

基于双解码 U 型卷积神经网络的胰腺分割^{*}

毕秀丽, 陆 猛, 肖 斌, 李伟生

(图像认知重庆市重点实验室, 重庆 400065)

通信作者: 肖斌, E-mail: xiaobin@cqupt.edu.cn



摘 要: 计算机断层成像 (computed tomography, CT) 中, 胰腺分割作为医学图像分析中最具挑战的任务之一, 由于其体积小、形状多变的特点, 导致传统的自动分割方法无法达到理想的分割精度. 利用高级语义特征指导低级特征的思想, 提出一种基于双解码 U 型卷积神经网络的单阶段胰腺分割模型. 模型由一个编码器和两个解码器构成, 两个解码器利用不同编码深度的特征将低级空间信息与高级语义信息有效结合, 加强分割网络对特征信息的高效利用, 能够对未裁剪、未降低分辨率的 CT 切片实现高精度度的分割. 实验结果表明, 方法能够在全尺寸的输入下实现较好的分割性能, 在公开胰腺数据集上的分割效果优于现有单阶段胰腺分割方法.

关键词: 医学图像; 胰腺分割; 卷积神经网络; 单阶段分割模型; 双解码 U-Net

中图法分类号: TP391

中文引用格式: 毕秀丽, 陆猛, 肖斌, 李伟生. 基于双解码U型卷积神经网络的胰腺分割. 软件学报, 2022, 33(5): 1947–1958. <http://www.jos.org.cn/1000-9825/6198.htm>

英文引用格式: Bi XL, Lu M, Xiao B, Li WS. Pancreas Segmentation Based on Dual-decoding U-Net. Ruan Jian Xue Bao/Journal of Software, 2022, 33(5): 1947–1958 (in Chinese). <http://www.jos.org.cn/1000-9825/6198.htm>

Pancreas Segmentation Based on Dual-decoding U-Net

BI Xiu-Li, LU Meng, XIAO Bin, LI Wei-Sheng

(Chongqing Key Laboratory of Image Cognition, Chongqing 400065, China)

Abstract: Pancreas segmentation in computed tomography (CT) is one of the most challenging tasks in medical image analysis. Due to small size and changeable shape, the traditional automatic segmentation methods can not achieve the acceptable segmentation accuracy. By using the idea of high-level semantic features to guide low-level features, this study proposes a single-stage pancreas segmentation model based on dual-decoding U-net. The proposed architecture consists of one encoder and two decoders, which can effectively combine low-level spatial information with high-level semantic information using the features of different coding depths to improve the segmentation accuracy of CT slices without clipping and resolution reduction. The experimental results show that this method can achieve better segmentation performance under full-size input. Moreover, the segmentation result by the proposed method is superior to the single-stage methods on the open dataset for pancreas segmentation tasks.

Key words: medical image; pancreas segmentation; convolutional neural networks (CNN); single-stage model; dual-decoding U-Net (DDU-Net)

计算机辅助诊断可以通过计算机技术的自动化操作, 实现传统诊断方法中复杂繁琐的人工操作, 减少人工成本, 在许多临床情况下辅助医生实现快速诊断. 基于计算机断层成像 (computed tomography, CT) 的计算机辅助诊断方法一般需要人工标注 CT 影像中的各个器官, 这对标注人员有一定的经验要求, 并且是一项重复量大的工作. 医学图像分割成为近年来的计算机辅助诊断的研究热点之一, 由于医学图像的复杂性高、缺少简单的线性特征、数据标注不统一、标注样本量少等原因, 直接将自然图像的语义分割方法^[1-3]应用到医学图像并不适用. 对于医学

^{*} 基金项目: 国家自然科学基金 (61806032, 61976031); 国家重点研发计划 (2016YFC1000307-3); 重庆市基础与前沿项目 (cstc2018jcyjAX0117); 重庆市教委科学技术研究计划 (KJZD-K201800601, KJQN201800611)

收稿时间: 2020-05-16; 修改时间: 2020-08-29; 采用时间: 2020-10-30

图像分割问题,研究者通常针对不同的分割任务提出了对应的医学分割模型,常见的分割区域包括肝脏^[4,5]、脾脏^[6]、肾脏^[7,8]、胰腺^[9,10]等.近年来,随着深度学习技术在图像分类^[11-13]、目标检测^[2,14]和语义分割^[10,15]等计算机视觉任务中的广泛应用,深度学习技术^[16]也逐渐被用于解决医学图像分割问题.得益于卷积神经网络^[17]的强鲁棒性、易于训练和优化的特点,在某些器官的分割效果上,戴斯相似系数 (Dice similarity coefficients, DSC) 可以达到 90%^[9,18,19],推动了医学图像分割向体积更小、更复杂器官分割方向发展.

胰腺分割作为医学图像分割领域最具代表性的任务之一,因为胰腺器官在形状、大小以及位置上差异大,且仅占输入的 CT 图像的比例小 (<0.5%),使得分割任务难度加大.因为用于分割的模型很容易被占据更大比例且包含复杂多变内容的背景区域所破坏,这对分割模型的性能提出了更高的要求.目前,针对胰腺的分割方法主要分为基于传统机器学习和基于深度学习的胰腺分割方法.在基于传统机器学习的胰腺分割方法中,基于体积的多重地图集登记方法^[20]被用在胰腺分割任务上,取得了不错的分割效果.Wang 等人^[21,22]提出了标签融合的方法优化器官上每个像素的标注,这种器官分割策略后来也被广泛应用到其他分割器官问题^[21,23,24].近年来,Farag 等人提出了一种新的自下而上的胰腺分割方法^[25],先将图像区域分割成多个像素块,再使用随机森林把这些像素块的胰腺和非胰腺进行分类,最终通过连通区域分析得到最终的分割图.总体而言,基于传统机器学习的胰腺分割方法虽然操作简单且不需要训练复杂的模型,但分割精度不高、效率低、泛化能力不够.

