

基于双视角注意力的小样本医学图像分割

祖春晓, 郑贤彬, 郭 锴

上海理工大学健康科学与工程学院, 上海

收稿日期: 2025年7月22日; 录用日期: 2025年8月14日; 发布日期: 2025年8月25日

摘 要

专业医学图像的采集成本与标注要求高, 导致可用标注数据量不足, 小样本医学图像分割成为当前解决该问题的重要方向。传统的方法通常从支撑图像中学习原型, 并通过最近邻搜索的方法分割查询图像。尽管这些算法展现出较高的性能, 但由于部分容积效应导致组织边界模糊; 灰度不均匀性造成同类组织表征差异; 前景和背景共享相似视觉特征, 使得模型易产生混杂错误分割。为此本文提出双视角协同增强网络(Dual-Perspective Collaborative Enhancement Network, DPCENet), 通过空间-通道双视角注意力机制实现协同精准特征筛选, 具体来说, 空间注意力精确定位器官空间分布; 通道注意力动态增强相关语义通道, 抑制无关噪声。考虑到医学图像前景和背景共享许多视觉特征, 模型采用双路径平衡设计, 并行处理前景与背景特征流。在三个公开的医学图像数据集上的实验表明, 该方法的性能优于现有方法, 详细的分析也验证了该模块的有效性。

关键词

小样本语义分割, 医学图像, 双视角协同增强

Few-Shot Medical Image Segmentation Based on Dual-Perspective Attention

Chunxiao Zu, Xianbin Zheng, Kai Guo

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Jul. 22nd, 2025; accepted: Aug. 14th, 2025; published: Aug. 25th, 2025

Abstract

The high acquisition cost and high annotation requirements of professional medical images lead to insufficient available annotated data. Small sample medical image segmentation has become an important direction to solve this problem. Traditional methods usually learn prototypes from supporting images and segment query images through nearest neighbor search. Although these

algorithms show high performance, the partial volume effect leads to blurred tissue boundaries; grayscale inhomogeneity causes differences in the representation of similar tissues; the foreground and background share similar visual features, which makes the model prone to mixed mis-segmentation. To this end, this paper proposes a dual-perspective collaborative enhancement network (DPCENet), which achieves collaborative and precise feature screening through a spatial-channel dual-perspective attention mechanism. Specifically, spatial attention accurately locates the spatial distribution of organs; channel attention dynamically enhances related semantic channels and suppresses irrelevant noise. Considering that the foreground and background of medical images share many visual features, the model adopts a dual-path balanced design to process the foreground and background feature streams in parallel. Experiments on three public medical image datasets show that the performance of this method is better than that of existing methods, and detailed analysis also verifies the effectiveness of this module.

Keywords

Few-Shot Semantic Segmentation, Medical Images, Dual-Perspective Collaborative Enhancement

Copyright © 2025 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

自动医学图像分割[1]是医学图像处理中的关键技术,它在临床诊断和治疗中发挥着重要作用。然而,传统的自动分割方法往往依赖于大规模且高质量的像素级注释数据集,这些数据集的创建通常需要专业医生进行详尽的标注工作,这一过程不仅耗时耗力,而且成本昂贵。鉴于此,医学成像领域对小样本学习(few-shot learning) [2]的兴趣日益增长。小样本学习的核心优势在于,它能够从极为有限的像素级注释示例(即支持集)中提取出具有区分性的特征表示,进而对未标记的样本(即查询集)进行精确的像素级标签预测。这种方法极大地减少了对大规模注释数据集的依赖,同时保持了分割任务的准确性和高效性,为医学图像分割领域带来了一种更为经济、高效的解决方案。

现有的小样本医学图像分割方法主要分为两种,一种是源于 SE-Net [3]的交互式方法,该方法在调节分支与分割器之间构建密集连接并引入挤压激励(squeeze-and-excitation)模块,强化了模块间的特征交互,从而显著提升分割性能。另一种方法是基于原型的方法[2] [4]-[8],现阶段基于原型的方法逐渐成为主流,其原理是从少量标注的支持图像中学习语义类别的原型,来指导对未标注查询图像的分割。许多研究者致力于获取更准确且更具代表性的原型,如 Q-net [4]采用双路径(dual-path)架构来捕获双重尺度的特征; CRAPNet [5]通过引入循环相似性注意力模块精细化地提取查询和支持图像特征之间的丰富关系信息; RPT [6]通过缩小查询图像与支撑图像之间的差异迭代的优化前景原型。尽管这些基于前景原型的处理方法具有很好的性能,但现有方法在局部细节、组织边界处理上存在缺陷。而且现阶段方法常常将前景平均池化为单一原型,这种操作不可避免地会导致有用信息丢失,并且单一原型不能完全代表类的整体分布,因此使用模糊原型预测最终结果可能会导致相当大的预测偏差。在医学图像分割中,单纯依赖前景处理可能导致模型将前景与背景混淆,错误地将背景元素识别为前景。与此相对,自然图像的背景通常较为简单,其组成元素相对独立,便于通过视觉感知进行区分。医学图像这种背景的差异性对分析算法设计和图像处理技术提出了更高的要求。

