

结合空洞特征增强与改进HR-Net的胰腺分割方法

刘新宇, 曹春萍

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月18日; 录用日期: 2024年5月14日; 发布日期: 2024年5月21日

摘要

在胰腺癌辅助诊断领域, 精确的胰腺分割是实现高效疾病诊断的基石, 同时也是医学图像处理领域中一项具有挑战性的任务。现有的主流胰腺分割方法大多依赖于U型的编码-解码路径, 然而该架构在上下采样的过程中, 容易导致关键空间信息的丢失, 影响分割准确率。针对上述问题, 本文提出了一种结合空洞特征增强与改进HR-Net的胰腺分割方法。首先, 对原有的二维HR-Net进行三维化改造, 以便更全面地挖掘三维医学图像内蕴的空间信息。其次, 设计多尺度空洞特征增强模块以重构网络深层结构, 并行捕获广泛的多尺度深度特征, 并采用通道加权对重要特征赋予更高的关注度。接着, 在模型多分辨率融合模块之后引入空间注意力机制, 使模型更加专注于融合特征图中的重要空间区域。最后, 本文提出分层渐进特征融合头, 以缓解传统分割头在大幅度上采样时遭受的信息丢失。实验结果表明, 在NIH-Pancreas数据集上测试的召回率、准确率和Dice相似系数分别为84.5%、86.3%、85.6%, 在分割性能上优于现有的主流方法。本文方法能有效保留胰腺细节特征, 并展现出较强的特征提取能力, 对腹部CT图像中的胰腺器官有较好的分割结果。

关键词

医学图像, 胰腺分割, 高分辨率网络, 空洞特征增强

Pancreas Segmentation Method Combining Atrous Feature Enhancement and Improved HR-Net

Xinyu Liu, Chunping Cao

School of Optical-Electrical & Computer Engineering, University of Shanghai for Science & Technology, Shanghai

Received: Apr. 18th, 2024; accepted: May. 14th, 2024; published: May. 21st, 2024

Abstract

In the field of pancreatic cancer auxiliary diagnosis, accurate pancreatic segmentation is the cornerstone for efficient disease diagnosis and also a challenging task in the field of medical image processing. The existing mainstream pancreatic segmentation methods mostly rely on a U-shaped encoding-decoding path; however, this architecture tends to lose critical spatial information during the upsampling and downsampling processes, affecting segmentation accuracy. To address these issues, this paper proposes a pancreatic segmentation method that combines atrous feature enhancement with an improved HR-Net. Firstly, the original two-dimensional HR-Net is transformed into three dimensions to more comprehensively mine the spatial information inherent in three-dimensional medical images. Secondly, a multi-scale atrous feature enhancement module is designed to reconstruct the deep structure of the network, capturing a wide range of multi-scale deep features in parallel and assigning greater attention to important features through channel weighting. Subsequently, a spatial attention mechanism is introduced after the model's multi-resolution fusion module, making the model more focused on important spatial regions in the merged feature maps. Lastly, this paper proposes a hierarchical progressive feature fusion head to alleviate the information loss suffered by traditional segmentation heads during significant upsampling. Experimental results show that on the NIH- Pancreas dataset, the recall rate, accuracy, and Dice similarity coefficient are 84.5%, 86.3%, and 85.6%, respectively, surpassing existing mainstream methods in segmentation performance. The method proposed in this paper can effectively preserve pancreatic detail features and demonstrates strong feature extraction capability, yielding better segmentation results for pancreatic organs in abdominal CT images.

Keywords

Medical Imaging, Pancreas Segmentation, High Resolution Networks, Atrous Feature Enhancement

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

伴随着医工交叉领域的技术革新, 医学图像处理技术已成为人工智能研究中的重要领域。将深度学习技术用于计算机断层成像(computed tomography, CT)的器官分割, 替代了传统诊断方法中复杂繁琐的人工步骤, 对患者的疾病诊断、术前规划以及预后评估等阶段具有显著的推动作用[1]。因此, 利用深度学习算法对 CT 图像进行器官分割成为了近年来的计算机辅助诊断的研究热点之一。近年来, 研究者们针对不同的器官分割任务, 提出了对应的医学分割模型[2] [3] [4]。其中, UNet [5]由于其轻量化的 U 型编解码模型结构与在大型器官分割任务中呈现出的优异分割性能, 成为目前主流分割模型之一。其在心脏、肾、肝脏等大型器官的 CT 图像分割效果上, 已初步满足了临床需求, 推动了医学图像分割向体积更小、边界更复杂的器官分割方向发展。

胰腺分割作为胰腺癌辅助诊断中的基础工作, 也是医学图像分割领域最具挑战性的任务之一, 由于胰腺解剖学特异性高, 不同个体之间的胰腺形状, 大小和位置存在一定的差异[6], 且胰腺区域在整个腹部 CT 图像的位置占比仅有不足 0.5% [7], 使得分割难度较大。这便要求分割模型具备更为卓越的性能表现, 从而能进行更充分的特征提取。然而, UNet 架构中的连续降采样过程在编码阶段容易导致细节与语

义信息的丢失,而在解码阶段,模型仅通过上采样操作来恢复高分辨率的语义特征,无法充分弥补先前的信息损失,从而限制了模型在分割任务上的精度。特别是对于形态和位置变化较大的胰腺器官,这一问题尤为突出。

因此,本研究提出了一种基于高分辨率网络(High-Resolution Net, HRNet) [8]架构的胰腺分割模型,HRNet 在整个分割过程中维持高分辨率特征,以缓解下采样过程中位置和细节信息丧失的问题。针对胰腺分割任务的具体需求,本文对 HRNet 的初始架构进行了优化,并在公开数据集上进行了评估。实验结果验证了所提方法相较于传统的编解码分割网络及原始高分辨率分割网络在解决胰腺分割问题上的有效性。

本文的主要贡献如下:

- 1) 将 HRNet 的基础架构拓展为三维分割网络,以深入挖掘三维 CT 图像内蕴的空间信息。
- 2) 针对 HRNet 初始的连续下采样导致图像位置和细节信息丢失问题,设计多尺度空洞特征增强模块对网络的架构进行优化。该模块在减少参数数量的基础上有效扩大了网络的感受野,从而实现对多尺度深度特征的捕获。
- 3) 针对胰腺器官在形态上的异质性,在各阶段多分辨率融合之后融入高效通道注意力模块,精细化胰腺的特征响应。
- 4) 针对原生分割头在对底层特征图上采样时跳跃较大,易造成信息丢失问题,提出了分层渐进特征融合头。通过层级化的上采样方法,使得来自网络各层的特征能够平滑且有效地融合,保证了最终分割预测的细节精度和整体一致性。

2. 相关工作

胰腺是人体腹腔内的一个软组织腺体,基于腹部 CT 图像的胰腺分割通常面临着如下挑战: 1) 胰腺解剖学特异性高,不同患者之间有着明显的形态差异。2) 腹部 CT 图像中,由于胰腺与周围组织在 X 射线密度上的相似性高,胰腺边界难以被清晰区分。3) 胰腺及其肿瘤在腹腔 CT 序列图像中占据的体积比例较小(约占 0.5%),分割模型易学习到负样本特征。因此,基于腹部 CT 图像的胰腺分割是一项艰巨的任务。