对于基于深度学习的胰腺分割方法,根据其是否进行二次分割操作,可以将这些分割方法大体分为两类.第一类是单阶段胰腺分割方法,该方法只对腹部 CT 切片进行一次分割就能得到最终的分割结果.其中最具代表性的是 Oktay 等人提出的 Attention U-Net^[26],该方法将注意力机制引入到基于 U-Net^[27]的 3D 分割模型中,利用注意力门的机制让网络自己学习需要关注的区域,相较于基于 U-Net 的方法有较大提升.但是由于胰腺在腹部 CT 图像中的占比较小,导致样本分布严重不平衡,使得 Attention U-Net 分割网络在特征的提取过程中被其他器官组织区域干扰,遗漏了大量的边缘信息,在一定程度上影响了分割效果.第二类方法是由粗到细的两阶段胰腺分割模型,该方法包括粗分割和细分割两个阶段.Roth 等人于 2016 年首次提出这种方法^[28],通过使用随机森林和空间聚合将深度神经网络学习到的内部与边界特征图的中层语义信息集成在一起,利用由粗到细的定位的分割不断地裁剪输入图像,使得最终结果优于他们在 2015 年提出的传统方法^[25],也成为后续两阶段胰腺分割模型的分割结果的对比标准.随后,Zhou 等人、Yu 等人首次提出了粗阶段定位与细阶段分割不断循环迭代的定点模型^[29,30],在粗分割阶段采用启发式裁剪策略处理 3 个轴的 2D 切片,在细分割阶段训练另一个分割网络,并以粗分割结果作为输入,最后输出细化分割的结果,这种方法在细分割阶段不断迭代后得到较好的分割结果.但由于该方法是基于 3 个轴的 2D 切片进行裁剪,两阶段的模型均是使用的 2D 模型,会丢失特征图之间的空间关联性.因此,Zhu 等人将连续的 3D 模型^[31]应用到由粗到细的两阶段胰腺分割上对输入图像进行反复裁剪,并采用残差连接的 Res-Net^[13]结构进一步提升了分割效果.相较于 2D 模型,该方法可以将 3D 特征中 3 个轴中丰富的空间信息完全利用,对分割结果的完整性有一定的提升.在近期最新的研究中,Asaturyan 等人^[32]提出了一种新的基于 Hausdorff 距离和正弦分量的损失函数来替换主流的 Dice Loss,并使用 3D 能量最小化算法对与粗定位融合后的输出进行精细分割.该方法可以有效的捕获局部边界信息,强化对边缘细节的检测,在连续 2D 切片间产生更大的空间平滑性和预测一致性,从而提高了在 2D 切片上逐片预测的精确度,在作者获取的两个 MRI 胰腺数据集上取得了优秀的分割结果.此外,Man 等人提出了以 DQN 驱动的可变形 U-Net 方法^[33],通过显式地与上下文信息交互,从胰腺中提取各向异性特征,以产生一个视觉上更紧、更精确的胰腺定位框输送到细分割进行精分割,以此提高复合模型在胰腺分割的整体精度.

对于基于深度学习的分割方法,两阶段分割方法虽然对输入的 CT 图像裁剪后进行训练测试,虽然裁剪后得到的感兴趣区域 (region of interest, ROI) 分辨率有所降低,但不断地迭代运算使得计算成本大幅度增加.此外,粗略裁剪在一定程度上造成了重要信息的丢失,影响分割图像的完整性,并且对胰腺区域进行预裁剪并不是真正意义上提高了分割效果,而是降低了分割难度,通过这类方法并不能提升网络对复杂器官的分割性能.其次,两阶段分割方法由于粗分割阶段决定着细分割阶段的输入,因此最终的分割性能在很大程度上取决于粗定位结果.如果在粗分割阶段遗漏了目标分割区域或生成了错的目标区域尺寸,就会给整个模型带来无法挽回的损失.此外,这类

由粗到细的两阶段分割方法在测试阶段需要大量地迭代来更新 ROI, 即使模型训练完成后也无法直接使用, 因为细阶段得到最终分割结果的前提是: 网络的输入是精准裁剪后的胰腺区域. 而这一过程是在测试中不断迭代确定的, 即还需要将结果不断地返回到模型输入进行多轮的 ROI 更新才能得到最终的输出. 这类方法不仅需要消耗较长的诊断时间而且增加了额外的计算成本, 提高了实际应用难度.

针对两阶段分割方法在实际临床诊断中存在的不足, 单阶段胰腺分割方法的临床应用前景更为宽广. 虽然目前单阶段分割方法的分割精度与多阶段模型还存在差距, 但模型结构简单、分割速度快、对设备计算能力要求低, 更符合当前的需要. 因此, 本文面向单阶段模型提出了一种全新的双解码 U 型网络结构 (dual-decoding U-Net, DDU-Net) 来加强对高级语义信息以及空间信息的利用, 以生成最终的预测图. 该网络结构由一个编码器和两个解码器构成, 解码器 1 通过深层解码结构提取高级语义特征, 解码器 2 得益于自身浅层解码的特性使得每层输出的特征包含丰富的空间特征, 最后将解码器 2 输出的中低级特征与解码器 1 得到的高级特征逐层融合, 能够提升特征提取精准性的同时保留更多的空间特征, 最终实现预测图的精确输出. 此外, 针对胰腺器官较小的问题, 本文并没有采用上述两阶段分割方法将原输入 CT 切片的 ROI 区域裁剪出来, 而是保持全尺寸 CT 切片的输入, 这样能够保证全尺寸下原 CT 切片的信息完整性, 网络可以利用胰腺周围器官的上下文信息来辅助胰腺区域的分割. 且增加了裁剪步骤, 对于采用人工方式进行裁剪不仅耗时还对裁剪人员有较高经验要求. 对于采用增加裁剪模型的二阶段分割方法, 在裁剪中不断地迭代大大增加了计算量. 此外, 在实际应用中可直接对病人腹部 CT 进行分割, 没有中间裁剪步骤, 是真正意义上的端到端模型. 最后, 本文对所提出网络的实际表现进行评估, 采用公开 TCIA 胰腺 CT-82 作为评价基准, 根据 4 折交叉验证得到的平均 DSC 结果表明: 本文提出的 DDU-Net 在全尺寸 CT 切片的输入下提升了分割效果, 相较于 U-Net 网络 DSC 有 3% 的提升, 在不需要多个阶段的卷积神经网络模型由粗到细对胰腺区域进行裁剪的情况下实现了准确分割, 具有一定的临床应用价值.

本文第 1 节将详细介绍本文提出的网络模型. 第 2 节首先验证本文提出网络的有效性, 然后通过多组实验验证和对比本文提出的单阶段分割模型与主流的单阶段模型方法. 最后, 第 3 节是全文的总结.

1 方 法

本文提出一种基于双解码 U 型卷积神经网络 DDU-Net 的单阶段胰腺分割方法, 具体的框架如图 1 所示.

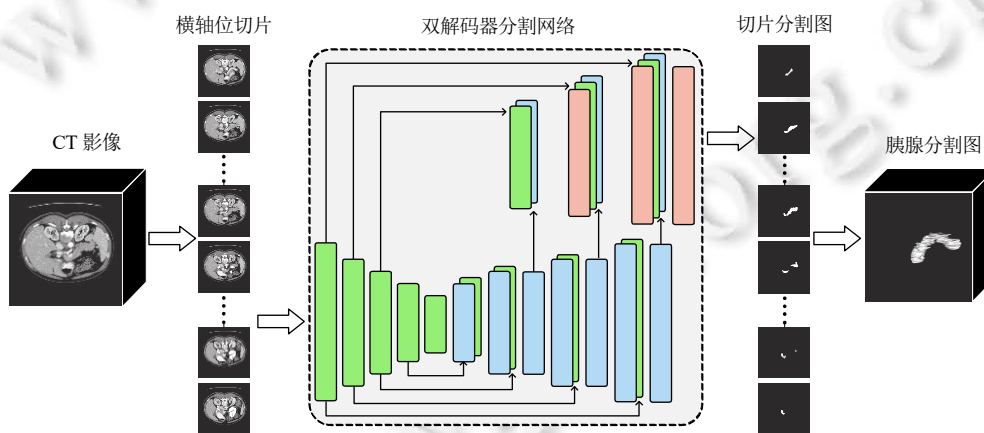


图 1 一种基于双解码 U 型卷积神经网络的胰腺分割方法的框架流程图

网络的输入是腹部 CT 横轴位的切片, 分割网络的主体结构中的编码器与 U-Net 网络结构类似. 在此基础上设计了两个不同的解码器分支, 解码器 1 通过多次卷积、上采样以及与编码器输出的特征进行通道合并, 将得到的每个尺度的特征图与解码器 2 逐层进行融合, 通过卷积和上采样得到胰腺部位的预测图切片, 最后将所有切片恢复成完整的三维 CT 分割图. 本文方法所提出的模型属于单阶段分割模型, 在网络的输入阶段无需对原 CT 切片进行裁剪或降低分辨率, 始终保持原始的输入, 不仅训练成本低, 模型胰腺区域预测速度快, 还保证了较高的分割

精度, 提升了分割图的完整性.