本文为适应医学图像特定需求,提出双视角协同增强网络,通过空间-通道双视角注意力机制的协

同作用实现特征精准筛选:

(1) 在空间维度,设计空间注意力模块,通过梯度引导的特征定位策略精确捕捉器官的空间分布,强化边界区域的响应,解决因局部模糊性导致的轮廓断裂问题;

(2) 在通道维度,引入通道权重分配机制,通过语义相关性评估目标器官的显著性通道(如脾脏特有的纹理编码),同时抑制背景噪声的干扰;

(3) 针对医学图像中前景与背景特征具有高度耦合性,模型采用双路径平衡架构,通过并行分支分别处理前景与背景特征流,并引入动态权重分配模块自适应地调节空间-通道注意力在前景和背景路径中的贡献度,从而避免传统单路径模型中因特征混淆导致的性能退化问题。

2. 小样本医学图像分割

2.1. 问题设定

小样本医学图像分割的目的是在解决传统深度学习需要依赖大量标注数据的问题,该方法仅依赖少量标注样本,即可实现未见类别的高效分割。数据集分为训练集 D_{tr} 和测试集 D_{te} 其标签分别为 Y_{train} 和 Y_{test} , 二者类别不重叠($Y_{train} \cap Y_{test} = \emptyset$)。我们采用随机抽样策略将数据集分为支撑集 S 和查询集 Q , 即 $D_{tr} = \{S_i, Q_i\}_{i=1}^{N_{tr}}$, $D_{te} = \{S_i, Q_i\}_{i=1}^{N_{te}}$, 每个数据集由 K 个注释的支撑图像和包含 N 个类别的查询图像组成。支撑集提供少量有标签样本,通过提取支撑图像的特征计算原型;查询集包含同一类别未标注的图像,对查询图像提取特征,并与支撑集原型进行比对,通过相似度计算(如余弦相似度)生成分割掩码。支撑集和查询集都是由图像-掩码对组成: $S = \{(I_s^i, M_s^i)\}_{i=1}^K$ 和 $Q = \{(I_q^i, M_q^i)\}_{i=1}^N$, 其中 I_s^i 和 M_s^i 分别代表第 i 个图像及对应的真实分割掩码, M_q^i 仅用于训练,上标的 K 和 N 是支撑集和查询集中每个类别里图像-掩码对的数量,下标 s 和 q 代表该图像或掩码属于支撑集还是查询集。小样本医学图像分割模型的输入是支撑集图像-掩码对和查询图像,输出是查询图像最终预测的二进制分割掩码。在本文中,我们遵循先前的工作[7]使用 $K = N = 1$ 的元学习任务以及 SSL-ALPNet [2]和 ADNet [9]的实验设定。

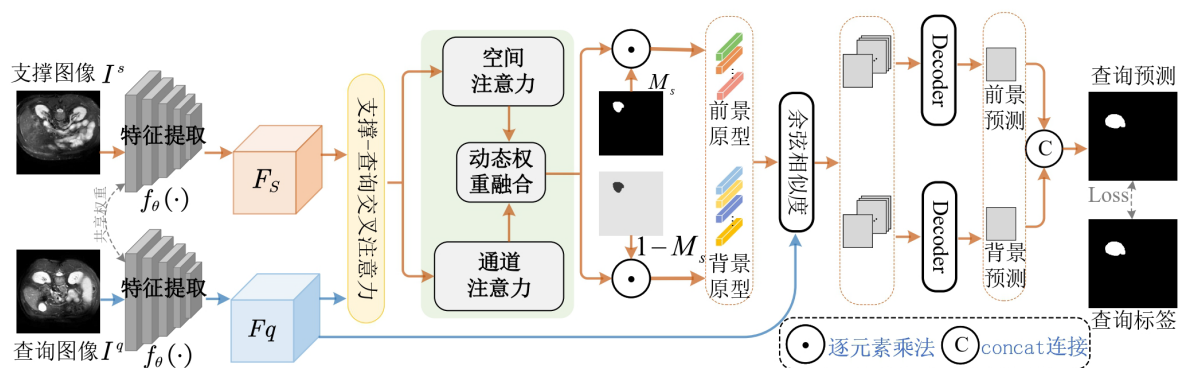


Figure 1. Overview of few-shot medical image segmentation method based on dual-perspective attention

图 1. 基于双视角注意力的小样本医学图像分割方法框架图

2.2. 模型框架

本文的整体框架如图 1 所示, 本文采取前景背景分离的双分支结构, 显示分离特征流, 使模型能更清晰地区分目标与非目标区域。具体分为三部分: (1) 从支撑图像中提取前景和背景特征, 并利用查询特征对支撑特征进行去噪实现支撑集到查询集的精准特征对齐; (2) 双视角协同增强网络模块, 通过两个独立的注意力网络分别提取不同维度信息的特征, 再利用自适应的动态权重融合机制提取最优前景原型和

