

# 基于循环显著性校准网络的胰腺分割方法

邱成健<sup>1</sup> 刘青山<sup>2,3,4</sup> 宋余庆<sup>1</sup> 刘哲<sup>1</sup>

**摘要** 胰腺的准确分割对于胰腺癌的识别和分析至关重要。研究者提出通过第一阶段粗分割掩码的位置信息缩小第二阶段细分割网络输入的由粗到细分割方法, 尽管极大地提升了分割精度, 但是在胰腺分割过程中对于上下文信息的利用却存在以下两个问题: 1) 粗分割和细分割阶段分开训练, 细分割阶段缺少粗分割阶段分割掩码信息, 抑制了阶段间上下文信息的流动, 导致部分细分割阶段结果无法比粗分割阶段更准确; 2) 粗分割和细分割阶段单批次相邻预测分割掩码之间缺少信息互监督, 丢失切片上下文信息, 增加了误分割风险。针对上述问题, 提出了一种基于循环显著性校准网络的胰腺分割方法。通过循环使用前一阶段输出的胰腺分割掩码作为当前阶段输入的空间权重, 进行两阶段联合训练, 实现阶段间上下文信息的有效利用; 提出卷积自注意力校准模块进行胰腺预测分割掩码切片上下文信息跨顺序互监督, 显著改善了相邻切片误分割现象。提出的方法在公开的数据集上进行了验证, 实验结果表明其改善误分割结果的同时提升了平均分割精度。

**关键词** 胰腺分割, 阶段上下文信息, 切片上下文信息, 卷积自注意力, 校准模块

**引用格式** 邱成健, 刘青山, 宋余庆, 刘哲. 基于循环显著性校准网络的胰腺分割方法. 自动化学报, 2022, 48(11): 2703–2717

**DOI** 10.16383/j.aas.c210865

## Pancreas Segmentation Based on Recurrent Saliency Calibration Network

QIU Cheng-Jian<sup>1</sup> LIU Qing-Shan<sup>2,3,4</sup> SONG Yu-Qing<sup>1</sup> LIU Zhe<sup>1</sup>

**Abstract** Accurate segmentation of the pancreas is very important for the identification and analysis of pancreatic cancer. The researchers proposed a coarse-to-fine segmentation method to reduce the input of the second-stage fine segmentation network through the position information of the first-stage coarse segmentation mask. Although the segmentation accuracy is greatly improved, however, the use of context information in the pancreas segmentation process has the following two problems: 1) The coarse segmentation and fine segmentation stages are trained individually, and the fine segmentation lacks the predicted mask information of the coarse segmentation, which suppresses the flow of context information between stages, resulting in part of the fine segmentation that cannot be more accurate than the coarse segmentation; 2) In the coarse and fine segmentation stage, there is a lack of mutual supervision information between the adjacent predicted masks of a single batch, which leads to the loss of inter-slice context information and increases the risk of false segmentation. To solve the above problems, a pancreas segmentation method based on the recurrent saliency calibration network is proposed. By recurrently using the previous stage output segmentation mask as the spatial weight of the current stage input and performing joint training, the context information between stages is effectively used. Besides, a convolutional self-attention calibration module is suggested, which performs cross-sequence supervision of inter-slice context information and significantly improves the false segmentation. The proposed method is verified on the public datasets, and the experimental results show that it improves the average segmentation accuracy while improving the results of false segmentation.

**Key words** Pancreas segmentation, stage context information, inter-slice context information, convolutional self-attention, calibration module

**Citation** Qiu Cheng-Jian, Liu Qing-Shan, Song Yu-Qing, Liu Zhe. Pancreas segmentation based on recurrent saliency calibration network. *Acta Automatica Sinica*, 2022, 48(11): 2703–2717

收稿日期 2021-09-09 录用日期 2022-03-13

Manuscript received September 9, 2021; accepted March 13, 2022

国家自然科学基金(61976106, 61772242, 61572239), 中国博士后科学基金(2017M611737), 江苏省六大人才高峰计划(DZXX-122), 江苏省研究生科研创新计划(KYCX21\_3374)资助

Supported by National Natural Science Foundation of China (61976106, 61772242, 61572239), China Postdoctoral Science Foundation (2017M611737), Six Talent Peaks Project in Jiangsu Province (DZXX-122), and Graduate Student Scientific Research Innovation Projects in Jiangsu Province (KYCX21\_3374)

本文责任编辑 黄华

Recommended by Associate Editor HUANG Hua

1. 江苏大学计算机科学与通信工程学院 镇江 212013 2. 南京信息工程大学自动化学院 南京 210044 3. 江苏省大数据分析技术重点实验室 南京 210044 4. 大气环境与装备技术协同创新中心 南京 210044

1. School of Computer Science and Communications Engineering, Jiangsu University, Zhenjiang 212013 2. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044 3. Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing 210044 4. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing 210044

胰腺癌具有侵袭性强、转移早、恶性程度高、发展较快、预后较差等特征,根据美国癌症协会报道,其5年生存率低于10%,死亡率非常高<sup>[1]</sup>.胰腺癌已成为严重威胁人类健康的重要疾病,并对临床医学构成巨大挑战.胰腺的准确分割对胰腺癌检测识别等任务起着至关重要的作用.胰腺处于人体后腹部的解剖位置,其脏器影像常被遮挡不易识别,且其形状和空间位置多变,在腹部CT图像中所占比例较小,其准确分割问题亟待解决.

近年来,由于神经网络的发展以及全卷积网络(Fully convolutional network, FCN)<sup>[2]</sup>的出现,医学图像分割准确率取得了较大提升.针对不同患者间胰腺形态差异性较大的解剖特征,基于单阶段深度学习分割算法极易受其较大背景区域影响,导致分割准确率下降.现阶段常用解决方法是基于由粗到细的分割算法<sup>[3-6]</sup>,通过粗分割阶段输出掩码进行定位,只保留胰腺及其周围部分区域作为细分阶段网络输入,减小背景区域对目标区域影响,提高分割精度.由粗到细的分割算法虽然减少了腹部影像背景区域对目标区域的干扰,但是针对形态和空间位置多变的胰腺小器官增强前景区域同样重要.同时粗分割阶段仅保留了定位框的位置信息,却丢失了胰腺输出分割掩码的先验特征信息,从而细分阶段缺少粗分割阶段上下文信息,有时会获得相比粗分割阶段更差的分割结果,如图1所示.此外,由于在CT影像中胰腺与邻近器官密度较为接近、组织重叠部分界限分辨困难,未合理利用相邻切片预测分割掩码上下文信息常导致误分割现象,如图2所示.结合相邻预测分割掩码容易看出,中间切片存在误分割区域(红色部分),合理利用预测分割掩码切片上下文信息能够校准误分割区域.

针对胰腺细分阶段缺少粗分割阶段上下文信息的问题,文献[3]提出了固定点的分割方法.训练阶段使用胰腺标注数据训练粗分割网络,然后使用

粗分割网络的预测结果对原CT图像进行定位、剪裁,只保留胰腺及其周围部分区域作为细分阶段网络输入,通过反向传播,优化细分阶段结果.测试阶段,固定细分阶段网络参数,使用细分阶段网络预测掩码获得定位框并剪裁CT图像,再次输入细分阶段网络,迭代此过程获得优化的分割掩码,以此缓解缺少阶段上下文信息的问题.但是此分割方法本质上仅循环利用细分阶段定位框的位置信息,缺少对分割掩码的循环利用,缺少联合训练,导致分割效果提升有限.

针对如何合理利用切片上下文信息解决胰腺与邻近器官密度较为接近、组织重叠部分界限分辨困难导致的误分割问题,研究者提出了利用卷积长短期记忆网络(Convolutional long short-term memory, CLSTM)<sup>[7]</sup>和三维分割网络的方法<sup>[8-10]</sup>.文献[8]将相邻CT切片输入到卷积门控循环单元(Convolutional gated recurrent units, CGRU)<sup>[11]</sup>,使当前隐藏层输出信息融合到下一时序隐藏层中,通过前向传播,当前隐藏层可获得融合之前切片上下文信息的输出表示.文献[9]通过双向卷积循环神经网络,同时利用当前层前后切片上下文信息进行胰腺分割.但是目前大多数基于卷积循环神经网络的分割方法在利用切片上下文信息时,只能按照输入切片顺序、逆序或结合顺序和逆序的方式.这些方式严重依赖输入序列顺序,并且相隔越远的切片在前向传播过程中能够共享的上下文信息越少.与前述方法不同,文献[10]将邻近切片输入三维卷积神经网络,有效利用切片上下文信息,改善了分割结果.但基于三维卷积神经网络的分割方法,受限于三维训练数据量过少和显存消耗过大,大多数方法是基于局部三维块的分割.虽然局部块中切片上下文信息得到了合理利用,但是全局三维信息却缺乏连续性,导致分割掩码存在过多噪点.相比于三维图像分割方法需要解决三维图像数据量过少及参数量过多带来的显存问题,基于卷积循环神经网络

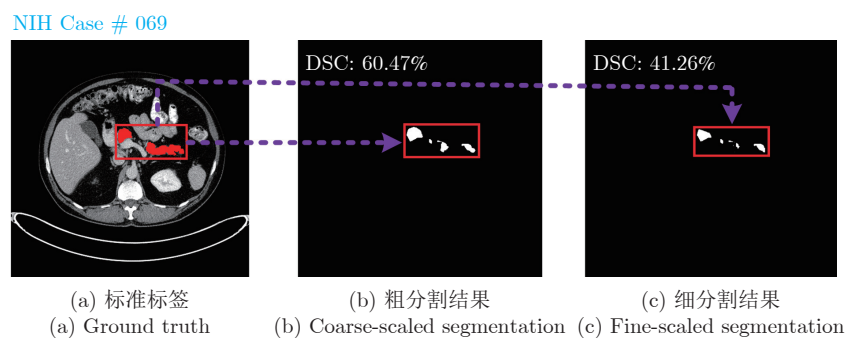


图1 粗细分割存在问题示例

Fig.1 A failure case of the coarse-to-fine pancreas segmentation approach

NIH Case # 068

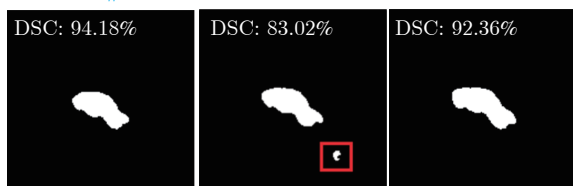


图 2 误分割示例

Fig. 2 An example of false segmentation

络的二维图像分割方法存在的问题可以通过设计算法解决。

根据以上分析, 本文针对现有基于由粗到细的二维胰腺分割方法中存在的问题, 设计了循环显著性校准网络, 其结合更多的阶段上下文信息和切片上下文信息. 通过设计的卷积自注意力校准模块跨顺序利用切片上下文信息校准每一阶段的胰腺分割掩码, 循环使用前一阶段的胰腺分割掩码定位目标区域, 增强当前阶段的网络输入, 完成分割任务的联合优化. 提出的方法在公开数据集上进行了实验验证, 结果表明其有效地解决了上述胰腺分割任务中存在的问题. 本文的主要贡献如下:

1) 提出循环显著性校准网络, 循环利用前一阶段胰腺分割掩码显著性增强当前阶段胰腺区域特征, 通过联合训练获取更多的阶段上下文信息.

