Wan C, et al. MAU-Net: a retinal vessels segmentation method [C]//IEEE Engineering in Medicine and Biology Society. EMBC, 2020.
- [14] Tang X, Peng J, Zhong B, et al. Introducing frequency representation into convolution neural networks for medical image segmentation via twin-Kernel Fourier convolution[J]. Comput Meth Prog Bio, 2021, 205(1): 106110.
- [15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. arXiv: 1706. 03762, 2017.
- [16] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: transformers for image recognition at scale[J]. arXiv: 2010. 11929, 2020.
- [17] Wang H, Cao P, Wang J, et al. UCTransNet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer [J]. arXiv: 2109. 04335, 2021.
- [18] Chen J, Lu Y, Yu Q, et al. TransUNet: transformers make strong encoders for medical image segmentation[J]. arXiv: 2102. 04306, 2021.
- [19] 王振华, 刘阳星, 赵晓雨, 等. 结合上下文和注意力机制改进的视盘分割模型[J]. 中国医学物理学杂志, 2023, 40(1): 47-53.
Wang ZH, Liu YX, Zhao XY, et al. Optic disc segmentation model improved by contextual information and attention mechanism [J]. Chinese Journal of Medical Physics, 2023, 40(1): 47-53.
- [20] Hatamizadeh A, Yang D, Roth H, et al. UNETR: transformers for 3D medical image segmentation[J]. arXiv. 2103. 10504, 2021.
- [21] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE/CVF, 2021: 13708-13717.
- [22] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. 2018: 3-19.
- [23] Armato SG, 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]. Med Phys, 2011, 38(2): 915-931.
- [24] Hancock MC, Magnan JF. Lung nodule malignancy classification using only radiologist-quantified image features as inputs to statistical learning algorithms: probing the lung image database consortium dataset with two statistical learning methods[J]. J Med Imaging, 2016, 3(4): 044504.
- [25] 钟思华, 郭兴明, 郑伊能. 改进U-Net网络的肺结节分割方法[J]. 计算机工程与应用, 2020, 56(17): 203-209.
Zhong SH, Guo XM, Zheng YN. Improved U-Net network for lung nodule segmentation[J]. Computer Engineering and Applications, 2020, 56(17): 203-209.
- [26] Tyagi S, Talbar SN. CSE-GAN: a 3D conditional generative adversarial network with concurrent squeeze-and-excitation blocks for lung nodule segmentation[J]. Comput Biol Med, 2022, 147: 105781.
- [27] Wu Z, Zhou Q, Wang F. Coarse-to-fine lung nodule segmentation in CT images with image enhancement and dual-branch network[J]. IEEE Access, 2021(9): 9.

(编辑:谭斯允)