1.1 双解码 U 型网络 DDU-Net

U-Net^[27]是由 Ronneberger 等人在 2015 年提出, 采用对称的编码-解码结构在医学图像分割领域已经被证实效果显著, 尤其通过跳跃连接将编码器特征与解码器特征在通道上进行特征融合, 这种融合方法有效地了解码器在上采样过程中特征信息不足的问题, 大幅度提升了分割网络性能. 但对于一些小器官分割任务, 仅使用跳跃连接的方式加强对特征的重用对分割效果提升有限, 尤其在网络的解码器部分, U-Net 通过编码器得到高级语义特征区域较小, 在逐层上采样的过程中与内容复杂的低级特征的融合时特征之间跳跃较大, 导致低级特征没有得到有效利用, 因此小器官分割的任务中效果不佳. 针对这一问题, 本文对 U-Net 进行了重新设计, 提出了一个包含两个解码器分支的 DDU-Net, 通过不同网络深度的解码器提取的高级特征信息来加强对浅层低级特征信息的高效融合, 稳定提升对小器官的分割精确程度. 结构如图 2 所示 (其中, EnBr 表示编码器, DeBr1 与 DeBr2 分别代表解码器分支 1 与解码器分支 2).

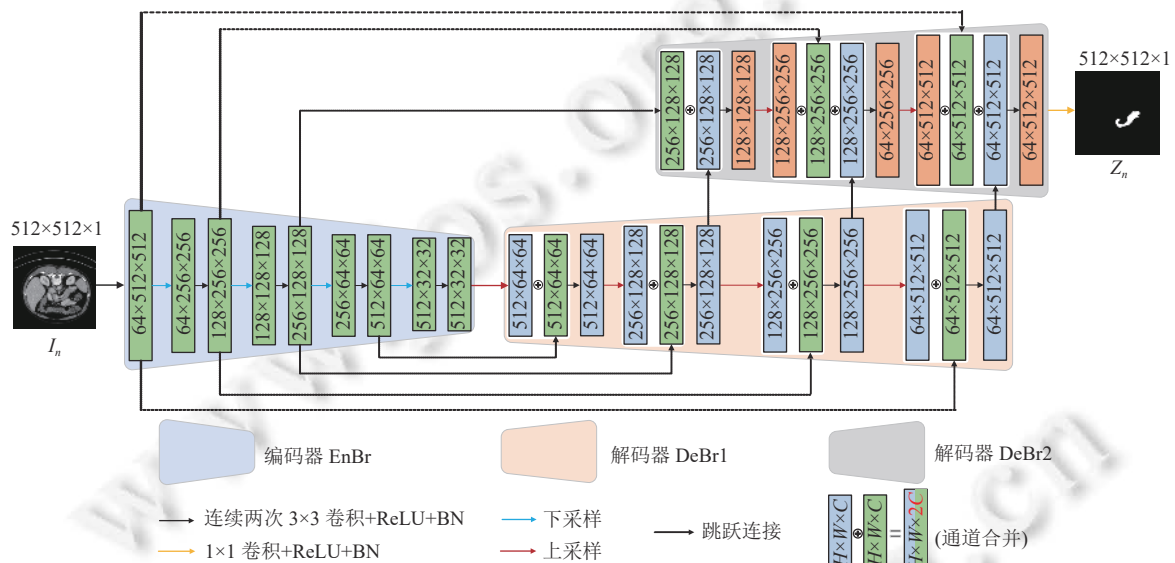


图 2 双解码 U 型网络

本文的网络结构由一个编码器 EnBr 和两个解码器 DeBr1, DeBr2 构成, 在编码器部分与标准 U-Net 结构一致, 包含相同的卷积层数与下采样次数, 每层由两次连续的 3x3 卷积组成, 在每个卷积后都有一个线性修正单元 (ReLU) 和批标准化 (BN) 操作, 每一层结束后通过步长为 2 的最大池化进行下采样, 将特征图的长和宽分别降低到原来的一半; 同时, 通过对通道进行翻倍来降低下采样操作造成的信息损失. 解码器部分由两个分支构成, 其中, 解码器 DeBr1 对编码器最后一层输出的特征图逐级进行上采样; 同时, 特征图的通道数减半, 得到尺寸为 64x64x512 的特征图. 为了增加解码器 DeBr1 的特征信息, 通过跳跃连接将编码器 EnBr 中对应层的特征图在通道上叠加 (图 2 中 EnBr 与 DeBr1 之间的虚线连接). 在解码器 DeBr2 部分, 将编码器对应层输出的特征图 (128x128x256) 与解码器 DeBr1 对应层输出的特征图在通道上进行合并作为解码器 DeBr2 的开始输入. 最后, 依次进行两次连续卷积、ReLU、批标准化以上采样操作得到每层的输出, 再与对应层编码器的输出、解码器 DeBr1 的输出在特征通道上合并, 后面的每层重复该项操作, 得到最终的分割图.

1.2 双解码器 DeBr1 与 DeBr2

人体腹部 CT 中包含了多个器官和复杂的组织, 如图 3 所示, 胰腺区域占整个 CT 切片比例较小.

基于胰腺器的这一特点, 不同于利用多阶段模型拆分分割任务, 本文将利用卷积神经网络本身的特点实现胰

腺分割. 为了使分割网络集中提取目标区域而尽可能不引入其他区域特征, 传统的做法是通过加深网络的深度. 在一定范围内, 层数越深提取的特征越高级, 通过多次的下采样, 能逐层丢弃复杂的冗余信息, 提取的目标区域特征就越准确. 但同时, 过多的下采样使得浅层的边缘等空间信息不断被丢弃, 虽然提高了胰腺区域的定位精度, 但会使最终的预测图丢失大量边缘信息 (如图 4(a) 中特征图 D3 所示).