背景原型；(3) 查询预测，将得到的前景 - 背景原型与查询特征计算余弦相似度生成初步预测，利用解码器对多尺度前景 - 背景预测结果进行特征融合，最终将整合后的预测与真实标签计算损失。

3. 基于双视角注意力的小样本医学图像分割

3.1. 特征提取

对于每个训练任务，首先冻结特征提取器 $f_\theta(\cdot)$ 的参数(不参与梯度更新)，利用该提取器分别对支撑图像 I^s 和查询图像 I^q 进行特征提取 $f_\theta(I^s)$ 和 $f_\theta(I^q)$ ，得到支撑特征 F_s 和查询图像特征 $F_q \in \mathbb{R}^{H \times W \times D}$ ， $f_\theta(\cdot)$ 是使用共享权重的特征提取器，该提取器是以 ResNet-101 [9] 作为主干网络，并且该网络已经在 MS-COCO 数据集[10]上进行预训练从而获取更好的性能。H、W、D 分别代表特征图的高度、宽度和通道深度。

由于支撑集仅提供少量标注样本，而查询图像可能因视角差异或个体解剖结构差异等影响与支撑图像存在显著的数据分布偏移，为了缓解支撑特征和查询特征之间的域差异(Domain Gap)对分割性能的影响，本文引入一种基于支撑 - 查询交叉注意力(Support-Query Cross-Attention, SQCA)的特征对齐网络。该模块采用经典交叉注意力网络架构[11]，将支撑特征和查询特征输入到网络中进行特征交互学习，支撑特征能够捕捉并吸收查询特征中的关键信息，从而增强其对未见类别的泛化能力。增强支撑图像中与查询图像目标区域最相关的特征，抑制无关背景干扰。用公式表示为：

$$A = \text{Softmax} \left(\frac{(W_q F_s)(W_k F_q)^T}{\sqrt{d}} \right) \quad (1.1)$$

$$F_s^* = F_s + \alpha \cdot (A \cdot W_v F_q) \quad (1.2)$$

其中 W_q 、 W_k 和 W_v 是可学习的线性变换矩阵， d 为特征维度， $A \in \mathbb{R}^{HW \times HW}$ 是支撑特征与查询特征的相似性权重， α 是可学习的缩放系数。

3.2. 双视角协同增强网络

前景和背景原型采用相类似的操作，通过特征提取阶段得到的支撑特征 $F_s^* \in \mathbb{R}^{H \times W \times D}$ 并行输入空间注意力模块和通道注意力模块，分别获得两个维度的加权特征，通过动态权重融合机制自动计算最优权重，自适应地调节空间和通道注意力在前景和背景路径中的不同贡献度。

3.2.1. 空间注意力

对于医学图像分割，部分容积效应会导致器官边界模糊，空间注意力通过计算全局自相似性获取整体依赖关系，能够有效增强边界区分能力，以此解决部分容积效应导致的边界模糊问题。具体来说，将经过特征提取后的支撑图像特征 F_s^* 分别输入到空间注意力机制和通道注意力机制中，如图 2 所示。对于空间注意力部分，首先将支撑图像特征形变为大小是 $HW \times D$ 的二维 query 张量，再转置即得到大小为 $D \times HW$ 二维 key 张量，query 和 key 相乘得到自注意力相似性矩阵形状为 $HW \times HW$ ，即计算特征图中所有空间位置之间的相关性，用公式表示为：

$$S_{ij,kl} = Q_{ij} K_{kl}^T \in \mathbb{R}^{HW \times HW} \quad (1.3)$$

$$A_{ij,kl} = \frac{\exp(S_{ij,kl})}{\sum_{kl} \exp(S_{ij,kl})} \in \mathbb{R}^{HW \times HW} \quad (1.4)$$

对相似性矩阵进行归一化操作，得到归一化权重矩阵，即对当前位置来说，其他所有位置对该位置

的贡献程度，值从 0 到 1。

$$F_{\text{spatial}} = A_{ij,kl} \cdot V_{ij} + V_{ij} \in \mathbb{R}^{HW \times D} \quad (1.5)$$

将该相似性矩阵与 **query** 相乘再通过残差连接即可得到由空间注意力加权后的支撑特征 F_{spatial} 。局部卷积难以区分两组织之间的边界，全局注意力允许边界像素直接参考远处同类区域的语义。同时，模型会学习到左右肾的对称性或肝脾的空间相对位置关系。即使局部边界模糊，注意力机制也能通过全局结构约束(如“左肾通常位于脊柱左侧”)纠正错误分割。