随着深度学习技术的迅猛发展,基于卷积神经网络(CNN)的分割方法逐渐取代传统机器学习的方法,提升了胰腺分割的精准度。由于胰腺在器官腹腔 CT 序列中占比较小,为解决类不平衡问题,Roth 等人 [9]提出了由粗到细的两阶段分割模型,在粗分割阶段采用超像素方法生成胰腺的候选区域,然后将候选区域导入分割网络进行局部剪枝细化。但由于超像素区域的不规则性,在进入细分网络时需要每个候选区域进行变形处理,这可能会引入失真并导致胰腺统一特征更难以学习。为了减少网络的学习参数和类不平衡问题的影响,Zheng 等人 [10]提出通过阴影集的分析来确定胰腺图像分割的不确定性区域。Liu 等人 [11]利用肝脏、脾脏和肾脏的位置信息来对胰腺器官进行粗定位。Huang 等人 [12]直接以中心点为基础剪裁图像作为细分网络的输入。然而,上述多阶段的分割方法无法实现端到端的联合优化,且训练困难,易失去全局信息,影响分割精度。

考虑到两阶段分割方法在实际临床诊断中存在的不足之处,单阶段胰腺分割方法在临床应用中具有更广泛的前景。2015年,Ronneberger 等人 [5]在 MICCAI 大会上提出了 UNet,该模型主要用于解决医学图像中像素级别的分类问题。由于其轻量化的 U 型编解码结构通过跳跃连接,实现了编码器和解码器特征的通道级特征融合,有效地解决了上采样过程中特征信息不足的问题,其在大型器官中展现出的优异分割性能,引领了医学图像分割的技术变革。Navab 等人 [13]针对三维医学图像的处理在 UNet 的基础上提出了 V-Net 模型,为了在单阶段的分割模型中有效抑制负样本区域的影响,Schlemper 等人 [14]在 3D

UNet 中引入注意力机制, 提出了 Attention-UNet 来进行胰腺分割, 削减了无关背景的计算权重, 以此来提升胰腺分割的准确性。Nishio 等人[15]利用数据增强方法对数据集进行预处理, 并利用深度 UNet 对目标数据集进行训练, 得出了较于 UNet 更准确的分割结果。Li 等人[16]提出了一种新的具有对抗性机制的多级金字塔池化残差 U 型网络。Tian 等人[17]提出了一种马尔可夫链蒙特卡罗(MCMC)引导的卷积神经网络。Ma 等人[18]通过通道剪枝技术实现了对 UNet 模型的压缩, 在保持分割精度的前提下有效减少了模型参数量。

然而, 基于 UNet 的分割模型应用于胰腺分割时, 却呈现出以下不足: 由于胰腺器官的边界模糊, 其边缘特征难以捕捉。而 UNet 下采样的池化操作, 使其编码阶段易丢失细节语义信息。解码阶段又仅通过上采样恢复高分辨率语义特征, 无法充分补偿先前的信息损失。Lu 等人[19]针对这一问题提出了一种新的注意力模块 CBAM (Channel-wise and Spatial), 该模块分别对通道特征和空间特征进行加权, 在下采样过程中提取的特征经过 CBAM 注意力模块的筛选后融合到上采样, 以确保胰腺的高分辨率细节特征得以保留。该方法在一定程度上缓解了胰腺器官边界模糊带来的影响, 但未对 UNet 串行结构导致的跨层特征融合问题加以改善, 仍会造成部分特征信息的丢失。2019 年, Sun 等人[8]提出了高分辨率网络, 其网络结构如图 1 所示。HRNet 为减少网络占用的内存消耗, 在图像进入主干网络之前, 进行两次连续的跨步卷积下采样。之后的主干网络包含四个分辨率分支与四个特征提取阶段, 其中, 第一个分辨率分支保持高分辨率表示完整保留了特征的位置和细节信息。每个特征提取阶段由多个顺序连接的残差卷积单元组成, 每个卷积单元由 2 个普通卷积和 1 个残差连接构成。每个特征提取阶段之后, 对底层特征图进行两倍下采样, 并将所有的分辨率分支进行特征融合, 高分辨率图使分割结果在空间中表征更精准, 低分辨率图使语义特征提取更充分。最后, HRNet 的分割头将所有分支直接上采样至最大分辨率后进行融合, 并输出最终分割结果。

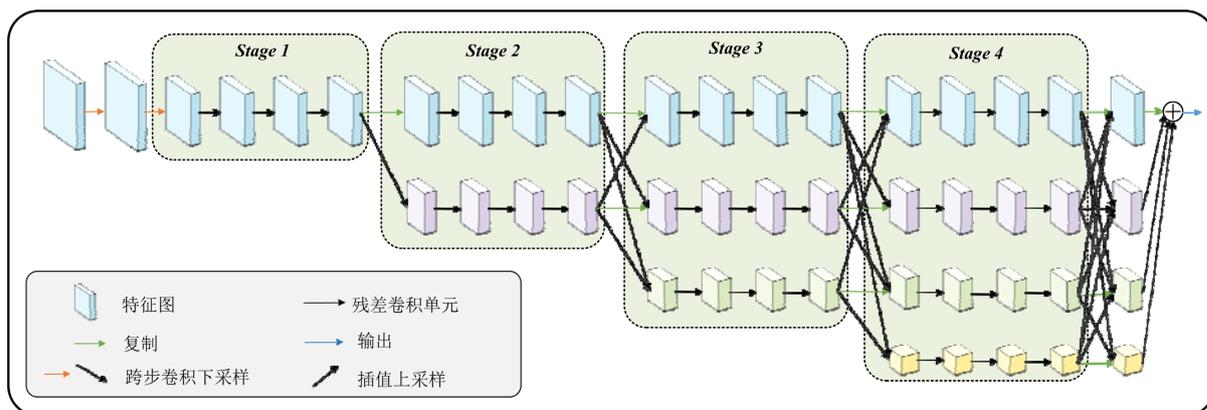


Figure 1. Diagram of HRNet structure

图 1. HRNet 结构图

Li 等人[20]首次将高分辨率网络应用于胰腺分割任务中, 验证了其并行多尺度特征融合结构在该领域相较于 UNet 串行编解码结构的先进性。然而, 鉴于胰腺分割任务的复杂性, HR-Net 在执行胰腺分割任务时, 仍存在以下不足: 1) HRNet 主要基于二维设计, 而 CT 图像本质上是三维数据, 二维分割网络难以充分利用三维图像中蕴含的空间上下文信息, 限制了模型在深层次特征理解方面的能力。2) 输入图像在送入主体结构之前需经历两次 2 倍跨步卷积下采样, 此操作虽有助于减少计算负担, 但也导致了胰腺的位置和细节信息的丢失, 从而影响了分割性能。3) HRNet 的原生分割头中, 各级特征图被直接上采样至最高分辨率, 导致深层特征与浅层特征之间差异明显, 二者进行特征融合时跳跃较大, 深层特征无