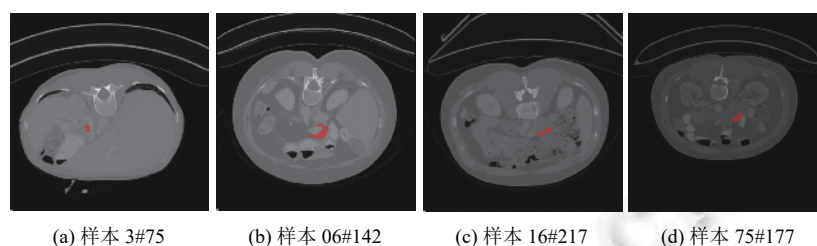


图 3 腹部 CT 的横轴位切片图, 图中红色区域为胰腺部位

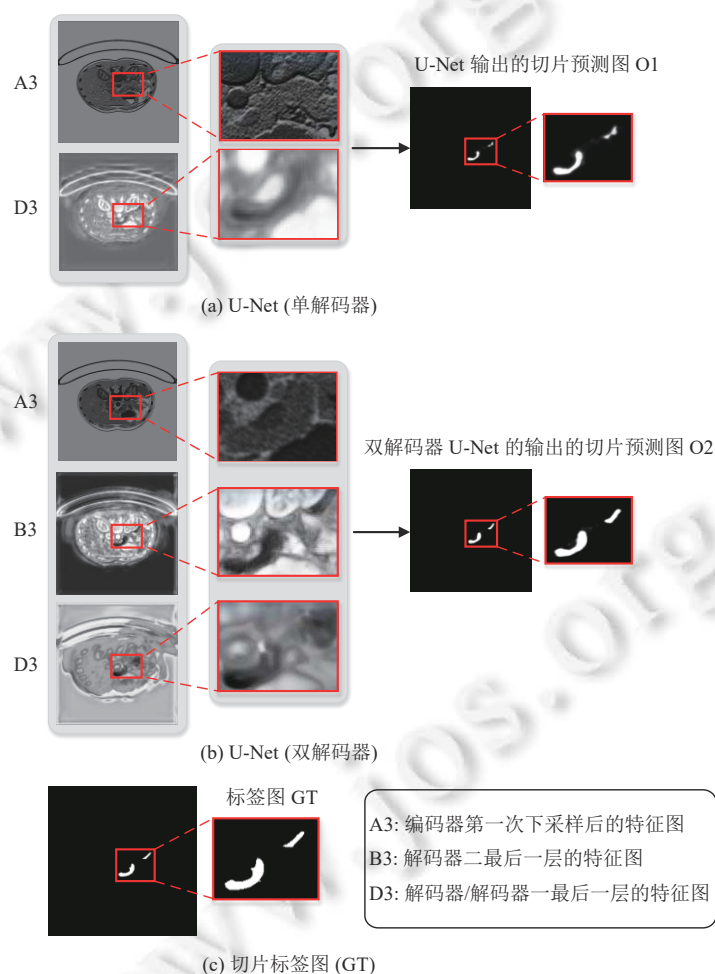


图 4 单解码器与双解码器最后一层的特征可视化, 图中红色区域为胰腺部位

U-Net 在编码器与解码器之间通过跳跃连接能够增加浅层的边缘信息 (如图 4(a) 中特征图 A3 与特征图 D3 融合), 但通过图 4(a) 中特征图 A3 可以发现: 编码器阶段的浅层特征依然伴有较多的其他器官组织信息, 虽然进

行了特征融合,但高低特征融合时特征图之间存在较大跳跃影响了融合效果,即编码器每层得到的高级特征区域过小,而编码器通过跳跃连接与其融合的浅层特征包含的特征区域过多,最终导致分割效果不佳(如图4(a)中预测图O1)。因此,从分割网络自身特点出发,本文提出在网络中增加一个解码器,在浅层特征与高层特征中间通过特征融合生成一组介于两者之间的特征信息(如图4(b)中B3),该解码器DeBr2在每层对编码器与解码器DeBr1输出的特征进行融合的同时,也作为独立的分支,即使在分割较小目标时,也能使特征融合更加完整(如图4(b)中预测图O2)。

在本文提出的DDU-Net中,编码器EnBr与解码器DeBr1构成的小网络结构可以看作是标准U-Net,而解码器DeBr2相较于解码器DeBr1则是更浅一层的解码器。DeBr2在逐层融合解码器对应的低级特征的同时还融合来自解码器DeBr1的深层高级语义特征,最后通过DeBr2生成分割网络的预测图。通过这种结构,解码器DeBr2除去开始层外,每层都包括三种不同语义级别的特征进行特征融合。其中,编码器DeBr2的特征图参与了通道间合并,较浅的编码器与解码器组合含有更丰富的浅层特征,能保留更多的空间信息,加入了DeBr1后,与更深一层的编码器输出相连,每层只逐级向下传递更高层的语义特征。在解码器DeBr2分支上的逐层合并增强了高低特征之间融合效果,这样使得网络在保留足够浅层特征的同时,还能够获取到更高层语义特征,实现在分割较小胰腺区域时既能精确定位到胰腺部位又能保留足够多的边缘特征,最终提升小目标器官分割的精度。

1.3 双解码器的特征融合方式

解码器部分由DeBr1与DeBr2构成,其中,解码器DeBr1每次上采样得到的特征图都通过跳跃连接的方式在通道上与DeBr2中对应层的特征图合并。图5(a)即是本文提出的双解码器特征融合方式。为了验证DeBr1每层特征的有效性以及与编码器逐层特征融合的必要性,分别进行如图5中描述的4组实验。

