

基于深度学习的三维肿瘤及器官分割

顾德¹, 王宁¹, 张寅斌², 刘乐³

1. 江南大学物联网工程学院, 江苏 无锡 214122; 2. 西安交通大学第二附属医院肿瘤科, 陕西 西安 710004; 3. 西安交通大学第二附属医院医学影像科, 陕西 西安 710004

【摘要】针对三维医学图像中由于肿瘤或器官的形状、尺度差异较大导致分割精度较低的问题,提出一种端到端的三维全卷积分割模型。首先,设计空洞立方集成模块在不同分辨率阶段实现多尺度集成,增强复杂边界上的识别能力;其次,引入跨阶段上下文融合模块融合浅层和深层特征,促进收敛并更准确地定位目标对象;最后,解码器对来自编码器的特征进行拼接以实现分割。在脑肿瘤分割数据集上,平均Dice相似性系数达到85.37%;在腹部器官分割数据集上,平均Dice相似性系数达到83.99%。实验结果表明所提模型在三维肿瘤和器官的分割上具有较高精度。

【关键词】肿瘤分割;器官分割;三维卷积神经网络;空洞立方集成模块;跨阶段上下文融合模块

【中图分类号】R318;TP183

【文献标志码】A

【文章编号】1005-202X(2024)09-1122-07

Three-dimensional tumor and organ segmentation based on deep learning

GU De¹, WANG Ning¹, ZHANG Yinbin², LIU Le³

1. School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; 2. Department of Oncology, the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710004, China; 3. Department of Medical Imaging, the Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710004, China

Abstract: In response to the challenge posed by the significant shape and scale variations of tumors and organs in three-dimensional medical images, which often results in low segmentation accuracy, an end-to-end three-dimensional fully convolutional segmentation model is introduced. A dilated cubic integration module is designed to achieve multi-scale integration at different resolution stages, thereby enhancing the recognition capability on complex boundaries. Subsequently, a cross-stage context fusion module is incorporated to merge shallow and deep features, thereby facilitating convergence and more precise localization of the target objects. Finally, features from the encoder are concatenated by the decoder to realize segmentation. The average Dice similarity coefficients reach 85.37% on the brain tumor segmentation dataset and 83.99% on the abdominal organ segmentation dataset. Experimental results indicate that the proposed model exhibits high accuracy in three-dimensional tumor and organ segmentation.

Keywords: tumor segmentation; organ segmentation; three-dimensional convolutional neural network; dilated cubic integration module; cross-stage context fusion module

前言

医学图像分割是医学影像学 and 深度学习交叉领域的一个热点。以计算机断层扫描(Computed Tomography, CT)和磁共振成像(Magnetic Resonance Imaging, MRI)为主的三维(Three-dimensional, 3D)医学成像模式提供丰富的解剖结构和病变信息,对

各种疾病的诊断和治疗至关重要。因此,作为计算机辅助诊断系统的一个组成部分,基于深度学习的肿瘤和器官分割有助于提高临床医生解读医学图像的效率 and 准确性^[1]。在过去的几十年里,传统的分割方法,如边缘检测^[2]、基于随机游走的联合分割^[3]和机器学习方法^[4]已被广泛使用,这些方法采用数学和统计模型,依靠边缘、颜色和纹理等特征来识别和分割图像中的结构或区域。然而,随着医学成像技术的不断进步,图像细节的复杂性和个体差异显著增加。基于手动或启发式规则的特征提取方法无法满足当前图像分割的复杂需求,对分割模型的准确性提出挑战^[5]。

【收稿日期】2024-04-02

【基金项目】江苏省自然科学基金(BK20180594, BK20231036)

