

ARNet:基于Transformer细化的自适应区域原型医学小样本语义分割网络

史晓龙, 瞿绍军, 石 在

(湖南师范大学 信息科学与工程学院, 长沙 410081)

E-mail: qshj@hunnu.edu.cn

摘要: 在医学图像语义分割领域, 小样本学习因其高效的数据利用率和优秀的泛化能力, 受到广泛关注. 为应对医学图像中前景与背景的极度不平衡问题, 目前基于区域原型的方法已成为主流. 然而, 现有的区域原型方法在应对数据样本中普遍存在的内类差异和类间差异方面, 分割效果有限. 为此, 本文提出一种基于Transformer细化的自适应区域原型网络 (ARNet), 该方法能够根据前景的大小、形状及特征分布自适应生成区域原型, 以更精确地学习多个子区域表示来提升模型的分割性能. 为进一步弥补区域原型方法所存在的全局类别信息丢失问题, 本文提出原型细化模块. 该模块由原型增强和原型对齐两部分组成, 原型增强部分基于Transformer结构和掩码注意力机制, 在增强局部信息的同时补充全局信息; 原型对齐部分则通过区域原型与查询特征的交互, 实现对查询目标的精确对齐. 实验表明, 所提出的方法在CHAOS和SABS两个公开数据集上均能取得优异的性能.

关键词: 深度学习; 小样本学习; 医学图像分割; 卷积神经网络; Transformer

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)04-1016-09

ARNet: Adaptive Region Prototypes with Transformer-based Refinement for Few-shot Medical Image Segmentation

SHI Xiaolong, QU Shaojun, SHI Zai

(College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China)

Abstract: In the field of semantic segmentation of medical images, Few-Shot learning has received widespread attention due to its efficient data utilization and excellent generalization ability. Region prototype-based methods have now become mainstream to cope with the problem of extreme imbalance between foreground and background in medical images. However, existing region prototype methods have limited segmentation effectiveness in dealing with the intra-class and inter-class variations prevalent in the data samples. To address this, we propose an Adaptive Region Prototype Network (ARNet) based on Transformer refinement, which can adaptively generate region prototypes based on the size, shape, and feature distribution of the foreground to learn multiple subregion representations more accurately to improve the segmentation performance of the model. We introduce a prototype refinement module to mitigate the loss of global category information inherent in the region prototype approach. The module consists of two components: the prototype enhancement component, which leverages the Transformer architecture and mask attention mechanism to enhance local information while supplementing global information, and the prototype alignment component, which achieves precise alignment of the query target through the interaction between region prototypes and query features. The proposed method achieves excellent performance on CHAOS and SABS two public datasets.

Keywords: deep learning; few-shot learning; medical image segmentation; convolutional neural network; Transformer

0 引言

医学图像分割作为医学影像分析中的核心任务, 广泛应用于疾病诊断和治疗规划中. 随着深度学习技术^[1]的飞速发展, 卷积神经网络 (Convolutional Neural Network, CNN) 已成为医学图像分割领域的重要工具. 尽管这些模型在性能上取得了显著进展, 但它们仍然依赖于大量标注数据进行训练, 这在医学图像领域尤为困难. 医学图像的标注不仅需要大量的

时间和人力成本, 还受到数据隐私和伦理问题的制约, 导致数据稀缺问题更加突出. 因此如何在有限的标注数据下有效地训练深度学习模型, 并确保其在实际临床应用中的可靠性和泛化能力已成为了当前医学图像分割技术面临的重大挑战.

为应对数据稀缺问题, 小样本学习 (Few-Shot Learning, FSL)^[2]作为一种高效的数据利用策略, 近年来受到广泛关注. FSL 首先被引入到目标检测领域^[3], 随后引入到自然图像分割领域, 提出了小样本语义分割任务 (Few-Shot Semantic

Segmentation, FSS)^[4,5],旨在学习一个模型能够仅通过少量标注数据完成测试阶段新类别的分割.在这两个领域中FSL技术通过最大限度地提高数据的利用效率减少了对大规模标注数据集的依赖,使模型展现出了优异的泛化能力和灵活的适应性.让研究者认识到小样本分割技术在医学图像处理中的巨大应用潜力,小样本医学分割任务(Few-Shot Medical Image Segmentation, FSMIS)应运而生.

目前,基于原型网络的方法^[6-8]已逐渐成为FSMIS研究中的主流方向.然而与FSS领域类似,这类模型同样面临类内差异对模型性能的制约.尤其在医学图像中,通常存在极端的前景-背景不平衡问题,进一步加剧了挑战.此外,由于医学图像的特殊性质,类内差异往往源自特定器官固有的多样性,这些多样性表现在大小、形状、轮廓等特征上,并且在不同患者或采集协议下可能发生变化.特别是在某些病变类别中常常会出现一定数量的差异区域.这些差异与自然图像中的物体颜色或形态差异有所不同,自然图像中的差异通常仍能保持类别信息的连续性,而医学图像中的差异则表现为突发性的变化.这使得传统的掩码平均池化(Mask Average Pooling, MAP)方法生成的原型在FSMIS任务中往往表现出低效和不精确的问题.

为了应对这一问题,相关研究^[9,10]提出通过提取每个类别的局部表示集合(即区域原型)以保留更多的类内局部细节信息,从而有效应对可能出现的类内差异.然而,目前大多数区域原型方法要么通过均匀划分前景区域生成,要么通过固定数量的随机初始化种子点聚合生成.这些方法未能充分考虑前景区域的大小、形状以及特征分布,存在两个主要问题:首先均匀划分前景区域或随机初始化种子点都可能导致生成的原型过于接近,甚至可能集中在边界区域,而降低区域原型代表性;其次医学数据集中不同类别之间通常存在显著的尺寸和形状差异,采用固定数量的子区域生成多个原型会限制其表达能力.

针对上述问题,本文提出一种自适应区域原型生成方法,该方法旨在根据前景像素的大小和形状自适应地将医学图像前景划分为多个子区域,从而减轻不同子区域间差异对模型性能的影响,更有效地学习多个精确的子区域表示.并基于此提出一种前置网络(Preliminary Network, Pre-Net),该网络结合SSPNet^[11]中的自支持和自匹配策略,将高置信查询信息融入区域原型中,有效缓解类内差异对模型性能的负面影响.此外,考虑到区域原型虽然能够捕获更多的类内局部信息,但也会显著丢失类别的全局信息,本文提出一个原型细化模块.该模块基于Transformer结构和掩码注意力机制,由原型增强和原型对齐两部分组成.原型增强部分在增强区域原型细节信息的同时,有效补充全局信息;原型对齐部分则进一步对齐支持原型和查询目标,提升区域原型的表达能力.结合前置网络与原型细化模块,本文提出基于Transformer细化的自适应区域原型小样本医学图像分割网络(ARPNNet),在多个公开医学图像分割数据集上均取得优异的效果.

综上所述,本文的主要贡献包括:

1) 提出一种自适应区域原型生成方法,这是一种灵活的种子点初始化方法,可适应不同的目标大小,形状,并有效应对医学图像所存在的类内突发性变化.

