DOI:10.16016/j.1000-5404.202103046

基于 2-3D 混合卷积神经网络的胰腺 CT 图像自动分割算法研究

宁民权,文 静,黄 意,王 翊 400044 重庆,重庆大学计算机学院

[摘要] **目的** 探讨 CT 图像中胰腺组织自动分割的深度学习算法。**方法** 以 2D 胰腺分割网络和 3D 胰腺分割网络为基础,利用 2D 分割网络提取的判别性语义信息,3D 分割网络提取出三维空间信息,最后将判别性的语义信息和三维空间信息进行融合实现胰腺分割。**结果** 本研究在 NIH 数据集上的进行了验证,平均 DICE 系数达到 83.2%(最大值 90.65%,最小值 67.04%)。分别超出了 2D 基准方法和 3D 基准方法 0.77%和 1.38%。其胰腺分割结果细节准确,边缘平滑。2-3D 方法分割出的胰腺组织相比于 2D 方法和 3D 方法分割出的胰腺,在轮廓形态上与对应的手工标注图吻合度更高,更准确。相比于 2D 方法,2-3D 方法有效综合了三维空间信息,对胰腺边缘位置进行了补足,避免了在胰腺边缘出现的漏分割情况。相比于 3D 方法,2-3D 方法有效地利用了 2D 方法提供的语义信息,规避了在胰腺边缘出现的错误分割的情况。**结论** 该方法融合了 2D 分割网络和 3D 分割网络的优点,能够有效地应用于对胰腺的准确分割。

[关键词] 胰腺分割;深度学习,计算机辅助诊断;2-3D 网络融合
[中图法分类号] R319; R322.491; R814.42 [文献标志码] A

Automatic segmentation algorithm of pancreatic CT images based on 2-3D mixed convolutional neural network

NING Minquan, WEN Jing, HUANG Yi, WANG Yi College of Computer Science, Chongqing University, Chongqing, 400044, China

[**Abstract**] Objective To investigate the deep learning algorithm for automatic segmentation of pancreatic tissue from CT images. Methods Based on both 2D and 3D pancreatic segmentation networks, which extracted discriminative semantic information and three-dimensional spatial information from 82 CT images of pancreas samples (taken from NIH database) respectively, a novel hybrid network was designed for automatic segmentation of pancreatic tissues. Results The proposed method was validated using NIH datasets, and the results of pancreas segmentation demonstrated accurately in detail and smooth in edge. The average Dice coefficient reached 83.2%, which exceeded the 2D and 3D benchmark methods by 0.77% and 1.38%, respectively. Conclusion This method combines both advantages of 2D and 3D segmentation networks, and thus can be effectively applied to the accurate segmentation of pancreas, providing a quantitative reference for subsequent clinical diagnosis and treatment.

[**Key words**] pancreas segmentation; deep learning; computer-aided diagnosis; 2-3D hybrid network

Supported by the General Program of National Natural Science Foundations of China (61672120) and the Fundamental Research Funds for Central Universities (2020CDCGJSJ043, 2020CDCGSL054,2019CDYGZD004). Corresponding author: WEN Jing, E-mail: wj@cqu.edu.cn

胰腺癌早期诊断困难,病程短,发病速度快,是预后最差的恶性肿瘤之一。胰腺癌 IV 期患者确诊后生存期小于 1 年,5 年平均生存率小于 1%[1]。图像分割在影像组学、三维可视化技术胰腺癌的早期诊断、鉴别

[基金项目] 国家自然科学基金面上项目(61672120);中央高校基本业务费(2020CDCGJSJ043,2020CDCGSL054,2019CDYGZD004)
[通信作者] 文 静,E-mail: wj@cqu.edu.cn

1000-5404/Copyright © 2021, 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/)

诊断、评估手术获益、疗效评估及预后预测等多个临床诊疗环节均发挥重要作用[2-3]。由于胰颈及胰体部重要血管多,复杂的毗邻结构使图像边界难以准确区分,目前仍多采用手动或半自动的分割方法、耗时费劲且准确性有待提高,而以卷积神经网络为代表的自动分割方法则成为未来进行胰腺分割的首选方法。