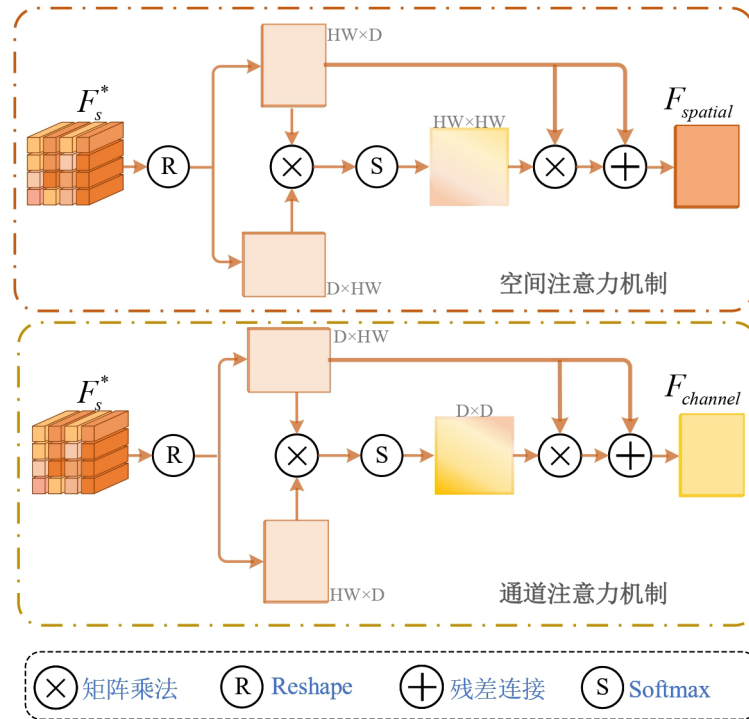


Figure 2. Dual-perspective collaborative attention module structure
图 2. 双视角协同注意力模型结构

3.2.2. 通道注意力

通道注意力用相似的方法，先将支撑特征形变为大小是 $D \times HW$ 的 **query** 张量和大小是 $HW \times D$ 的 **key** 张量，二者相乘计算相似性矩阵大小为 $D \times D$ ，即可得到当前通道相对于其他所有通道的相关性。归一化操作后得到通道间的贡献程度，最后相似性矩阵乘 **query** 后经过残差连接并转置为 $HW \times D$ 的通道注意力加权特征 F_{channel} 。对关键通道相应增加权重，如分割脾脏时，脾脏纹理通道权重接近 1，背景通道接近 0。这是因为在医学图像中不同器官共享相似的灰度和纹理等，模型容易将相似组织混淆，所以需要通道注意力增强前景背景的区分度。

3.2.3. 动态权重融合

最后将空间注意力和通道注意力的特征融合。先将空间注意力特征和通道注意力特征形变为同样大小 $\mathbb{R}^{D \times H \times W}$ ，两特征求和后做平均值池化得到融合表示 P ，再经过两层包含 **ReLU** 激活函数的 **MLP** 将 P 映射为二维权重 w ，最后归一化操作，以此保证 $w_{\text{spatial}} + w_{\text{channel}} = 1$ 。

$$P = \text{Avgpool}(F_{\text{spatial}} + F_{\text{channel}}) \in \mathbb{R}^{D \times 1 \times 1} \quad (1.6)$$

$$w = \text{Softmax}(W_2 \cdot \text{ReLU}(W_1 \cdot P)) \in \mathbb{R}^{2 \times 1 \times 1} \quad (1.7)$$

$$w_{\text{spatial}}, w_{\text{channel}} = w[0], w[1] \quad (1.8)$$

其中 $W_1 \in \mathbb{R}^{D/8 \times D}$ 和 $W_2 \in \mathbb{R}^{2 \times D/8}$ 是可学习参数。经过融合后的输出为:

$$\tilde{F}_i = w_{\text{spatial}} \cdot F_{\text{spatial}} + w_{\text{channel}} \cdot F_{\text{channel}} \quad (1.9)$$

$i \in \{f, b\}$, f 和 b 分别代表前景和背景, 对前景和背景采用相同的操作, 模型自动优化权重值。动态权重融合自适应权重分配, 使空间注意力和通道注意力在医学图像分割中实现协同优化, 尤其在器官交界处, 通道注意力抑制背景器官相关通道, 通过空间注意力精确定位目标器官轮廓。

3.3. 查询预测

在原型生成阶段, 我们通过掩码加权的方式分别计算前景和背景原型。对于前景原型, 直接将支撑特征和二进制支撑掩码进行逐元素相乘; 对于背景原型, 将支撑特征和支撑掩码相乘得到前景原型, 用 1 减去支撑掩码计算出背景掩码, 背景掩码和支撑特征相乘得到背景原型。具体用公式表示为:

$$P_{fg} = \frac{\sum_{i,j} M_s(i,j) \cdot \tilde{F}_f(i,j,:)}{\max(\sum_{i,j} M_s(i,j), \varepsilon)}, P_{bg} = \frac{\sum_{i,j} (1 - M_s(i,j)) \cdot \tilde{F}_b(i,j,:)}{\max(\sum_{i,j} (1 - M_s(i,j)), \varepsilon)} \quad (1.10)$$

我们遵循传统的小样本医学图像分割的方法, 首先分别计算前景、背景原型与查询特征的余弦相似性, 再通过简单的解码器[12]进行融合得到前景预测和背景预测, 该解码器是一个轻量级的卷积网络。最后, 我们通过拼接(concat)操作整合最终预测的概率图, 并基于此计算主交叉熵损失。