法被有效利用。

纵览现有文献, 基于深度学习技术的胰腺分割领域已取得显著研究进展。然而, 这些方法尚存在不容忽视的局限性。本文深入分析了基于编码-解码架构的主流胰腺分割网络的关键不足, 并针对高分辨率网络进行胰腺分割时遭遇的挑战提出一系列改进措施。

3. 方法

针对基于 U 型编解码网络和基于高分辨率网络分割胰腺存在的问题, 本文提出结合空洞特征增强与改进 HR-Net 的胰腺分割方法, 具体分割流程如图 2 所示。

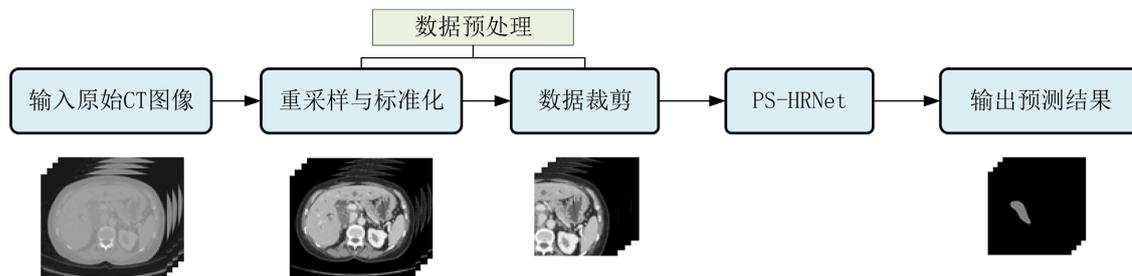


Figure 2. Flowchart of the segmentation method
图 2. 本文分割方法流程图

3.1. 数据预处理

医学图像预处理是医学图像分析流程中至关重要的一步, 主要目的是提高图像数据的质量, 以便更准确、有效地进行后续的图像分割任务, 本文的数据预处理主要包括以下几个步骤:

1) 数据重采样: 为减少由于 CT 图像数据在采样间隔上的不一致性导致的实验偏差, 通过三线性插值算法将图像沿 x 轴、y 轴、和 z 轴三个维度的采样间隔分别调整至 0.85 mm, 0.85 mm, 1.0 mm。

2) 数据标准化: 在处理腹部 CT 图像时, CT 值的标准化是一项重要的预处理步骤, 旨在突出胰腺区域的结构。胰腺在 CT 图像中的 CT 值大约在 30 到 40 HU 左右, 但为了充分考虑周围组织和结构的影响, 以及增强胰腺与周围组织的对比度, 本文将 CT 值标准化到-100 HU 到+200 HU 的范围内, 使得胰腺及周围的主要组织的细节特征更加明显。

3) 数据裁剪: 鉴于胰腺在 CT 图像中仅占较小的比例, 本文依据预标注的标签信息, 对所有输入图像执行了区域裁剪, 将其裁剪至包含胰腺区域的 $128 \times 128 \times 64$ 体素块, 旨在确保输入数据的统一性。

3.2. PS-HRNet 模型

针对基于编码-解码网络和基于高分辨率网络分割胰腺存在的问题, 本文提出结合空洞特征增强与改进 HR-Net 的胰腺分割方法, 并将其命名为胰腺高分辨率分割网络(Pancreatic Segmentation High Resolution Network, PS-HRNet), 其网络架构如图 3 所示。PS-HRNet 对基线模型做了如下改进:

1) 将 HRNet 中的二维卷积操作拓展为三维卷积, 以便有效捕获 CT 图像中的三维特征。

2) 考虑到胰腺体积较小, 本文将 HRNet 初始的二次跨步卷积下采样操作, 调整为单次下采样, 保留了高分辨率特征图的空间信息。

3) 设计多尺度空洞特征增强模块, 以缓解密集特征融合产生的大量计算成本, 并有效扩大模型感受野, 融合深度特征信息。

4) 在模型多分辨率融合模块之后引入高效通道注意力模块, 让网络更加精细地调整不同分辨率的特

征信息。

5) 提出分层渐进特征融合头, 改善原生分割头无法有效利用各级分辨率特征问题。

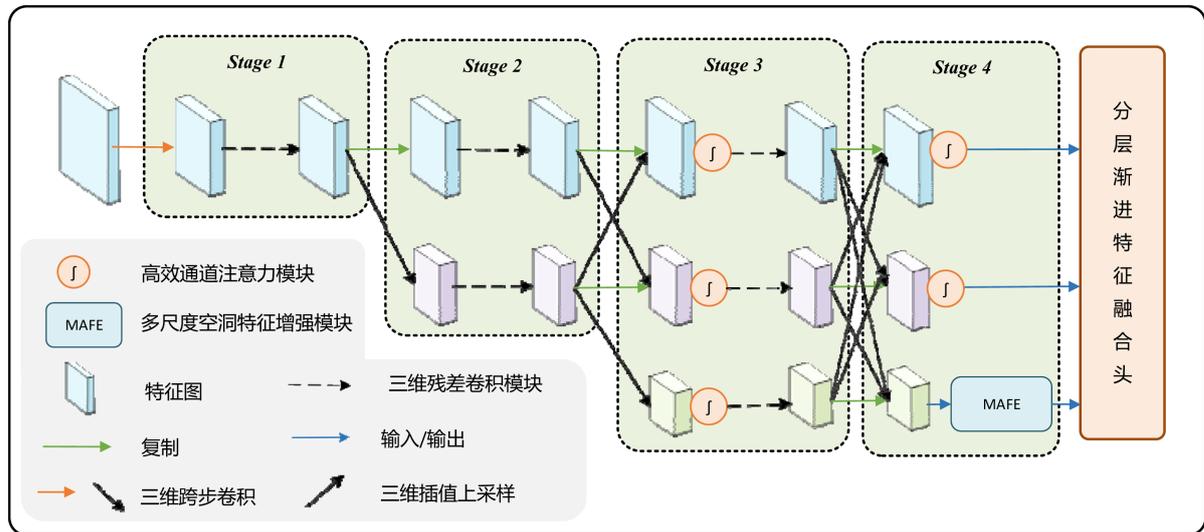


Figure 3. Diagram of PS-HRNet structure

图 3. PS-HRNet 结构图

3.3. 多尺度空间空洞特征增强模块(MAFE)

在 HRNet 中, 输入图像进入主干网络前, 会进行连续两次的 2 倍跨步卷积下采样, 达到降低 GPU 显存消耗的目的, 但由于胰腺在腹部 CT 图像中占比较小, 这种操作会导致严重的空间信息丢失。本文方法将其调整为 2 倍的下采样, 以保留胰腺细节信息。然而, HRNet 复杂的并行子网结构增加了模型的空间复杂度, 多个阶段的密集特征融合带来了大量重复或冗余的信息, 这会导致模型训练时产生大量的计算成本, 限制了网络深度的扩展。因此, 本文设计多尺度空洞特征增强模块(Multiscale Atrous Feature Enhancement, MAFE)取代第四阶段底层特征图的下采样操作, 通过堆叠不同膨胀率的空洞卷积, 在降低网络参数的同时扩展了模型的感受野, 并通过加权机制增强了关键信息的表达, 进一步提升了网络性能。