2) 设计了卷积自注意力模块, 使得胰腺所有输入切片预测分割掩码之间可以平行地进行跨顺序上下文信息互监督, 校准预测分割掩码.

3) 在 NIH (National institutes of health) 和 MSD (Medical segmentation decathlon) 胰腺数据集进行了大量实验, 实验结果验证了提出方法的有效性及其先进性.

## 1 相关工作

**由粗到细的两阶段分割方法.** 由粗到细的两阶段分割方法主要分为两类: 基于传统算法和基于深度学习的方法. 前者主要使用如超像素、图谱等传统算法获得粗分割结果, 再通过随机森林、Graphcut 等方法获得细分割结果<sup>[12-13]</sup>; 后者主要是基于深度学习的粗细分割方法<sup>[14]</sup>, 基于数据驱动、自动化学习模型参数, 进行像素级别分类, 因其高精度和稳定性, 逐渐取代传统由粗到细的分割方法.

基于深度学习的粗细分割方法在粗分割训练阶段, 输入 CT 切片  $M^C$ , 经过粗分割卷积神经网络  $f(M^C, \theta^C)$ , 预测结果记为  $N^C$ , 与真实标签 (Ground truth)  $Y$  进行损失计算, 通过反向传播优化粗分割结果. 在细分割训练阶段, 针对粗分割网络预测结

果, 使用最小外接矩形算法获得胰腺位置坐标  $(p_x, p_y, w, h)$ , 对 CT 输入  $M^C$  进行剪裁, 获得感兴趣区域  $M^F$  作为细分割网络  $f(M^F, \theta^F)$  的输入; 获得细分割网络输出预测结果  $N^F$  并还原图像大小, 记为  $Y^P$ , 与真实标签 (Ground truth)  $Y$  进行损失计算, 通过反向传播优化细分割结果. 其中, 上标参数  $C$ 、 $F$  分别表示粗分割阶段和细分割阶段;  $(p_x, p_y)$ ,  $w$ ,  $h$  分别表示外接矩形框的左上角坐标, 宽和高;  $\theta^C$ ,  $\theta^F$  分别表示粗、细分割网络参数.

在测试阶段, 将 CT 切片输入训练好的粗细分割网络即可获得测试结果. 粗细分割方法中使用的网络主要是基于 UNet<sup>[15]</sup>, FCN<sup>[2]</sup> 以及基于这两个基础结构的改进网络. 本文在粗细分割方法的基础上针对胰腺解剖性质提出基于循环显著性校准网络的胰腺分割方法, 均使用 UNet<sup>[15]</sup> 作为基础骨干网络.

**胰腺分割方法.** 传统医学图像分割常用方法有水平集<sup>[16]</sup>、混合概率图模型<sup>[17]</sup> 和活动轮廓模型<sup>[18]</sup> 等. 随着深度学习的发展, 基于卷积神经网络 (Convolutional neural network, CNN) 的分割方法由于其较高的精确度和较好的泛化性逐渐取代传统方法. 目前大多数基于深度学习的胰腺分割方法核心思想来源于 FCN<sup>[2]</sup>, FCN 改进了卷积神经网络, 用卷积层替代最后的全连接层, 同时将浅层语义特征通过上采样与深层特征相融合, 补充分割目标的位置信息, 提高了分割准确率. 另一种常用于胰腺分割的方法采用“编码器-解码器”结构<sup>[15]</sup>, 编码器负责逐层提取渐进的高级语义特征, 解码器通过反卷积或上采样的方法逐层恢复图像分辨率至原图大小, 同一层次编码器和解码器通过跳跃结构相连接.

由于胰腺形状、大小和位置多变, 上述单个阶段基于 FCN 或“编码器-解码器”结构的分割网络难以获得准确分割结果. 文献 [3] 首先提出了基于卷积神经网络由粗到细的两阶段分割算法, 使用粗分割掩码的位置信息剪裁细分割阶段网络的输入, 减小背景区域对胰腺区域分割的影响. 相比于文献 [3], 文献 [19] 更进一步, 在使用由粗到细的两阶段分割算法的同时, 通过设计轻量化模块减少了粗细分割阶段模型的参数; 而文献 [20] 则直接以中心点为基础剪裁图像作为细分割网络的输入. 文献 [21] 利用肝脏、脾脏和肾脏的位置信息定位胰腺器官, 这不同于上述直接通过粗分割定位胰腺器官的方法. 文献 [22] 提出基于由下至上的方法, 首先使用超像素分块进行粗分割, 然后基于超像素块集成分割结果. 文献 [23] 使用最大池化方法融合 CT 切片三个轴信息, 获取候选区域, 在候选区域中从边缘至内部聚合分割结果. 文献 [24] 提出基于图谱的粗

细分割方法,改善了分割结果. 近来,文献[25]提出了一种基于强化学习的两阶段分割算法. 首先,使用 DQN (Deep Q network) 回归胰腺坐标位置,剪裁只保留胰腺及其周围部分区域;然后,细分割阶段使用可行变卷积网络获得分割结果.

以上方法均取得了较为准确的分割结果,但是将胰腺分割粗细两阶段分开训练,细分割阶段缺少粗分割阶段上下文信息的问题,依然难以用有效的方式处理. 文献[3]在测试阶段使用固定点算法,固定细分割模型参数,循环利用当前阶段预测分割掩码获取定位框位置信息作为下一阶段输入的先验,以此达到使用之前阶段上下文信息的效果. 此方法本质上只迭代使用细分割定位框位置信息,缺乏粗分割输出分割掩码的有效利用.

合理使用切片上下文信息解决胰腺误分割,同样至关重要. 文献[8, 26]首先使用卷积神经网络提取特征,然后利用卷积长短期记忆网络<sup>[7]</sup>提取切片上下文信息分割胰腺,但切片上下文信息不能够跨顺序、平行化共享,并且前向传播存在信息丢失的问题. 文献[27]使用对抗学习思想,分别使用两个判别器约束主分割网络,捕获空间语义信息和切片上下文信息,但对抗网络的不稳定性使得训练和测试结果波动性较大. 文献[28]使用相邻切片局部块作为输入,编码器部分使用三维卷积以递进的方式逐层融合切片上下文信息,解码器部分使用二维转置卷积输出中间切片分割掩码. 由于 CT 切片之间层厚和层间距的差异性,且局部块输入未使用任何插值方法,捕捉到的三维切片上下文信息具有不一致性和局部性. 文献[10, 29–30]使用三维分割方法获取切片上下文信息,受限于显存和三维数据量,全局三维信息缺乏连续性.

针对现有胰腺分割方法中缺少阶段上下文信息的问题,以及在使用循环卷积神经网络分割胰腺的过程中,利用相邻切片上下文信息存在顺序依赖并且随着相邻切片间隔距离的增加导致全局信息正相关减少的问题,本文提出了一种循环利用阶段上下文信息和切片上下文信息的二维胰腺图像分割网络. 首先,将相邻 CT 切片作为粗分割网络输入,获得粗分割掩码;然后,通过最小矩形框算法对获得的粗分割掩码进行胰腺区域坐标定位;接着,使用粗分割掩码作为权重增强细分割阶段输入切片的胰腺区域特征,获取细分割掩码,细分割掩码以同样的方式增强下一阶段分割网络的输入;最后,循环迭代上述过程直到达到指定停止条件. 通过此方法,有效降低了分割平均误差,提高了分割方法的稳定性.

## 2 本文方法

针对当前由粗到细的两阶段胰腺分割算法利用阶段上下文信息和切片上下文信息存在的问题,本文提出了循环显著性校准网络. 其采用 UNet 和卷积自注意力校准模块作为骨干网络,接受相邻横断位胰腺 CT 切片输入;当前阶段卷积自注意力校准模块利用切片上下文信息校准 UNet 输出掩码的同时,利用自身输出掩码显著性增强下一阶段 UNet 网络输入;循环 UNet 和卷积自注意力校准模块,联合阶段上下文信息和切片上下文信息提升分割性能,整体网络架构如图 3 所示.

### 2.1 循环显著性校准网络

本文聚焦于联合阶段间和胰腺序列图像切片上下文信息提升分割准确率. 为了合理利用当前阶段胰腺分割掩码的位置和形状等先验信息,显著增强下一阶段分割网络的输入;同时,通过平行、跨顺序直接利用相邻切片分割掩码改善自身明显误分割现象,提出循环显著性校准网络,其分割迭代过程如图 4 所示. 循环显著性指每个阶段的校准分割掩码  $P$  经增强模块  $g(P, \varphi)$  特征提取后获得像素矩阵,此像素矩阵为胰腺前景相关矩阵. 使用此像素矩阵和下一阶段输入图像  $M$  进行像素对像素相乘,显著增强胰腺区域,抑制背景区域.

选择 UNet 基础分割网络模型  $f(\cdot, \theta)$  作为骨干网络,该模型的输入为胰腺的相邻 CT 切片,记为  $X$ ,通过基础分割网络模型推断出输出掩码  $N$ . 由于胰腺与邻近器官密度较为接近、组织重叠部分界限分辨困难,容易导致基础分割网络出现误分割现象. 因此,本文基于切片上下文信息设计了卷积自注意力校准模块  $a(\cdot, \eta)$ ,校准基础分割网络输出的分割掩码  $N$ ,卷积自注意力校准模块的输出表示为  $P$ ;为了能够获取更加准确的胰腺位置,设置了固定分割掩码像素阈值 0.5,来二值化  $P$ ,其输出表示如式 (1) 所示.

$$Z = \begin{cases} 1, & P_{ij} \geq 0.5 \\ 0, & P_{ij} < 0.5 \end{cases} \quad (1)$$

其中,  $i$ 、 $j$  为分割掩码中像素值位置坐标.