3.4. 损失函数

在小样本医学图像分割任务中, 损失函数的目标是逐步缩小预测和真实掩码标签之间的距离, 在本文中, 我们采用交叉熵损失函数来衡量查询预测分割掩码和其对应的真实掩码之间的损失, 公式表示为:

$$L_{\text{seg}} = -\frac{1}{HW} \sum_{h,w} m_q^{fg}(h,w) \odot \log(\tilde{m}_q^{fg}(h,w)) + m_q^{bg}(h,w) \odot \log(\tilde{m}_q^{bg}(h,w)) \quad (1.11)$$

其中 H 、 W 代表空间位置点的总数, m_q^{fg} 代表查询图像的真实掩码标签, m_q^{bg} 是通过 $1 - m_q^{fg}$ 计算得来, $\tilde{m}_q^{fg}(h,w)$ 和 $\tilde{m}_q^{bg}(h,w)$ 分别代表前景预测结果和背景预测。此外, 我们沿用了另一种反向学习的方法[2], 用查询图像作为支持集来预测支持图像的标签, 具体来说将查询图像和预测得到的查询掩码计算原型, 再与支撑图像算余弦相似性, 得到的支撑预测掩码和支撑真实掩码标签计算损失, 用公式表示为:

$$L_{\text{reg}} = -\frac{1}{HW} \sum_{h,w} m_s^{fg}(h,w) \odot \log(\tilde{m}_s^{fg}(h,w)) + m_s^{bg}(h,w) \odot \log(\tilde{m}_s^{bg}(h,w)) \quad (1.12)$$

最后将所有损失相加得到总损失:

$$L_{\text{total}} = L_{\text{seg}} + L_{\text{reg}} \quad (1.13)$$

4. 实验结果与分析

4.1. 实验准备

4.1.1. 数据集

本文, 我们在四个公开数据集中评估我们方法的有效性, 包括:

(1) 腹部 MRI 分割(Abd-MRI) [13], 是从 ISBI 2019 联合健康腹部器官分割挑战赛中获得的腹部 MRI

数据集。该数据集包括 20 例完整的 3D T2-SPiR MRI 扫描序列, 每个扫描平均约 36 个连续切片(slice), 覆盖完整的腹部解剖结构, 分割部位同腹部 MRI 分割。

(2) 腹部 CT 分割(Abd-CT) [14]是来自 MICCAI 2015 Multi-Atlas 腹部器官分割挑战赛, 包含 30 例临床采集的 3D 腹部 CT 扫描数据。本文需分割的器官包括左肾、右肾、脾脏和肝脏。

(3) 心脏 MRI 分割(Card-MRI) [15]源自 MICCAI 2019 多序列心脏 MRI 分割挑战赛。拥有 35 例完整的 3D 心脏 MRI 扫描, 每个扫描平均分为 13 个连续切片(slice)。并且提供了三个心脏关键结构的专家标注: 左心室血池(LV-BP)、左心室心肌(LV-MYO)和右心室心肌(RV)。

4.1.2. 实验设置

为了确保模型评估的公平性和可比较性, 本文严格遵循大多数小样本医学图像分割的设定[2] [9], 采用两种设置分为 Setting-1 和 Setting-2。Setting-1 是指, 测试阶段需分割的目标器官允许出现在训练图像的背景区域中, 但在训练过程中不对该器官进行任何标注或分割监督。该设置模拟了临床实践中常见的情况, 即某些器官可能在扫描图像中同时存在但未被标注。Setting-2 是更严格的设定, 要求测试待分割的目标器官不会出现在任何训练集上。对于腹部数据集, 将左肾和右肾分为下腹部组, 肝脏和脾脏分为上腹部组, 假设训练分割左肾和右肾, 测试分割肝脏和脾脏, 那么所有用作训练的图片中不能出现肝脏和脾脏两器官。但是在心脏数据集中, 单个切片通常同时包含所有心腔结构, 难以从单个切片中排除测试类别, 因此 Setting-2 不能应用于该数据集。

在本研究中, 对三维医学图像进行了标准化处理, 以适应深度学习模型的输入要求。具体来说, 所有 3D 扫描图像首先被转换成 2D 切片, 并调整至统一的分辨率, 即 256×256 像素, 以便于网络训练和评估过程。每个二维切片在通道维度上进行了三次重复。我们使用传统的 Dice 评分作为衡量指标[2], 公式如下:

$$DSC(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|} \times 100\% \quad (1.14)$$

其中 X 和 Y 是预测掩码和真实标签, 最终结果在 0 到 100 之间, 越接近 100 表示模型性能越好。

该模型被训练了 100k 次迭代, 批量大小设置为 1。在训练过程中, 初始学习率设置为 0.001。每 1000 次迭代的步进衰减率为 0.98。

4.2. 结果分析

4.2.1. 与现有方法比较

为了证明我们方法的有效性, 本文通过与以下几个现有方法进行比较: PANet [7]、ADNet [9]、RPTNet [6]、SSL-ALPNet [2]、Q-Net [4]、CAT-Net [16]和 GMRD [12]。为确保实验对比的公平性和可重复性, 我们从官方渠道获取所有对比方法的原始代码, 并与 DPCENet 一起在同一环境下运行所有实验。