多尺度空间空洞特征增强模块的结构如图 4 所示。首先, 将 $F \in R^{H \times W \times D \times C}$ (H 为高度, W 为宽度, D 为深度, C 为通道数) 作为该模块的输入, 输入特征图 F 被并行地送入三个不同配置的卷积层: 一个尺寸为 $3 \times 3 \times 3$ 的标准卷积层以捕获局部特征信息, 其输出记为 x_1 ; 另外两个分别采用 $3 \times 3 \times 3$ 尺寸, 膨胀率为 2 和 4 的空洞卷积层, 旨在通过较大的感受野抽象出更加深层的特征信息, 这两层输出分别记作 x_2 与 x_3 。随后, 这三个分支的输出通过拼接操作融合为一个联合特征 x_c , 以便综合不同感受野下的空间信息, 获得富含深层多尺度信息的联合特征表示。考虑到经卷积提取后的特征图中各通道所蕴含信息的差异可能对分割结果产生不同影响, 本研究进一步引入了一个注意力向量以指导特征学习过程, 优化通道间信息的重要性评估, 为每个通道分配适当权重, 以精炼融合后的特征表示。具体地, 首先通过一个 $3 \times 3 \times 3$ 的标准卷积层提取的信息, 接着利用全局平均池化(Global Average Pooling, GAP)和 $1 \times 1 \times 1$ 的卷积操作计算得到通道维度上的注意力权重, 随后通过批归一化(Batch Normalization, BN)层进行规范化, 并采用 Sigmoid 激活函数进行激活, 最终, 将得到的注意力权重与原始联合特征进行加权乘法操作, 从而赋予重要特征更高的注意力分数, 同时抑制不关键的信息, 增强了特征表达的精度, 实现了特征的精细化。MAFE 模块的数学表达式如式(1)所示。

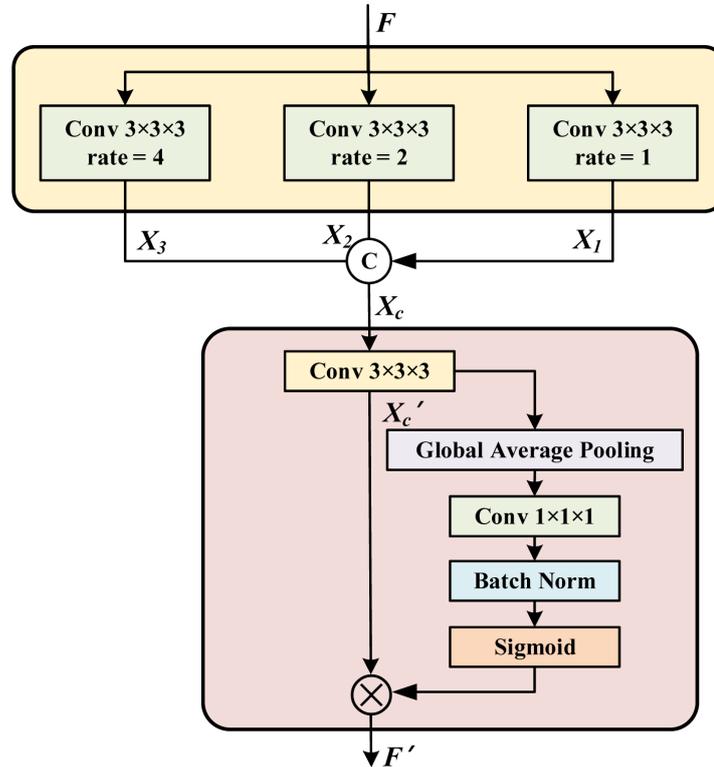


Figure 4. Diagram of multiscale atrous feature enhancement module
图 4. 多尺度空间空洞特征增强模块结构图

$$F' = x'_c \otimes \sigma \left(F_{BN} \left(Conv_{1 \times 1} \left(F_{GAP} \left(Conv_{3 \times 3} (x_c) \right) \right) \right) \right) \quad (1)$$

其中, $Conv$ 表示卷积操作, \otimes 表示逐元素相乘, σ 表示 Sigmoid 函数, F_{BN} 表示批归一化, F_{GAP} 表示全局池化。

3.4. 高效通道注意力模块(ECA)

HRNet 通过在网络中维持并融合多个分辨率的特征图, 实现对不同尺度特征的综合利用。通道注意力机制[21] (Squeeze-and-Excitation, SE)能够为输入信息根据其重要性赋予不同的权重。为了进一步强调各层的重要特征, 本文在每个多分辨率特征融合节点后引入高效通道注意力模块(Efficient Channel Attention, ECA)。ECA 模块在 SE 机制的基础上进行了优化, 去除 SE 中的全连接层, 并采用一维卷积来捕捉相邻通道间的依赖, 以轻量化的架构实现了特征引导。

ECA 模块操作流程如下: 首先, 采用全局平均池化(GAP)对融合后的多分辨率特征进行空间压缩, 随后通过一维卷积处理压缩后的特征, 以学习通道间的关系。通过 Sigmoid 激活函数生成最终的通道权重, 并将这些权重与原始的多分辨率特征进行逐通道的加权乘法, 得到加强的特征表示。ECA 模块的工作原理如图 5 所示, 通过简化的结构减少了模型的参数量, 同时也减少了因维度缩减导致的信息损失, 有效增强了网络对关键特征的识别能力, 从而提取到更加精准的胰腺细节特征信息。其中, 给定通道维度 C 卷积核大小 k 可自适应确定, 其计算公式为式(2)。

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma_{odd}} \right\rfloor \quad (2)$$

其中, $|e|_{odd}$ 表示为最接近 e 的奇数, γ 与 b 皆为常数, 以维系特征通道数量与一维卷积大小的比例。

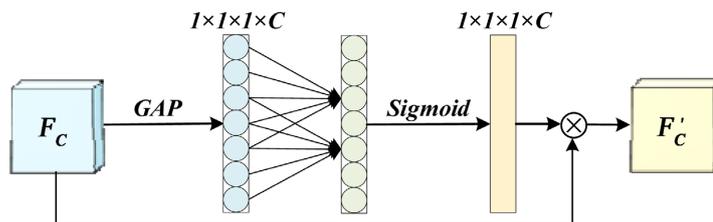


Figure 5. Diagram of efficient channel attention module
图 5. 高效通道注意力模块模块结构图

3.5. 分层渐进特征融合头(LPFF)

在 HRNet 原生分割头中, 各个不同分辨率的特征图通过上采样操作被统一调整至最高分辨率, 并通过拼接操作合成最终的特征表示。然而, 这一直接上采样方法在处理深层特征图时, 由于尺度跳跃过大, 容易导致信息的丢失。同时, 深层与浅层特征之间的显著差异也加剧了在特征融合过程中的困难, 限制了深层特征的有效利用。鉴于此, 本研究提出了一种分层渐进特征融合头(Layered Progressive Feature Fusion, LPFF), 旨在通过渐进和平滑的方式实现多尺度特征的有效整合。

