

基于双路信息聚合原型度量网络的小样本遥感图像语义分割

整体框架

一、项目背景与研究动机

小样本语义分割是遥感图像分析中的关键难题之一，因其标注成本高、数据获取困难，传统全监督方法在样本不足时易过拟合、泛化性差。此前的研究（如 MSAPNet、MLBFPNet）已在多尺度特征融合与多原型度量方面取得进展，但仍存在以下核心挑战：

- 类内特征空间膨胀问题**：支持集与查询集中的同类目标因拍摄角度、分辨率、光照等差异，可能在外观上相差较大，造成原型生成不一致，影响分割精度；
- 支持集标签利用不足**：支持集的真值掩码仅用于原型生成阶段，未参与监督学习过程，削弱了模型的训练引导能力。

为解决上述问题，本项目提出 **DIAPNet (Dual-path Information Aggregation Prototype Network)**，引入双路监督结构、跨分支特征交互机制与动态引导型多尺度解码模块，从特征一致性建模、监督机制设计和上下文融合三个维度全面提升小样本遥感图像语义分割性能。

二、主要贡献与创新点

1. 双路监督结构 (Dual Supervision)

在传统“支持集→查询集”监督路径的基础上，引入“查询集→支持集”反向监督通道，构建双向监督结构。该结构使得支持集标签既用于原型生成，又直接参与支持集预测的监督，从而提升了类内特征一致性与原型鲁棒性。

2. 分支信息交互模块 (Cross-branch Interaction Module)

设计基于 Non-local 和 Self-Attention 的交互模块，支持集与查询集之间实现双向全局语义信息传递，强化模型对共性目标的建模能力。同时引入 SKNet 子模块，实现多尺度通道注意力加权，提升模型在不同尺度下的泛化表现。

3. 动态层次特征引导的解码结构 (Dynamic Hierarchical Decoding)

对传统 ASPP 模块进行升级，引入多项改进：

- 任务自适应路径引导机制 (Dynamic Routing Module)**：根据输入特征复杂度动态激活不同空洞率路径，适应不同目标密度场景（如密集城区与稀疏乡村）；
- 前背景引导残差融合模块 (Foreground-Background Residual Guidance)**：利用支持集中前景与背景特征，对 ASPP 分支进行上下文校准，增强目标边界建模能力；
- 分层残差连接机制 (Hierarchical Residual Merging)**：从高膨胀率分支逐级向低膨胀率路径传递上下文信息，缓解原始 ASPP 分支独立处理所导致的信息割裂；
- 最终融合 IRM (Information Reallocation Module) 进一步整合多尺度特征，增强分割头的语义表达能力与结构一致性。

三、网络结构总览与处理流程

1. **输入处理**：将支持集与查询集图像输入多尺度注意力特征提取网络，提取语义特征；
2. **特征交互**：在提取特征基础上，通过交互模块完成支持-查询集的双向信息融合；
3. **原型生成**：利用 PMMs 多原型模块从支持集提取多个类别原型，并与查询特征进行前/背景度量嵌入；
4. **双路监督**：一条路径输出查询集预测结果，另一条路径输出支持集预测结果，均参与监督训练；
5. **分割解码**：通过动态层次引导的 ASPP 解码结构提取多尺度上下文，输出最终掩码；
6. **损失函数设计**：采用交叉熵与尺度感知损失联合训练，提高整体精度与小目标检测能力。

四、实验设计与评估结果

1. 消融实验

对三个关键模块分别进行消融分析：

- 去除任一模块均导致平均 mIoU 下降 1.5%–2.5%，验证各模块对整体性能的显著贡献；
- 三模块联合使用效果最佳，体现了结构设计的互补性。

2. 模块优化探索

- 交互模块中，SKNet 嵌入于第4分支时性能最优，mIoU 提升约 1.4%；
- ASPP 解码模块采用“右向左”残差传递结构，在所有子集上均取得最佳上下文融合效果。