【作者简介】顾德, 博士, 副教授, 研究方向: 基于图像的生物医学信息识别, E-mail: gude@jiangnan.edu.cn

最近,深度学习由于能从训练数据中自主学习层次特征而引起人们的极大关注,这些方法在医学图像分割中得到成功的应用,包括全卷积网络^[6]、级联网络^[7]、Mask RCNN^[8]、注意力机制^[9]等,与传统方法相比,这些方法取得了显著的效果。Milletari等^[10]率先将深度学习引入3D医学图像分割,利用全卷积网络学习多阶段特征和残差函数进行分割。深度神经网络存在着梯度消失和退化的问题,许多研究已经针对这些问题进行优化。Chen等^[11]开发了一种体素残差结构(VoxRes),将残差学习的优势扩展到3D网络。Wang等^[12]设计递归残差块,该块包含多个残差连接,以缓解特征退化。Ibtehaz等^[13]设计一个具有残差连接的卷积层链,其中1×1卷积核伴随残差连接,以缓解编码器-解码器特征之间的差异。Zhang等^[14]引入dense-inception模块,使网络既深又宽,而不会受限于梯度消失或冗余计算。Zhou等^[15]使用3D ShuffleNetV2作为编码器,然后引入一种具有残差块的解码器来防止退化。Wang等^[16]设计一种具有编码和解码功能的改进残差模块,提高网络的训练速度和性能。Isensee等^[17]引入一种基于U-Net的多级监督机制,并对预处理和后处理进行标准化。

多尺度特征提取一直是语义分割中的一个关键主题,尤其是在3D医学图像领域,不同个体间肿瘤和器官的形态、大小差异明显。为捕获多尺度信息,Zhou等^[18]使用3D膨胀卷积在主分支的末端形成特征金字塔,但其网络分辨率较低。Sun等^[19]在多个阶段的第一层使用膨胀卷积来提取不同特征的感受野,但使用单一的、不变的膨胀率。Zhang等^[20]使用包含高分辨率分支和并行多分辨率分支的分层设计来提取多尺度特征。Chen等^[21]在编码器中使用多线程膨胀卷积,但其最大膨胀系数为2,表明提取多尺

度特征的能力相对不足,在解码器中使用具有不同内核大小的普通卷积,大幅增加参数的数量和计算负载。Chang等^[22]将编码器设计为具有不同数量的3×3×3卷积的两条路径,并仅在编码器的最后一层使用3D膨胀卷积来提取多尺度信息。

在以上工作的启发下,本研究提出在编码过程的每一个分辨率阶段使用空洞立方集成(Dilated Cubic Integration, DCI)模块集成不同尺度大小的特征,膨胀率更丰富,多尺度提取能力更强;此外,引入跨阶段上下文融合(Cross-stage Context Fusion, CCF)模块融合不同分辨率阶段的特征,有效地融合低级和高级语义特征,提取更丰富的上下文信息,并且可以作为阶段内残差连接的补充,为长距离梯度传播提供高效的路径。

1 端到端分割方法

1.1 模型总体结构

本研究模型的总体结构见图1,由编码器和解码器组成。对于多模态图像,每个模态对应于每个通道的3D输入,对于单模态图像则被视为单通道输入。在编码器内,下采样层将每个阶段的特征采样到不同的空间维度,然后由卷积层、DCI模块和CCF模块进行处理,以进行详细的层次特征学习。具体来说,假设编码器的输入大小是子体积 $X \in R^{H \times W \times D \times C_{in}}$,每个阶段的下采样层将分辨率降低一半,同时将通道数量增加一倍。因此,对于阶段 i ,特征大小被重塑为 $\frac{H}{2^i} \times \frac{W}{2^i} \times \frac{D}{2^i} \times 2^i C_0 (i \in \{1, 2, 3, 4\})$ 。解码器使用转置卷积对瓶颈特征进行上采样,随后将它们与编码器的特征拼接,然后通过卷积层进行融合。默认情况下,所有卷积层都带有实例标准化和ReLU激活函数。