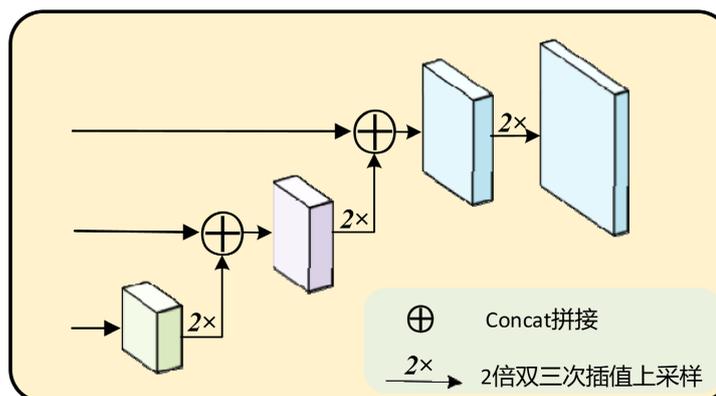


Figure 6. Diagram of layered progressive feature fusion head
图 6. 分层渐进特征融合头结构图

分层渐进特征融合头结构如图 6 所示, 该模块按照特征图分辨率自低至高的顺序, 逐对进行特征融合, 最后总体上将特征图采样至原图大小以实现分割。此外, 原生分割头在上采样过程中采用的双线性插值法虽能实现尺寸调整, 但可能会损害对胰腺分割任务至关重要的边缘信息, 导致图像不连续性, 从而影响特征融合的质量。为此, 本文引入双三次插值法对上采样步骤进行优化。双三次插值是通过在两个方向上应用三次插值来实现。具体来说, 它不仅考虑了目标像素点最近的四个像素(双线性插值法所做的), 而是考虑了更广的 16 个像素点, 即在每个方向上各取 4 个点。插值函数是一个三次多项式, 其参数由邻近的像素值决定。这使得双三次插值能够更准确地预测目标像素点的值, 其数学表达式如公式(3)所示

$$f(x, y) = \sum_{i=-1}^2 \sum_{j=-1}^2 a_{ij} x^i y^j \quad (3)$$

其中, x 和 y 是相对于目标像素位置的坐标偏移, a_{ij} 是根据邻近像素值通过插值计算得到的系数。

3.6. 损失函数

由于胰腺在腹部 CT 图像中占比较小, 因此在损失函数的设计上不光需要考虑边界细节的准确性判定还要考虑类别不平衡问题。Dice 损失函数通过优化预测分割与真实标注的重叠区域来改善类别不平衡问题, 而二元交叉熵(Binary Cross-Entropy, BCE)损失函数则关注于像素级的分类准确性。本文将二者结合, 提出混合损失函数, 以更全面地指导模型学习, 通过捕捉不同方面的误差, 促进模型在胰腺等小器官的精确分割上取得更好的性能。Dice 损失函数、二元交叉熵损失函数与本文混合损失函数的数学表达式如公式(4)~(6)所示:

$$Loss_{(Dice)} = -\frac{2\sum_{i=1}^N x_i \cdot \hat{x}_i + s}{\sum_{i=1}^N x_i + \sum_{i=1}^N \hat{x}_i + s} \quad (4)$$

$$Loss_{(BCE)} = -\sum_{i=1}^N [x_i \log(\hat{x}_i)] + (1 - x_i) \log(1 - \hat{x}_i) \quad (5)$$

$$Loss = \alpha Loss_{(BCE)} + (1 - \alpha) Loss_{(Dice)} \quad (6)$$

其中, N 表示输入体素个数, x_i , \hat{x}_i 分别代表了预测体素与标签体素, 以 α 设定不同损失函数的权重, 本文设置 $\alpha = 0.5$ 。

4. 实验及分析

为了验证本文算法对于胰腺分割任务的有效性, 本节将其与现有的胰腺分割方法进行分析比对。为保证验证实验的公平性与可靠性, 本文所采用的对比方法均使用原作者的开源代码, 并在当前公开的最大的、亦是应用最广泛的 NIH 胰腺数据集上进行训练与验证。

4.1. 实验数据集

本文采用美国国立卫生研究院(National institutes of health, NIH)公开的胰腺数据集[22] (NIH-Pancreas)验证本文算法的有效性。NIH-Pancreas 数据集是迄今为止最大规模的公开胰腺分割数据集, 其中包含 82 例腹部增强 CT 扫描序列和相应人工标注的胰腺标签, CT 序列体积大小为 $512 \times 512 \times D$, 其中 $D \in [181, 466]$, 表示横断面方向的切片样本数量。医务工作者对 CT 序列逐层进行手动标记, 并由经验丰富的放射科医生进行检验修正, 确保了数据集的可靠性。本文实验中按照 7:2:1 的比例, 将数据集划分为 57 例训练集、16 例验证集、9 例测试集。

4.2. 实验细节

本文在 Linux 系统上, 以 Python3.10 与 PyTorch 2.1.0 为基础搭建实验环境, 使用 NVIDIA GeForce RTX 3080 显卡训练模型。实验的参数设置对分割效果具有至关重要的影响, 本文实验参数设置详见表 1。

Table 1. Experimental parameter setting

表 1. 实验参数设置

优化器	批大小	学习率	迭代次数	动量参数
Adam	2	0.0001	500	$\beta_1 = 0.9$ $\beta_2 = 0.999$

4.3. 评价指标

本文采用医学图像分割常用的戴斯相似性系数(Dice Similarity Coefficient, DSC)、召回率(Recall)、精确率(Precision)作为分割结果的主要评价指标, 其数学表达式如公式(7)~(9)所示:

$$DSC(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|} \quad (7)$$

公式(7)中, X 为胰腺掩码标签像素集合, $X \in \{0, 1\}^{d \times h \times w}$ 。 Y 为胰腺掩码的预测像素集合 $Y \in \{0, 1\}^{d \times h \times w}$ 。 d, h, w 分别指胰腺 CT 序列在冠状面、横断面、矢状面的大小。戴斯相似性系数被广泛用于衡量医学图像分割中分割结果与真实标签的相似性, DSC 越接近于 1, 说明分割结果愈准确。

$$\text{recall} = \frac{TP}{TP + FN} \quad (8)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (9)$$

其中, TP 表示真阳性, 即正确地将胰腺像素点分类为胰腺的数量, FP 表示假阳性, 即被错误地分类为背景的像素点数量, FN 表示假阴性, 即被错误地分类为胰腺的像素点数量。召回率越大表示网络对于胰腺特征越敏感, 精确率越大表示网络对于胰腺像素判定越准确。

4.3. 训练过程变化曲线

在训练过程中, 研究者们特别关注损失值(Loss)的变化, 以此作为评估模型学习状态的重要指标。若在训练阶段, 损失值随着迭代次数的增加而下降, 这表明模型的学习状态良好, 正向着收敛方向逐步调整学习参数。本章绘制了训练过程中的训练集(蓝色)和验证集(红色)损失值随迭代次数(Epoch)变化的曲线图, 如图 7 所示。