2) 提出一种简单高效的前置网络,通过引入高置信查询信息有效缓解了类内差异对模型性能的影响.该网络不仅能够独立完成分割任务,其分割效果也非常稳定.

3) 提出一种原型细化模块,利用注意力机制进一步增强区域原型,并实现与查询目标的精确对齐,从而显著提升分割性能.

1 相关工作

1.1 医学图像分割

随着深度学习技术的快速发展,基于卷积神经网络(CNN)的方法已成为医学图像分割领域的主流技术,最具代表性的是U-Net^[12]架构.U-Net采用对称的编码器-解码器结构,其对称设计使得U-Net在细粒度分割任务中表现出色.其中解码器的逐步恢复过程有助于图像细节的重建,编码器则确保了语义信息的完整性.此外,U-Net通过引入跳跃连接,将编码器中的特征图直接传递至解码器的相应层中,从而有效弥补编码过程中丢失的空间信息,显著提高了细节和边缘的恢复能力,在保留高分辨率特征的同时避免了信息的过度抽象.自U-Net提出以来众多改进版本相继问世,进一步提升了其性能,如U-Net 3D^[13]、V-Net^[14]、Y-Net^[15]、Attention U-Net^[16]和nnU-Net^[17]等.近年来Transformer架构的成功应用已成为计算机视觉领域的重要发展趋势.许多研究者将注意力机制与U-Net相结合取得了显著的进展,如DS-TransU-Net^[18]、Swin-Unet^[19]和TransUNet^[20]等.然而这些基于CNN的模型通常依赖于大量标注数据进行训练,并且在面对未见类别时缺乏有效的泛化能力.而在医学图像处理中数据稀缺问题尤为突出,且潜在的分割目标(如不同的解剖结构、各类病变等)种类繁多、数量庞大,这使得该类模型在实际应用中面临诸多挑战.

1.2 小样本语义分割

为了解决传统语义分割模型对大规模标注数据依赖的问题,OSLSM^[21]引入小样本方法到语义分割任务.该方法采用双分支结构,通过利用支持图像中的类别信息对查询图像进行分割,开创了基于元学习的小样本语义分割新范式,为FSS任务的进一步研究奠定了基础.而在ProtoNet^[6]的推动下,PL^[7]首次将原型学习方法应用于FSS并取得了显著的效果.此后基于度量学习的元学习方法逐渐成为解决FSS问题的主要研究方向.

目前的FSS模型主要可分为基于原型和基于匹配两大类.其中原型方法因其较低的计算复杂度和较强的泛化能力,成为该领域最早引入的技术.早期的研究主要集中于提升单一原型的表达能力,采用更为精细的特征提取与对齐策略以期通过原型实现对每个类别的全面且精确的表示,如PANet^[22]、CANet^[23]和PFENet^[24]等方法.随后,由于单一原型在许多情况下难以充分覆盖目标区域并易引发歧义,PPNet^[25]、PMM^[26]和ASGNet^[27]等方法提出了局部原型的概念,旨在通过生成多个局部原型来更好地保留细粒度特征.如今随着原型表达能力的提升,研究逐渐聚焦于类间关系和类内变化对原型网络的影响,如IPRNet^[28]和SSPNet^[11]等方法.然而在原型方法的发展过程中,许多研究者发现其在信息利用方面存在一定局限,尤其在处理类内差异时存在严重的信息丢

失问题.为了突破这一限制,一些研究者基于 Transformer 结构及其注意力机制提出了基于匹配的方法.这类方法,如 CyCTR^[29]、VAT^[30]和 HDMNet^[31],主要通过捕捉像素级别的相关性进行分割,有效克服了原型方法在信息利用方面的不足,从而实现了更优的分割效果.然而这些方法在分割过程中往往面临较大的计算开销问题.

1.3 小样本医学图像分割

由于医学图像在数据采集过程中受到法律、伦理和隐私等多重因素的制约,使得数据稀缺问题尤为突出.近年来 FSS 领域在解决传统分割模型对大规模标注数据集依赖的问题上取得了长足进步,研究者们逐渐认识到 FSL 技术在医学图像处理中的巨大应用潜力,进而催生了小样本医学图像分割(FSMIS)任务.

SENet^[32]首次将 FSL 应用于医学图像处理并提出了双臂结构.其“调节器臂”通过处理带标注的支持图像生成任务表示,并传递给分割器臂进行查询图像分割.为增强两者协同作用,SENet 引入了“挤压与激励块”,通过促进大量交互来实现更优异的分割效果.GCN-DE^[33]则提出了一种高效的全局相关性模块,旨在捕捉支持图像与查询图像之间的相关关系以提升分割性能,并通过引入区分嵌入损失,优化同类特征聚类并增大不同器官间在特征空间中的距离.MPrNet^[34]提出了一种基于交互学习的测试时间优化算法(Interactive Learning-based Test Time Optimization Algorithm, IL-TTOA),旨在以交互且可控的方式实时提升模型性能.上述方法侧重于构建创新的连接机制,考虑到器官和病变通常位于腹部和肿瘤特定位置,研究者通过新的机制促进支持与查询图像间的有效交互,常通过多种注意力机制融合于交互模块.

与此同时,基于 FSS 原型网络的技术原理衍生出一系列基于度量学习的 FSMIS 方法.其中 SSL-ALPNet^[35]考虑到医学图像处理领域数据匮乏的训练环境,提出了一种新的医学图像自监督小样本分割框架.在该框架中,训练过程无需依赖标注数据,而是通过生成基于超像素的伪标签来提供监督,开创了 FSMIS 研究的新方向.鉴于原型是基于有限数量的切片,ADNet^[36]指出现有的原型生成方法难以充分捕捉背景特

征.为此 ADNet 避免了对大背景的建模,提出了一种新的异常检测启发式方法来完成小样本医学图像分割.后来 Q-Net^[37]在 ADNet 的基础上设计了双路径结构,并引入了两个基于查询信息的计算模块:查询信息阈值自适应模块和查询原型细化模块.该设计通过动态调整阈值以更好地提取查询信息,从而更新语义类原型,并结合不同深度的特征分割结果,进一步优化模型性能.而为应对医学图像中的类内变化,PAMI^[38]提出了一种从支持医学图像中提取多个代表性子区域的方法,以保留更多的局部细节信息,并引入 Transformer 结构提出了原型去偏模块,以去除区域原型中可能存在的噪声,将模型的性能提升到了一个全新的水平.DSSNet^[39]则采用前景与背景双路径并行优化原型策略,并提出自匹配动态优化策略,通过高置信度原型进行校准减小了支持样本与查询样本间的类内差异,也取得了非常不错的模型性能.