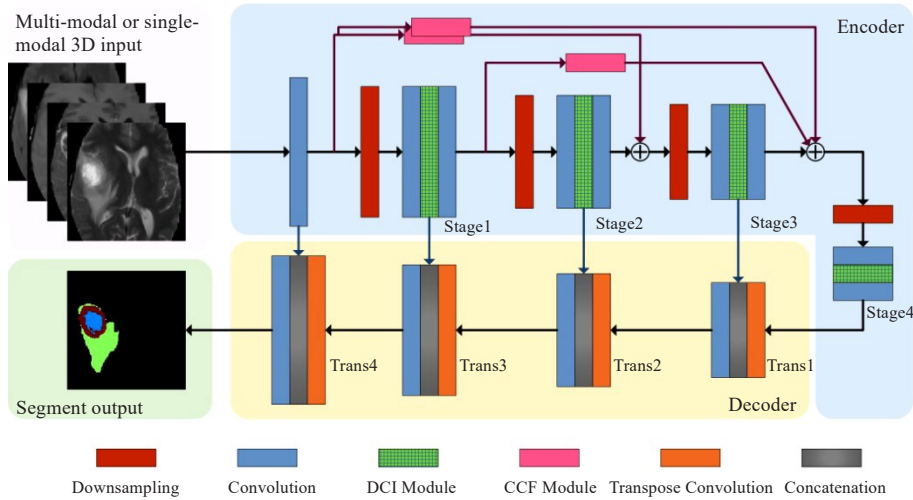


图1 提出模型的总体结构图

Figure 1 Overall structure of the proposed model

1.2 DCI模块

通过对在单个尺度上提取的卷积特征进行重采样,膨胀卷积可以准确地检测任何大小的区域。该技术有效地拓宽了卷积核的感受野,以涵盖更大的上下文,而不会增加参数数量或计算负载^[23]。如图2a所示,膨胀率为 r 的膨胀卷积在连续的卷积核之间引入 $r-1$ 个零,将卷积核大小从 $k \times k \times k$ 扩大到 $k_e = k + (k-1)(r-1)$ 。例如,在步长为1且不考虑填充的情况下,膨胀率为4的 $3 \times 3 \times 3$ 膨胀卷积具有 $9 \times 9 \times 9$

的感受野。用标准的 $3 \times 3 \times 3$ 卷积实现相同的感受野大小需要4次连续的卷积运算,从而显著增加参数和计算负载。

虽然许多研究在网络的特定阶段或层中结合了膨胀卷积,但为进一步提高网络识别复杂的目标能力,本研究模型的每个分辨率阶段都采用多尺度融合。如图2b所示,具有不同膨胀率 $r_i (i \in \{1, 2, 3, 4\})$ 的4个膨胀卷积同时对前一层的特征图进行重采样,在单独归一化之后,对它们求和,然后进行非线性激活。

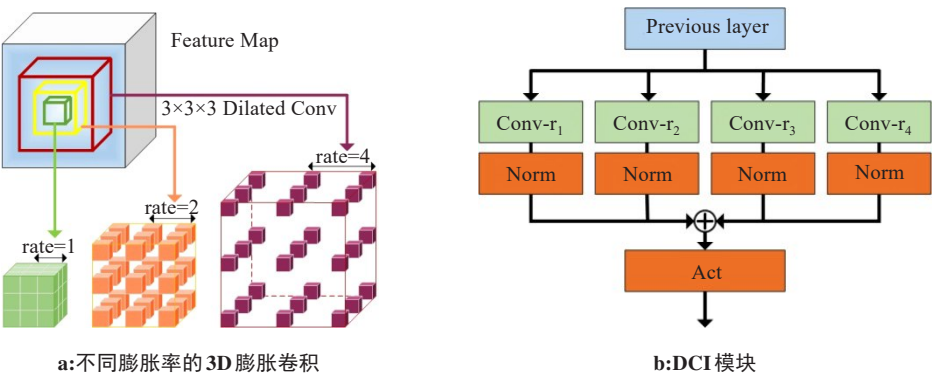


图2 3D膨胀卷积和DCI模块
Figure 2 3D dilated convolution and dilated cubic integration module