随着大规模器官分割标注数据集合的出现和相关技术的发展[4-5],卷积神经网络已经在胰腺周边器官(肝脏、脾脏、左右肾脏等器官)上取得令人满意的效果,但由于胰腺周边复杂的毗邻结构、不同病例胰腺形状的多变性和 CT 图像中胰腺模糊的边界,造成了现存的基于卷积神经网络的自动分割算法难以精准的分割胰腺。

目前基于卷积神经网络的自动分割算法,大致可以分为两类:2D 方法和 3D 方法。其中 2D 方法以单张切片为输入,UNet[6] 是最为经典的 2D 方法,Attention UNet[7] 是 UNet 的改进,其使用上层特征监督下层参数的训练,约束参数优化方向。同时对轴状面、冠状面、矢状面 3 个面进行分割,使用投票的方式融合分割结果[8]。3D 方法则以体块作为输入,3D-UNet[9] 将 UNet[6] 中的 2D 卷积替换为 3D 卷积,VNet[10] 在 3D-UNet 的基础之上引入了残差结构[11],将级联的思想引入 3D 网络,使用级联 3D 网络的形式提升分割效果[12]。使用预测-微调的策略进一步提升了 3D 网络的分割效果。

2D 方法的优点在于可以对单张切片中的特征进行深度挖掘,但其缺点在于仅关注了单张切片的信息却忽视了三维空间信息。3D 方法以体数据块作为输入,优点在于可以有效的利用三维空间信息,缺点在于 3D 方法仅能使用小的切块作为输入,其网络深度也受到了限制,这使得 3D 方法无法有效的挖掘单张切片的信息,在单张切片上的分割效果差于 2D 方法。

鉴于目前 2D 方法和 3D 方法存在问题,本研究提出了一种可以有效结合 2D 方法和 3D 方法优点的 2-3D 混合网络,综合利用 2D 方法提取的单张切片的语义信息和 3D 方法提取的上下文语义信息取得更好的胰腺分割效果。

1 资料与方法

1.1 数据集来源

本实验使用的数据集是来自美国国立卫生研究院的 NIH[13] 胰腺分割公开数据集,其中包含 82 例腹部 CT 增强扫描,以及相应体数据的绝对位置、数据密度等信息。扫描分辨率为 512×512 像素,切片数在 181~466 之间,层厚为 1.5~2.5 mm。

1.2 数据预处理

数据预处理步骤包含下述步骤:①窗口范围调整。②数据进行离差标准化,将像素值映射到[0,1]之间。③数据裁剪和缩放。

具体操作步骤如下:

①在原始腹部 CT 图像中,各组织的 HU 值范围不同。其中常见的胰腺 HU 值范围为 30~55,根据像素值-HU 值转换方程,计算出胰腺的窗口范围,再对其进行适度延伸,最终得出的胰腺窗口范围为[-300,300]。其中像素值-HU 值转换方程如下所示,其中 Pixel 表示像素值,ScaleFactor 表示放缩倍率,Bias 表示偏差值。

$$HU = Pixel * ScaleFactor + Bias$$

②依照得出的胰腺窗口范围对其进行截断,再对截断后的数据进行离差标准化,其中离差标准化的公式如下所示:

$$Y = \frac{X - x^{min}}{x^{max} - x^{min}}$$

③由于胰腺位置相对固定,我们首先对胰腺的位置进行统计,再根据统计的位置信息对胰腺进行裁剪,将图像从 512×512 裁剪为 256×256,再使用二次线性差值将图像从 256×256 下采样至 128×128。

1.3 使用 2-3D 混合网络进行胰腺分割

预处理之后,我们已经在一定程度上排除了其余器官对胰腺的干扰,并根据统计的胰腺位置信息对图像进行了裁剪,在此基础之上对其使用 2-3D 混合卷积神经网络进行分割可以得到更好的效果。