通过对式 (1) 的输出  $Z$  应用最小矩形框算法获得包围胰腺分割掩码框的位置坐标  $(p_x, p_y, w, h)$ ,位置坐标获得过程如图 5 所示. 其中,蓝色框为单个连通区域分割结果的定位,绿色框为整合多段分割结果的定位.

为改善下一阶段分割过程中缺少当前阶段上下文信息的问题,使用校准模块输出分割掩码  $P$  作为

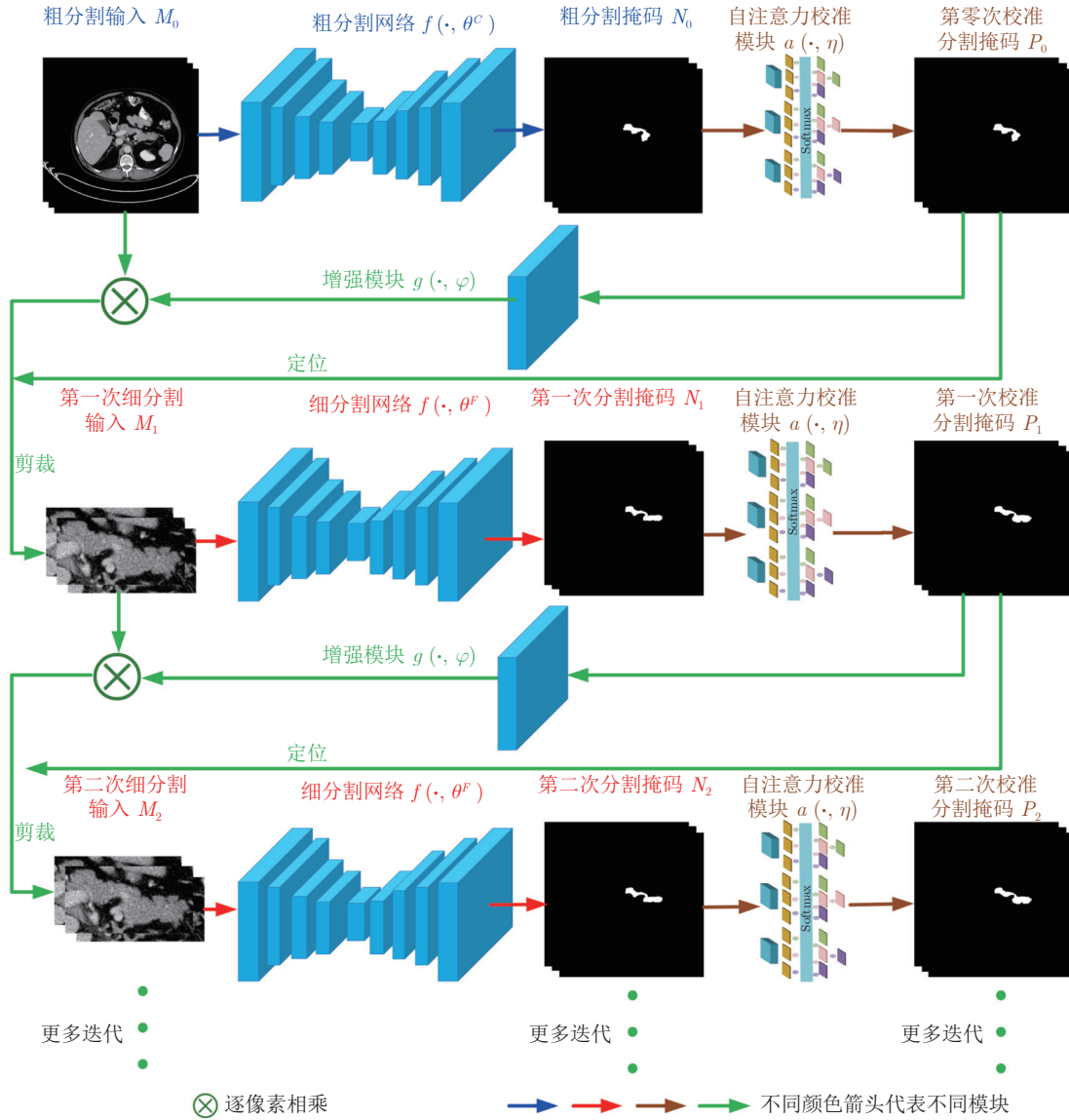


图 3 循环显著性校准网络

Fig.3 Recurrent saliency calibration network

潜在变量输入到显著性增强模块  $g(P, \varphi)$ , 提取特征概率作为下一阶段分割网络输入  $X$  的先验空间权重, 并结合上述定位坐标  $(p_x, p_y, w, h)$  增强并缩小下一阶段网络的输入, 显著减小背景区域对分割的影响. 对于在整个腹部图像中区域占比较小, 形状和位置多变的胰腺器官来说, 此过程极为重要, 其显著增强了胰腺区域, 弱化了不相关区域. 过程如式 (2) 所示.

$$M = \text{Crop}(X \otimes g(P, \varphi)) \quad (2)$$

其中,  $M$  为增强并缩小的下一阶段输入;  $\otimes$  表示对应像素点相乘;  $\text{Crop}$  表示利用定位坐标  $(p_x, p_y, w, h)$  对各阶段输入做剪裁.  $\theta, \eta, \varphi$  为相应模块共

享网络参数.

图 4 右图是图 4 左图的展开形式, 其中  $M_0$  作为胰腺初始输入图像和  $X$  相同, 其大小远大于其他阶段的网络输入  $M_t (t > 0)$ , 所以第一次粗分割阶段和其余分割阶段网络参数  $\theta$  应加以区分, 分别使用  $\theta^C$  和  $\theta^F$  表示. 在循环迭代过程中, 由于各分割阶段输入  $X$  不变, 并且输入  $X$  需与显著性增强模块  $g(P, \varphi)$  输出作逐像素相乘, 为了保持  $g(P, \varphi)$  是一个输入输出同大小的模块, 设置卷积核大小为  $3 \times 3$ , 步长为 1, 填充为 1.

整个循环迭代分割过程如式 (3) 所示.

$$P_t = a(f(\text{Crop}(X \otimes g(P_{t-1}, \varphi)), \theta), \eta) \quad (3)$$

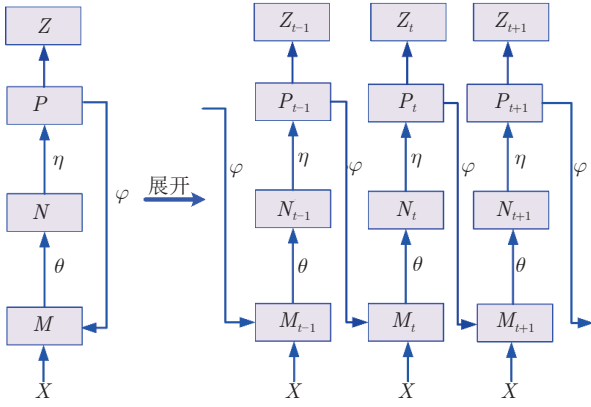


图 4 迭代过程

Fig. 4 Iteration process

NIH Case # 075

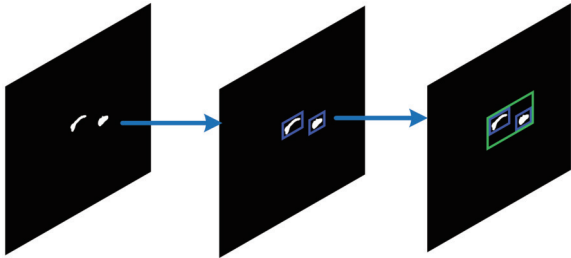


图 5 基于最小矩形框的定位过程

Fig. 5 The process of localization based on minimum rectangle algorithm

根据以上分析可以看出, 整个网络运算过程是可微的, 结合所有阶段损失函数进行联合训练. 本文采用 DSC (Dice-Sørensen coefficient) 作为损失函数, 如式 (4) 所示.

$$\mathcal{L}(Y, P) = 1 - \frac{2 \sum YP}{\sum Y + \sum P} \quad (4)$$

其中,  $Y$  是真实标签,  $P$  为各阶段预测分割掩码.

结合各阶段分割网络和卷积自注意力校准模块, DSC 联合损失函数如式 (5) 所示.

$$\mathcal{L} = \sum_{i=0}^T \lambda_i [\mathcal{L}(Y, N_i) + \mathcal{L}(Y, P_i)] \quad (5)$$

其中,  $T$  为循环分割次数停止阈值. 由于粗分割阶段和其余分割阶段胰腺切片输入大小不一致, 粗分割阶段主要用于获取胰腺的初步定位和粗分割掩码, 故设置较小的权重参数, 且满足  $3\lambda_0 = \lambda_1 = \lambda_2 = \dots = \lambda_T = 3/(3T + 1)$ .

## 2.2 卷积自注意力校准模块

针对胰腺与邻近器官密度较为接近、组织重叠

部分界限分辨困难而导致的误分割问题, 本文提出在循环显著性校准网络每个分割阶段嵌入卷积自注意力校准模块, 其合理利用切片上下文信息校准胰腺相邻切片误分割区域.

本文设计的卷积自注意力校准模块基于自注意力机制<sup>[31]</sup>. 自注意力机制在处理序列信息输入时, 能够跨顺序、平行化地与序列输入中其他时间点输入进行直接交互. 但由于自注意力机制使用的线性变换忽略了图像像素之间的空间关系, 本文提出卷积自注意力校准模块, 使用卷积操作替换线性变换. 卷积自注意力校准模块如图 6 左图所示 (图中以批次大小 3 为例), 图 6 右图为图 6 左图中获得单张校准分割掩码的计算过程, 其余两张切片校准分割掩码获得过程计算方式类似, 图中  $P$  表示预测分割掩码的热力图显示, 中间是颜色条, 其取值范围在  $[0, 1.0]$  之间. 为了描述和解释的方便, 将卷积自注意力校准模块中所有标量以向量或者矩阵的形式表示如下.

$$N_t = \begin{pmatrix} N_t^1 \\ N_t^2 \\ N_t^3 \end{pmatrix}, Q = \begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix},$$

$$K = \begin{pmatrix} k^1 \\ k^2 \\ k^3 \end{pmatrix}, V = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix},$$

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix},$$

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}, P_t = \begin{pmatrix} P_t^1 \\ P_t^2 \\ P_t^3 \end{pmatrix}$$

其中,  $N_t$  为分割网络  $f(M_t, \theta)$  输出的相邻分割掩码;  $Q$  为查询特征向量;  $K, V$  为键值对特征向量;  $B$  为相似度度量矩阵;  $C$  为像素权重矩阵;  $P_t$  为校准模块输出掩码.