1.3 CCF模块

随着信息流深入卷积神经网络,梯度消失和退化问题不可避免,这导致细节信息的丢失和定位精度的下降。残差连接有助于在一个阶段内不同层之间直接传输信息,有助于保存和恢复一些丢失的细节。大多数梯度优化方法主要侧重于在单个分辨率阶段内添加或增强快捷连接,进一步地将这种快捷的路径扩展到不同的分辨率阶段,通过长距离上下文融合优化梯度,从而提高定位准确性。如图3所示,起始阶段的特征首先通过卷积层进行通道扩充,随后并行通过跨步卷积层、平均池化和最大池化,将高分辨率特征图下采样到目标空间维度,最后使用带有标准化和激活的卷积层实现特征融合。

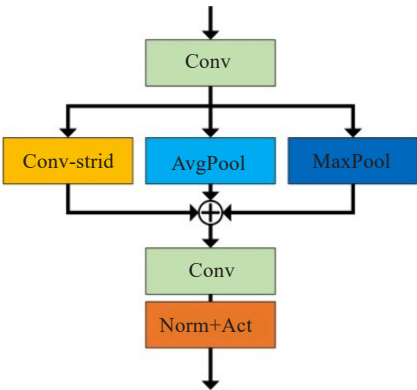


图3 CCF模块
Figure 3 Cross-stage context fusion module

1.4 损失函数

损失函数定义了模型预测与实际结果之间的差异。交叉熵损失函数(Cross Entropy Loss, CE)使得网络的优化更稳定,而Dice损失函数对类别不平衡非常敏感,因此网络的损失函数结合二者,在稳定训练的同时确保对小肿瘤或器官区域给予足够关注。定义如下:

$$L_{total} = L_{CE} + L_{Dice} \tag{1}$$

$$L_{CE} = - \sum_{k=1}^K y_k \log \hat{y}_k \tag{2}$$

$$L_{Dice} = 1 - \frac{2|y_k \cap \hat{y}_k|}{|y_k| + |\hat{y}_k|} \tag{3}$$

其中, y_k 和 \hat{y}_k 分别表示第 k 个分割类别的真实值和预测值。

2 实验与结果分析

2.1 数据集与数据增广

采用两个公开数据集对本研究模型进行实验。第一个为脑肿瘤分割数据集(BraTS2018)^[24],包含4种MRI序列: T_1 、 T_1ce 、 T_2 和FLAIR,它们以4通道的形

式输入网络。该数据集具有3种不同的肿瘤标签类别:坏死和非增强肿瘤(Non-Enhancing Tumor, NCR/NET)、增强肿瘤(Enhancing Tumor, ET)和瘤周水肿(Peritumoral Edema, ED)。分割任务需要识别3个区域:整个肿瘤(WT)区域,包括NCR/NET、ET和ED;肿瘤核心(TC)区域,包括NCR/NET和ET;以及增强肿瘤(ET)区域。第二个为CT腹部多器官分割数据

集(Synapse)^[25],以单通道的形式输入网络,需要分割的器官有:主动脉(Aor.)、胆囊(Gall.)、脾脏(Spl.)、左肾(Kid.(L))、右肾(Kid.(R))、肝脏(Liv.)、胰腺(Pan.)和胃(Sto.)。图4展示了两个数据集的原始图像、2D切片注释和3D注释,其中,图4a~c中的颜色代表强度值变化;图4d及e中的不同颜色代表不同注释。