2 方法

2.1 问题定义

在 FSS 任务中尽管训练集 D_{train} 和测试集 D_{test} 都来自同一数据源,但它们的类别并不重叠.FSS 模型的目标是从 D_{train} 中学习元知识,并将其泛化到包含稀疏标注图像中的未见类别.在此设置中,将 D_{train} 中的类别定义为已见类别 C_{seen} ,而将 D_{test} 中的类别定义为未见类别 C_{unseen} ,且约束条件为 $C_{seen} \cap C_{unseen} = \emptyset$.在训练过程中,通过从 D_{train} 中多次采样来构建 FSS 任务.每项任务都从 C_{seen} 中随机抽取 N 个类别,并采样 $K+1$ 个切片-掩码对,其中 K 个切片-掩码对构成支持集 $S = \{(I_s^i, M_s^i)\}$,其余构成查询集 $Q = \{(I_q^j, M_q^j)\}$.这里 I 和 M 代表三维医学图像数据的二维切片和其相应的二进制掩码.在每个任务中模型利用支持集中 K 个切片-掩码对所蕴含的类别信息完成对查询集的预测.在测试过程中,任务构建方式与训练阶段相同,但通过从 D_{test} 中采样来评估模型的性能.

2.2 网络结构

本文提出的 ARPNet 模型框架如图 1 所示.该框架主要

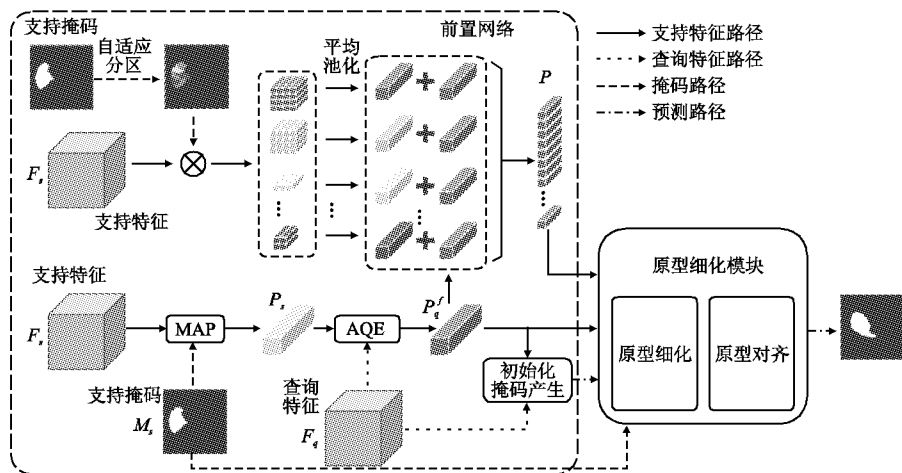


图 1 ARPNet 的整体结构,包括前置网络和原型细化模块

Fig. 1 Overall structure of our ARPNet, including the Pre-Net and prototype refinement transformer

由前置网络和原型细化模块两部分组成.前置网络作为模型

的前半部分负责为后续的原型细化模块提供必要的输入,并

且能够独立完成分割任务。原型细化模块在前置网络生成的输入基础上进一步增强原型信息,并对查询目标进行对齐,从而提升预测结果的准确性。具体来说,对于给定的支持图像 I_s 和查询图像 I_q ,使用两个共享权重的卷积神经网络 $f(x)$ 分别提取支持特征 $F_s = f(I_s) \in R^{h \times w \times c}$ 和查询特征 $F_q = f(I_q) \in R^{h \times w \times c}$,其中 s 和 q 分别表示支持和查询, $h \times w$ 表示空间维度, c 表示通道数。前置网络由两个并行计算分支组成:粗略原型计算分支和区域原型计算分支。在粗略原型计算分支中,首先通过 MAP 操作将 F_s 转化为粗略原型 p_s ,随后通过融合高置信度查询特征生成改进后的原型 p'_s ,并利用该原型生成初始预测掩码 M_{ini} 。在区域原型计算分支中采用本文提出的自适应区域原型方法将 F_s 转化为区域原型集 P ,并在此过程中将粗略原型计算分支生成的 p'_s 融入其中,以进一步补充查询信息和全局信息。在后续过程中,在 M_{ini} 的监督下区域原型集 P 经过原型细化模块处理后生成增强区域原型集 \tilde{P} 。最终通过 \tilde{P} 与查询特征 F_q 之间的余弦相似度度量得到最终的预测结果 y_{final} 。

2.3 前置网络(Pre-Net)

2.3.1 特征提取

本文借鉴了 Q-Net^[37] 的设计,采用在 MSCOCO 数据集上预训练的 ResNet-101 作为特征提取的骨干网络。该网络结构如图 2 所示,包含步长为 2 的卷积层、最大池化层、4 个残差块以及多个由 3 层瓶颈子块组成的模块。与 Q-Net 不同的是,本文未采用其双路径特征提取结构,而是仅利用 Block4 生

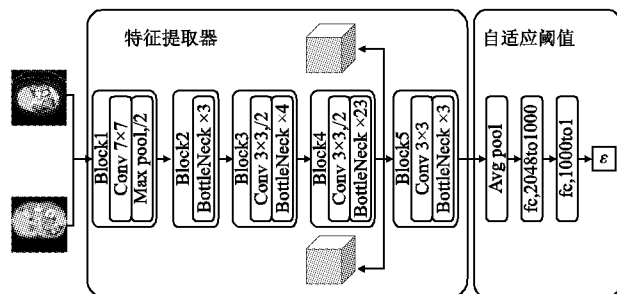


图 2 主干网络的结构以及自适应阈值提取过程,

其中 ε 表示自适应阈值

Fig. 2 Structure of the backbone network and the adaptive threshold extraction process, where ε represents the adaptive threshold

成的特征进行图像分割。因此,在 Block5 阶段,省略了下采样操作,使得 Block5 后的特征分辨率与 Block4 保持一致。此设计有助于后续基于查询的阈值自适应结构将 Block5 提取的特征映射为一维阈值,从而更好地支持 Block4 后续图像分割任务。

2.3.2 区域原型学习

如图 1 所示,前置网络由两个并行计算的分支构成:粗略原型计算分支和区域原型计算分支。粗略原型计算分支采用传统的原型生成方法 MAP,生成粗略的支持原型。得益于 MAP 方法通过对所有前景特征进行平均加权来生成原型,能够有效整合特征信息从而表现出较强的类别特异性和鲁棒性。尽管该方法可能丢失一定的细节信息,但其能够高效提取

查询特征中的共有类别特征,从而为区域原型计算分支补充查询信息和全局信息。

首先通过公式(1)对支持特征 F_s 应用 MAP 操作生成粗略的支持原型 p_s :

$$p_s = \frac{\sum_{x,y} F_s^{(x,y)} \zeta[M_s^{(x,y)} = c]}{\sum_{x,y} \zeta[M_s^{(x,y)} = c]} \quad (1)$$