表 1 和表 2 分别列举了 ABD-MRI 和 ABD-CT 两数据集在 setting-1 设定下和 setting-2 设定下的 Dice 评分结果。实验结果表明, DPCENet 在两种设定下都优于之前的结果, 在 ABD-MRI 数据集的 setting-1 设定下, 与之前最好的方法 Q-Net [4]相比提升了 0.83, 其中左肾这一部位比最高的方法 SSL-ALPNet [2]提高了 4.58; 在 ABD-MRI 数据集的 setting-2 设定下, 与之前最好的方法 GMRD [12]提高了 1.23。ABD-CT 数据集在 setting-1 和 setting-2 设定下比之前最好方法 SSL-ALPNet [2]分别提高了 1.09 和 2.15。表 3 显示心脏数据集的分割效果, 我们的方法相比之前最优的方法 SSL-ALPNet [2]提高了 0.91。综合三个数据集的实验结果来看, 我们的方法在肝脾、肾脏和心脏分割任务中均表现出稳定且显著的性能提升, 验证了本方法在多器官、多模态医学图像分割中的普适性和鲁棒性。

为了便于客观比较，图 3 和图 4 展示了本文方法与现有方法在三个医学数据集上的可视化结果对比。我们的模型三个数据集上的性能都得到了显著改善，表现出分割清晰边界的卓越能力。

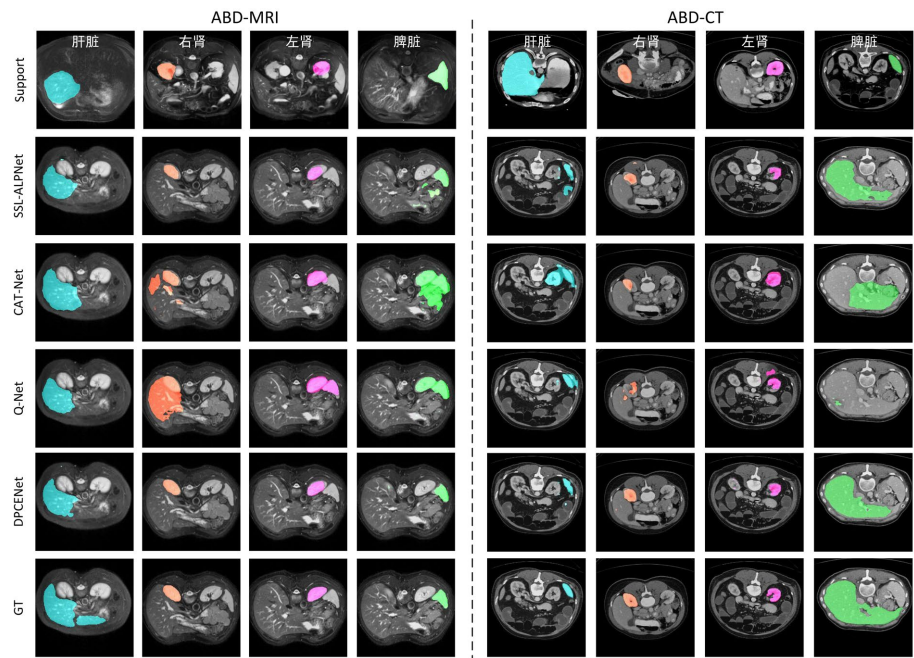


Figure 3. The visualization results in the ABD-MRI dataset (the left side) and the ABD-CT dataset (the right side) under setting-2
图 3. 在 setting-2 下，ABD-MRI 数据集(左侧)和 ABD-CT 数据集(右侧)的可视化结果

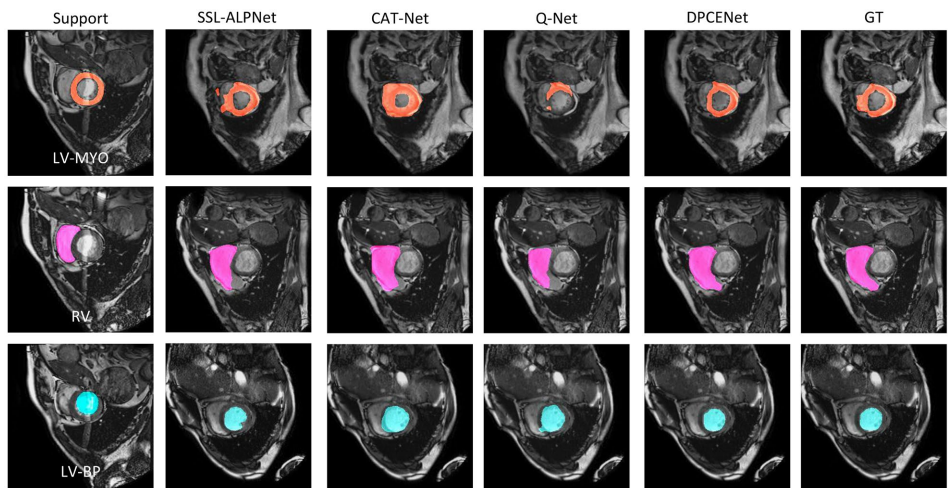


Figure 4. The visualization results in the Card-MRI dataset under setting-1
图 4. 在 setting-1 设定下 Card-MRI 数据集的可视化结果