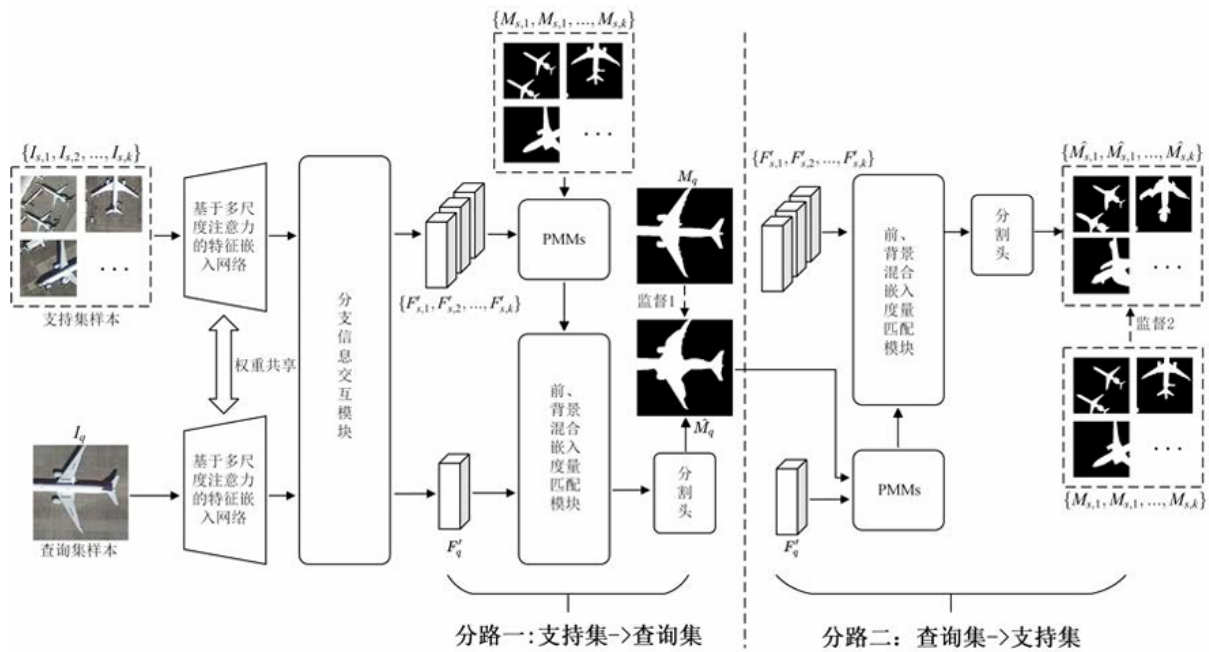
3. 方法对比实验

- 数据集：iSAID-5i，采用五折划分，分别在 1-shot 和 5-shot 场景下进行测试；
- 相比 PANet、CANet、PFENet、PMMs、SDM、HMRE 等主流方法：
 - DIAPNet 在所有子集上取得最优平均 mIoU（1-shot: **37.09%**，5-shot: **42.59%**）；
 - 相较 MLBFNet 显著提升（1-shot 提升 **5.03%**，5-shot 提升 **1.77%**）。

4. 可视化结果分析

- DIAPNet 在小目标（如储罐）与外观多变目标（如棒球场）上的边界分割更清晰；
- 减少了背景误分与漏检现象，展现出强抗噪能力与跨样本一致性。

分模块



1 双路监督结构 (Dual Supervision)

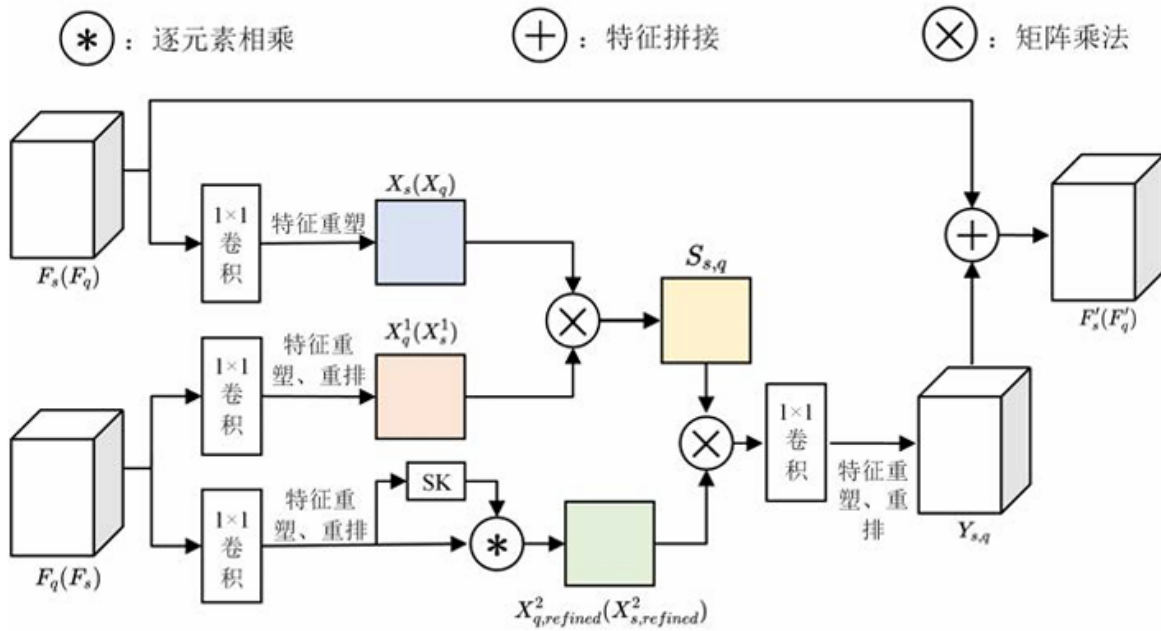
在传统小样本分割方法中，支持集的标签掩码通常仅在原型生成阶段被使用，用作先验信息参与构建类原型，而未直接参与损失计算。这种单向监督策略忽略了支持集中蕴含的真实标签所能提供的监督信号，导致模型对支持集的语义理解不充分，也限制了原型表达的判别性与稳定性。