为了提取多样性的特征表示, 首先将相邻分割掩码  $N_t$  中的每一个元素分别通过  $3 \times 3$  卷积, 获得输出  $\alpha_1, \alpha_2, \alpha_3$ ; 然后将输出特征  $\alpha_1$  分别通过 3 个不同的  $1 \times 1$  卷积, 获得  $N_t^1$  的查询特征向量  $q^1$  和键值对特征向量  $k^1, v^1$ ; 输出特征  $\alpha_2, \alpha_3$  以与  $\alpha_1$  相同的方式获得  $(q^2, k^2, v^2)$  和  $(q^3, k^3, v^3)$ , 其中所有  $3 \times 3$  和  $1 \times 1$  卷积都保持输出和输入大小一致. 下面通过  $Q, K, V$  来表示输出掩码  $P_t$  获得的过程, 如式 (6) ~ 式 (8).

$$B = QK^T \quad (6)$$

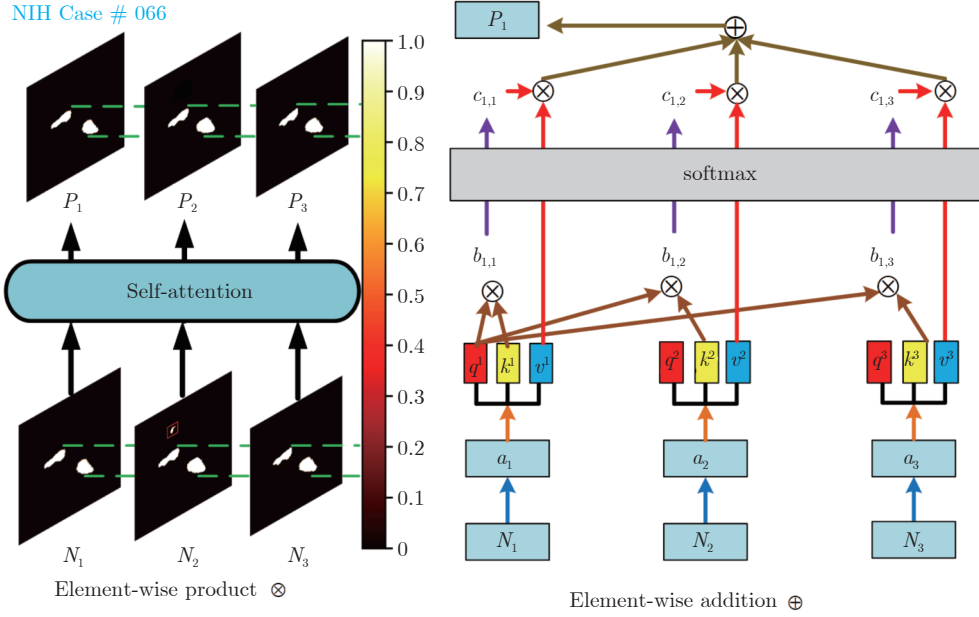


图6 卷积自注意力校准模块网络图

Fig.6 Network of convolution self-attention calibration module

$$C = \text{softmax} \left( \frac{B}{\sqrt{d_k}}, \dim = -2 \right) \quad (7)$$

$$P_t = CV \quad (8)$$

$B$  为  $N_t$  中的每个元素的特征查询向量分别与其他元素的键向量作相似度度量获得的矩阵, 表示其他元素对当前元素的影响程度, 这里相似度度量是像素对像素的乘法操作,  $d_k$  为键向量维度.  $C$  通过  $\text{softmax}$  函数对相似度度量矩阵进行归一化,  $\dim = -2$  表示对倒数第二个维度进行归一化; 获得像素权重矩阵以后, 与  $N_t$  中的每个元素的值向量  $v$  像素对像素相乘, 再进行融合获得最终输出表示  $P_t$ . 分别用式 (9) ~ 式 (15) 表示详细计算过程, 其中卷积操作都拥有不同的参数.

$$\alpha_1, \alpha_2, \alpha_3 = \text{conv}_{3 \times 3} (N_t^1), \text{conv}_{3 \times 3} (N_t^2), \text{conv}_{3 \times 3} (N_t^3) \quad (9)$$

$$(q^1, k^1, v^1) = \text{conv}_{1 \times 1} (\alpha_1), \text{conv}_{1 \times 1} (\alpha_1), \text{conv}_{1 \times 1} (\alpha_1) \quad (10)$$

$$(q^2, k^2, v^2) = \text{conv}_{1 \times 1} (\alpha_2), \text{conv}_{1 \times 1} (\alpha_2), \text{conv}_{1 \times 1} (\alpha_2) \quad (11)$$

$$(q^3, k^3, v^3) = \text{conv}_{1 \times 1} (\alpha_3), \text{conv}_{1 \times 1} (\alpha_3), \text{conv}_{1 \times 1} (\alpha_3) \quad (12)$$

$$b_{i,1}, b_{i,2}, b_{i,3} = q^i k^1, q^i k^2, q^i k^3 \quad (i = 1, 2, 3) \quad (13)$$

$$(c_{j,1}, c_{j,2}, c_{j,3}) = \text{softmax} \left( \frac{b_{j,1}}{\sqrt{d_k}}, \frac{b_{j,2}}{\sqrt{d_k}}, \frac{b_{j,3}}{\sqrt{d_k}} \right) \quad (j = 1, 2, 3) \quad (14)$$

$$P_t^1, P_t^2, P_t^3 = \sum_{k=1}^3 c_{1,k} v^k, \sum_{k=1}^3 c_{2,k} v^k, \sum_{k=1}^3 c_{3,k} v^k \quad (15)$$

### 3 实验结果与分析项

#### 3.1 数据集及预处理

本文实验使用 NIH<sup>[22]</sup> 胰腺分割数据集和 MSD<sup>[32]</sup> 胰腺分割数据集. NIH 数据集总共包含 82 位受试者的 CT 样本, 每位受试者样本中 CT 切片数量最少 181 张, 最多 466 张, 每一张切片大小为  $512 \times 512$  像素, 切片厚度在 0.5 mm 到 1.0 mm 之间; MSD 数据集总共包含 281 位受试者的 CT 样本, 每位受试者样本中 CT 切片数量在 37 到 751 之间, 每一张切片大小为  $512 \times 512$  像素.

在本文实验中, 所有 CT 切片的 HU (Hounsfield unit) 值根据统计结果被限制在  $[-120, 340]$ , 并把 CT 切片及其对应标签归一化到  $[0, 1]$  之间, 同时随机做  $[-15^\circ, 15^\circ]$  随机旋转.

#### 3.2 实验方法细节及评价指标

实验使用 Pytorch 1.2.0 版本, 在 Ubuntu

16.04 操作系统的 2 块 RTX 2080ti 独立显卡进行训练, 训练时批次大小设置为 3, 使用 ReLU<sup>[33]</sup> 作为激活函数, Adam<sup>[34]</sup> 作为优化方法, 学习率  $lr = 1.0 \times 10^{-4}$ , 受限于显存大小, 训练阶段最大迭代次数  $T$  设置为 4. 实验使用 4 折交叉验证方法确保结果的鲁棒性. 数据集被平均分为 4 份, 每次选择其中的 3 份作为训练集, 另 1 份作为验证集. 共实验 4 次, 计算平均 DSC 准确率, 作为最终结果.

**训练过程.** 训练过程中需要通过反向传播算法最小化损失函数 (式 (5)). 值得注意的是: 训练过程前期, 由于网络参数的随机初始化, 各阶段产生了错误分割掩码, 所以训练初始阶段使用标准标签作为上下文先验, 增强并剪裁下一分割阶段网络的输入.

**测试过程.** 测试过程和训练过程不同, 测试阶段缺少标准标签, 所以使用各阶段分割掩码作为先验信息, 增强并缩小下一阶段的输入; 同样, 测试过程中不需要优化参数, 对于中间结果可以丢弃, 所以迭代次数的阈值不再限于 GPU 显存, 理论上可以无上限. 本文设定测试的循环分割次数停止阈值  $T$  为 6, 因为实验的观察结果表明, 当迭代次数较大时分割准确率提升有限.

实验采用 DSC 作为评价指标, 如式 (16) 所示, 真实标签和分割掩码交集的两倍与真实标签和分割掩码并集的比值. 其中,  $Y$  是真实标签,  $P$  为预测分割掩码.

$$DSC(Y, P) = \frac{2 \sum YP}{\sum Y + \sum P} \quad (16)$$

表 1 粗细分割分开训练、联合训练和循环显著性联合训练分割结果

Table 1 Segmentation of coarse-to-fine separate training, joint training and recurrent saliency joint training

方法	平均 DSC (%) $\pm$ Std (%)		最大 DSC (%)		最小 DSC (%)	
	NIH	MSD	NIH	MSD	NIH	MSD
粗细分割分开训练	81.96 $\pm$ 5.79	78.92 $\pm$ 9.61	89.58	89.91	48.39	51.23
粗细分割联合训练	83.08 $\pm$ 5.47	80.80 $\pm$ 8.79	90.58	91.13	49.94	52.79
循环显著性网络联合训练	85.56 $\pm$ 4.79	83.24 $\pm$ 5.93	91.14	92.80	62.82	64.47

表 2 循环显著性网络测试结果

Table 2 Test results of recurrent saliency network segmentation

迭代次数	平均 DSC (%) $\pm$ Std (%)		最大 DSC (%)		最小 DSC (%)	
	NIH	MSD	NIH	MSD	NIH	MSD
第 0 次迭代 (粗分割)	76.81 $\pm$ 9.68	73.46 $\pm$ 11.73	87.94	88.67	40.12	47.76
第 1 次迭代	84.89 $\pm$ 5.14	81.67 $\pm$ 8.05	91.02	91.89	50.36	52.90
第 2 次迭代	83.34 $\pm$ 5.07	82.23 $\pm$ 7.57	90.96	91.94	53.73	56.81
第 3 次迭代	85.63 $\pm$ 4.96	82.78 $\pm$ 6.83	91.08	92.32	57.96	58.04
第 4 次迭代	85.79 $\pm$ 4.83	82.94 $\pm$ 6.46	91.15	92.56	62.97	63.73
第 5 次迭代	85.82 $\pm$ 4.82	83.15 $\pm$ 6.04	91.20	92.77	62.85	63.99
第 6 次迭代	85.86 $\pm$ 4.79	83.24 $\pm$ 5.93	91.14	92.80	62.82	64.47

### 3.3 实验对比分析

本节基于公开数据集 (NIH 和 MSD) 设置不同实验对照组, 验证基于循环显著性校准网络的胰腺分割方法. 主要分为 5 部分: 1) 阶段上下文信息有效性分析; 2) 切片上下文信息有效性分析; 3) 结合阶段上下文信息和切片上下文信息的循环显著性校准网络有效性分析; 4) 输入切片数目对分割结果的影响; 5) 网络模型参数量及时间消耗.