其中 c 表示类别, M_s 表示支持掩码,而 (x,y) 表示特征的空间位置索引。指示函数 ζ 用于确定参数是否为真,若为真则返回值为 1 否则返回值为 0。

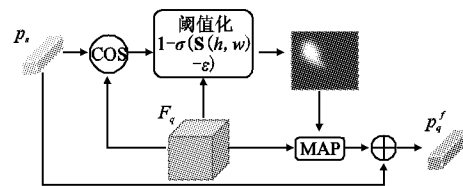


图 3 AQE 模块

Fig. 3 AQE module

为了更有效地提取高置信度的查询信息,本文结合 Q-Net^[37] 中的自适应阈值策略与 SSPNet^[11] 中的自支持前景原型生成方法 (Self-Support Foreground Prototype, SSFP),提出了自适应查询信息提取模块 (Adaptive Query Extraction, AQE)。其具体步骤如下:首先计算 p_s 与查询特征 F_q 之间的像素级余弦相似度生成相似度图谱 M_{sim} ;然后依据公式(2)生成粗略的查询掩码 M_q ,并通过 MAP 方法将 F_q 转化为查询原型 p_q ;最后应用公式(3)进行信息融合得到最终原型 p'_q 。该过程的可视化效果如图 3 所示。

$$M_q = 1 - \sigma(M_{sim} - \varepsilon) \quad (2)$$

$$p'_q = \omega_1 p_s + \omega_2 p_q \quad (3)$$

其中参数 ε 表示自适应阈值,如图 2 所示, ε 为 Block5 生成的特征通过一个 MLP 层得到的一维映射。符号 σ 表示 Sigmoid 激活函数。参数 ω_1 和 ω_2 为用于调节信息融合比例的权重,在实验中均设置为 0.5。

区域原型计算分支采用本文所提出的自适应分区方法,以生成包含更多局部信息的区域原型。首先根据当前图像的前景特征数量自适应地确定分割区域所需的种子点数量。具体而言,通过设定一个最小分区特征数 n ,并根据前景特征数量 N 与该最小值的比值来确定种子点的数量 l ,计算如公式(4)所示:

$$l = \frac{N}{n} \quad (4)$$

接着,如图 4 所示,根据该数量 l 迭代生成初始化种子点。假设边界特征数量为 e ,第 1 步对于每一个前景特征都需要通过公式(5)计算其与所有边界特征的欧式距离产生一个集合 $D = [d_1, d_2, \dots, d_e]$,然后取每一集合中的最小值组成一个集合 $D_{\min} = [d_1^{\min}, d_2^{\min}, \dots, d_N^{\min}]$ 。

$$d(f_1, f_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5)$$

第 2 步,选择 D_{\min} 中最大值所对应的点存储为种子点;第 3 步,将该点视为边界点并从第 1 步重新开始,直至生成所有种子点组成一个集合 $S = [s_1, s_2, \dots, s_l]$ 。

区域掩码的生成过程与上述类似。在此过程中,考虑到种

子点有可能位于突变区域的边界位置而导致区域划分不准确. 为避免仅依靠欧式距离分区时可能出现的区域原型代表性下降问题, 还引入了每个像素位置的特征向量与种子点位置特征向量之间的内积相似性度量. 通过结合这种相似性信息, 可以更精确地调整每个像素的分配. 以某一个前景特征为例, 首先通过公式(5)计算其与所有种子点之间的欧式距离

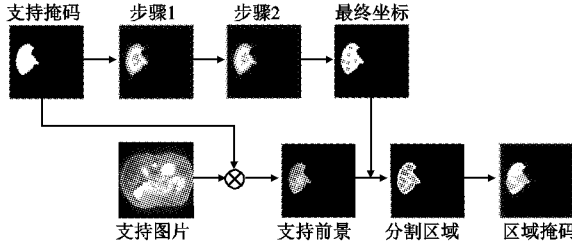


图4 自适应区域原型生成过程

Fig. 4 Adaptive region prototype generation process

产生一个集合 $D_s = [d_1, d_2, \dots, d_l]$, 然后计算其与所有种子点之间的余弦相似度产生一个集合 $D_{\cos} = [m_1, m_2, \dots, m_l]$, 将两个集合相加得到 $D_r = D_s + D_{\cos}$. 然后将该特征分配到 D_r 中最大值所对应的种子点区域中. 最终生成的子区域数量与初始种子点的数量保持一致. 对每个子区域内的所有特征向量进行加权求和取平均后得到区域原型集 P .

2.3.3 特征激活

为了充分利用网络各阶段生成的原型类别信息, 受到 IPMT^[40] 的启发, 采用特征激活 (Feature Activation, FA) 方法将这些信息传递到特征图中, 并激活目标区域以产生增强支持特征和查询特征, 以便后续模块能有效利用, 计算过程如公式(6)所示:

$$\bar{F}_s = FA(F_s \odot p'_q), \bar{F}_q = FA(F_q \odot p'_q) \quad (6)$$

其中, \bar{F}_s 和 \bar{F}_q 分别表示通过 FA 过程获得的增强支持特征和查询特征, 这些特征作为后续细化模块的输入, 符号 \odot 表示特征拼接操作. FA 过程由一个 1×1 的卷积层以及两个 3×3 的卷积层组成.

鉴于先前在分割任务中成功应用 FA 的相关研究^[23], 本文通过计算 p'_q 与 \bar{F}_q 之间的余弦相似度生成初始掩码 M_{ini} , 该掩码将用于指导后续细化模块的分割过程.

2.4 原型细化模块

前置网络生成的区域原型在保留局部区域信息方面表现出色. 然而当某一类别的前景区域被划分为多个子区域时, 每个子区域势必会丢失该类别的全局信息. 在实际的分割任务中, 局部细节信息有助于识别类内变化区域, 而全局信息则确保原型具有足够的类别特异性, 从而更有效地识别支持集与查询集之间的共享特征, 并减少背景区域的误识别. 然而现有的区域原型生成方法在突出局部区域信息的同时, 难以有效保留全局信息, 而仅依靠区域原型又无法充分保证分割性能.

因此, 为进一步增强局部区域信息并尽可能保留全局信息, 本文提出了原型细化模块. 该模块通过引入 Transformer 结构和掩码注意力机制, 构建区域原型与支持前景特征之间的全局相关关系, 以实现前景特征的重新加权聚合. 通过这种方式获得的支持信息将更加侧重于与区域原型内积相似度较高的支持特征. 同时, 由于整个过程对所有前景特征进行了注

意力加权聚合, 因此也能够有效地保留一定的全局信息. 此外, 原型细化模块中还引入了原型对齐机制, 通过与查询特征的交互能够实现区域原型与查询特征中相似区域的对齐, 从而实现精确的特征对齐与区域识别.