本研究使用的 2-3D 混合网络主要分为 2D 部分和 3D 部分。其中 2D 部分使用的是主干网络为 ResNet34[14] 的 ResUNet,编码器通道数设置为(64,128,256,512,1 024),每个残差组使用的残差单元数量为(3,4,6,3,2)。解码器部分每个残差组使用的残差单元数量为(1,1,1,1),步长为 2 的卷积层和反卷积层被应用于下采样和上采样,具体设置参照 ResNet[14]。3D 部分使用的是主干网络为 ResNet18-3D 的 VNet[10] 网络,解码器通道数设置为(32,64,128,256),每个残差组使用的残差单元数量为(2,2,2,2)。解码器部分每个残差组使用的残差单元数量为(1,1,1),步长为 2 的 3D 卷积层和 3D 反卷积层被应用于下采样和上采样,具体设置参照 VNet[10]。

具体分割流程包含下述步骤:①对预处理之后的 CT 图像进行重新组织,获取 2D 网络输入和 3D 网络原始输入。②将 2D 网络输入送入 2D 模型获取 2D 分割结果和 2D 多尺度语义特征。其中 2D 多尺度语义特征指的是将图像 CT 特征输入卷积神经网络后,经

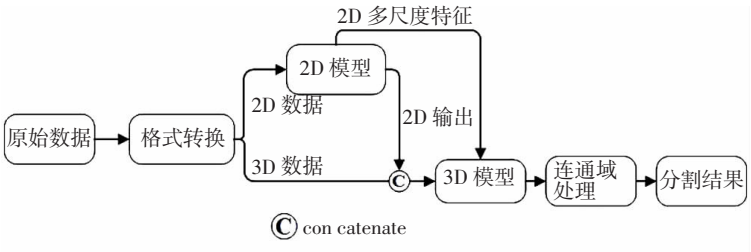


图 1 胰腺分割算法的网络构架

过不同数量的卷积和下采样层得到的不同分辨率的特征信息。③对 2D 分割结果进行转换,转换成为 3D 形式并同 3D 网络原始输入进行结合得到 3D 网络输入。④将 2D 多尺度语义特征进行转换,转换为 3D 形式的特征。⑤将步骤③得到的 3D 网络输入和步骤④得到的 3D 特征输入 3D 模型,获取 3D 分割结果。⑥对步骤⑤得到的 3D 分割结果进行 3D 最大连通域处理,得到最终的分割结果。2-3D 混合网络算法的网络架构图如图 1 所示。

1.4 模型训练方法

胰腺 CT 图像数据集来自于公开的胰腺分割数据集:NIH[13]胰腺分割数据集。我们对其进行了五折交叉验证,每折中选取其中的 61(60)套作为训练集,5 套作为验证集,16(17)套作为测试集。因为 2D 特征同 3D 特征相差较大,直接训练容易导致模型无法收敛,从而导致分割结果的精度较低,故而将训练过程分为 3 个阶段:①单独训练 2-3D 混合模型中的 2D 网络。②固定 2D 网络参数,单独训练 3D 网络。③对 2D 网络和 3D 网络进行联合训练,对参数进行微调。学习率为 5E-5 的 Adam 优化器和下述损失函数被应用于以上 3 个阶段:

$$L(y,\hat{y})=[E(y,\hat{y})+Dice(y,\hat{y})]$$

其中交叉熵损失(CE)和 DICE 损失具体公式如下:

$$CE(y,\hat{y})=\frac{1}{N}\sum_{i=1}^{N}[y_i\log\hat{y}_i+(1-y_i)(1-log\hat{y})]$$