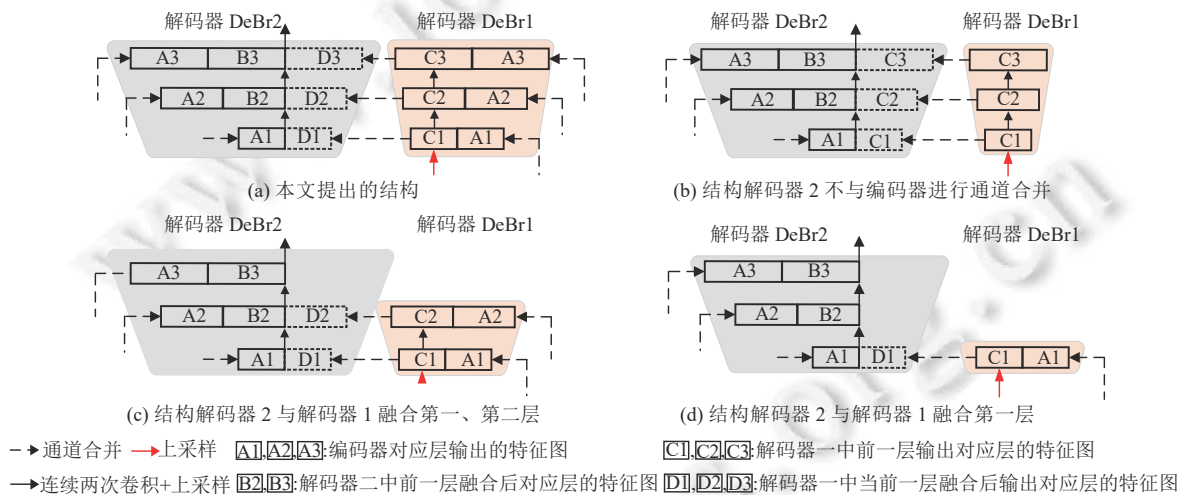


图5 双解码器不同连接方式结构图

在图5列出的结构中,可将实验分为两组:第1组包含结构(a)、结构(b)两个结构,验证在解码器DeBr1中逐层融合编码器特征的必要性,其中,结构(a)为本文方法采用的结构。从表1的实验结果可以得出:结构(a)的分割效果优于(b)结构,进一步说明解码器DeBr1在对编码的特征进行解码时融合对应编码器层的特征有助于分割性能的提升。在此基础上进行第2组实验,通过结构(a)~结构(c)这3种结构来逐一判断解码器DeBr1中所有输出层的有效性。从表1可以看出:分割网络的性能与解码器DeBr2与DeBr1的融合层数呈正相关,即解码器DeBr2与DeBr1每层特征都进行融合时网络的分割性能最佳,同时也验证了解码器DeBr1每层得到的特征对整体分割网络都有贡献,一定程度对解码器DeBr2有信息补充作用。为了更为直观地验证这一结论,在结构(a)的基础上,分别将网络中A1, D1, B2, D2, B3, D3层的特征图输出,通过特征图可视化的方法得到各级特征的可视图如图6所示。通过图6可以发现:解码器DeBr1中的D1, D2, D3特征包含更多的高级语义特征,对目标区域的指向性

更好, 而编码器部分的特征低级特征丰富, 例如图 6(a). 因此, 解码器 DeBr2 中每层特征都进行高低层特征融合, 尤其在最后两层融合了 3 种不同的特征, 包括编码器中的低级特征、上一层编码的得到的混合特征、解码器 DeBr1 的高级特征, 这些多级特征的融合提升了最终分割图的完整性与准确性.

表 1 双解码器不同连接方式的实验结果

连接方式	DSC	Precision	Recall
方式(a)	0.751±0.12	0.788	0.760
方式(b)	0.719±0.17	0.789	0.710
方式(c)	0.743±0.14	0.799	0.736
方式(d)	0.738±0.15	0.781	0.740

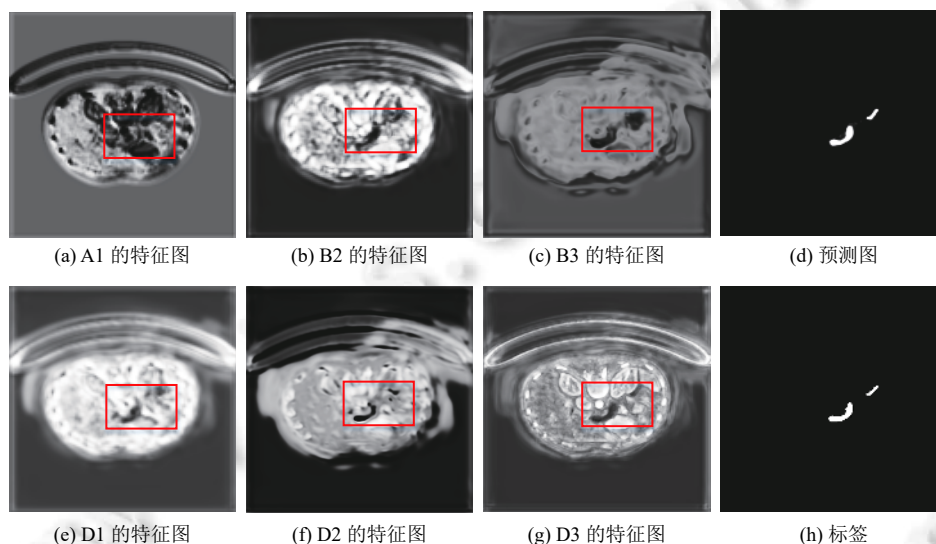


图 6 双解码器中特征可视化结果, 图中红色框区域表示胰腺所在部位

1.4 损失函数

小器官分割一直是医学图像分割中的难题, 主要挑战在于正负样本分布不平衡, 使用普通的自然图像分割损失函数效果表现不佳. 在本文的胰腺分割任务中, 采用医学图像分割中表现较好的 Dice Loss. 对于任意输入的胰腺 CT 切片 I_n , 经过分割网络的编码器与逐层特征融合的双解码器最终得到预测的分割图 $\hat{Z}_n$. 损失函数如公式 (1)^[27]所示:

$$L_{\text{Dice}}(Z_n, \hat{Z}_n) = -\frac{1}{N} \sum_{n=1}^N \left(\frac{2 \cdot Z_n \cdot \hat{Z}_n}{Z_n + \hat{Z}_n} \right) \quad (1)$$

其中, $L_{\text{Dice}}(Z_n, \hat{Z}_n)$ 表示损失函数, Z_n , $\hat{Z}_n$ 分别为第 n 张切片的标签图与的预测图的一维展开向量, N 表示批大小, n 表示切片编号.

2 实验结果与分析

为了证明双分支解码器对于分割任务的有效性, 并将本文提出的分割网络与已有的单阶段分割方法以及经典的二阶段方法进行对比和分析, 本节将设计实验分别进行验证. 为了确保实验对比的公平性和可信度, 本节所涉及的一阶段对比方法均使用作者公开的原始代码, 并且在目前公开的最大的胰腺数据集上进行训练和验证.

2.1 实验数据

本文数据集是由美国国立卫生研究院 (National Institutes of Health, NIH) 公开的胰腺数据集, 该数据集是目前

公开的最大规模胰腺分割数据集, 共包含 82 例腹部增强 CT 和对应的由人工逐层标注的标签, CT 的分辨率为 $512 \times 512 \times L$, L 是沿着 CT 横轴面切片的数量, 其中, $L \in [181, 466]$. 在对数据的处理中, 首先将 82 例 CT 样本随机分为 A (20 例), B (20 例), C (21 例), D (21 例) 4 组进行四折交叉验证, 并沿横轴面将每折 CT 划分成 512×512 大小的切片. 最后, 将不含有标注的切片剔除只保留包含胰腺区域的有效切片作为原始数据. 此外, 为了降低数据集样本量较小的影响, 本文在每折训练样本中随机抽取 2 000 张切片分别进行高斯模糊与高斯噪声数据增广, 详细的数据集划分与增广操作如表 2 所示. 最终, 本文实验数据集由 11 003 张切片构成.

表 2 实验数据

数据划分	训练数据						测试数据			
	组合方式	CT例数	切片数	数据增广		总训练切片数	组合方式	CT例数	有效切片数	总数据
				高斯噪声: $Ksize=3$ $SigmaX=2$ $SigmaY=2$	高斯模糊: $Mean=0$ $Variance=0.01$					
第1折	B+C+D	62	5 304	2 000	2 000	9 304	A	20	1 699	11 003
第2折	A+C+D	62	5 276	2 000	2 000	9 276	B	20	1 727	
第3折	A+B+D	61	5 084	2 000	2 000	9 084	C	21	1 919	
第4折	A+B+C	61	5 345	2 000	2 000	9 345	D	21	1 658	

2.2 评价指标

本文的实验部分利用 DSC 作为评价指标, 验证分割方法的有效性, 其公式如公式 (2) 所示:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \tag{2}$$

其中, X, Y 分别表示预测图像像素集合与标签图像像素集合. 此外, 精确值 (precision) 和召回率 (recall) 也作为辅助参考指标: 精确率为预测为胰腺区域中真实胰腺区域的比例值, 召回率为真实胰腺区域中被预测为胰腺区域的比例值.