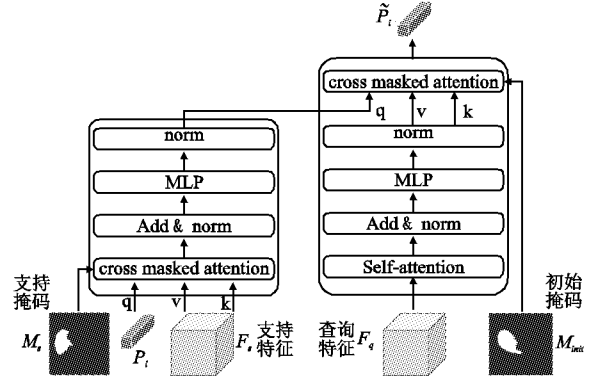


图5 原型细化模型, 左侧为原型增强部分, 右侧为原型对齐部分

Fig. 5 Prototype refinement model, where the left side illustrates the prototype enhancement component and the right side depicts the prototype alignment component

如图5所示, 原型细化模块主要由原型增强与原型对齐两部分构成. 原型增强操作的实现过程如下: 区域原型 $P_i \in P$ 作为查询 (query), 而支持特征 \bar{F}_s 则作为键 (key) 和值 (value) 输入到掩码交叉注意力块中. 在该注意力模块中首先需要计算查询与键之间的注意力矩阵, 这实际上是通过计算区域原型与每个前景支持特征的内积相似度从而衡量它们之间的相关性. 该注意力矩阵通过公式(7)进行计算:

$$A(P_i, F_s) = \frac{(P_i)^T F_s}{\sqrt{d}} \quad (7)$$

其中 d 是输入序列的维数. 通过对二进制支持掩码 M_s 进行平坦化操作, 以便在注意力图中有效地限制前景区域的关注范围. 通过这一操作, 掩码能够对前景区域进行精确定位, 从而引导注意力机制关注于具有显著信息的区域. 掩码交叉注意力 (MaskCross) 的计算公式如式(8)所示:

$$MaskCross(P_i, \bar{F}_s, M_s) = \text{sum}(M_s \cdot \bar{F}_s (\text{softmax}(A(P_i, \bar{F}_s)))) \quad (8)$$

其中, \bar{F}_s 表示支持特征, M_s 表示支持掩码.

如图5所示, 在经过残差结构处理后, 输出还需要通过两个归一化 (norm) 层和一个多层感知机 (MLP) 层, MLP 由两个线性变换层组成, 来恢复特征的尺度和分布, 以保证特征信息的稳定性和一致性.

在原型对齐操作中, 为了有效捕获查询特征中每个像素的全局上下文信息, 首先将初始特征进行展平处理并输入到自注意力块中. 该自注意力模块由一个多头注意力 (MHA) 层和一个 MLP 层构成. 给定输入序列 S , MHA 层首先将该序列投影为 3 个具有不同权重的序列 K, Q 和 V . 接着, 通过公式(9)计算注意力矩阵 A :

$$A(Q, K) = \frac{QK^T}{\sqrt{d}} \quad (9)$$

然后,注意力矩阵通过 softmax 函数进行归一化并乘以值序列 V 以获得输出序列 O ,具体计算如公式(10)所示:

$$O = \text{softmax}(A)V \quad (10)$$

输出序列在经过两个归一化层和一个 MLP 层后进入掩码交叉注意力层,这里的操作与原型增强部分一致,区别在于键和值被查询特征 F_q 所替代,监督掩码也被前置网络生成的初始掩码 M_{init} 所替代。

2.5 预测与训练损失

假设生成了 s 个区域原型,首先计算每个区域原型与查询特征之间的余弦相似度并拼接,得到一个相似度矩阵 $M \in R^{s \times h \times w}$ 。然后,最终的查询预测可以通过公式(11)得到:

$$y_{final} = 1 - \sigma(\text{argmax}(M) - \varepsilon) \quad (11)$$

本文采用二元交叉熵损失函数来衡量预测掩码与标准真实值之间的误差,从而对前置网络及其后续细化模块进行监督,分别如公式(12)和公式(13)所示:

$$L_{pre} = -\frac{1}{HW} \sum_h \sum_w y_q(h,w) \log(M_{init}(h,w)) \quad (12)$$

$$L_{main} = -\frac{1}{HW} \sum_h \sum_w y_q(h,w) \log(y_{final}(x,y)) \quad (13)$$

借鉴文献[22]的思路,通过使用查询图像作为支持集来反向预测支持图像的标签,进而构建原型对齐正则化项,具体如公式(14)所示:

$$L_{align} = -\frac{1}{HW} \sum_h \sum_w y_s(h,w) \log(y_{final}^s(x,y)) \quad (14)$$

综上所述,每个训练轮次的损失函数定义如公式(15)所示:

$$L = L_{pre} + L_{main} + L_{align} \quad (15)$$

3 实验

3.1 实验数据及评价准则

本文在两个具有不同模态和解剖结构的公开数据集上全面评估了所提出的方法,具体包括:(a) CHAOS^[41]:该数据集来自 ISBI 2019 联合健康腹部器官分割挑战,包含腹部 MRI 图像;(b) SABS^[42]:该数据集来自 MICCAI 2015 MultiAtlas 腹部标记挑战,包含腹部 CT 图像;CHAOS 和 SABS 数据集共享相同的标签类别,涵盖肝脏(Liver)、脾脏(Spleen)、左肾(LK)和右肾(RK)。

在模型训练之前,根据文献[35]中的超体素聚类方法生成用于训练的伪掩码。为了进行元学习方法中的情景训练,遵循文献[36]中的策略以选择支持切片和查询切片。对于查询集和支持集中的每一张三维图像,将感兴趣区域系统地细分为 3 个大小相等的片段。每个查询片段对应的支持样本由所有支持扫描中与之匹配的片段的中心切片组成。训练过程中按照文献[36]中的设定将实验数据分成 5 个部分进行交叉验证。

本文提出的方法在 PyTorch^[43] 框架下实现,采用预训练的 ResNet-101^[1] 作为主干网络,该网络在部分 MS-COCO 数据集上进行了预训练。数据预处理参考了文献[35]中的方法。具体地,每个单通道二维切片将被复制 3 次,并沿着级联方向排列以适应卷积网络的输入格式。在训练过程中采用 SGD 优化器,在 1-way, 1-shot 设置下进行训练,训练总迭代

次数超过 5 万次。训练阶段的初始学习率设定为 1×10^{-3} ,并在每 1000 次迭代后根据 0.8 的衰减因子对学习率进行更新。

为了评估模型性能,采用 Dice 相似系数(Dice Similarity Coefficient, DSC)来比较模型预测结果与标准真实值的相似度。它通过公式(16)衡量预测掩码 A 与真实掩码 B 的重叠比率:

$$DSC(A,B) = \frac{2A \cap B}{A+B} \times 100\% \quad (16)$$

此外,本文通过两种设置来考察所提出方法在推广到新数据时的能力。设置 1:在训练阶段,部分切片可能包含测试类的样本;设置 2:在训练阶段,完全排除包含测试类的切片。

3.2 与其他先进算法比较

本文将所提出的 ARNet 与现有先进方法进行了比较,包括 SENet^[32]、PANet^[22]、SSL-ALPNet^[35]、ADNet^[36]、AAS-DCL^[44]、SRCL^[45]、CRAPNet^[46]、Q-Net^[37] 和 PAMI^[38]。其中, PANet 是自然图像领域典型的原型小样本分割方法,而 SENet、SSL-ALPNet、ADNet、AAS-DCL、SRCL、CRAPNet、Q-Net 和 PAMI 则是小样本医学图像分割任务中的代表性工作。实验结果如表 1 所示,ARNet 在两个数据集上,无论是在设置 1 还是设置 2 下,表现都优于目前表现最好的方法 PAMI。在设置 1 下,ARNet 在 CHAOS 数据集上取得了 82.45% 的 DSC 分数,在 SABS 数据集上为 78.75%,分别超越 PAMI 方法 0.07% 和 1.06%。在设置 2 下,ARNet 在 CHAOS 数据集上获得了 80.01% 的 DSC 分数,在 SABS 数据集上为 72.31%,分别超越 PAMI 方法 0.48% 和 0.82%。综合来看,所提出的 ARNet 方法在两个数据集和两种评估设置下均持续优于目前表现最好的方法,进一步验证了其在提高分割精度方面的有效性及较强的泛化能力。