本研究引入“双路监督结构”，将支持集和查询集的标签信息同时引入到损失函数中，实现对模型的双向引导与优化。该结构包含两个信息流路径：

- **第一条路径是从支持集至查询集的正向监督通道。**支持集图像首先通过特征提取模块提取特征，再结合其真实标签掩码生成前景与背景类别原型。随后，这些原型与查询集的特征进行匹配，生成查询集的预测掩码。该预测掩码会与查询集的真值掩码进行比对，并计算监督损失，从而对模型进行优化。
- **第二条路径是从查询集至支持集的反向监督通道。**该路径将前一条路径中生成的查询集预测掩码与查询集特征一起用来反向构造新的原型。随后，这些基于查询集的信息会与支持集特征进行匹配，生成支持集的预测掩码，并利用支持集真实标签进行监督。

通过以上双向流程，模型在训练阶段能够同时从支持集和查询集中学习语义信息，使生成的原型表示更加稳定、鲁棒性更强。尤其在小样本环境下，该结构显著提升了特征表征的一致性与任务间的泛化能力。

2 分支信息交互模块 (Cross-branch Interaction Module)



小样本任务中的支持集和查询集虽然目标类别一致，但由于来源图像在成像条件（如光照、尺度、纹理等）上存在差异，导致模型对同类目标的表征可能存在显著偏差。这种类内特征差异不仅会降低原型匹配的准确性，还会使模型在跨任务间难以保持一致的语义理解。

为了解决这一问题，本章设计了“分支信息交互模块”，通过引入非局部注意力机制，实现支持集与查询集之间的语义对齐和深层次信息交互。

具体而言，支持集与查询集图像首先分别经过主干网络提取多尺度特征图。随后，这两个特征图被同时送入交互模块，在该模块中，分别通过多个卷积操作将特征映射投影到注意力空间，并进行空间维度上的对齐与重排。通过相似性计算，模块能够获取支持集与查询集中相应位置之间的相关性权重，从而生成一个跨分支的注意力图。

为了提升注意力的表达能力，模块进一步引入了通道选择机制，即引入轻量级的选择性注意机制（如SKNet），以强化模型对关键语义通道的关注，并自适应抑制冗余信息。

生成的注意力图随后与查询集的增强特征进行融合，得到非局部响应特征。这一响应特征最终通过维度调整与通道恢复，与原始支持集特征进行残差连接，从而形成语义交互增强后的支持集特征。查询集特征的处理流程与其对称。

在 k-shot 场景下，所有支持集样本的特征将同时与查询集进行交互处理，分别生成增强后的多个支持特征。这种跨支路的语义对齐方式有效缓解了类内差异过大的问题，提升了模型对同一类别目标的一致性理解能力。

3 动态层次特征引导的 ASPP 解码结构 (Dynamic Hierarchical Decoding)

空洞空间金字塔池化模块（ASPP）是当前语义分割中常见的多尺度上下文建模方法。然而，其多分支并行结构在遥感图像场景中仍存在若干局限：一是各路径之间独立处理，导致上下文信息难以联动融合；二是对小目标响应不敏感，容易造成漏检；三是在复杂地物结构下分割边界不清晰。

为了解决上述问题，本章提出了一种“动态层次特征引导的 ASPP 解码结构”，从三方面进行改进：

(1) 任务自适应路径激活

传统 ASPP 固定采用多个膨胀率路径对输入进行处理，缺乏对图像复杂度的动态适应能力。本结构引入路径激活控制机制，通过轻量级通道注意力模块，根据输入特征图的整体语义内容动态调整各路径的激活强度。例如，当输入图像中包含大量细粒度目标时，低膨胀率路径将被强化；反之，则增强大膨胀率路径的全局感受能力。该机制提升了模型对不同类型目标结构的感知能力与处理灵活性。