2.3 实现细节

网络的参数设定对分割的效果起着至关重要的作用, 本文提出的 DDU-Net 参数设置见表 3.

表 3 超参数设置

类型	优化器Adam	迭代次数	输出图像尺寸	批大小	卷积	最大池化	上采样
超参数	$Lr=0.0001$ $Beta_1=0.9$ $Beta_2=0.99$	50	512×512	4	卷积核为 3×3 步长为1	滤波器为 2×2 步长为2	双线性插值

本文所有实验的梯度优化均使用 Adam 优化器, 批大小设置为 4, 总迭代次数为 50, 学习率为 0.000 1. 最后, 训练模型的硬件平台是 Tesla V100 服务器, 使用一块 V100 GPU 训练一次模型需要花费 5 小时.

2.4 本文方法与其他分割网络方法的对比

在验证网络模型的分割结果时, 本文使用标准的 U-Net^[27], Attention U-Net^[26]以及 U-Net++^[34]作为主要对比的分割网络模型, 分别从 3 个不同的角度将本文提出的网络模型与以上 3 个分割网络进行对比: 首先, U-Net 作为医学图像经典的分割网络在多个分割场景中都有较好的效果, 且近几年提出的多个分割网络也都是基于 U-Net 进行创新, 因此作为实验的基础标准; 其次, Attention U-Net 作为 U-Net 的改进模型在分割网络中成功的将注意力机制与分割网络相结合来提升模型的分割性能, 同时也是针对胰腺分割而提出, 与本文提出的方法的分割任务一致; 而 U-Net++在结构上有着更大的创新, 可以看作具有多个分支编码器与解码器组合的 U-Net 网络; 最后, 同时也引入了经典的二阶段分割方法 FPM^[29], 采用定点模型由粗到细分成两个阶段训练模型, 大幅度提升了胰腺分割效果.

在胰腺分割任务中, 将 U-Net 作为基准的对比模型, 不同分割方法对应的 DSC 见表 4. 可以看出: U-Net 作为

基准网络模型表现出了稳定的分割性能, 不仅模型轻量化, 训练时长与计算量都保持最小, 但在面对小型器官分割所得到的召回率较低, 即丢失了大量的分割区域的完整性, 分割效果最差; U-Net++虽然相较于 U-Net 有一定的提升, 但是在分割结果上存在同样的问题, 召回率依然较低, 且模型复杂度增加了接近一倍; Attention U-Net 在各方面比较均衡, 注意力机制的加入, 使得在全尺寸 CT 图像作为输入下能够精确地关注到目标分割区域, 但整体略低于本文分割效果; 本文分割效果在双解码器的作用下, 加强了高级特征对低级特征的指导 使得模型对目标分割区域的指向性大大提升, 且模型的计算成本仅小幅度增加. 需要说明的是: 所有的结果都是在全尺寸、未裁剪的 CT 图像作为输入的情况下得到的. 对于二阶段分割方法 FPM, 将分割任务拆分为第 1 阶段的粗裁剪与第 2 阶段的细分割, 在分割结果上较一阶段分割方法有较大幅度提升, 但模型的训练时长、生成切片预测图时长与计算量均有多倍的增加, 尤其每张切片的预测图的生成时长是一阶段分割方法的 10 倍. 二阶段分割方法计算量大的主要原因在于需要训练两个模型, 且细分割模型在测试时需要反复训练迭代 10 次才能得到较好的精度. 此外, 该方法采用了 CT 图像 3 个轴的切片每阶段同时训练 3 个模型, 导致计算量成倍增加. 虽然该二阶段方法对分割效果有一定贡献, 但巨大的计算量对于临床应用场景局限性太大.

表 4 本文方法与其他分割方法的对比实验结果

方法		DSC	Precision	Recall	训练时长	生成每张预测图 时长 (ms)	计算量 (GFLOPs)
二阶段	FPM ^[29]	0.802±0.09	0.814	0.831	14 h 15 m	577	635.5
	U-Net ^[27]	0.719±0.18	0.763	0.712	3 h 35 m	35	111.3
一阶段	U-Net++ ^[34]	0.727±0.14	0.779	0.724	5 h 39 m	71	222.7
	Attention U-Net ^[26]	0.748±0.09	0.770	0.768	3 h 57 m	39	142.1
	DDU-Net	0.751±0.12	0.788	0.760	4 h 55 m	46	165.8

为了进一步与其他方法对比分割效果, 在图 7 中进行了的多个样本的分割效果图比较 (图中每行代表不同的切片样本及其分割结果; (a) 列分别表示对应编号的切片图; (b) 列表示对应编号切片的标签图; (c)–(g) 列分别表示 5 组不同方法在对应编号的切片图分割后的预测图).