#### 3.3.1 阶段上下文信息有效性分析

本节对比实验展示阶段上下文信息对于分割性能的影响, 分别进行了两部分实验.

1) 针对粗细分割分开训练、粗细分割联合训练以及循环显著性网络联合训练进行对比实验. 粗细分割联合训练以及循环显著性网络联合训练都使用了显著性增强模块利用阶段上下文信息, 其中每一分割阶段都去掉了卷积自注意力校准模块. 实验结果如表 1 所示. 其中粗细分割联合训练相比于粗细分割分开训练在两个数据集上都展示了更高的平均分割准确率和更低的标准差, 其主要因为显著性增强模块显著性增强胰腺区域并联合粗细分割阶段上下文信息进行联合优化; 而循环显著性网络联合训练相比于粗细分割联合训练带来的分割效果提升, 来源于使用更多的阶段上下文信息联合训练. 由上述分析可知, 更多的阶段上下文信息对于胰腺分割准确率提升有重要贡献.

2) 针对循环显著性网络测试阶段进行分析, 如表 2 所示. 随着第 1 次迭代, 在 NIH 数据集上, 胰



腺的平均 DSC 准确率从 76.81% 上升到 84.89%, 标准差从 9.68% 降到 5.14%; 在 MSD 数据集上, 胰腺的平均 DSC 准确率从 73.46% 上升到 81.67%, 标准差从 11.73% 降到 8.05%. 由于粗分割阶段分割掩码上下文信息的引入, 平均 DSC 准确率和稳定性都有较大的提升. 但是, 后续迭代过程中, 由于分割掩码先验信息对于较准确分割结果作用的减少, 在两个数据集上平均 DSC 准确率和标准差仅仅小幅度上升和下降; 但对于最小 DSC 分割准确率提升明显, 分别从 40.12% 上升到最高的 62.82%、47.76% 上升到最高的 64.47%, 有效提升了胰腺分割困难样本的 DSC 分割准确率.

### 3.3.2 切片上下文信息有效性分析

本节对比实验展示切片上下文信息对于胰腺分割性能的影响, 分别进行了两部分实验.

1) 针对粗细分割以及循环显著性网络联合训练在添加和未添加卷积自注意力校准模块利用切片上下文信息情况下, 进行实验结果分析, 如表 3 所示. 相对于未添加卷积自注意力校准模块的粗细分割联合训练, 添加了卷积自注意力校准模块的粗细分割联合训练在 NIH 数据集上, 胰腺平均 DSC 准确率提升了 1.64%, 标准差下降了 0.40%; 在 MSD 数据集上, 胰腺平均 DSC 准确率提升了 1.29%, 标准差下降了 0.88%. 在两个数据集上, 胰腺最小 DSC 分割准确率也有所上升. 同样, 循环显著性网络联合训练在添加卷积自注意力校准模块 (本文方法) 时, 相比于未添加卷积自注意力校准模块, 其分割性能在分割准确率和稳定性上均提升明显. 由此可

以看出, 卷积自注意力校准模块能够利用切片上下文信息改善胰腺分割结果.

2) 针对本文方法中校准模块分别基于卷积自注意力或者基于卷积循环神经网络在分割胰腺时进行实验对比, 如表 4 所示. 将本文方法框架中卷积自注意力校准模块分别换成单层卷积长短期记忆循环神经网络 (CLSTM)<sup>[7]</sup>、单层卷积门控单元 (ConvGRU)<sup>[11]</sup> 和单层轨迹门控循环单元 (TrajGRU)<sup>[35]</sup> 等卷积循环神经网络, 进行实验对比. 从两个数据集的实验结果可以看出, 基于卷积自注意力的校准模块不管是在胰腺平均 DSC 分割准确率、标准差或者最大、最小分割准确率上都要好于部分基于卷积循环神经网络的校准模块<sup>[7, 11, 35]</sup>.

### 3.3.3 循环显著性校准网络有效性分析

为进一步说明本文所提方法在胰腺分割方法中的优势, 本文方法与当前具有代表性的方法进行了比较.

NIH 胰腺数据集上实验结果如表 5 所示, 本文方法与其他具有代表性的胰腺基准分割方法进行了比较. 相比于其他二维胰腺分割方法<sup>[3, 19-23, 26, 36]</sup>, 在以下两方面改进: 1) 联合训练利用更多的阶段上下文信息; 2) 使用卷积自注意力校准模块校准每一阶段胰腺分割掩码. 平均 DSC 分割准确率从最高的 85.40% 提升到 87.11%, 显著改善了胰腺平均分割结果; 最大分割准确率从最高的 91.46% 上升到 92.57%. 相比于三维胰腺分割方法<sup>[5, 10, 29-30, 37]</sup>, 本文提出的卷积自注意力校准模块充分利用切片上下文信息, 显著减少参数量 (GPU 显存消耗) 的同时, 达

表 3 添加校准模块结果对比  
Table 3 Comparison results of adding calibration module

方法	平均 DSC (%) $\pm$ Std (%)		最大 DSC (%)		最小 DSC (%)	
	NIH	MSD	NIH	MSD	NIH	MSD
粗细分割联合训练未添加校准模块	83.08 $\pm$ 5.47	80.80 $\pm$ 8.79	90.58	91.13	49.94	52.79
粗细分割联合训练添加校准模块	84.72 $\pm$ 5.07	82.09 $\pm$ 7.91	90.98	92.90	50.27	53.35
循环显著性网络未添加校准模块	85.86 $\pm$ 4.79	83.24 $\pm$ 5.93	91.14	92.80	62.82	64.47
循环显著性网络添加校准模块	87.11 $\pm$ 4.02	85.13 $\pm$ 5.17	92.57	94.48	67.30	68.24

表 4 胰腺分割基于 CLSTM 和自注意力结果对比  
Table 4 Comparison results based on CLSTM and self-attention mechanism in pancreas segmentation

方法	平均 DSC (%) $\pm$ Std (%)		最大 DSC (%)		最小 DSC (%)	
	NIH	MSD	NIH	MSD	NIH	MSD
基于 CLSTM 校准模块	86.13 $\pm$ 4.54	84.21 $\pm$ 5.80	91.20	93.47	63.18	64.76
基于 ConvGRU 校准模块	86.34 $\pm$ 4.21	84.41 $\pm$ 5.62	92.31	94.05	65.73	66.02
基于 TrajGRU 校准模块	86.96 $\pm$ 4.14	84.87 $\pm$ 5.22	92.49	94.32	67.20	67.93
基于卷积自注意力校准模块	87.11 $\pm$ 4.02	85.13 $\pm$ 5.17	92.57	94.48	67.30	68.24

表 5 NIH 数据集上不同分割方法结果比较 (“—”表示文献中缺少参数说明)

Table 5 Comparison of different segmentation methods on NIH dataset (“—” indicates a lack of reference in the literature)

方法	分割维度	平均 DSC (%) ± Std (%)	最大 DSC (%)	最小 DSC (%)
文献 [22]	2D	71.80 ± 10.70	86.90	25.00
文献 [23]	2D	81.27 ± 6.27	88.96	50.69
文献 [36]	2D	82.40 ± 6.70	90.10	60.00
文献 [3]	2D	82.37 ± 5.68	90.85	62.43
文献 [37]	3D	84.59 ± 4.86	91.45	69.62
文献 [10]	3D	85.99 ± 4.51	91.20	57.20
文献 [5]	3D	85.93 ± 3.42	91.48	75.01
文献 [29]	3D	82.47 ± 5.50	91.17	62.36
文献 [20]	2D	82.87 ± 1.00	87.67	81.18
文献 [19]	2D	84.90 ± —	91.46	61.82
文献 [26]	2D	85.35 ± 4.13	91.05	71.36
文献 [21]	2D	85.40 ± 1.60	—	—
文献 [30]	3D	86.19 ± —	91.90	69.17
本文方法	2D	<b>87.11 ± 4.02</b>	<b>92.57</b>	<b>67.30</b>

到三维分割同等效果,提高了运算效率,并且将胰腺平均分割准确率从最高的 86.19% 提升到 87.11%,最大分割准确率从最高的 91.90% 上升到 92.57%.

MSD 胰腺数据集上实验结果如表 6 所示,本文方法与具有代表性的胰腺基准分割方法进行了比较.相比于二维分割方法<sup>[28]</sup>,平均 DSC 分割准确率从 84.71% 提升到 85.13%,标准差从 7.13% 降到 5.17%,显著提升了胰腺分割方法的稳定性;最小分割准确率从 58.62% 上升到 68.24%,提高了困难样本的分割准确率.相比于三维胰腺分割方法<sup>[38-40]</sup>,平均 DSC 分割准确率从最高的 84.22% 提升到 85.13%;最大和最小分割准确率均有所提升.

本文方法在 NIH 及 MSD 胰腺数据集上箱线图如图 7 所示.本文对部分结果进行了展示,如图 8、图 9 所示.选取了 5 个受试者样本,同一行为同一个受试者不同切片的胰腺分割结果.蓝色实线代表预测结果,红色实线代表真实标签.从图中可以看出,本文方法分割结果和真实标签非常接近.