(2) 前背景引导融合机制

考虑到支持集已经包含了显式的前景与背景类别信息，解码结构进一步引入了一种前背景引导策略。通过利用支持集所生成的前景与背景原型分别对 ASPP 输出特征进行加权融合，强调模型对关键语义区域的注意力，并抑制背景区域的干扰。这一过程有效提升了模型在结构复杂区域（如建筑边缘、林地边界）的分割精度。

(3) 层间特征信息传递机制

为了打破传统 ASPP 各路径间的独立性，本结构设计了“分层残差信息融合机制”。具体而言，具有更大感受野的路径（即较高膨胀率）会将其提取的全局特征信息传递至下一层路径中，以帮助其更好地融合上下文。这种从全局到局部的信息流引导方式增强了模型对多尺度目标的建模能力，尤其在处理多目标并存、尺度变化剧烈的遥感图像时效果更为显著。

为了融合不同路径的特征，本结构进一步引入“信息重分配模块”（IRM）。该模块首先对来自不同尺度的特征进行拼接，通过多层卷积与残差结构进行特征混合，最终输出统一维度的特征图，用于生成最终的预测掩码。

实验：

实验环境与设置：

模型训练采用元学习框架，共进行 50000 次元任务迭代，训练批大小设置为 4。优化器方面使用 Adam，初始学习率为 0.001，衰减策略为 step-based。损失函数中的两个权重系数（双路监督损失权重与尺度损失权重）均设置为 0.5。小样本实验分别在 1-shot 与 5-shot 场景下进行评估，覆盖 iSAID-5i 数据集中全部五折子集，以衡量模型在不同样本数下的鲁棒性。

在完成各项消融实验、结构优化与参数配置后，该结构包含三个关键模块：双路监督机制、分支信息交互模块，以及动态层次特征引导的 ASPP 解码结构。其中，SKNet 模块嵌入于交互模块的第 4 支路，同时 ASPP 模块采用从高膨胀率至低膨胀率的“Right → Left”信息传递策略。

在 iSAID-5i 数据集上进行的五折交叉实验中，最终评估指标如下：

1-shot 小样本分割性能（单位：%）

子集	子集1	子集2	子集3	平均 mIoU
mIoU	39.63	26.66	44.98	37.09

5-shot 小样本分割性能（单位：%）

子集	子集1	子集2	子集3	平均 mIoU
mIoU	45.44	33.27	49.06	42.59

项目问答

1. 你在项目里边主要做了哪些工作？

我主要负责三个核心部分的设计与实现：（1）**双路监督机制**的构建，包括支持集和查询集监督路径的设计、反向原型生成策略及损失整合；（2）**分支信息交互模块**的搭建，基于非局部注意机制完成支持/查询特征间的跨支路信息融合；（3）**ASPP模块的层次增强改进**，我设计了跨空洞率路径的信息传递策略，并通过实验验证“右向左”特征传导方式的优越性。此外，我还承担了**消融实验、结果可视化与论文撰写**工作。

2. 你这个项目主要是做什么的？

该项目旨在解决**遥感图像小样本语义分割**任务中存在的类内差异大、样本不足等问题。通过引入三大创新模块，提升模型对**同类地物的跨样本语义一致性建模能力**，从而在只有少量标注样本的条件下，也能实现精确的目标分割。

3. 相关的背景技术有一个大概的了解，xxx领域常用的算法是如何处理的？

遥感小样本分割通常借助**元学习（meta-learning）+原型度量（prototypical matching）**方式建模，其中代表方法有 PANet、CANet、PFENet 等。它们通过支持集生成类别原型，然后对查询集进行相似度度量预测。然而，它们存在对类内变化鲁棒性不足、支持集监督信息利用不充分的问题。

4. 项目网络结构是如何设计的？分为哪几个模块部分？

整体网络结构由五部分构成：

- 1. **特征提取模块**：使用多尺度注意力结构提取支持集与查询集的中高层语义特征；
- 2. **分支信息交互模块**：非对称双分支非局部交互增强，减少类内差异；
- 3. **多原型生成模块（PMMs）**：生成前背景多个原型进行度量；
- 4. **双路监督结构**：支持→查询、查询→支持双路径原型监督；
- 5. **动态ASPP解码器**：多尺度上下文建模，含特征引导与路径选择机制。

5. Xxx和&&&的区别是什么？为什么用这个xxx不用&&&？

比如我们采用的是“**双路监督结构**”，而不是传统的“单向监督结构”，原因是后者仅利用支持集标签用于原型生成，而不参与损失，信息利用不足。相比之下，双路结构引入**查询集→支持集反向路径**，使得支持集标签也能作为监督信号直接参与模型优化，提升原型稳定性。