3.3 消融实验

本节中,通过消融实验分析所提出模型中各个组件的有效性。所有消融实验均在设置 1 下的 CHAOS 数据集上进行。

3.3.1 前置网络(Pre-Net)有效性

为了全面深入地分析前置网络的有效性,首先对各组成部分的作用进行了细致地探讨和分析。在这些实验中,特别需要强调的是,在训练过程中并未引入原型细化模块,而是直接使用前置网络所生成的区域原型进行预测。通过这一实验设置,能够更为清晰地了解自适应区域原型在未使用增强模块下的表现,并有效揭示前置网络本身的基本性能特征,并为后续模块的优化效果提供了有力的参考。具体而言,针对自适应区域原型的有效性,对比了采用 5 个随机种子点的前置网络性能。此外,为进一步验证所提出的 AQE 模块的有效性,对比了前置网络在采用未改进的 SSFP 模块时的性能表现。

如表 2 所示,在未启用“AQE”时,采用的是原本的 SSFP 模块。两者的主要差异在于是否引入自适应阈值来评估高置信度查询特征。当“自适应”未勾选时,区域原型的生成采用 5 个随机种子点;而勾选“自适应”后,则采用本文提出的自适应生成方法构建区域原型。表 2 的实验结果表明,采用自适应区域原型生成方法的模型性能优于使用固定数量随机种子点方法的模型。这充分验证了自适应区域原型生成方法在提升模型性能方面的有效性。同时,当启用 AQE 模块时,模型性能优于仅使用传统 SSFP 模块的情况。进一步地,启用 AQE 模

块和自适应生成方法的组合,相较于两者均未启用的情况, DSC 系数提升了 5.66%。这一结果充分体现了自适应区域原

型生成方法与 AQE 模块的协同作用,显著提升了模型的整体性能。此外表 3 展示了近年来表现优异的几种模型的比较结

表 1 设置 1 和设置 2 下不同方法在 CHAOS 和 SABS 数据集上的 DSC 分数(%)比较

Table 1 Comparison of DSC scores (%) of different methods on the CHAOS and SABS datasets under setting 1 and setting 2

设定	method	CHAOS					SABS				
		左肾	右肾	脾脏	肝脏	平均值	左肾	右肾	脾脏	肝脏	平均值
设置 1	PANet	30.99	32.19	40.58	50.40	38.53	20.67	21.19	36.04	49.55	32.86
	SENet	45.78	47.96	47.30	29.02	42.51	24.42	12.51	43.66	35.42	29.00
	SSL-ALPNet	81.92	85.18	72.18	76.10	78.84	72.36	71.81	70.96	78.29	73.35
	ADNet	73.86	85.80	72.29	82.11	78.51	72.13	79.06	63.48	77.24	77.24
	AAS-DCL	80.37	86.11	76.24	72.33	78.76	74.58	73.19	72.30	78.04	74.52
	SRCL	79.34	87.42	76.01	80.23	80.77	73.45	71.22	73.41	76.06	73.53
	CRAPNet	81.95	86.42	74.32	76.46	79.79	74.69	74.18	70.37	75.41	73.66
	Q-Net	78.36	87.98	75.99	81.74	81.02	-	-	-	-	-
	PAMI	81.83	88.73	76.37	82.59	82.38	76.52	80.57	72.38	81.32	77.69
	本文	81.52	90.74	74.14	83.42	82.45	77.21	81.33	74.11	82.34	78.75
	设置 2	PANet	53.45	38.64	50.90	42.26	46.33	32.34	17.37	29.59	38.42
SENet		62.11	61.32	51.80	27.43	50.66	32.83	14.34	0.23	0.27	11.91
SSL-ALPNet		73.63	78.39	67.02	73.05	73.02	63.34	54.82	60.25	73.65	63.02
ADNet		59.64	56.68	59.44	77.03	63.20	48.41	40.52	50.97	70.63	52.63
AAS-DCL		76.90	83.75	74.84	69.94	76.36	64.71	69.95	66.36	71.61	68.16
SRCL		77.07	84.24	73.73	75.55	77.65	67.39	63.37	67.36	73.63	67.94
CRAPNet		74.66	82.77	70.82	73.82	75.52	70.91	67.33	70.17	70.45	69.72
Q-Net		64.81	65.94	65.37	78.25	68.59	-	-	-	-	-
PAMI		74.51	86.73	75.80	81.09	79.53	72.36	67.54	71.95	74.13	71.49
本文		76.52	88.63	73.02	81.88	80.01	73.44	66.78	72.56	76.47	72.31

果,尽管本文 Pre-Net 在性能上略逊于这些模型,但其结构简洁且未附加任何额外模块,即可与这些模型在性能上接近,进一步验证了所提出方法的有效性。

表 2 AQE 和自适应区域原型的有效性消融实验

Table 2 Ablation study on the effectiveness of the AQE and adaptive region prototype

AQE	自适应	DSC
		72.88
✓		73.41
	✓	77.66
✓	✓	78.54

表 3 前置网络与其他方法的 DSC(%)分数比较

Table 3 Comparison of DSC (%) scores between the Pre-Net and other methods

方法	DSC
AAS-DCL	78.76
SRCL	80.77
CRAPNet	79.79
Q-Net	81.02
Pre-Net	78.54

图 6 和图 7 展示了 Pre-Net 与近年来在同类任务中表现优异的 PAMI^[38]之间的定性比较。通过观察可见,Pre-Net 与 PAMI 在预测准确度上差异较小,主要区别体现在边界预测的模糊程度上。该结果进一步验证了 Pre-Net 在整体预测性能上的有效性和优势。值得注意的是,Pre-Net 未引入任何复杂的细化模块,其预测结果仅通过一个简单的基于余弦相似度

的分类器获得;而 PAMI 则采用了更为复杂的网络结构和额外模块。这一对比突显了 Pre-Net 在保持高效性能的同时,能够维持简洁的结构设计。

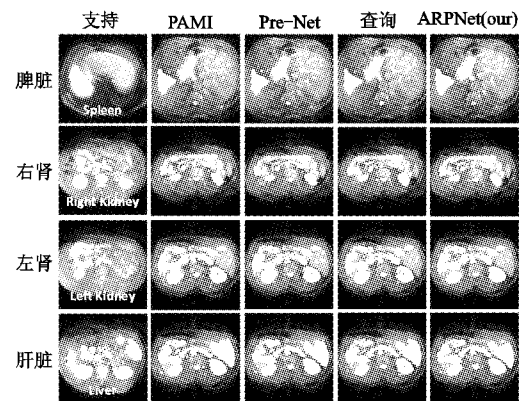


图 6 ARPNet、Pre-Net 和 PAMI 在 CHAOS 数据集上的分割结果

Fig. 6 Segmentation results of ARPNet, Pre-Net and PAMI on the CHAOS dataset

3.3.2 原型细化模块的有效性

原型细化模块主要包括“原型增强”和“原型对齐”两个部分。本节分析了这两部分对模型性能的独立影响,并进一步探讨了它们之间的协同作用。如表 4 所示,在未启用“原型增强”和“原型对齐”模块时,模型表现是 Pre-Net 的分割性能,旨在为对比分析提供基准,从而更清晰地验证该模块的有效性。