### 3.3.4 输入切片数目对分割结果的影响

为进一步说明胰腺输入切片数目对本文方法的影响,将切片数目输入分别设置为 3、5、7 进行实验比较,如表 7 和表 8 所示.随着胰腺输入切片数目的增加,平均 DSC 分割准确率和最大 DSC 分割准确率均有小幅度提升,最小 DSC 分割准确率提升更为明显.可以看出,增加切片数目对于分割困难样本具有较大的帮助.对于胰腺器官边界模糊的

表 6 MSD 数据集上不同分割方法结果比较

Table 6 Comparison of different segmentation methods on MSD dataset

方法	分割维度	平均 DSC (%) ± Std (%)	最大 DSC (%)	最小 DSC (%)
文献 [39]	3D	79.98 ± 7.71	93.73	61.64
文献 [38]	3D	82.37 ± 5.68	90.85	62.43
文献 [28]	2D	84.71 ± 7.13	95.54	58.62
文献 [40]	3D	84.22 ± 5.91	92.75	66.58
本文方法	2D	<b>85.13 ± 5.17</b>	<b>94.48</b>	<b>68.24</b>

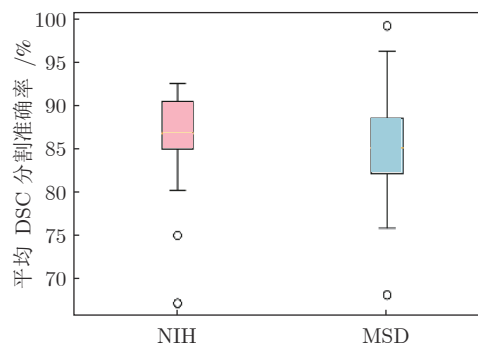


图 7 本文方法在 NIH 数据集及 MSD 数据集上箱线图

Fig. 7 Box plot of the method in this paper on NIH dataset and MSD dataset

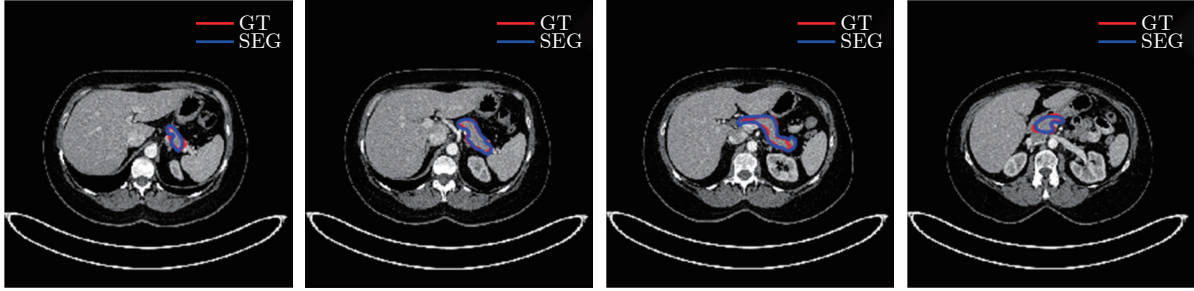
困难样本、胰腺周围脂肪与十二指肠灰度分布较为接近的困难样本以及切片中分割目标较小的困难样本,结合更多的切片数目能够明显提升目标分割精度.

### 3.3.5 网络参数量及时间消耗

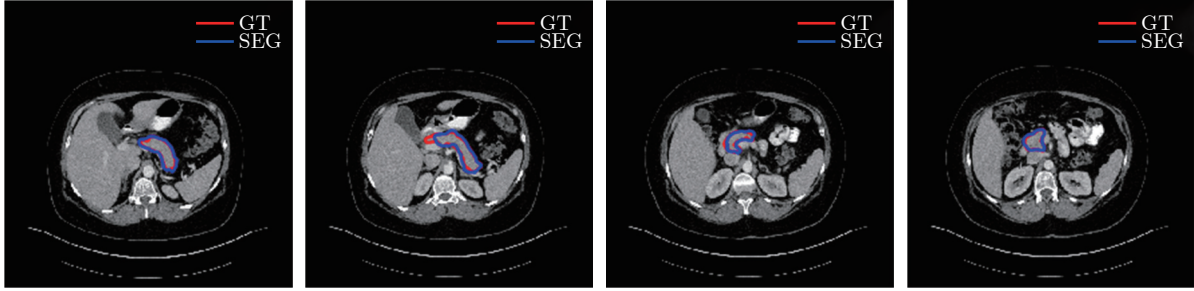
本文使用 UNet 作为基础骨干网络,除粗分割阶段,后续分割阶段共享网络参数,减少了参数量.相比于 FCN<sup>[2]</sup> 经典分割网络,提出的分割模型具有更少的参数量,如表 9 所示.虽然相比于单阶段的 UNet<sup>[15]</sup>, 3D UNet<sup>[41]</sup>, AttentionUNet<sup>[42]</sup> 和 UNet++<sup>[43]</sup> 等分割网络,参数量有所增加,但是单阶段的分割方法分割精度较低;相比于 Fix-point<sup>[3]</sup> 使用 FCN 作为骨干网络并且利用三个轴状面分别训练模型分割胰腺,本文参数量显著减少;相比于 GG-PFN<sup>[28]</sup>,虽然参数量有所增加,但是分割精度有所提升.

文献 [44] 使用胰腺器官的三个轴状面作为输入训练模型,并且分割阶段使用额外两个模型融合视觉特征,增加了时间消耗,如表 10 所示.文献 [23] 在三个轴状面上分别进行定位、分割,在 Titan X (12 GB) GPU 上训练了 9~12 个小时.文献 [22] 使用由下至上的方法,首先使用超像素分块,然后基于超像素分块集成分割结果,每个阶段分开训练.

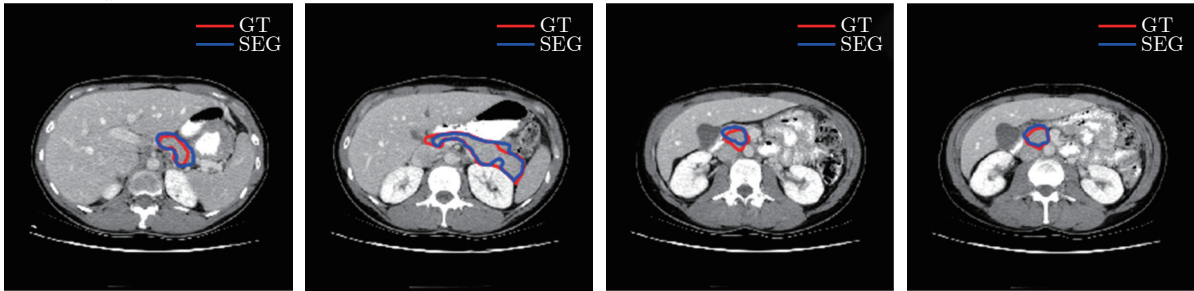
NIH Case # 064



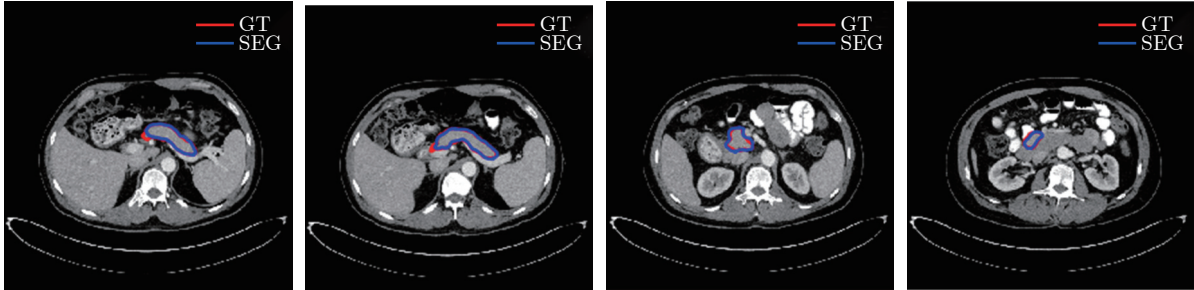
NIH Case # 067



NIH Case # 073



NIH Case # 075



NIH Case # 080

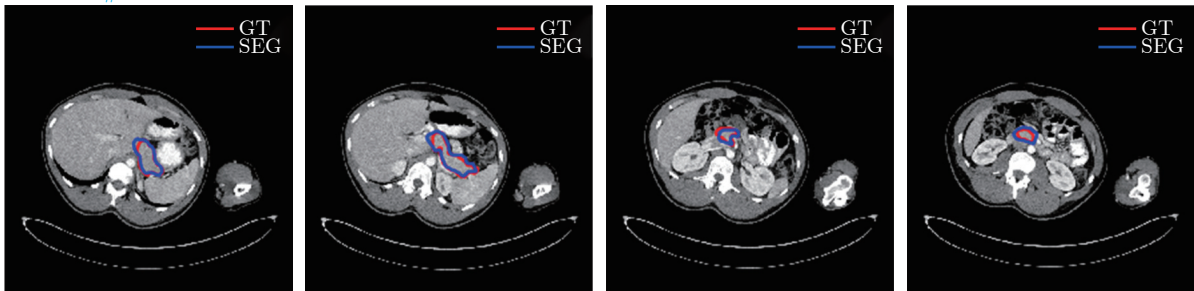


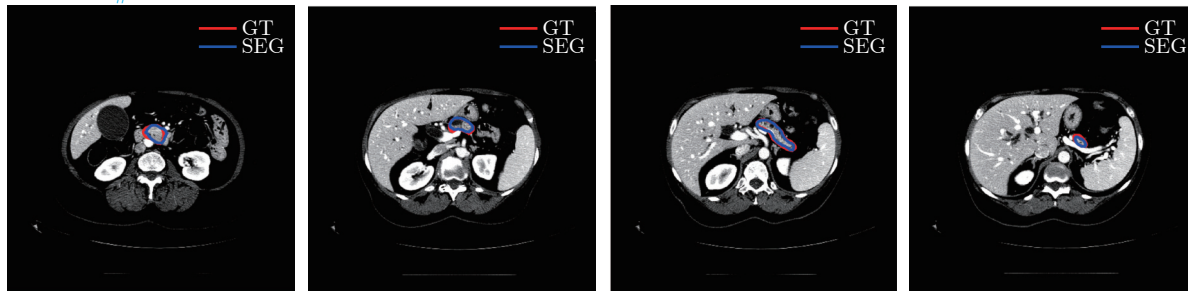
图 8 NIH 数据集分割结果对比

Fig.8 Comparison of segmentation results on NIH dataset

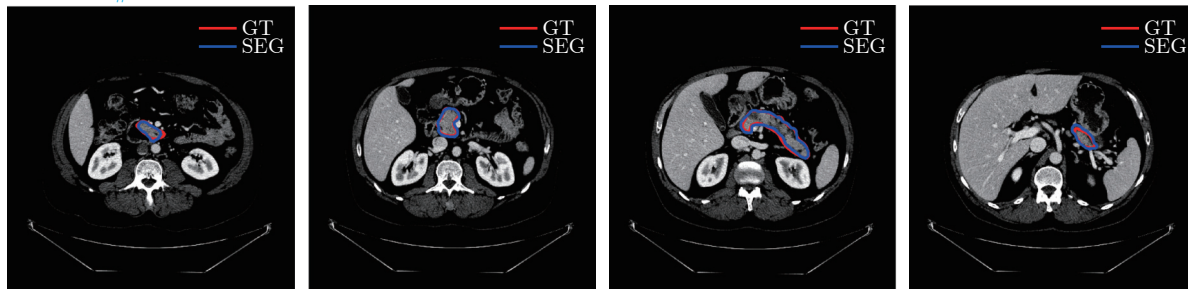
而本文胰腺分割方法使用端到端的训练方法, 降低了每个病例的平均测试时间. 文献 [3] 使用固定点

