

基于PM-UNet的腹部肝脏分割网络

王智信

北方工业大学 信息学院 北京 100000

摘要: 传统卷积神经网络在捕捉全局上下文信息和多尺度特征提取上存在局限,容易忽视重要特征。为此,提出一种结合并行注意力和多尺度卷积交叉融合的网络。并行注意力机制有效捕捉全局信息和通道关系,抑制无关特征;多尺度卷积模块通过不同感受野融合特征,增强表达能力。实验结果表明,该方法在3DIRCADb数据集上显著优于现有分割网络,分割结果更接近真实值。

关键词: 深度学习; 肝脏分割

引言

肝硬化和肝癌是全球范围内前20位的死亡原因之一^[1]。在此背景下,自动分割腹部CT图像中的肝脏区域对诊断和治疗肝脏疾病至关重要。

最初为图像分类任务设计的全卷积网络^[2]启发了Ronneberger等人提出的UNet模型^[3],该模型专为医学图像分割而生。UNet通过U形结构有效结合上下文信息,提升分割精度,表现优异,训练快速且数据需求少,广泛应用于该领域。Zhou等人提出了一种名为UNet++^[4]通过嵌套的密集跳跃连接缩小了编码器与解码器的语义差距。Oktay等人基于CNN提出了Attention U-Net^[5]引入空间和通道注意力门,增强了跳跃连接中对不同形状目标的处理能力。除此之外,还有AGRes-UNet^[6]、R2U-Net^[7]等UNet变体。这些方法改进了卷积神经网络,但卷积局限于局部区域,难以捕捉全局信息和长距离依赖,因此在处理广泛上下文关联时表现不足。TransUnet^[8]结合CNN提取多层次特征,并通过Transformer建模长距离依赖。然而,CNN传递给Transformer的单一尺度特征限制了局部信息的有效交互,尽管能建模长距离依赖,但局部特征未能充分融合。

针对上述问题,提出一种改进的UNet网络,命名为PM-UNet(Parallel Attention and Multi-scale Convolution Fusion UNet)。结合并行注意力机制(parallel attention module, PAM)和多尺度卷积融合模块(multi scale convolution cross fusion block, MSCCFB),提升肝脏分割的准确性与鲁棒性,优化细节捕捉和分割性能。

1 方法

PM-UNet结构如图1所示,保留了传统UNet的主干,同时改进了编码器和解码器。从第三层起,引入PAM和

MSCCFB,以提升特征提取效果和网络性能,更好捕捉关键特征。

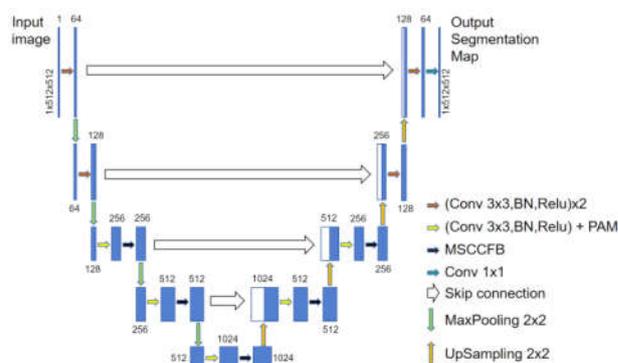


图1 PM-UNet网络架构

1.1 PAM

PAM由通道注意力模块(channel attention module, CAM)和空间注意力模块(spatial attention module, SAM)组成。

CAM通过为各通道分配权重,突出关键特征,提升特征表示能力。CAM先通过全局平均池化和最大池化提取全局信息,再通过全连接层生成通道权重并相加,最后使用Sigmoid函数规范化权重,应用于特征图,增强模型对重要信息的关注,提升性能。

SAM通过加权特征图的空间维度,聚焦重要区域。SAM先计算特征图在通道维度的平均值和最大值,生成两个特征图并拼接,经过卷积层处理后,使用Sigmoid函数得到空间注意力权重,帮助模型动态关注关键区域,提升性能。