从表 4 的结果可见,启用原型增强模块后,相比基线模型

DSC 系数提升了 0.53%。这一结果表明原型增强部分在一定程度上有效增强了区域原型的的信息,然而提升幅度较小,主要原因在于支持特征与查询特征之间普遍存在较大差异,导致

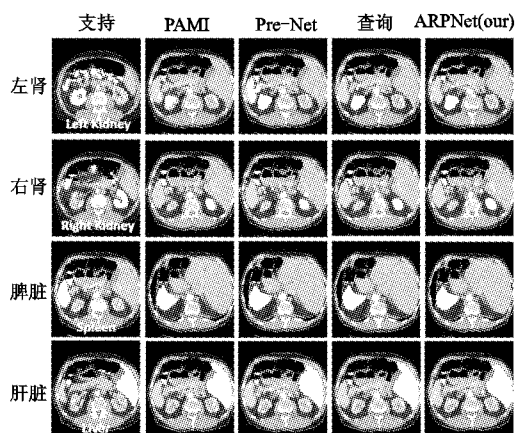


图 7 ARPNet、Pre-Net 和 PAMI 在 SABS 数据集上的分割结果

Fig. 7 Segmentation results of ARPNet, Pre-Net and PAMI on the SABS dataset

有助于查询分割的支持信息本身较为有限。因此尽管进行了增强处理,但其效果仍然相对有限。另一方面,单独启用原型对齐模块时,DSC 系数提升了 1.85%,显著高于单独启用原型增强部分,进一步验证了原型对齐在模型中的重要性。尽管支持信息有限,但这些信息能够提取出部分高置信度的查询信息,从而完成原型对齐并显著提升模型性能。与此同时当两个模块同时启用时,模型的 DSC 系数达到了 82.45%,相较于基线模型提升了 3.91%。这表明原型增强模块不仅有效补充了全局信息,突出局部信息,还促进了原型对齐部分在查询信息提取上的表现。两者的协同作用进一步提升了模型的整体性能。

表 4 原型细化模块的有效性消融实验
Table 4 Ablation study on the effectiveness of the prototype refinement module

原型增强	原型对齐	DSC
		78.54
✓		79.07
	✓	80.39
✓	✓	82.45

为了进一步验证原型细化模块的有效性,图 6 和图 7 展示了 ARPNet 与 Pre-Net 的定性比较,其中 Pre-Net 是没有添加原型细化模块的模型性能,而 APRNet 是在 Pre-Net 后面添加原型细化模块后的模型性能。从结果中可以明显看出,ARPNet 在预测准确度和边界预测精度上都表现出了显著提升,验证了原型细化模块在增强区域原型局部细节信息方面的有效性。此外 ARPNet 在减少错误预测像素数量方面也取得了显著进展,进一步证明了原型细化模块在有效对齐支持查询目标方面的作用,从而提高了预测质量。

4 总结

本文针对医学图像中前景与背景比例严重失衡、内类变

化突发性强,以及类间差异较大等特殊性质对小样本分割模型性能的影响,提出了一种自适应区域原型生成方法,该方法根据目标物体的大小自适应生成原型数量,并根据目标物体的形状生成聚合种子点,从而生成更具代表性且能保留更多局部细节信息的区域原型。鉴于区域原型在保留局部信息的同时可能丧失全局信息,基于 Transformer 结构和掩码注意力机制提出了原型细化模块。该模块包含原型增强和原型对齐两个部分。原型增强部分通过注意力机制进一步强化区域原型的局部信息,并补充全局信息;原型对齐部分则通过区域原型与查询特征的交互,引入查询信息,实现区域原型与查询目标的精确对齐。在两个医学小样本分割数据集上结合两种不同的实验设置,本文提出的 ARPNet 在性能上都表现出了优异的结果。然而在实验过程中,原型细化模块中的原型对齐部分由于缺乏真实查询掩码的监督,不可避免地引入了噪声。此外当前细化模块的结构设计仍较为简单。因此未来的工作将着重于优化原型细化模块,以期进一步提升模型性能。

References:

- [1] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770-778.
- [2] Wang Y Q, Yao Q M, Kwok J T, et al. Generalizing from a few examples: a survey on few-shot learning[J]. Association for Computing Machinery Computing Surveys, 2020, 53(3):1-34.
- [3] LI L F, FAN X Y. Few-shot object detection with meta-learning and multi-scale feature fusion[J]. Journal of Chinese Computer Systems, 2023, 44(12):2804-2811.
- [4] WEI T, LI X L, LIU H. Survey on image semantic segmentation in dilemma of few-shot[J]. Computer Engineering and Applications, 2023, 59(2):1-11.
- [5] CHEN Q, YANG Y, HUANG T L, et al. A survey on few-shot image semantic segmentation[J]. Frontiers of Data and Computing, 2021, 3(6):17-34.
- [6] Dong N Q, Xing E P. Few-shot semantic segmentation with prototype learning[C]//British Machine Vision Conference (BMVC), 2018:79-92.
- [7] Snell J, Kevin S, Richard Z. Prototypical networks for few-shot learning[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems(NeurIPS), 2017:4080-4090.
- [8] WU J, QU S J. Self-guided and feature-guided enhancement network for few-shot segmentation[J/OL]. Journal of Chinese Computer Systems, 2025: 1-12, <http://kns.cnki.net/kcms/detail/21.1106.TP.20240926.1734.029.html>.
- [9] Ouyang C, Biffi C, Chen C, et al. Self-supervision with superpixels: Training few-shot medical image segmentation without annotation [C]//European Conference on Computer Vision (ECCV), 2020:762-780.
- [10] Zhu Y Z, Wang S D, Xin T, et al. Partition-a-medical-image: extracting multiple representative sub-regions for few-shot medical image segmentation[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73:1-12, doi:10.1109/TIM.2024.3381715.
- [11] Fan Q, Pei W J, Tai Y W, et al. Self-support few-shot semantic segmentation[C]//European Conference on Computer Vision (ECCV), 2022:701-719.
- [12] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention International Conference (MICCAI), 2015:234-241.
- [13] Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-net: learning dense volumetric segmentation from sparse annotation[C]//Medi-