方法, 分别使用三个 FCN 训练胰腺的三个轴状面输入图像, 循环使用细分割掩码位置信息优化分割

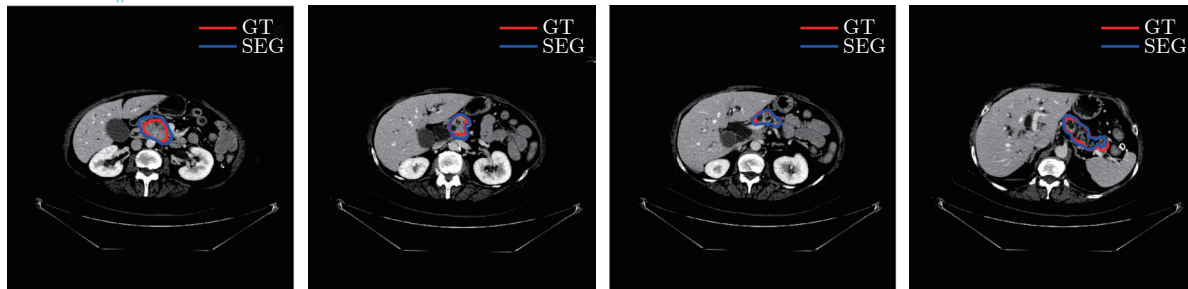
MSD Case # 291



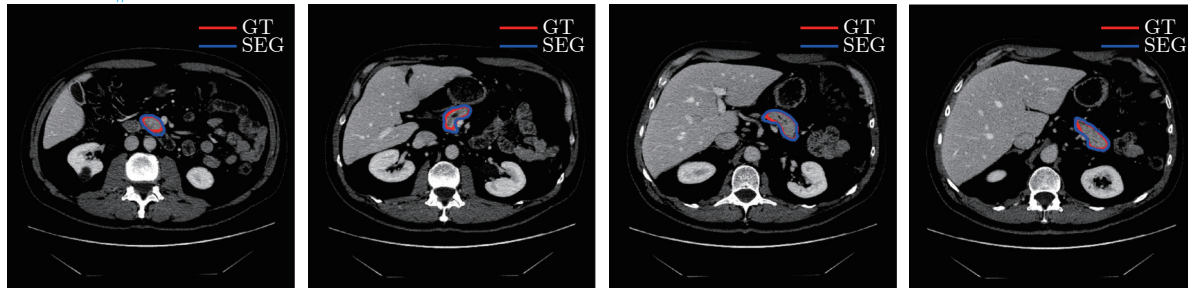
MSD Case # 302



MSD Case # 323



MSD Case # 344



MSD Case # 348

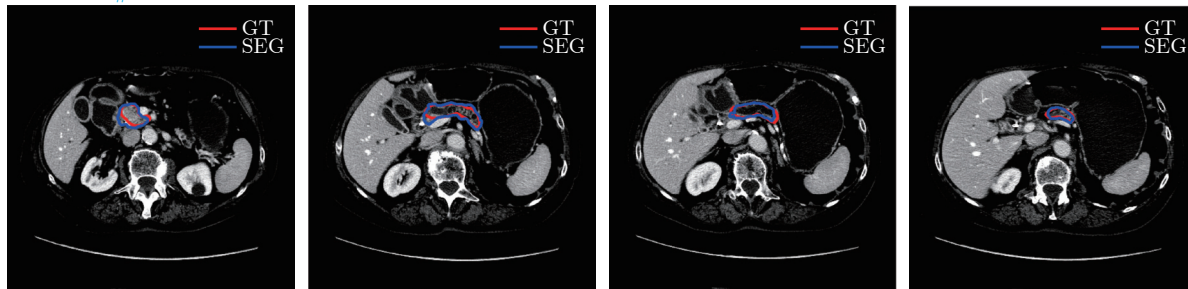


图 9 MSD 数据集分割结果对比

Fig.9 Comparison of segmentation results on MSD dataset

掩码, 显著增加了时间消耗. 相比于上述方法, 本文胰腺分割方法虽然增加了循环显著性模块和校准模

块, 但循环显著性模块和校准模块设计简单并且基于矩阵运算, 运算时间增加不明显, 并且本文方法

表 7 NIH 数据集不同网络输入切片数目分割结果比较  
Table 7 Comparison of the segmentation of different network input slices on NIH dataset

网络输入切片数目	分割维度	平均 DSC (%) ± Std (%)	最大 DSC (%)	最小 DSC (%)
3	2D	87.11 ± 4.02	92.57	67.30
5	2D	87.53 ± 3.74	92.69	69.32
7	2D	87.96 ± 3.25	92.94	71.91

表 8 MSD 数据集不同网络输入切片数目分割结果比较  
Table 8 Comparison of the segmentation of different network input slices on MSD dataset

网络输入切片数目	分割维度	平均 DSC (%) ± Std (%)	最大 DSC (%)	最小 DSC (%)
3	2D	85.13 ± 5.17	94.48	68.24
5	2D	85.86 ± 5.01	94.75	70.31
7	2D	86.29 ± 4.80	95.01	73.07

表 9 不同分割方法参数量比较  
Table 9 Comparison of the number of parameters of different segmentation methods

方法	分割维度	参数量
FCN <sup>[2]</sup>	2D	134.26 M
UNet <sup>[15]</sup>	2D	28.34 M
3D UNet <sup>[44]</sup>	3D	16.31 M
AttentionUNet <sup>[42]</sup>	2D	35.06 M
UNet++ <sup>[48]</sup>	2D	36.74 M
Fix-point <sup>[9]</sup>	2D	807.93 M
GGPFN <sup>[28]</sup>	2D + 3D	42.00 M
本文方法	<b>2D</b>	<b>59.47 M</b>

表 10 不同分割方法时间消耗比较  
(“—”表示文献中缺少参数说明)

Table 10 Comparison of time consumption of different segmentation methods (“—” indicates a lack of reference in the literature)

方法	分割维度	每个病例平均测试时间 (min)	训练时间 (h)	设备
文献 [44]	2D	2 ~ 3	—	—
文献 [22]	2D	1 ~ 3	~ 55	GTX Titan Z (12 GB)
文献 [23]	2D	2 ~ 3	9 ~ 12	Titan X (12 GB)
文献 [3]	2D	~ 3	—	—
文献 [36]	2D	—	~ 3	GTX Titan X (12 GB)
本文方法	2D	<b>1.1</b>	~ 8	<b>RTX 2080ti (11 GB)</b>

仅基于横断面作为输入, 使用 UNet<sup>[14]</sup> 而非 FCN<sup>[2]</sup> 作为分割骨干网络, 显著减少参数量及前馈传播时间. 相比于文献 [36] 基于循环卷积神经网络使用多

切片作为输入, 本文方法训练及测试时间有所增加, 但分割精确度提升明显.

## 4 总结与展望

针对胰腺分割面临的问题, 本文提出了基于循环显著性校准网络的胰腺分割方法. 其主要贡献在于: 1) 利用更多的阶段上下文信息联合训练, 改善了传统由粗到细胰腺分割方法仅使用粗分割阶段输出掩码定位框坐标信息作为细分网络输入的先验, 导致缺少阶段上下文信息的问题; 2) 使用卷积自注意力校准模块跨顺序、平行化利用相邻切片上下文信息的同时, 自动校准每一分割阶段输出掩码, 解决了胰腺与邻近器官密度较为接近、组织重叠部分界限分辨困难导致的误分割问题. 和其他胰腺分割方法相比, 本文方法显著提高了样本平均 DSC 分割准确率并改善了困难样本分割结果. 本文方法可用于辅助医疗诊断, 后续研究将考虑如何利用更多的阶段上下文信息及切片上下文信息改善分割结果的同时, 使用模型蒸馏方法轻量化模型框架.