PAM通过结合CAM和SAM增强特征图的关键特征关注,如图2所示。首先,对输入特征图应用通道和空间注意力,得到一维和二维注意力图。然后,将两者与原始特征图逐元素相乘生成加权特征图,并通过残差连接与

作者简介: 王智信(2004-),男,北京,本科生,主要研究方向为医疗影像分割。

原始特征图相加，防止梯度消失，确保信息有效传递，提升性能和表达能力。

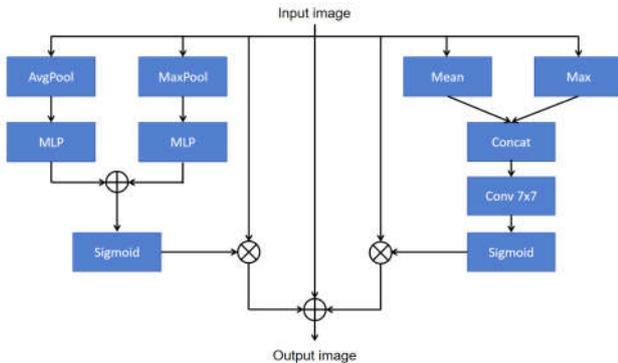


图2 PAM示意图

1.2 MSCCFB

为解决单一尺寸卷积核感受野固定导致的特征提取困难，提出了MSCCFB模块，利用多尺度卷积更有效捕捉上下文信息（如图3所示）。MSCCFB通过1x1、3x3、5x5三个卷积核提取局部特征，并将三个分支特征拼接后，通过1x1卷积降维整合，增强特征多样性。最终，模块合并三个分支输出，提升特征提取能力和网络对不同尺度信息的适应性。

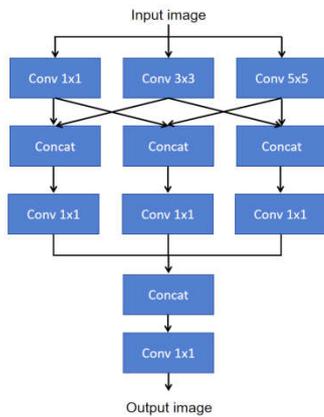


图3 MSCCFB示意图

2 实验与分析

2.1 数据集

为评估PM-UNet方法在腹部CT图像中的肝脏分割性能，我们在3DIRCADb数据集上进行了5折交叉验证，并与现有方法进行对比。数据集包含20例腹部CT图像，每张图像分辨率为512×512像素，层间分辨率为1.25至4.00毫米，层内分辨率为0.56至0.87毫米。每次验证使用16个训练案例和4个测试案例，评估算法的性能。

2.2 评价指标

通过Dice相似系数（DSC）、Jaccard系数、Precision和Recall等量化指标全面评估了不同分割方法的性能。

DSC衡量模型预测与真实标签的相似度，取值范围为0到1，值越大表示重叠度越高，分割效果越好。公式（1）定义了DSC，其中A为分割结果，B为真实标签。较高的DSC值表明分割一致性较好。

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

Jaccard系数衡量两个样本集的相似度，计算为交集与并集的比值（公式2）。A和B分别为预测结果和真实值，值接近1表示高重叠、性能优良，接近0则表示低重叠、性能较差。

$$Jaccard = \frac{A \cap B}{A \cup B} \quad (2)$$

精确率衡量被预测为正样本中实际为正样本的比例（公式3），其中TP为真阳性，FP为假阳性。

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

召回率表示实际正样本中被正确标记为正样本的比例（公式4），其中TP为真阳性，FN为假阴性。

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

2.3 实验结果

为评估PM-UNet在腹部肝脏分割任务中的优势，我们与UNet、UNet++、Attention U-Net和TransUNet进行了比较。所有模型在相同实验条件下测试，使用3DIRCADb测试集。结果显示，PM-UNet相较于UNet在Dice、Jaccard、Precision和Recall上分别提高了0.8%、1.32%、0.78%和0.66%。与其他模型相比，PM-UNet也有所提升，证明其在肝脏分割任务中的鲁棒性。