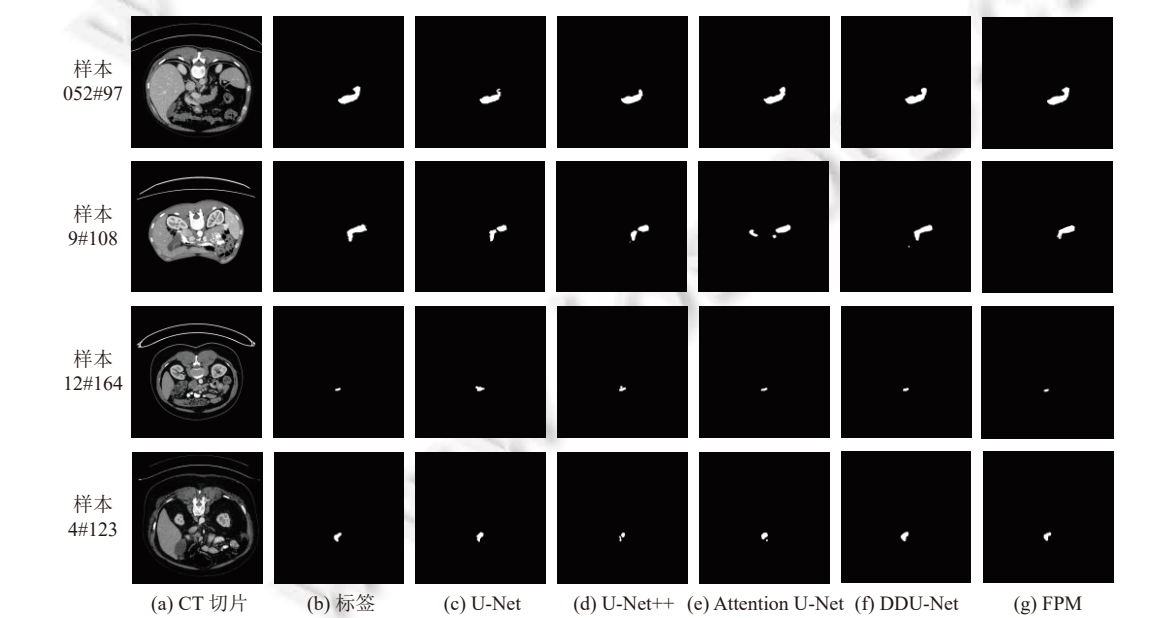


图 7 本文方法与其他分割方法的分割结果

通过分割图的对比发现: 当胰腺区域较大且相较于周围组织界限明显的情况下, 如图 7 中第 1 行样本 052#97 (在 NIH 数据集中第 52 号病人的第 97 张切片图), 5 种分割网络都保持了较高的分割精度, 说明 5 种分割方法对此类样本都能有效分割; 但在目标分割区域与周围组织成像相近时, 如图 7 样本 9#108, 5 种方法的分割效果都有所下降, 其中, Attention U-Net 误分割出多余的其他区域, 而本文提出的 DDU-Net 方法依然维持了分割图的结构完整性. 最后, 对于胰腺分割任务, 目标区域小的样本在 3D CT 的切片中占大部分, 这类样本在不被裁剪的情况下, 对分割网络模型有着较高的性能要求, 如图 7 中样本 12#164 与 4#123. 在这种情况下, 普通的 U-Net 网络分割效果较差, 在分割图的形状上与标签差距较大, 虽然 U-Net++ 要比 U-Net 的分割效果略有提升, 但与标签相比依然误差较大. 另外, Attention U-Net 在目标分割区域形状简单的情况下能够实现精确分割 (如图 7 中样本 12#164 对应的 (e) 列), 但是在形状复杂的样本下会将目标分割区域断开 (如图 7 中样本 4#123 对应的 (e) 列). 本文提出的 DDU-Net 在小目标分割区域下能够保证较高的分割完整性, 且目标区域形状与标签相近. 最后, 二阶段分割方法 FPM 分割效果如图 7 中 (g) 列所示, 分割效果比一阶段方法略有提升.

3 总 结

本文提出了一种基于双解码 U 型卷积神经网络的胰腺分割方法, 在 U 型分割网络的基础上增加了双解码器分支, 分别提取不同网络深度的特征信息, 然后在对应尺度下对特征进行融合, 加强了高级语义特征对低级特征的指导, 丰富了不同尺度下的特征信息, 达到对腹部 CT 影像中胰腺区域的精准分割. 该方法在单阶段模型以及未裁剪 CT 作为网络输入的前提下具有较强的分割能力, 适用于复杂组织下的小器官区域分割. 本文提出的分割方法属于单阶段方法, 在保持较高的分割精度的情况下, 实现了在完整 CT 图像作为输入下对胰腺进行高精度分割的端到端分割, 不仅无需对输入 CT 影像进行裁剪, 且模型训练测试简单、对测试设备要求低. 在未来, 我们将在该分割网络结构下融入其他模块, 实现针对更多器官进行分割的方法.

References:

- [1] Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834–848. [doi: [10.1109/tpami.2017.2699184](https://doi.org/10.1109/tpami.2017.2699184)]
- [2] Ren SQ, He KM, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. [doi: [10.1109/tpami.2016.2577031](https://doi.org/10.1109/tpami.2016.2577031)]
- [3] He KM, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. In: *Proc. of 2017 IEEE Int'l Conf. on Computer Vision*. Venice: IEEE, 2017. 2980–2988. [doi: [10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322)]
- [4] Ling HB, Zhou SK, Zheng YF, Georgescu B, Suehling M, Comaniciu D. Hierarchical, learning-based automatic liver segmentation. In: *Proc. of the 2008 IEEE Conf. on Computer Vision and Pattern Recognition*. Anchorage: IEEE, 2008. 1–8. [doi: [10.1109/cvpr.2008.4587393](https://doi.org/10.1109/cvpr.2008.4587393)]
- [5] Heimann T, van Ginneken B, Styner MA, et al. Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans. on Medical Imaging*, 2009, 28(8): 1251–1265. [doi: [10.1109/TMI.2009.2013851](https://doi.org/10.1109/TMI.2009.2013851)]
- [6] Linguraru MG, Sandberg JK, Li ZX, Shah F, Summers RM. Automated segmentation and quantification of liver and spleen from CT images using normalized probabilistic atlases and enhancement estimation. *Medical Physics*, 2010, 37(2): 771–783. [doi: [10.1118/1.3284530](https://doi.org/10.1118/1.3284530)]
- [7] Lin DT, Lei CC, Hung SW. Computer-Aided kidney segmentation on abdominal CT images. *IEEE Trans. on Information Technology in Biomedicine*, 2006, 10(1): 59–65. [doi: [10.1109/titb.2005.855561](https://doi.org/10.1109/titb.2005.855561)]
- [8] Ali AM, Farag AA, El-Baz AS. Graph cuts framework for kidney segmentation with prior shape constraints. In: *Proc. of the 10th Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*. Brisbane: Springer, 2007. 384–392. [doi: [10.1007/978-3-540-75757-3_47](https://doi.org/10.1007/978-3-540-75757-3_47)]
- [9] Chu CW, Oda M, Kitasaka T, Misawa K, Fujiwara M, Hayashi Y, Nimura Y, Rueckert D, Mori K. Multi-organ segmentation based on spatially-divided probabilistic atlas from 3D abdominal CT images. In: *Proc. of the 16th Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention*. Nagoya: Springer, 2013. 165–172. [doi: [10.1007/978-3-642-40763-5_21](https://doi.org/10.1007/978-3-642-40763-5_21)]
- [10] Roth H, Oda M, Shimizu N, Oda H, Hayashi Y, Kitasaka T, Fujiwara M, Misawa K, Mori K. Towards dense volumetric pancreas