## References

- Siegel R, Miller K, Fuchs H, Jemal A. Cancer statistics 2021. *CA: A Cancer Journal for Clinicians*, 2021, **71**(1): 7–33
- Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015. 3431–3440
- Zhou Y, Xie L, Shen W, Wang Y, Fishman E K, Yuille A L. A fixed-point model for pancreas segmentation in abdominal CT scans. In: Proceedings of the 2017 International Conference on Medical Image Computing and Computer-assisted Intervention. Quebec, Canada: Springer, 2017. 693–701
- Zhang Y, Wu J, Liu Y, Chen Y F, Chen W, Wu E X, et al. A deep learning framework for pancreas segmentation with multi-atlas registration and 3D level-set. *Medical Image Analysis*, 2021, **68**(2): 101884–101889
- Wang W, Song Q, Feng R, Chen T, Chen J, Chen D Z, et al. A fully 3D cascaded framework for pancreas segmentation. In: Proceedings of the 2020 IEEE International Symposium on Biomedical Imaging (ISBI). Iowa, USA: IEEE, 2020. 207–211
- Ma Chao, Liu Ya-Shu, Luo Gong-Ning, Wang Kuan-Quan. Combining concatenated random forests and active contour for the 3D MR images segmentation. *Acta Automatica Sinica*, 2019, **45**(5): 1004–1014  
(马超, 刘亚淑, 骆功宁, 王宽全. 基于级联随机森林与活动轮廓的 3D MR 图像分割. *自动化学报*, 2019, **45**(5): 1004–1014)
- Shi X, Chen Z, Wang H, Yeung D Y, Wong W K, Woo W. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In: Proceedings of the 2015 Advances in Neural Information Processing Systems. Quebec, Canada: MIT, 2015. 802–810
- Yang Z, Zhang L, Zhang M, Feng J, Wu Z, Ren F, et al. Pancreas segmentation in abdominal CT scans using inter-/intra-slice contextual information with a cascade neural network. In: Proceedings of the 2019 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

- Berlin, Germany: IEEE, 2019. 5937–5940
- 9 Li H, Li J, Lin X, Qian X. A model-driven stack-based fully convolutional network for pancreas segmentation. In: Proceedings of the 5th International Conference on Communication, Image and Signal Processing (CCISP). Chengdu, China: IEEE, 2020. 288–293
- 10 Zhao N, Tong N, Ruan D, Sheng K. Fully automated pancreas segmentation with two-stage 3D convolutional neural networks. In: Proceedings of the 2019 International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China: Springer, 2019. 201–209
- 11 Ballas N, Yao Y, Pal C, Courville A. Delving deeper into convolutional networks for learning video representations. arXiv preprint arXiv: 1511.06432, 2015.
- 12 Farag A, Lu L, Turkbey E, Liu J, Summers R M. A bottom-up approach for automatic pancreas segmentation in abdominal CT scans. In: Proceedings of the 2014 International Conference on Medical Image Computing and Computer-Assisted Intervention Workshop on Computational and Clinical Challenges in Abdominal Imaging. Cambridge, USA: Springer, 2014. 103–113
- 13 Kitasaka T, Oda M, Nimura Y, Hayashi Y, Fujiwara M, Misawa K, et al. Structure specific atlas generation and its application to pancreas segmentation from contrasted abdominal CT volumes. In: Proceedings of the 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention Workshop on Medical Computer Vision. Munich, Germany: Springer, 2015. 47–56
- 14 Deng G, Wu Z, Wang C, Xu M, Zhong Y. CCANet: Class-constraint coarse-to-fine attentional deep network for subdecimeter aerial image semantic segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, **60**(1): 1–20
- 15 Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of the 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015. 234–241
- 16 Li Yang, Zhao Yu-Qian, Liao Miao, Liao Sheng-Hui, Yang Zhen. Automatic liver segmentation from CT volumes based on level set and shape descriptor. *Acta Automatica Sinica*, 2021, **47**(2): 327–337  
(李阳, 赵于前, 廖苗, 廖胜辉, 杨振. 基于水平集和形状描述符的腹部 CT 序列肝脏自动分割. *自动化学报*, 2021, **47**(2): 327–337)
- 17 Xia Ping, Shi Yu, Lei Bang-Jun, Gong Guo-Qiang, Hu Rong, Shi Dong-Xia. Ultrasound medical image segmentation based on hybrid probabilistic graphical model in complex-wavelet domain. *Acta Automatica Sinica*, 2021, **47**(1): 185–196  
(夏平, 施宇, 雷帮军, 龚国强, 胡蓉, 师冬霞. 复小波域混合概率图模型的超声医学图像分割. *自动化学报*, 2021, **47**(1): 185–196)
- 18 Feng Bao, Chen Ye-Hang, Liu Zhuang-Sheng, Li Zhi, Song Rong, Long Wan-Sheng. Segmentation of breast cancer on DCE-MRI images with MRF energy and fuzzy speed function. *Acta Automatica Sinica*, 2020, **46**(6): 1188–1199  
(冯宝, 陈业航, 刘壮盛, 李智, 宋嵘, 龙晚生. 结合 MRF 能量和模糊速度的乳腺癌图像分割方法. *自动化学报*, 2020, **46**(6): 1188–1199)
- 19 Zhang D, Zhang J, Zhang Q, Han J, Zhang S, Han J. Automatic pancreas segmentation based on lightweight DCNN modules and spatial prior propagation. *Pattern Recognition*, 2021, **114**(6): Article No. 107762
- 20 Huang M L, Wu Z Y. Semantic segmentation of pancreatic medical images by using convolutional neural network. *Biomedical Signal Processing and Control*, 2022, **73**(3): 103458–103470
- 21 Liu Z, Su J, Wang R H, Jiang R, Song Y Q, Zhang D Y, et al. Pancreas Co-segmentation based on dynamic ROI extraction and VGGU-Net. *Expert Systems With Applications*, 2022, **192**(6): 116444–116453
- 22 Roth H R, Lu L, Farag A, Shin H C, Liu J M, Turkbey E B, et al. Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In: Proceedings of the 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015. 556–564
- 23 Roth H R, Lu L, Lay N, Harrison A P, Farag A, Sohn A, et al. Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation. *Medical Image Analysis*, 2018, **45**(3): 94–107
- 24 Karasawa K, Oda M, Kitasaka T, Misawa K, Fujiwara M, Chu C, et al. Multi-atlas pancreas segmentation: Atlas selection based on vessel structure. *Medical Image Analysis*, 2017, **39**(5): 18–28
- 25 Man Y, Huang Y, Feng J, Li X, Wu F. Deep Q learning driven CT pancreas segmentation with geometry-aware U-Net. *IEEE Transactions on Medical Imaging*, 2019, **38**(8): 1971–1980
- 26 Li J, Lin X, Che H, Li H, Qian X. Pancreas segmentation with probabilistic map guided bi-directional recurrent UNet. *Physics in Medicine and Biology*, 2021, **66**(11): 115010–115026
- 27 Khosravan N, Mortazi A, Wallace M, Bagci U. Pan: Projective adversarial network for medical image segmentation. In: Proceedings of the 2019 International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China: Springer, 2019. 68–76
- 28 Fang C, Li G, Pan C, Li Y, Yu Y. Globally guided progressive fusion network for 3D pancreas segmentation. In: Proceedings of the 2019 International Conference on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China: Springer, 2019. 210–218
- 29 Mo J, Zhang L, Wang Y, Huang H. Iterative 3D feature enhancement network for pancreas segmentation from CT images. *Neural Computing and Applications*, 2020, **32**(16): 12535–12546
- 30 Wang Y, Zhang J, Cui H, Zhang Y, Xia Y. View adaptive learning for pancreas segmentation. *Biomedical Signal Processing and Control*, 2021, **66**(4): 102347–102361
- 31 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. In: Proceedings of the 2017 Advances in Neural Information Processing Systems. California, USA: MIT, 2017. 5998–6008
- 32 Simpson A L, Antonelli M, Bakas S, Bilello M, Farahani K, Van G. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv: 1902.09063, 2019.
- 33 Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. Florida, USA: ML Research, 2017. 315–323
- 34 Kingma D P, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv: 1412.6980, 2014.
- 35 Shi X J, Gao Z H, Lausen L, Wang H, Yeung D Y, Wong W K, et al. Deep learning for precipitation nowcasting: A benchmark and a new model. In: Proceedings of the 2017 Advances in Neural Information Processing Systems. California, USA: MIT, 2017. 5617–5627
- 36 Cai J, Lu L, Xie Y, Xing F, Yang Y. Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function. arXiv preprint arXiv: 1707.04912, 2017.
- 37 Zhu Z, Xia Y, Shen W, Fishman E, Yuille A. A 3D coarse-to-fine framework for volumetric medical image segmentation. In: Proceedings of the 2018 International Conference on 3D Vision (3DV). Verona, Italy: IEEE, 2018. 682–690

- 38 Roth H R, Shen C, Oda H, Oda M, Hayashi Y, Misawa K, et al. Deep learning and its application to medical image segmentation. *Medical Imaging Technology*, 2018, **36**(2): 63–71
- 39 Roth H R, Oda H, Zhou X, Shimizu N, Yang Y, Hayashi Y, et al. An application of cascaded 3D fully convolutional networks for medical image segmentation. *Computerized Medical Imaging and Graphics*, 2018, **66**(4): 90–99
- 40 Li W, Wu X, Hu Y, Wang L, He Z, Du J. High-resolution recurrent gated fusion network for 3D pancreas segmentation. In: Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN). Shenzhen, China: IEEE, 2021. 1–7
- 41 Cicek O, Abdulkadir A, Lienkamp S S, Brox T, Ronneberger O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Proceedings of the 2016 International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece: Springer, 2016. 424–432
- 42 Oktay O, Schlemper J, Folgoc L L, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: Learning where to look for the pancreas. arXiv preprint arXiv: 1804.03999, 2018.
- 43 Zhou Z, Siddiquee R M M, Tajbakhsh N, Liang J. Unet++: A nested U-Net architecture for medical image segmentation. In: Proceedings of the 2018 Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada, Spain: Springer, 2018. 3–11
- 44 Roth H, Lu L, Farag A, Sohn A, Summers R. Spatial aggregation of holistically-nested networks for automated pancreas segmentation. In: Proceedings of the 2016 International Conference on Medical Image Computing and Computer-Assisted Intervention. Athens, Greece: Springer, 2016. 451–459



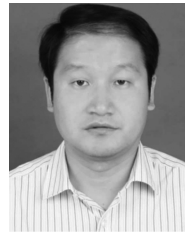
**邱成健** 江苏大学计算机科学与通信工程学院博士博士研究生。主要研究方向为医学图像分割。

E-mail: 2111908005@stmail.ujs.edu.cn

**(QIU Cheng-Jian** Ph.D. candidate at the School of Computer Science

and Communication Engineering, Jiangsu University.

His main research interest is medical image segmentation.)



**刘青山** 南京信息工程大学自动化学术院教授。主要研究方向为视频内容分析与理解。E-mail: qslu@nuist.edu.cn

**(LIU Qing-Shan** Professor at the School of Automation, Nanjing University of Information Science and Technology. His main research interest is video content analysis and understanding.)



**宋余庆** 江苏大学计算机科学与通信工程学院教授。主要研究方向为医学图像分析, 数据挖掘。

E-mail: yqsong@ujs.edu.cn

**(SONG Yu-Qing** Professor at the School of Computer Science and Communication Engineering, Ji-

angsu University. His research interest covers medical image analysis and data mining.)



**刘哲** 江苏大学计算机科学与通信工程学院教授。主要研究方向为数据智能处理, 医学图像分析。本文通信作者。E-mail: lzhe@ujs.edu.cn

**(LIU Zhe** Professor at the School of Computer Science and Commu-

nication Engineering, Jiangsu University. Her research interest covers intelligent data processing and medical image analysis. Corresponding author of this paper.)