6. Xxx模块是什么？项目处理流程是什么？

以**分支信息交互模块**为例：它是在支持集和查询集之间建立一种非对称、双向的注意力交互机制。项目完整流程为：特征提取 → 交互模块 → 原型生成 → 前背景度量 → 双路监督 → ASPP解码 → 输出分割掩码。

7. Xxx是怎么识别的？怎么检测的？怎么分割的？

模型通过原型度量实现目标区域分割。首先通过支持集图像及标签构建前背景原型；然后查询图像特征与原型进行匹配，输出相似度图；最后经过分割头（ASPP）和软最大化操作生成最终掩码，从而实现目标识别与像素级分割。

8. 为什么用xx模块，Xxx模块这样说设计的效果如何？xx设计的作用？

例如：**动态层次ASPP模块**之所以引入，是为了应对遥感图像中目标尺度差异大、边界复杂的问题。相比传统ASPP，该模块引入了跨路径信息流动（如从大膨胀率向小膨胀率），并通过前背景引导机制增强语义一致性。实验中平均 mIoU 提升达 1%-1.5%。

9. Xx算法是什么？你这个项目的xx算法描述下？（算法步骤/网络输入输出流程是什么？）

整体算法遵循**任务级训练流程**：

1. 每次随机采样一组支持集和查询集；
 2. 提取特征后进行分支交互；
 3. 构建原型并完成双向监督；
 4. 查询特征与原型匹配，生成预测；
 5. 输入 ASPP 解码模块；
 6. 输出预测掩码并根据查询/支持标签计算损失。
-

10. 项目算法未来还有哪些需要改进的地方？你们做了哪些改进？（项目创新点）

未来可以加强对原型的多样性建模、支持跨域遥感图像适配等方向。我们当前已做的三点主要改进为：

1. **引入双路监督**：充分利用支持集标签；
2. **设计跨分支交互机制**：解决类内特征膨胀；

3. 提出动态层次解码器：多尺度上下文联合建模。

11. 为什么使用xx算法，不使用xx算法？（项目使用的网络的好处）

我们采用的是基于原型度量的结构，而不是基于全监督的 encoder-decoder，是因为小样本场景标注极少，原型度量可实现任务适配与快速泛化。传统全监督方法在样本稀缺下极易过拟合。

12. 项目的数据集哪来的？简述下数据集（项目使用的什么数据集？）

我们使用的是公开遥感分割数据集 **iSAID-5i**，它是对大规模 iSAID 数据集的五折划分版本，共包含 15 个类别，每个类别在不同折中轮流作为测试类，适合小样本设置下的交叉验证。

13. 数据是如何处理的？

数据预处理包括：统一图像大小为 512×512，归一化操作，随机裁剪与镜像增强。标签掩码进行 one-hot 编码以适配多类分割任务。

14. 算法的识别效果、精度怎么样？

在 iSAID-5i 上，模型在最终设置下达到了：

- **1-shot 平均 mIoU：37.09%**
 - **5-shot 平均 mIoU：42.59%**
- 模型在边界清晰度、小目标检测与背景抑制方面效果优异。
-

15. 数据集的评价指标代表什么含义？

我们主要使用 **mIoU（mean Intersection over Union）**。它表示预测区域与真实标签区域的交集与并集之比，衡量分割结果的精度和完整性，是语义分割任务的主流指标。

16. 损失函数是什么？

我们使用的是**双路交叉熵损失 + 尺度感知损失**：

- 支持→查询和查询→支持路径各自计算 cross-entropy；
 - 尺度感知损失提升模型对小目标、边界区域的精度。
-

17. 实验设置的参数？模型训练了多久？

- 批大小：4；
- 迭代次数：50000 个元任务；
- 学习率初始设为 0.001；

- 优化器：Adam；
 - 模型训练约耗时 **2-3 天（单卡）**，整体项目周期约 3-4 个月。
-

18. 做了哪些实验？

主要实验包括：

- 三大核心模块的消融实验；
 - SKNet 嵌入位置探索；
 - ASPP 模块结构方向对比；
 - 全模型可视化分析；
 - 五折交叉验证下的平均性能评估。
-

19. 项目中遇到的困难？

初期最大的问题是支持集→查询集反向监督中，查询生成的原型稳定性不够，容易过拟合。我通过增加前背景混合策略，并调整原型正则项，稳定了预测路径；同时，在交互模块中加入通道注意力提升了关键特征引导能力。