其中 $\hat{y}$ 表示预测结果,$\hat{y}$ 表示手工标注的结果。

1.5 评价指标

本文利用医学图像分割中常用的 Dice 系数来评估模型的分割效果,Dice 系数公式如下:

$$DiceCoef(y,\hat{y})=\frac{2\sum_i y_i\hat{y}_i}{\sum_i y_i\hat{y}_i}$$

其中 $\hat{y}$ 表示预测结果,y 表示手工标注的结果。

2 结果

利用 pytorch 深度学习框架对 2D 卷积神经网络、3D 卷积神经网路和 2-3D 卷积神经网络的胰腺分割算法编程实现。其中 2D 网络指的是本研究提出的 2-3D 混合网络中的 2D 部分,3D 网络指的是本研究提出的 2-3D 混合网络中的 3D 部分。

本研究在 NIH[13]数据集上对 2D、3D 和 2-3D 方法进行了测试,结果如表 1 所示,2-3D 方法在平均值、最小值、最大值上均高于 2D 方法和 3D 方法,同时 2-3D 方法方差小于 2D 方法,远小于 3D 方法,这证明 2-3D 方法更加稳定可靠。除此之外,本研究还与主流的医学图像分割方法进行了比较,结果如表 2 所示,其中带“*”的方法表示该方法使用了与本研究同样的数据集进行重新训练。

表 1 2D、3D 方法和 2-3D 混合方法在 NIH 胰腺分割数据集上的表现

组别	2D	3D	2-3D
最大值	90.25%	89.75%	90.65%
最小值	63.94%	56.39%	67.04%
平均值+方差	82.43%±5.56	81.82%±7.65	83.2%±5.08

表 2 各种方法在在 NIH 胰腺分割数据集上的表现

方法	平均值	最大值	最小值
UNet *[6]	80.6%	88.18%	64.07%
FCN *[15]	79.3%	87.64%	63.94%
ResNet34[14]	82.43%	90.25%	63.94%
3D UNet-Patch *[16]	77.86%	89.25%	21.83%
3D UNet-Full *[5]	81.03%	88.6%	49.06%
Attention UNet[7]	83.1%	-	-
2-3D	83.2%	90.65%	67.04%

准确的胰腺分割对手术计划的制定、放射治疗具有重要意义。错误的胰腺分割对后续的治疗具有极为恶劣的影响。图 2 给出了 2D 网络、3D 网络以及本文提出的 2-3D 混合网络在单张切片上的分割结果。如图 2 第 1 行代表在当前切片胰腺目标极小时的情况:2D 分割方法在当前切片胰腺目标较小时难以进行准确的分割,出现漏检的情况,3D 方法虽然成功的分割

出了胰腺,但是分割结果同手工标注有较大差异,2-3D方法成功分割出胰腺并取得了较好的分割效果。如图2第2、3行代表在当前切片胰腺目标较大时的情况:2D、3D和本研究提出的2-3D方法均能较好的分割出胰腺,但2-3D方法分割出的胰腺组织相比于2D方法和3D方法分割出的胰腺,在轮廓形态上与对应的手工标注图吻合度更高,更准确。

同时我们可对胰腺分割结果进行进一步的处理,处理后的数据可以应用于解剖教学和3D打印。图3展示了2D、3D、2-3D方法得出的胰腺分割结果和手工标注胰腺的三维重建效果展示图。相比于手工标注的胰腺,2-3D方法得出的胰腺分割结果轮廓更为平滑、细腻。相比于2D方法,2-3D方法有效综合了三维空间信息,对胰腺边缘位置进行了补足,避免了在胰腺边缘出现的漏分割情况。相比于3D方法,2-3D方法有效地利用了2D方法提供的语义信息,规避了在胰腺边缘出现的错误分割的情况。在临床应用中,使用2-3D方法得出胰腺分割结果进行3D打印,得出的胰腺模型不会出现突出和截断情况,其形状更加接近病人的真实胰腺。