表1 不同模型在数据集3DIRCADb上的对比实验

Module	DSC	Jaccard	Precision	Recall
UNet	93.52	88.22	93.24	93.95
UNet++	94.28	89.24	93.86	94.57
Attention U-Net	94.22	89.21	93.81	94.55
TransUNet	93.89	88.61	93.63	94.32
PM-UNet	94.32	89.54	94.02	94.61

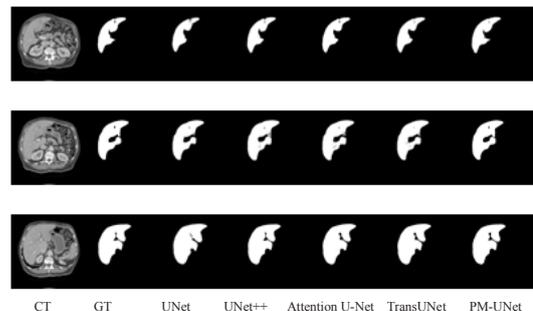


图4 不同模型分割结果图

图4展示了不同模型在3DIRCADb数据集上的分割结果。各模型能有效分割肝脏,但在形状不规则、与周围器官接近或边界模糊时,均存在欠分割或过分割现象。UNet在识别微小目标时较弱,UNet++、Attention U-Net和TransUNet分割的边缘轮廓较粗糙。相比之下,PM-UNet在分割精度和边缘清晰度上表现最佳,与真实标签更为相似。

2.4 消融实验

为验证PM-UNet各模块的有效性,进行了消融实验(见表2)。在基线模型UNet上分别引入MSCCFB和PAM。引入MSCCFB后,Dice分数、Jaccard系数、Precision和Recall分别提高了0.22%、0.08%、0.21%和0.15%;引入PAM后,分别提高了0.35%、0.37%、0.34%和0.29%。实验结果表明, MSCCFB和PAM均能提升网络分割性能。

表2 消融实验结果

Module	DSC	Jaccard	Precision	Recall
UNet	93.52	88.22	93.24	93.95
UNet+MSCCFB	93.74	88.30	93.45	94.10
UNet+PAM	93.87	88.59	93.58	94.24

结束语

PM-UNet通过PAM有效聚焦目标区域,减少干扰;MSCCFB增强了特征提取的多样性,提升了对复杂特征的理解。实验结果显示,PM-UNet在3DIRCADb数据集上优于现有方法。未来,我们将扩展数据集、改进预处理以增强对比度,提升模型泛化能力,并优化算法与模型以提高分割准确性、计算效率,减少参数与计算复杂度。

参考文献

[1] Jiang, Chun, et al. "Radiomics in the diagnosis and

treatment of hepatocellular carcinoma". *Hepatobiliary & Pancreatic Diseases International* 22(4), 346-351(2023).

[2] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition(CVPR).Boston:IEEE, 2015: 3431-3440.

[3] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention- MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.

[4] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHSH N, et al. UNet++: a nested U-Net architecture for medical image segmentation[J]. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018, 11045: 3-11.

[5] Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: Learning where to look for the pancreas. arXiv:1804.03999,2018.

[6] ZHANG J X,JIANG Z K,DONG J,et al.Attention gate resu-net for automatic MRI brain tumor segmentation[J]. *IEEE Access*,2020,8:58533-58545.

[7] Alom MZ, Yakopcic C, Hasan M, et al. Recurrent residual U-Net for medical image segmentation[J]. *J Med Imaging (Bellingham)*, 2019, 6(1): 014006.

[8] Chen J,Lu Y,Yu Q,et al.Transunet:Transformers make strong encoders for medical image segmentation[J].arXiv preprint arXiv:2102.04306, 2021.