Table 1. The results of different models on the ABD-MRI and ABD-CT datasets under setting-1
表 1. 在 setting-1 设定下不同模型在 ABD-MRI 和 ABD-CT 数据集上的结果

方法	ABD-MRI					ABD-CT				
	肝脏	右肾	左肾	脾脏	均值	肝脏	右肾	左肾	脾脏	均值
PANet	27.73	30.41	34.96	47.37	35.11	55.72	50.42	56.52	60.86	57.88

续表

ADNet	49.74	84.21	62.97	76.79	68.42	41.17	32.86	31.27	77.26	45.67
RPTNet	73.72	86.83	65.58	60.32	71.61	68.48	60.52	84.37	64.51	69.47
SSL-ALPNet	67.78	84.88	79.61	74.32	76.65	70.11	72.62	76.35	67.29	71.59
Q-Net	76.24	86.40	72.13	72.47	76.81	56.82	55.63	69.39	68.65	62.63
CAT-Net	70.54	83.00	75.30	70.59	74.86	66.98	47.83	69.09	66.24	62.54
GMRD	65.44	89.95	75.97	73.65	76.25	56.48	62.27	79.92	63.06	65.43
DPCENet	73.72	81.50	84.19	71.13	77.64	67.35	76.05	78.79	68.53	72.68

Table 2. The results of different models on the ABD-MRI and ABD-CT datasets under setting-2**表 2.** 在 setting-2 设定下不同模型在 ABD-MRI 和 ABD-CT 数据集上的结果

方法	ABD-MRI					ABD-CT				
	肝脏	右肾	左肾	脾脏	均值	肝脏	右肾	左肾	脾脏	均值
PANet	61.25	66.94	63.17	69.37	65.68	43.73	34.69	37.58	61.71	44.42
ADNet	59.44	59.64	56.68	77.03	63.19	50.97	48.41	40.52	70.63	52.63
RPTNet	75.15	60.11	66.27	67.45	67.25	60.12	53.84	82.28	54.24	62.62
SSL-ALPNet	55.35	76.38	73.24	68.38	68.34	61.97	59.05	64.18	69.14	63.59
Q-Net	74.71	74.71	64.15	71.52	71.27	37.87	41.75	66.21	64.44	52.57
CAT-Net	60.23	60.23	78.57	77.45	69.12	46.73	46.87	65.01	52.53	52.79
GMRD	73.80	70.25	69.37	74.85	72.07	64.16	55.35	72.46	60.88	63.21
DPCENet	67.82	74.97	78.04	72.39	73.30	69.36	64.04	64.94	64.63	65.74

Table 3. The results of different models on the Card-MRI datasets under setting-1**表 3.** 在 setting-1 设定下不同模型在 Card-MRI 数据集上的结果

方法	Card-MRI			
	RV	LV-MYO	LV-BP	均值
PANet	57.13	44.76	72.77	58.20
ADNet0	65.37	82.29	58.86	68.84
RPTNet	76.63	80.15	58.81	71.86
SSL-ALPNet	77.59	63.29	85.36	75.41
Q-Net	67.99	52.09	86.21	68.76
CAT-Net	69.37	48.81	81.33	66.51
GMRD	80.82	73.65	60.83	71.77
DPCENet	76.52	63.69	88.75	76.32

4.2.2. 消融实验

本文对 setting-2 设定下的 ABD-CT 数据集上采用控制变量方法对 SQCA 和 DPCE 两部分进行了消融实验,如表 4 所示。当 SQCA 被移除时,模型对未知类别的泛化能力降低,导致性能下降;当 DPCE 被移除时,性能有更明显下降,说明模型对目标器官的定位能力和细节上都明显降低。因此本文提出两个模块对整体性能上有不同程度的提升。

Table 4. Ablation study experimental results
表 4. 消融实验结果

SQCA	DPCE	肝脏	右肾	左肾	脾脏	均值
√	√	69.36	64.04	64.94	64.63	65.74
√	×	63.56	59.48	63.13	66.61	63.20
×	√	58.21	53.94	73.32	70.37	63.96
×	×	54.02	61.56	72.02	64.85	63.11