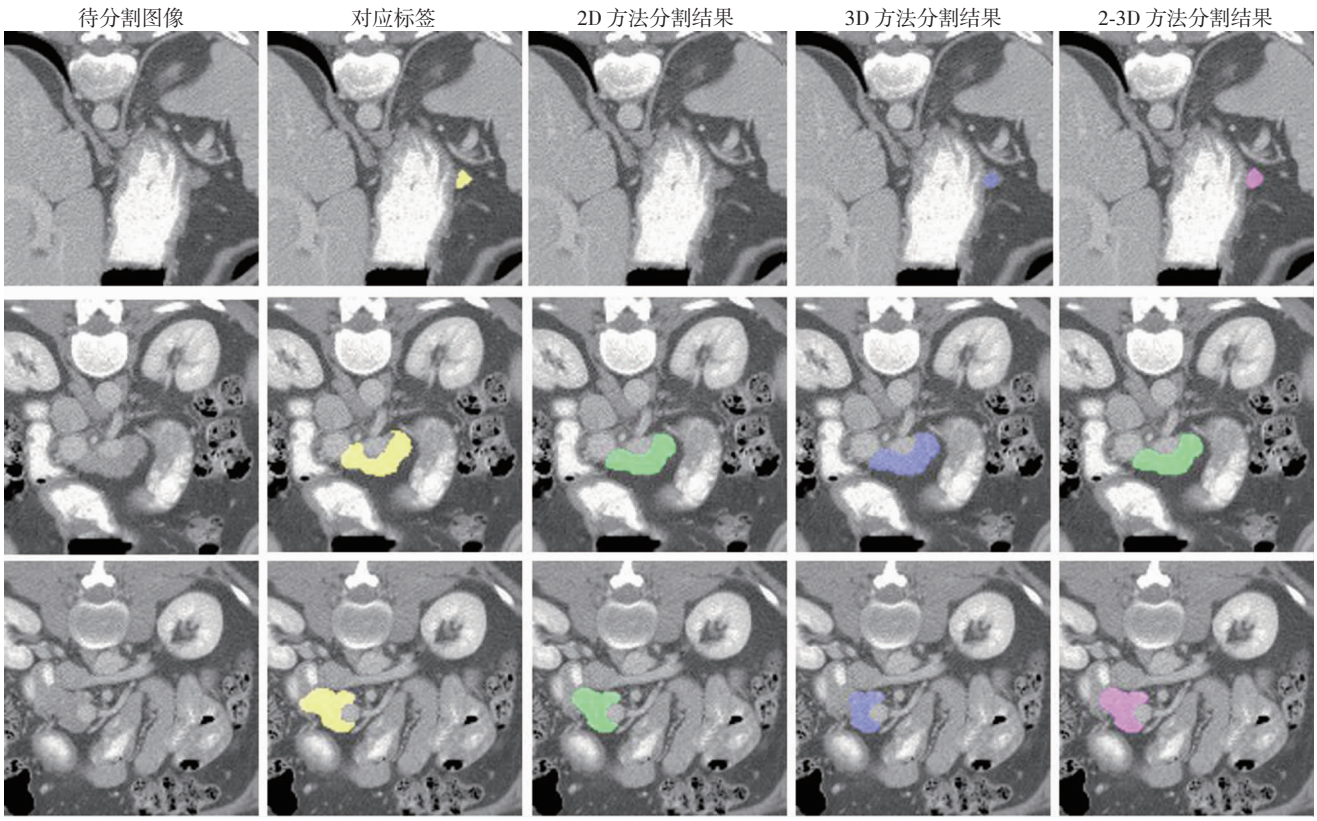
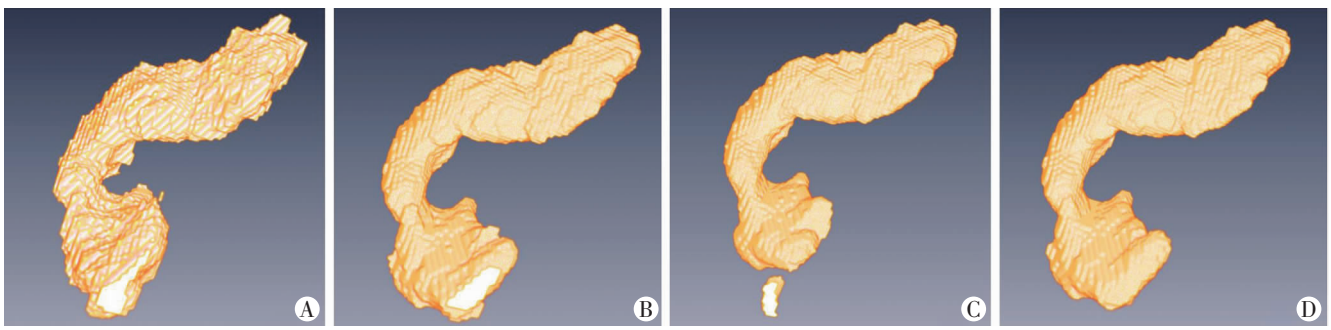