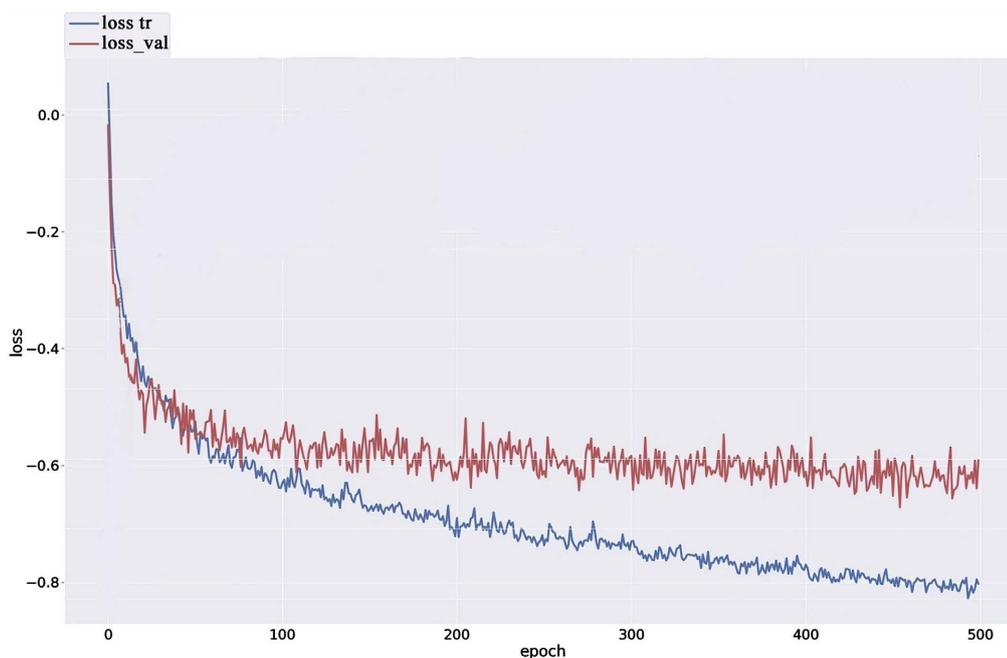


Figure 7. PS-HRNet model Loss variation curves
图 7. PS-HRNet 模型 Loss 变化曲线

4.4. 量化评估

为了验证本文分割方法的有效性, 本文将其与近年来主流的胰腺分割模型进行性能对比。具体包括前文提及的 Ronneberger 等人提出的医学图像分割模型 UNet [5]、Sun 等人提出并行多尺度特征融合结构的 HR-Net [8]、Navab 针对 3D 医学图像分割提出的 V-Net [13]以及 Oktay 等人针对胰腺分割提出的 Attention UNet [14]。表 2 展示了本文方法与主流的胰腺分割模型在 NIH-Pancreas 数据集上的分割结果量化对比, 最优结果加粗表示。

Table 2. Comparative results of mainstream pancreas segmentation models

表 2. 主流胰腺分割模型的对比结果表

模型	卷积维度	召回率	准确率	DSC
UNet	2D	71.9%	76.1%	73.2%
HR-Net	2D	75.2%	77.9%	76.5%
V-Net	3D	81.4%	83.4%	82.7%
Attention-UNet	3D	82.8%	83.2%	83.3%
Ours	3D	84.5%	86.3%	85.6%

如表 2 所示, UNet 在面对胰腺这类小器官分割时召回率较低, 即存在大量过分割现象。本文方法在召回率、准确率和 Dice 相似系数分别提高了 11.6%, 10.2%, 12.4%。HR-Net 的多尺度信息融合使其更好地捕捉胰腺的细节和形态特征, 本文方法在召回率、准确率和 Dice 相似系数较之分别提高了 8.3%, 8.4%, 9.1%。V-Net 作为三维分割模型, 在实际分割精度上明显高于前者, 本文方法在召回率、准确率和 Dice 相似系数较之分别提高了 2.1%, 3.1%, 2.9%。Attention UNet 作为 UNet 针对胰腺分割任务的改进模型, 注意力机制的引入使得其能迅速捕捉胰腺图像中的重要特征, 本文方法在召回率、准确率和 Dice 相似系数较之分别提高了 1.7%, 2.1%, 2.3%。

表 3 展示了先进的胰腺分割模型在 NIH-Pancreas 数据集上的分割结果进行量化对比, 所对比的先进胰腺分割模型在本文相关工作部分均有介绍, 最优结果加粗表示。

Table 3. Comparison results of advanced pancreas segmentation models

表 3. 先进胰腺分割模型的对比结果表

模型	卷积维度	召回率	准确率	DSC
文献[15] (2020)	2D	/	/	78.9%
文献[10] (2020)	2D	86.3%	83.1%	84.4%
文献[16] (2021)	3D	83.8%	83.2%	83.0%
文献[17] (2021)	2D	82.7%	74.6%	78.1%
文献[20] (2021)	3D	/	/	85.5%
文献[11] (2022)	3D	84.6%	86.1%	85.4%
文献[12] (2022)	2D	77.4%	85.3%	82.9%
文献[18] (2022)	2D	/	/	81.4%
本文方法	3D	84.5%	86.3%	85.6%

从表中可知, 基于 3D 卷积的分割模型在不同评估指标下的分割精度普遍高于基于 2D 卷积的分割模型。本文方法在准确率和 Dice 相似系数这两个关键评估标准下取得了最优的成绩, 这进一步证明了本文方法的有效性和优越性。

4.5. 定性评价

图 8 展示了多个样本的横截面切片分割效果比较图, 图中每列代表不同的切片样本及其分割结果。图 8 的可视化对比表明: 当胰腺面积较大且边界特征明显时(#53-Slice103), 五种分割模型都可以得到较准确的分割结果。当待分割区域边界复杂(#18-Slice98)或是有背景器官干扰(#82-Slice113)时, 分割准确度都有所下降, 其中 UNet 受到的影响最大, 其分割预测图多处出现过分割现象。从预测结果的形态和边界的拟合来看, 本文提出的方法在局部结构不明晰的情况下取得了更好的分割性能。