- cal Image Computing and Computer-Assisted Intervention International Conference (MICCAI), 2016:424-432.
- [14] Milletari F, Navab N, Ahmadi S A. V-net: fully convolutional neural networks for volumetric medical image segmentation [C]//4th International Conference on 3D Vision (3DV), 2016:565-571.
- [15] Mehta S, Mercan E, Bartlett J, et al. Y-Net: joint segmentation and classification for diagnosis of breast biopsy images [C]//Medical Image Computing and Computer-Assisted Intervention International Conference (MICCAI), 2018:893-901.
- [16] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: learning where to look for the pancreas [J]. arXiv preprint arXiv:1804.03999, 2018.
- [17] Isensee F, Jaeger P F, Kohl S A A, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation [J]. Nature Methods, 2021, 18(2):203-211.
- [18] Lin A L, Chen B Z, Xu J Y, et al. Ds-transunet: dual swin transformer u-net for medical image segmentation [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:1-15.
- [19] Cao H, Wang Y Y, Chen J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation [C]//European Conference on Computer Vision (ECCV), 2022:205-218.
- [20] Chen J N, Lu Y Y, Yu Q H, et al. Transunet: transformers make strong encoders for medical image segmentation [J]. arXiv preprint arXiv:2102.04306, 2021.
- [21] Shaban A, Bansal S, Liu Z, et al. One-shot learning for semantic segmentation [J]. arXiv preprint arXiv:1709.03410, 2017.
- [22] Wang K X, Liew J H, Zou Y T, et al. Panet: few-shot image semantic segmentation with prototype alignment [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019:9197-9206.
- [23] Zhang C, Lin G H, Liu F Y, et al. Canet: class-agnostic segmentation networks with iterative refinement and attentive few-shot learning [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019:5217-5226.
- [24] Tian Z T, Zhao H S, Shu M, et al. Prior guided feature enrichment network for few-shot segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(2):1050-1065.
- [25] Liu Y F, Zhang X Y, Zhang S Y, et al. Part-aware prototype network for few-shot semantic segmentation [C]//European Conference on Computer Vision (ECCV), 2020:142-158.
- [26] Yang B Y, Liu C, Li B H, et al. Prototype mixture models for few-shot semantic segmentation [C]//European Conference on Computer Vision (ECCV), 2020:763-778.
- [27] Li G, Jampani V, Sevilla Lara L, et al. Adaptive prototype learning and allocation for few-shot segmentation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021:8334-8343.
- [28] Okazawa A. Interclass prototype relation for few-shot segmentation [C]//European Conference on Computer Vision (ECCV), 2022:362-378.
- [29] Zhang G W, Kang G L, Yang Y, et al. Few-shot segmentation via cycle-consistent transformer [C]//Advances in Neural Information Processing Systems (NeurIPS), 2021:21984-21996.
- [30] Hong S, Cho S, Nam J, et al. Cost aggregation with 4d convolutional swin transformer for few-shot segmentation [C]//European Conference on Computer Vision (ECCV), 2022:108-126.
- [31] Peng B H, Tian Z T, Wu X Y, et al. Hierarchical dense correlation distillation for few-shot segmentation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023:23641-23651.
- [32] Roy A G, Siddiqui S, Pölsterl S, et al. "Squeeze & excite" guided few-shot segmentation of volumetric images [J]. Medical Image Analysis, 2020, 59:101587, doi:10.1016/j.media.2019.101587.
- [33] Sun L Y, Li C X, Ding X H, et al. Few-shot medical image segmentation using a global correlation network with discriminative embedding [J]. Computers in Biology and Medicine, 2022, 140:105067, doi:10.1016/j.compbiomed.2021.105067.
- [34] Feng R W, Zheng X S, Gao T X, et al. Interactive few-shot learning: limited supervision, better medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2021, 40(10):2575-2588.
- [35] Ouyang C, Biffi C, Chen C, et al. Self-supervision with superpixels: Training few-shot medical image segmentation without annotation [C]//European Conference on Computer Vision (ECCV), 2020:762-780.
- [36] Hansen S, Gautam S, Jenssen R, et al. Anomaly detection-inspired few-shot medical image segmentation through self-supervision with supervoxels [J]. Medical Image Analysis, 2022, 78:102385, doi:10.48550/arXiv.2203.02048.
- [37] Shen Q Q, Li Y N, Jin J Y, et al. Q-net: query-informed few-shot medical image segmentation [C]//Proceedings of SAI Intelligent Systems Conference, 2023:610-628.
- [38] Zhu Y Z, Wang S D, Xin T, et al. Partition-a-medical-image: Extracting multiple representative sub-regions for few-shot medical image segmentation [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73:1-12.
- [39] GAO F, JIANG A L, JI W, et al. A dual-branch self-optimizing segmentation network for few-shot medical image segmentation [J/OL]. Computer Engineering and Applications, 2025:1-14, http://kns.cnki.net/kcms/detail/11.2127.tp.20250121.1527.002.html.
- [40] Liu Y W, Liu N, Yao X W, et al. Intermediate prototype mining transformer for few-shot semantic segmentation [C]//Advances in Neural Information Processing Systems (NeurIPS), 2022:38020-38031.
- [41] Kavur A E, Gezer N S, Barış M, et al. CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation [J]. Medical Image Analysis, 2021, 69:101950, doi:10.1016/j.media.2020.101950.
- [42] Landman B, Xu Z, Igelsias J, et al. Miccai multi-atlas labeling beyond the cranial vault-workshop and challenge [C]//Proceedings of Medical Image Computing and Computer-Assisted Intervention Multi-Atlas Labeling Beyond Cranial Vault-Workshop Challenge (MICCAI), 2015:12-13.
- [43] Paszke A. Pytorch: an imperative style, high-performance deep learning library [J]. arXiv preprint arXiv:1912.01703, 2019.
- [44] Kwon H, Jeong S, Kim S, et al. Dual prototypical contrastive learning for few-shot semantic segmentation [J]. arXiv preprint arXiv:2111.04982, 2021.
- [45] Wang R Z, Zhou Q, Zheng G Y. Few-shot medical image segmentation regularized with self-reference and contrastive learning [C]//Medical Image Computing and Computer-Assisted Intervention International Conference (MICCAI), 2022:514-523.
- [46] Ding H, Sun C C, Tang H, et al. Few-shot medical image segmentation with cycle-resemblance attention [C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2023:2488-2497.

附中文参考文献:

- [3] 李丽芬, 范新焯. 元学习与多尺度特征融合的小样本目标检测 [J]. 小型微型计算机系统, 2023, 44(12):2804-2811.
- [4] 韦婷, 李馨蕾, 刘慧. 小样本困境下的图像语义分割综述 [J]. 计算机工程与应用, 2023, 59(2):1-11.
- [5] 陈琼, 杨咏, 黄天林, 等. 小样本图像语义分割综述 [J]. 数据与发展前沿, 2021, 3(6):17-34.
- [8] 吴杰, 瞿绍军. 自我引导和特征引导增强的小样本分割网络 [J/OL]. 小型微型计算机系统, 2025:1-12, http://kns.cnki.net/kcms/detail/21.1106.TP.20240926.1734.029.html.
- [39] 高峰, 降爱莲, 冀伟, 等. 双路自优化小样本医学图像分割网络 [J/OL]. 计算机工程与应用, 2025:1-14, http://kns.cnki.net/kcms/detail/11.2127.tp.20250121.1527.002.html.