图2 CT图像的胰腺分割结果



A:手工标注胰腺的三维重建效果;B:2D部分得出的胰腺分割结果三维重建效果;C:3D部分得出的胰腺分割结果三维重建效果;D:2-3D混合卷积神经网络得出的胰腺分割结果三维重建效果

图3 胰腺分割结果的三维重建效果

3 讨论

精准的胰腺分割有助于医生进行简单、快速、准确的CT图像分析,可以有效地减轻医学专家的工作量,提高医生工作效率[17]。提升胰腺相关疾病的诊断效率和准确性。在课堂教学中利用胰腺分割结果进行的器官三维重建可以作为辅助教学手段提升教学效率,提高学生对相关知识的理解。在临床实践中,利用胰腺分割结果进行的器官三维重建可以有效辅助医生判断病灶、规划手术方案提升手术成功率。可见,准确的胰腺自动分割是进行胰腺相关疾病的早期诊断、影像组学分析、精准放疗、3D打印、课堂教学等任务的重要基础步骤。

针对2D方法和3D方法存在的问题:2D方法无法利用三维空间信息,3D方法无法深度挖掘单张切片信息。本研究提出了1种针对胰腺CT图像全自动分割方法,有效的综合了2D方法和3D方法的优点,并规避了2D方法和3D方法的缺点,创新性的将2D网络和3D网络进行有机结合。相比于单独2D方法或3D方法,2-3D混合方法可以有效地提升胰腺分割的准确度,从分割结果的三维效果展示图来看,2-3D方法得出三维效果展示图较之于2D方法更显平滑、细腻且在边缘位置仍能进行准确的分割,不存在漏分割的情况,较之于3D方法,其分割结果更为准确且减少了错分割的情况。

在后续的工作中,可以通过改变2-3D混合网络中2D部分和3D部分的网络结构来达到更好的分割效果,也可以从2-3D信息融合方式上下功夫,得到更有效的信息融合方式。可见,本研究提出的2-3D混合网络仍有巨大的改进空间,同时新颖的2-3D混合卷积神经网络结构对后续进一步提升胰腺分割的效果具有启发意义。

参考文献:

[1] CHAN-SENG-YUE M, KIM J C, WILSON G W, *et al*. Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution[J]. Nat Genet, 2020, 52(2): 231-240. DOI: 10.1038/s41588-019-0566-9.

[2] 何铭, 薛华丹, 金征宇. 影像组学在胰腺癌的应用及研究进展[J]. 中华放射学杂志, 2020, 54(10): 1029-1032. DOI: 10.3760/cma.j.cn112149-20190903-00738.
HE M, XUE H D, JIN Z Y. Radiomics in pancreatic cancer: the state of art[J]. Chin J Radiol, 2020, 54(10): 1029-1032. DOI: 10.3760/cma.j.cn112149-20190903-00738.