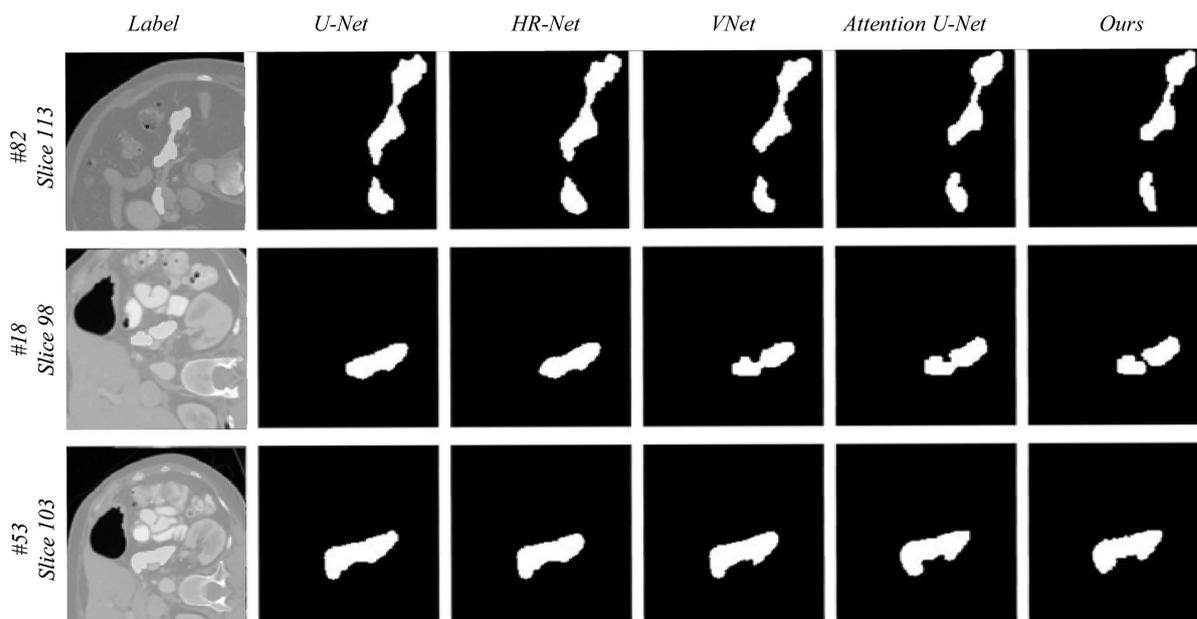


Figure 8. Comparison of cross-sectional segmentation effects for different networks

图 8. 不同网络的横截面分割效果对比

为了更直观地彰显不同网络的整体分割结果, 图 9 展示了不同网络的 3D 重建模型效果图, 从左往右分别是通过真实标记和五种不同的网络方法生成的 3D 模型。本文方法生成的 3D 模型在结构上更加完整, 与真实标记的 3D 模型相似度最高, 证明了本文方法分割效果的优越性。

4.6. 消融实验

通过分析不同初始下采样倍率的对比实验、主体网络模块的有效性以及分层渐进特征融合头的有效性分别验证本文方法的有效性。

4.6.1. 不同初始下采样策略的对比

本节旨在探究不同的初始下采样策略对 PS-HRNet 分割性能的具体影响。将 HRNet 采用的四倍下采样与本文采用的二倍下采样进行对比实验, 相应的实验结果如表 4 所示, 其中最优结果加粗标注。

根据表 4 的数据, 可以观察到初始的 2 倍下采样的分割精度要优于 4 倍下采样。其根本原因在于, 胰腺器官在腹部 CT 图像中占比小, 且其形态及位置具有较大的变异性。在采用四倍下采样时, 极易导

致关键的特征位置与细节信息的丢失, 进而降低了模型在胰腺分割任务上的准确性。

4.6.2. 网络模块的有效性

为了验证本文在基线网络 HR-Net 中所改进的多尺度空间空洞特征增强模块(MAFE)与高效通道注意力模块(ECA)的有效性, 本文设计了一系列消融实验, 实验结果如表 5 所示。

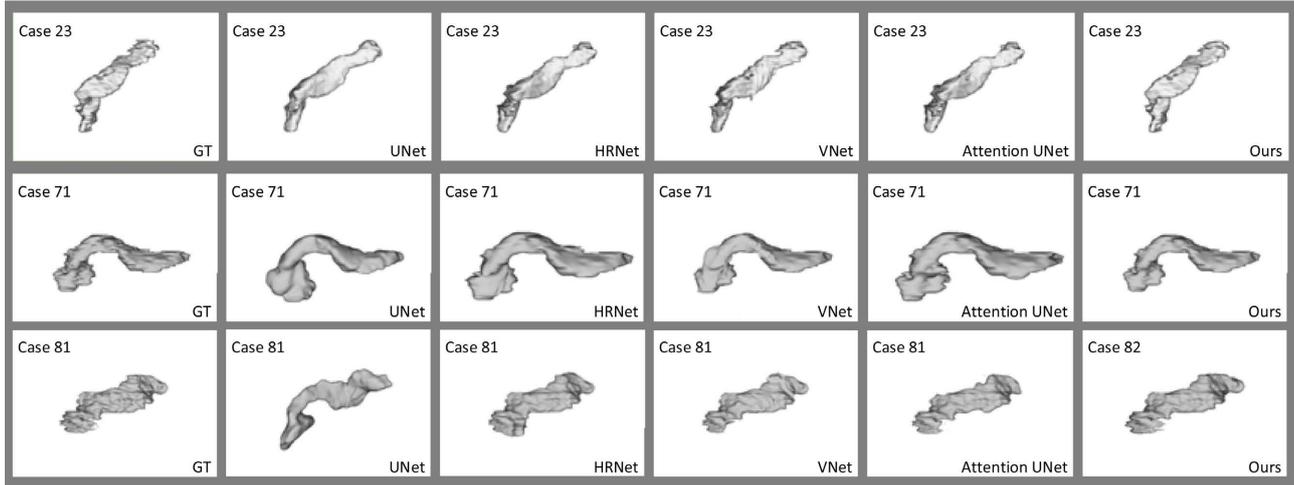


Figure 9. Comparison of 3D reconstructed models with different network segmentation results

图 9. 不同网络分割结果的 3D 重建模型对比

Table 4. Segmentation result data for down-sampling with different multiplicities

表 4. 不同倍率的下采样的分割结果对比

初始下采样倍率	召回率	准确率	DSC
4	82.4%	83.4%	82.7%
2	84.5%	86.3%	85.6%

Table 5. Comparative results of ablation experiments with different network structures

表 5. 不同网络结构的消融实验结果表

MAFE	ECA	召回率	准确率	DSC
×	×	82.5%	84.6%	83.2%
√	×	83.9%	85.2%	84.1%
×	√	83.3%	85.7%	84.5%
√	√	84.5%	86.3%	85.6%

从表 5 的实验结果可以看出, 经过多尺度空间空洞特征增强模块(MAFE)重构后的模型, 各项性能指标相较于基线模型均有提升, 在召回率、准确率和 Dice 相似系数的评分上分别提高了 1.4%, 0.6%, 0.9%。这一增益在召回率的提高中表现得尤为明显, 这表明 MAFE 模块聚焦于胰腺区域的结构特征, 抑制了假阳性结果, 有效减少了过分割现象的产生。在集成了高效通道注意力模块(ECA)的模型中, 相较于基线模型在召回率、准确率和 Dice 相似系数的评分上分别提高了 0.8%, 0.9%, 1.3%。最后, 网络中 MAFE 模块和 ECA 模块的组合表明, 这两种模块的结合产生互补效应, 共同促进了胰腺分割性能的提升。

4.6.3. 分层渐进特征融合头的有效性

为了评估分层渐进特征融合头对胰腺分割效能的贡献, 本文通过比较引入分层渐进特征融合头前后网络性能的差异来进行分析。实验结果如表 6 中所示。

Table 6. Comparative results of ablation experiments with layered progressive feature fusion head

表 6. 分层渐进特征融合头消融实验

分割头	召回率	准确率	DSC
原生分割头	82.8%	84.4%	84.2%
分层渐进特征融合头	84.5%	86.3%	85.6%

与原生分割头相比, 采用分层渐进特征融合头的网络在召回率、准确率及 Dice 系数上分别实现了 1.7%、1.9% 和 1.4% 的提升。这一结果表明, 通过引入分层渐进特征融合头, 以层级的上采样的策略进行多尺度特征的融合, 能够有效缓解不同层级的特征图融合时跨度较大的问题, 充分利用各分辨率层特征, 减少在特征融合过程中的信息丢失, 从而增强胰腺分割网络的性能。