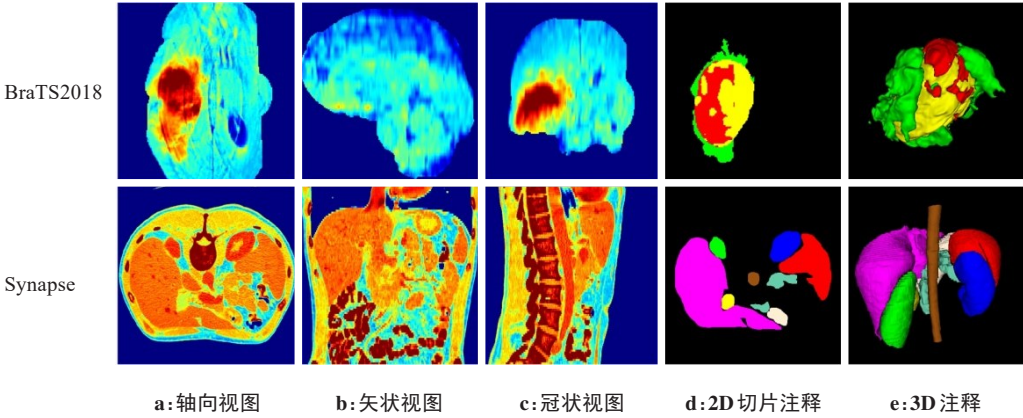


图4 数据集样本
Figure 4 Samples from dataset

在训练过程中应用数据增广技术来提高模型的泛化能力。这些技术包括随机翻转、随机旋转、强度缩放和强度偏移,以0.3的概率来执行。

2.2 评价指标

使用3D医学图像分割中常用的评价指标Dice相似性系数(Dice Similarity Coefficient, DSC)和95分位豪斯多夫距离(95% Hausdorff Distance, 95%HD)进行性能评价。DSC是衡量两个样本相似性的指标,较高的DSC值表示分割结果与真实标注相似度高,是主要评价指标。HD是衡量预测区域边缘和真实标注区域边缘之间差异的指标,较小的HD值表示二者边缘非常接近,由于对极端值的高敏感性,因此通常使用95%HD。定义如下:

DSC = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{4}

HD = \max \left\{ \max_{a \in A} \min_{b \in B} d(a, b), \max_{b \in B} \min_{a \in A} d(a, b) \right\} \tag{5}

其中, A 表示真实标注, B 表示预测分割, $|\cdot|$ 表示相应区域的体素数量, $d(a, b)$ 表示点 a 和 b 之间的欧里几得距离。

2.3 实验结果

2.3.1 消融研究 实验使用BraTS2018数据集,对单个模块进行删除或参数调整来评估它们的贡献,结果见表1。可以看出,CCF模块对ET区域的分割精度有很大的提升。因为ET区域包含于TC区域,二者差异很小所以分割困难,CCF模块通过长距离上下文融合技术减少信息丢失、提高定位精度,能有效区分相似度高的区域,并对梯度的传播产生有利的影响,促进网络收敛,损失曲线见图5。

表1 消融研究结果
Table 1 Ablation study results

方法	DSC/%				95%HD/mm			
	ET	TC	WT	均值	ET	TC	WT	均值
w/o CCF	77.56	85.06	91.34	84.65	3.67	6.63	3.92	4.74
w/o DCI	79.89	82.98	89.27	84.04	3.79	6.88	4.27	4.98
r=[1, 2, 4]	79.82	85.01	90.16	84.99	3.69	6.78	3.94	4.80
本文方法	79.76	85.28	91.07	85.37	3.77	6.78	3.93	4.82

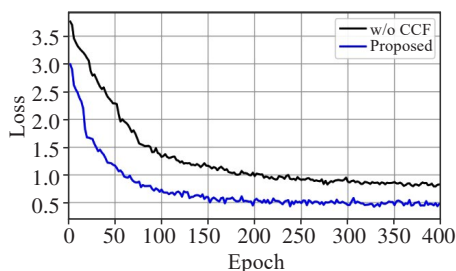


图5 CCF模块消融实验的损失曲线

Figure 5 Loss curves for the ablation study of cross-stage context fusion module

将DCI模块的4个膨胀卷积只用一个普通卷积替换来评估其贡献,结果表明,DCI模块对TC和WT区域的分割性能有显著的提升。TC和WT区域是面积较大、边缘复杂的区域,DCI模块通过对任意分辨

率阶段内的特征图进行多尺度集成,能精准识别分割边界,95%HD指标也反映了这一点。本文方法的DCI模块使用膨胀率为[1, 2, 4, 6]的4个膨胀卷积,相比使用膨胀率为[1, 2, 4]的3个膨胀卷积的总体分割精度更高。