[3] 李仓, 吴晓康, 厉学民, 等. 三维可视化技术结合3D腹腔镜在胰腺中段切除术中的临床应用[J]. 中华胰腺病杂志, 2020(6): 434-437. DOI: 10.3760/cma.j.cn115667-20200610-00090.
LI C, WU X K, LI X M, *et al*. Clinical application of 3D visualization technology combined with 3D laparoscopy in the resection of middle segment pancreatectomy[J]. Chin J Pancreatol, 2020(6): 434-437. DOI: 10.3760/cma.j.cn115667-20200610-00090

[4] ZHOU Y Y, XIE L X, SHEN W, *et al*. A fixed-point model for pancreas segmentation in abdominal CT scans[M]//Medical Image Computing and Computer Assisted Intervention - MICCAI 2017. Cham: Springer International Publishing, 2017: 693-701. DOI: 10.1007/978-3-319-66182-7_79

[5] ROTHH R, SHEN C, ODA H, *et al*. Deep learning and its application to medical image segmentation[J]. Med Imaging Technol, 2018, 36. DOI: 10.11409/mit.36.63

[6] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. MICCAI, 2015, 9351. 234-241. DOI: 10.1007/978-3-319-24574-4_28.

[7] OKTAYO, SCHLEMPER J, FOLGOC L L, *et al*. Attention U-net: learning where to look for the pancreas[EB/OL]. 2018.

[8] ZHOUX R, ITO T, TAKAYAMA R, *et al*. Three-dimensional CT image segmentation by combining 2D fully convolutional network with 3D majority voting[M]//Deep Learning and Data Labeling for Medical Applications. Cham: Springer International Publishing, 2016: 111-120. DOI: 10.1007/978-3-319-46976-8_12.

[9] ÇIÇEK Ö, ABDULKADIRA, LIENKAMP S S, *et al*. 3D U-net: learning dense volumetric segmentation from sparse annotation[C]//Med Image Comput Comput Assist Interv MICCAI, 2016: 424-432. DOI: 10.1007/978-3-319-46723-8_49.

[10] MILLETARIF, 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). October 25-28, 2016, Stanford, CA, USA. IEEE, 2016: 565-571. DOI: 10.1109/3DV.2016.79.

[11] ROTHH R, ODA H, ZHOU X R, *et al*. An application of cascaded 3D fully convolutional networks for medical image segmentation[J]. Comput Med Imaging Graph, 2018, 66: 90-99. DOI: 10.1016/j.compmedimag.2018.03.001.

[12] NWE T L, MIN O Z, GOPALAKRISHNAN S, *et al*. Improving 3D brain tumor segmentation with predict-refine mecha-

nism using saliency and feature maps[C]//2020 IEEE International Conference on Image Processing (ICIP). October 25-28, 2020, Abu Dhabi, United Arab Emirates. IEEE, 2020: 2671-2675. DOI: 10.1109/ICIP40778.2020.9190806.

[13] Roth H R, Lu L, Farag A, *et al*. DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 556-564. DOI: 10.1007/978-3-319-24553-9_68.

[14] HE K M, ZHANG X Y, REN S Q, *et al*. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, USA. IEEE, 2016: 770-778. DOI: 10.1109/CVPR.2016.90.

[15] SHELHAMERE, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(4): 640-651. DOI: 10.1109/TPAMI.2016.2572683.

[16] ROTH H R, ODA H, ZHOU X R, *et al*. An application of cascaded 3D fully convolutional networks for medical image segmentation[J]. Comput Med Imaging Graph, 2018, 66: 90-99. DOI: 10.1016/j.compmedimag.2018.03.001.

[17] 郑彩侠, 张同舟, 孙长江, 等. 图像分割方法在医学领域中的应用[J]. 中国医疗设备, 2018, 33(6): 1-5, 11. DOI: 10.3969/j.issn.1674-1633.2018.06.001.
ZHENG C X, ZHANG T Z, SUN C J, *et al*. Application of image segmentation methods in the medical field[J]. China Med Devices, 2018, 33(6): 1-5, 11. DOI: 10.3969/j.issn.1674-1633.2018.06.001.

(**收稿**:2021-03-04;**修回**:2021-05-17)

(编辑 王 红)