5. 总结

针对现有胰腺分割网络编码过程丢失位置和细节信息的问题, 本文提出了一种结合空洞特征增强与改进 HR-Net 的胰腺分割方法。提出多尺度空洞特征增强模块改进基线网络架构, 扩张模型感受野的同时减少密集特征融合带来的冗余计算。在各阶段多分辨率融合之后, 设计高效通道注意力模块, 精确调整胰腺特征的响应, 细化对胰腺结构的识别和分割。最后, 设计分层渐进特征融合头改进原生分割头, 缩小深浅层语义特征的差距, 充分利用各分辨率层特征。

在 NIH 胰腺数据集上的实验证明, 本文提出的方法可有效地克服复杂胰腺分割中的挑战, 从而实现高精度的分割结果, 其分割精度超越了目前主流的胰腺分割模型。然而, 本文仅依靠 CT 影像进行图像分割, 未来工作将探索如何结合患者的其他身份信息进行跨模态辅助分割, 以实现更为精准的分割, 值得我们思考。

基金项目

浙江省卫生健康委员会面上项目(2022KY122)资助; 浙江省中医药科技计划项目(2019ZA023)资助。

参考文献

- [1] Zheng, Y., Liu, D., Georgescu, B., *et al.* (2017) Deep Learning Based Automatic Segmentation of Pathological Kidney in CT: Local versus Global Image Context. In: Lu, L., Zheng, Y., Carneiro, G. and Yang, L., Eds., *Deep Learning and Convolutional Neural Networks for Medical Image Computing*, Springer, Cham, 241-255. https://doi.org/10.1007/978-3-319-42999-1_14
- [2] Wang, Z., Bhatia, K.K., Glocker, B., *et al.* (2014) Geodesic Patch-Based Segmentation. In: Golland, P., Hata, N., Barillot, C., Hornegger, J. and Howe, R., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014*, Springer, Cham, 666-673. https://doi.org/10.1007/978-3-319-10404-1_83
- [3] Wolz, R., Chu, C., Misawa, K., *et al.* (2013) Automated Abdominal Multi-Organ Segmentation with Subject-Specific Atlas Generation. *IEEE Transactions on Medical Imaging*, **32**, 1723-1730. <https://doi.org/10.1109/TMI.2013.2265805>
- [4] Chu, C., Oda, M., Kitasaka, T., *et al.* (2013) Multi-Organ Segmentation Based on Spatially-Divided Probabilistic Atlas from 3D Abdominal CT Images. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C. and Navab, N., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*, Springer, Berlin, 165-172. https://doi.org/10.1007/978-3-642-40763-5_21
- [5] Ronneberger, O., Fischer, P. and Brox, T. (2015) UNet: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Inter-*

- vention—MICCAI 2015, Springer, Cham, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [6] 曹路洋, 李建微. 面向深度学习的胰腺医学图像分割方法研究进展[J]. 小型微型计算机系统, 2022, 43(12): 2591-2604.
- [7] Gibson, E., Giganti, F., Hu, Y., *et al.* (2018) Automatic Multi-Organ Segmentation on Abdominal CT with Dense V-Networks. *IEEE Transactions on Medical Imaging*, **37**, 1822-1834. <https://doi.org/10.1109/TMI.2018.2806309>
- [8] Sun, K., Xiao, B., Liu, D., *et al.* (2019) Deep High-Resolution Representation Learning for Human Pose Estimation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 5686-5696. <https://doi.org/10.1109/CVPR.2019.00584>
- [9] Farag, A., Lu, L., Roth, H.R., *et al.* (2016) A Bottom-Up Approach for Pancreas Segmentation Using Cascaded Superpixels and (Deep) Image Patch Labeling. *IEEE Transactions on Image Processing*, **26**, 386-399. <https://doi.org/10.1109/TIP.2016.2624198>
- [10] Zheng, H., Chen, Y., Yue, X., *et al.* (2020) Deep Pancreas Segmentation with Uncertain Regions of Shadowed Sets. *Magnetic Resonance Imaging*, **68**, 45-52. <https://doi.org/10.1016/j.mri.2020.01.008>
- [11] Liu, Z., Su, J., Wang, R., *et al.* (2021) Pancreas Co-Segmentation Based on Dynamic ROI Extraction and VGGUNet. *Expert Systems with Applications*, **192**, Article ID: 116444. <https://doi.org/10.1016/j.eswa.2021.116444>
- [12] Huang, M.L. and Wu, Y.Z. (2022) Semantic Segmentation of Pancreatic Medical Images by Using Convolutional Neural Network. *Biomedical Signal Processing and Control*, **73**, Article ID: 103458. <https://doi.org/10.1016/j.bsp.2021.103458>
- [13] Milletari, F., Navab, N. and Ahmadi, S.A. (2016) V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 2016 *Fourth International Conference on 3D Vision (3DV)*, Stanford, 25-28 October 2016, 565-571. <https://doi.org/10.1109/3DV.2016.79>
- [14] Oktay, O., Schlemper, J., Folgoc, L.L., *et al.* (2018) Attention UNet: Learning Where to Look for the Pancreas. arXiv: 1804.03999.
- [15] Nishio, M., Noguchi, S. and Fujimoto, K. (2020) Automatic Pancreas Segmentation Using Coarse-Scaled 2D Model of Deep Learning: Usefulness of Data Augmentation and Deep UNet. *Applied Sciences*, **10**, Article 3360. <https://doi.org/10.3390/app10103360>
- [16] Li, M., Lian, F., Wang, C., *et al.* (2021) Accurate Pancreas Segmentation Using Multi-Level Pyramidal Pooling Residual UNet with Adversarial Mechanism. *BMC Medical Imaging*, **21**, Article No. 168. <https://doi.org/10.1186/s12880-021-00694-1>
- [17] Tian, M., He, J., Yu, X., *et al.* (2021) MCMC Guided CNN Training and Segmentation for Pancreas Extraction. *IEEE Access*, **9**, 90539-90554. <https://doi.org/10.1109/ACCESS.2021.3070391>
- [18] 马豪, 刘彦, 张俊然. 基于模型压缩与重构 UNet 的胰腺分割[J]. 计算机工程与设计, 2022(7): 43.
- [19] Lu, L., Jian, L., Luo, J., *et al.* (2019) Pancreatic Segmentation via Ringed Residual UNet. *IEEE Access*, **7**, 172871-172878. <https://doi.org/10.1109/ACCESS.2019.2956550>
- [20] Li, W., Wu, X., Hu, Y., *et al.* (2021) High-Resolution Recurrent Gated Fusion Network for 3d Pancreas Segmentation. 2021 *International Joint Conference on Neural Networks (IJCNN)*, Shenzhen, 18-22 July 2021, 1-7. <https://doi.org/10.1109/IJCNN52387.2021.9534042>
- [21] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-And-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [22] Roth, H.R., Lu, L., Farag, A., *et al.* (2015) Deeporgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference*, Munich, 5-9 October 2015, 556-564.