2.3.2 对比实验 为证实本研究模型的有效性,在两个数据集上与3D U-Net^[26]、V-Net^[10]和各自领域的两种先进模型进行比较。

对于脑肿瘤分割,通过表2可以看出,本文方法实现85.37%的最高平均DSC;TC和WT区域的DSC提升较大,分别比次优值高0.98%和0.97%;WT区域的95%HD提升较大,相比次优值高0.79 mm。从图6的可视化比较可以看出,本文方法在描绘大区域的边界上更精确,同时对难以分割的ET区域有较好的识别能力。

表2 BraTS2018数据集分割结果对比

Table 2 Comparison of segmentation results on the BraTS2018 dataset

网络模型	DSC/%				95%HD/mm			
	ET	TC	WT	均值	ET	TC	WT	均值
3D U-Net ^[26]	75.40	78.30	87.30	80.33	4.56	8.03	5.90	6.16
V-Net ^[10]	76.60	81.00	89.60	82.40	7.21	7.82	6.54	7.19
HMNet ^[20]	78.60	84.30	90.10	84.33	2.69	7.73	4.72	5.04
DPAFNet ^[21]	79.50	83.90	90.00	84.46	2.92	5.71	6.51	5.04
本文方法	79.76	85.28	91.07	85.37	3.77	6.78	3.93	4.82

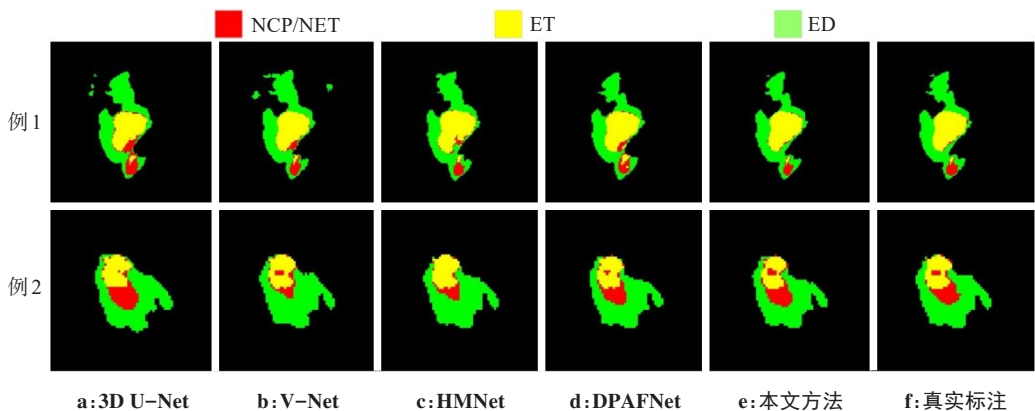


图6 BraTS2018数据集分割结果可视化对比

Figure 6 Visualization comparison of segmentation results on the BraTS2018 dataset

腹部器官分割的比较结果见表3,本文方法实现83.99%的最高平均DSC,且在5个器官(即主动脉、右肾、肝脏、胰腺和脾脏)上实现最优,分别比次优值提高1.68%、2.40%、0.07%、9.57%和3.78%,在胰腺和脾脏的分割方面表现出非凡的能力。虽然胰腺是所

考虑的8个器官中形状最不规则的,但利用多阶段多尺度集成和CCF的技术能实现高效的特征提取和精确分割,即使在复杂多变的区域也是如此。可视化比较见图7,与其他方法相比,本文方法在器官轮廓的描绘和定位方面表现出更好的结果。

表3 Synapse数据集分割结果对比(%)
Table 3 Comparison of segmentation results on the Synapse dataset (%)