- segmentation in CT using 3D fully convolutional networks. In: Proc. of the SPIE 10574, Medical Imaging 2018: Image Processing. Houston: SPIE, 2018. 105740B. [doi: [10.1117/12.2293499](https://doi.org/10.1117/12.2293499)]
- [11] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 2017, 60(6): 84–90. [doi: [10.1145/3065386](https://doi.org/10.1145/3065386)]
 - [12] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014.
 - [13] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90)]
 - [14] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proc. of the 2014 IEEE Conf. on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 580–587. [doi: [10.1109/cvpr.2014.81](https://doi.org/10.1109/cvpr.2014.81)]
 - [15] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 3431–3440. [doi: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965)]
 - [16] Tian X, Wang L, Ding Q. Review of image semantic segmentation based on deep learning. *Ruan Jian Xue Bao/Journal of Software*, 2019, 30(2): 440–468 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5659.htm> [doi: [10.13328/j.cnki.jos.005659](https://doi.org/10.13328/j.cnki.jos.005659)]
 - [17] Zhou FY, Jin LP, Dong J. Review of convolutional neural network. *Chinese Journal of Computers*, 2017, 40(6): 1229–1251 (in Chinese with English abstract). [doi: [10.11897/SP.J.1016.2017.01229](https://doi.org/10.11897/SP.J.1016.2017.01229)]
 - [18] Wang ZH, Bhatia KK, Glocker B, Marvao A, Dawes T, Misawa K, Mori K, Rueckert D. Geodesic patch-based segmentation. In: Proc. of the 17th Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention. Boston: Springer, 2014. 666–673. [doi: [10.1007/978-3-319-10404-1_83](https://doi.org/10.1007/978-3-319-10404-1_83)]
 - [19] Wolz R, Chu CW, Misawa K, Fujiwara M, Mori K, Rueckert D. Automated abdominal multi-organ segmentation with subject-specific atlas generation. *IEEE Trans. on Medical Imaging*, 2013, 32(9): 1723–1730. [doi: [10.1109/tmi.2013.2265805](https://doi.org/10.1109/tmi.2013.2265805)]
 - [20] Modat M, McClelland J, Ourselin S. Lung registration using the NiftyReg package. In: van Ginneken B, Murphy M, Heimann T, Pekar V, Deng X, eds. *Medical Image Analysis for the Clinic—A Grand Challenge*. CreateSpace Independent Publishing Platform, 2010. 33–42.
 - [21] Wang HZ, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA. Multi-atlas segmentation with joint label fusion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2013, 35(3): 611–623. [doi: [10.1109/TPAMI.2012.143](https://doi.org/10.1109/TPAMI.2012.143)]
 - [22] Wang L, Shi F, Li G, Gao YZ, Lin WL, Gilmore JH, Shen DG. Segmentation of neonatal brain MR images using patch-driven level sets. *NeuroImage*, 2014, 84: 141–158. [doi: [10.1016/j.neuroimage.2013.08.008](https://doi.org/10.1016/j.neuroimage.2013.08.008)]
 - [23] Bai WJ, Shi WZ, O'Regan DP, Tong T, Wang HY, Jamil-Copley S, Peters NS, Rueckert D. A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images. *IEEE Trans. on Medical Imaging*, 2013, 32(7): 1302–1315. [doi: [10.1109/tmi.2013.2256922](https://doi.org/10.1109/tmi.2013.2256922)]
 - [24] Murphy K, van Ginneken B, Reinhardt JM, et al. Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge. *IEEE Trans. on Medical Imaging*, 2011, 30(11): 1901–1920. [doi: [10.1109/TMI.2011.2158349](https://doi.org/10.1109/TMI.2011.2158349)]
 - [25] Roth HR, Lu L, Farag A, Shin HC, Liu JM, Turkbey EB, Summers RM. DeepOrgan: Multi-level deep convolutional networks for automated pancreas segmentation. In: Proc. of the 18th Int'l Conf. on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015. 556–564. [doi: [10.1007/978-3-319-24553-9_68](https://doi.org/10.1007/978-3-319-24553-9_68)]
 - [26] Oktay O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D. Attention U-Net: Learning where to look for the pancreas. arXiv: 1804.03999, 2018.
 - [27] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proc. of the 18th Int'l Conf. on Medical Image Computing and Computer-assisted Intervention. Munich: Springer, 2015. 234–241. [doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28)]
 - [28] Roth HR, Lu L, Farag A, Sohn A, Summers RM. Spatial aggregation of holistically-nested networks for automated pancreas segmentation. In: Proc. of the 19th Int'l Conf. on Medical Image Computing and Computer-assisted Intervention. Athens: Springer, 2016. 451–459. [doi: [10.1007/978-3-319-46723-8_52](https://doi.org/10.1007/978-3-319-46723-8_52)]
 - [29] Zhou YY, Xie LX, Shen W, Wang Y, Fishman EK, Yuille AL. A fixed-point model for pancreas segmentation in abdominal CT scans. In: Proc. of the 20th Int'l Conf. on Medical Image Computing and Computer-assisted Intervention. Quebec City: Springer, 2017. 693–701. [doi: [10.1007/978-3-319-66182-7_79](https://doi.org/10.1007/978-3-319-66182-7_79)]
 - [30] Yu QH, Xie LX, Wang Y, Zhou YY, Fishman EK, Yuille AL. Recurrent saliency transformation network: Incorporating multi-stage visual cues for small organ segmentation. In: Proc. of the 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 8280–8289. [doi: [10.1109/cvpr.2018.00864](https://doi.org/10.1109/cvpr.2018.00864)]
 - [31] Zhu ZT, Xia YD, Shen W, Fishman E, Yuille A. A 3D coarse-to-fine framework for volumetric medical image segmentation. In: Proc. of the 2018 Int'l Conf. on 3D Vision. Verona: IEEE, 2018. 682–690. [doi: [10.1109/3dv.2018.00083](https://doi.org/10.1109/3dv.2018.00083)]

- [32] Asaturyan H, Thomas EL, Fitzpatrick J, Bell JD, Villarini B. Advancing pancreas segmentation in multi-protocol MRI volumes using hausdorff-sine loss function. In: Proc. of the 10th Int'l Workshop on Machine Learning in Medical Imaging. Shenzhen: Springer, 2019. 27–35. [doi: [10.1007/978-3-030-32692-0_4](https://doi.org/10.1007/978-3-030-32692-0_4)]
- [33] Man YZ, Huang YSB, Feng JY, Li X, Wu F. Deep Q learning driven CT pancreas segmentation with geometry-aware U-Net. IEEE Trans. on Medical Imaging, 2019, 38(8): 1971–1980. [doi: [10.1109/tmi.2019.2911588](https://doi.org/10.1109/tmi.2019.2911588)]
- [34] Zhou ZW, Siddiquee MR, Tajbakhsh N, Liang JM. UNet++: A nested U-Net architecture for medical image segmentation. In: Proc. of the 4th Int'l Workshop on Deep Learning in Medical Image Analysis and the 8th Int'l Workshop on Multimodal Learning for Clinical Decision Support. Granada: Springer, 2018. 3–11. [doi: [10.1007/978-3-030-00889-5_1](https://doi.org/10.1007/978-3-030-00889-5_1)]

附中文参考文献:

- [16] 田莹, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述. 软件学报, 2019, 30(2): 440–468. <http://www.jos.org.cn/1000-9825/5659.htm> [doi: [10.13328/j.cnki.jos.005659](https://doi.org/10.13328/j.cnki.jos.005659)]
- [17] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述. 计算机学报, 2017, 40(6): 1229–1251. [doi: [10.11897/SP.J.1016.2017.01229](https://doi.org/10.11897/SP.J.1016.2017.01229)]



毕秀丽(1982—), 女, 博士, 副教授, CCF 专业会员, 主要研究领域为图像处理, 多媒体安全和图像取证, 医学图像处理.



肖斌(1982—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为图像增强与复原, 图像分析与识别, 数字水印, 信息隐藏, 医学图像处理.



陆猛(1992—), 男, 硕士, 主要研究领域为医学图像分割.



李伟生(1975—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为信道编码, 密码学, 无线和移动安全.