5. 结束语

本文针对医学图像特有的部分容积效应和灰度不均匀性等挑战性问题，提出了双视角协同增强网络 (DPCENet)，通过深入分析医学图像分割中的难点，我们设计了空间 - 通道双视角协同注意力网络：利用空间注意力建立全局长程依赖，精准捕捉器官的空间分布；通道注意力强化目标器官相关特征并抑制噪声干扰。同时网络结构上采用双路径策略并行处理前景和背景特征流，并引入动态权重融合机制自适应地为前景和背景分别分配不同注意力权重，有效解决了医学图像中组织特征高度耦合的难题。在三个具有挑战性的公开数据集上的评估表明，我们提出的方法能够准确分割多种复杂的器官，该技术有望辅助医生提升对罕见病变的诊断效率和准确性，为临床决策提供可靠的智能支持。

参考文献

- [1] Asgari Taghanaki, S., Abhishek, K., Cohen, J.P., Cohen-Adad, J. and Hamarneh, G. (2020) Deep Semantic Segmentation of Natural and Medical Images: A Review. *Artificial Intelligence Review*, **54**, 137-178. <https://doi.org/10.1007/s10462-020-09854-1>
- [2] Ouyang, C., Biffi, C., Chen, C., Kart, T., Qiu, H. and Rueckert, D. (2020) Self-Supervision with Superpixels: Training Few-Shot Medical Image Segmentation without Annotation. *Computer Vision-ECCV 2020: 16th European Conference*, Glasgow, 23-28 August 2020, 762-780. https://doi.org/10.1007/978-3-030-58526-6_45
- [3] Guha Roy, A., Siddiqui, S., Pölsterl, S., Navab, N. and Wachinger, C. (2020) “Squeeze & Excite” Guided Few-Shot Segmentation of Volumetric Images. *Medical Image Analysis*, **59**, Article ID: 101587. <https://doi.org/10.1016/j.media.2019.101587>
- [4] Shen, Q., Li, Y., Jin, J. and Liu, B. (2024) Q-net: Query-Informed Few-Shot Medical Image Segmentation. In: *Proceedings of SAI Intelligent Systems Conference*, Springer, 610-628. https://doi.org/10.1007/978-3-031-47724-9_40
- [5] Ding, H., Sun, C., Tang, H., Cai, D. and Yan, Y. (2023) Few-Shot Medical Image Segmentation with Cycle-Resemblance Attention. 2023 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 2-7 January 2023, 2488-2497. <https://doi.org/10.1109/wacv56688.2023.00252>
- [6] Zhu, Y., Wang, S., Xin, T. and Zhang, H. (2023) Few-Shot Medical Image Segmentation via a Region-Enhanced Prototypical Transformer. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 271-280. https://doi.org/10.1007/978-3-031-43901-8_26
- [7] Wang, K., Liew, J.H., Zou, Y., Zhou, D. and Feng, J. (2019) Panet: Few-Shot Image Semantic Segmentation with Prototype Alignment. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 9197-9206. <https://doi.org/10.1109/iccv.2019.00929>
- [8] 周志敏. 基于改进 U-Net 的医学图像分割算法研究[D]: [硕士学位论文]. 镇江: 江苏科技大学, 2024.
- [9] Hansen, S., Gautam, S., Jenssen, R. and Kampffmeyer, M. (2022) Anomaly Detection-Inspired Few-Shot Medical Image Segmentation through Self-Supervision with Supervoxels. *Medical Image Analysis*, **78**, Article ID: 102385. <https://doi.org/10.1016/j.media.2022.102385>
- [10] Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., *et al.* (2014) Microsoft COCO: Common Objects in Context. *Computer Vision-ECCV 2014: 13th European Conference*, Zurich, 6-12 September 2014, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48
- [11] Wang, X., Girshick, R., Gupta, A. and He, K. (2018) Non-Local Neural Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 7794-7803.

-
- <https://doi.org/10.1109/cvpr.2018.00813>
- [12] Cheng, Z., Wang, S., Xin, T., Zhou, T., Zhang, H. and Shao, L. (2024) Few-Shot Medical Image Segmentation via Generating Multiple Representative Descriptors. *IEEE Transactions on Medical Imaging*, **43**, 2202-2214. <https://doi.org/10.1109/tmi.2024.3358295>
- [13] Kavur, A.E., Gezer, N.S., Barış, M., Aslan, S., Conze, P., Groza, V., *et al.* (2021) CHAOS Challenge-Combined (CT-MR) Healthy Abdominal Organ Segmentation. *Medical Image Analysis*, **69**, Article ID: 101950. <https://doi.org/10.1016/j.media.2020.101950>
- [14] Landman, B., Xu, Z., Igelsias, J., Styner, M., Langerak, T. and Klein, A. (2015) Multi-Atlas Labeling beyond the Cranial Vault-Workshop and Challenge. *Proceedings MICCAI Multi-Atlas Labeling beyond Cranial Vault-Workshop Challenge*, Vol. 5, 12.
- [15] Zhuang, X. (2019) Multivariate Mixture Model for Myocardial Segmentation Combining Multi-Source Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**, 2933-2946. <https://doi.org/10.1109/tpami.2018.2869576>
- [16] Lin, Y., Chen, Y., Cheng, K. and Chen, H. (2023) Few Shot Medical Image Segmentation with Cross Attention Transformer. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 233-243. https://doi.org/10.1007/978-3-031-43895-0_22