网络模型	DSC								
	Aor.	Gall.	Kid.(L)	Kid.(R)	Liv.	Pan.	Spl.	Sto.	均值
3D U-Net ^[26]	87.06	60.50	82.09	76.17	94.10	51.35	87.16	74.04	76.55
V-Net ^[10]	85.35	61.87	87.10	80.75	87.84	60.05	88.56	76.98	78.56
UNETR ^[27]	84.99	63.13	84.19	83.90	95.39	62.25	87.66	74.21	79.46
Swin UNETR ^[28]	89.34	65.70	83.69	78.98	95.17	67.43	89.37	77.46	80.89
本文方法	91.02	64.85	86.99	86.30	95.46	77.00	93.15	77.17	83.99

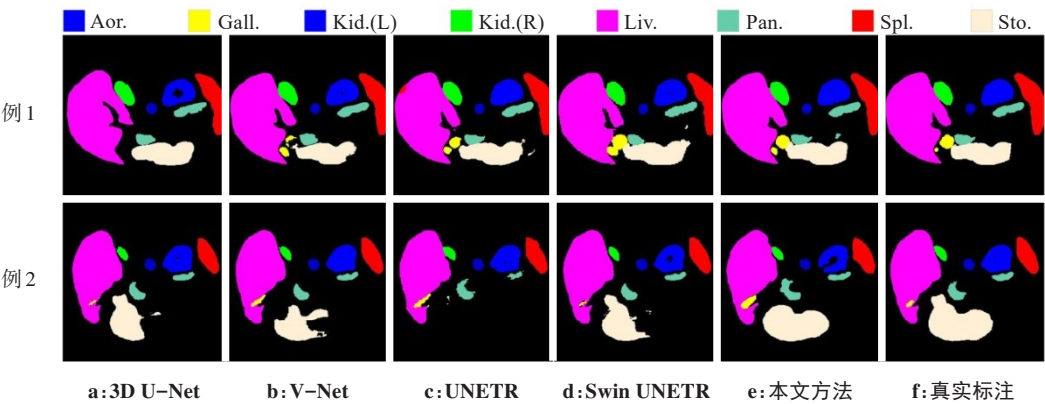


图7 Synapse数据集分割结果可视化对比
Figure 7 Visualization comparison of segmentation results on the Synapse dataset

3 结 论

本研究针对3D医学图像中由于肿瘤、器官的复杂性和多变性导致分割困难的问题,提出一种端到端3D医学图像分割模型。根据分割目标尺度跨度大的特性,设计DCI模块,并将其用于在网络的每个分辨率阶段,执行多阶段多尺度融合,该模块有效拓宽了感受野,增强了对复杂边界的识别能力。为更好地优化梯度,增强定位,设计了CCF模块,该模块促进收敛,提高对差异性较小的对象的识别能力。在脑肿瘤和腹部器官分割上的实验表明本文方法的性能优于经典模型和先进模型,这将成为辅助临床医生诊断和治疗的有利工具。鉴于本文方法在某些器官区域表现出良好的性能,在未来的工作中将尝试针对这些特定区域,如胰腺的分割上,展开进一步的研究。此外,尝试将卷积与注意力机制进行结合以获取更为广泛的上下文信息也是未来值得开展的研究。

【参考文献】

[1] Jiang HY, Diao ZS, Shi TY, et al. A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation[J]. Comput Biol Med, 2023, 157: 106726.

[2] Garcia Ugarriza L, Saber E, Vantaram SR, et al. Automatic image segmentation by dynamic region growth and multiresolution merging [J]. IEEE Trans Image Process, 2009, 18(10): 2275-2288.

[3] Dong CH, Zeng XY, Lin LF, et al. An improved random walker with Bayes model for volumetric medical image segmentation[J]. J Healthc Eng, 2017, 2017: 6506049.

[4] Erickson BJ, Korfiatis P, Akkus Z, et al. Machine learning for medical imaging[J]. Radiographics, 2017, 37(2): 505-515.

[5] Yu Y, Wang CP, Fu Q, et al. Techniques and challenges of image segmentation: a review[J]. Electronics, 2023, 12(5): 1199.

[6] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(4): 640-651.

[7] Shukla PK, Zakariah M, Hatamleh WA, et al. AI-DRIVEN novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network[J]. J Healthc Eng, 2022, 2022: 4277436.

[8] Zhou ZW, Siddiquee MM, Tajbakhsh N, et al. UNet++: redesigning skip connections to exploit multiscale features in image segmentation [J]. IEEE Trans Med Imaging, 2020, 39(6): 1856-1867.

[9] Zhou CH, Ding CX, Wang XC, et al. One-pass multi-task networks with cross-task guided attention for brain tumor segmentation [J]. IEEE Trans Image Process, 2020, 29: 4516-4529.

[10] Milletari F, Navab N, Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). Piscataway, NJ, USA: IEEE, 2016: 565-571.

[11] Chen H, Dou Q, Yu LQ, et al. VoxResNet: deep voxelwise residual networks for brain segmentation from 3D MR images [J]. Neuroimage, 2018, 170: 446-455.

[12] Wang LS, Xie C, Zeng NY. RP-net: a 3D convolutional neural network for brain segmentation from magnetic resonance imaging[J]. IEEE Access, 2019, 7: 39670-39679.

[13] Ibtehaz N, Rahman MS. MultiResUNet: rethinking the U-net

- architecture for multimodal biomedical image segmentation [J]. Neural Netw, 2020, 121: 74-87.
- [14] Zhang ZA, Wu CD, Coleman S, et al. DENSE-INception U-net for medical image segmentation[J]. Comput Methods Programs Biomed, 2020, 192: 105395.
- [15] Zhou XY, Li XY, Hu K, et al. ERV-Net: an efficient 3D residual neural network for brain tumor segmentation[J]. Expert Syst Appl, 2021, 170: 114566.
- [16] Wang JJ, Gao J, Ren JW, et al. DFP-ResUNet: convolutional neural network with a dilated convolutional feature pyramid for multimodal brain tumor segmentation[J]. Comput Methods Programs Biomed, 2021, 208: 106208.
- [17] Isensee F, Jaeger PF, Kohl SA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nat Methods, 2021, 18(2): 203-211.
- [18] Zhou ZX, He ZS, Shi MF, et al. 3D dense connectivity network with atrous convolutional feature pyramid for brain tumor segmentation in magnetic resonance imaging of human heads[J]. Comput Biol Med, 2020, 121: 103766.
- [19] Sun JD, Peng YJ, Guo YF, et al. Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN[J]. Neurocomputing, 2021, 423: 34-45.
- [20] Zhang RF, Jia SS, Adamu MJ, et al. HMNet: hierarchical multi-scale brain tumor segmentation network[J]. J Clin Med, 2023, 12(2): 538.
- [21] Chen WK, Zhou WF, Zhu L, et al. MTDCNet: a 3D multi-threading dilated convolutional network for brain tumor automatic segmentation [J]. J Biomed Inform, 2022, 133: 104173.
- [22] Chang YK, Zheng ZZ, Sun YW, et al. DPAFNet: a residual dual-path attention-fusion convolutional neural network for multimodal brain tumor segmentation[J]. Biomed Signal Process Control, 2023, 79(Part 1): 104037.
- [23] Chen LC, 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.
- [24] Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS)[J]. IEEE Trans Med Imaging, 2015, 34(10): 1993-2024.
- [25] Landman B, Xu ZB, Igelsias JE, et al. MICCAI multi-atlas labeling beyond the cranial vault-workshop and challenge[C]//MICCAI Multi-Atlas Labeling Beyond Cranial Vault-Workshop Challenge, 2015: 12.
- [26] 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.
- [27] Hatamizadeh A, Tang YC, Nath V, et al. UNETR: transformers for 3D medical image segmentation[C]//2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Piscataway, NJ, USA: IEEE, 2022: 1748-1758.
- [28] Tang YC, Yang D, Li WQ, et al. Self-supervised pre-training of Swin transformers for 3D medical image analysis[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 20698-20708.

(编辑:谭斯允)