



浙江大學
ZHEJIANG UNIVERSITY

Prior Guided Feature Enrichment Network for Few-Shot Segmentation

先验特征引导的特征增强网络

汇报人：王嘉恒



01 问题背景

02 论文创新点

03 网络框架

04 实验分析

05 总结

01

问题背景

当前小样本语义分割所存在的问题



问题背景

传统语义分割：

传统语义分割的框架如FCN，DeepLab，PSPNet等，训练模型都需要大量带标注的图像，存在着两个问题：

- ①标注成本高。
- ②对未知类别的推广能力很差（泛化能力）。



问题背景

小样本语义分割任务描述：

- ①数据分为支持集S和查询集Q。
- ②用一个预训练的backbone来提取支持图像和查询图像的特征。
- ③经过一系列特征的处理得到预测的结果。分类器，余弦相似度的计算或者卷积。
- ④训练时所学习的类和测试时的类是完全不相同的。

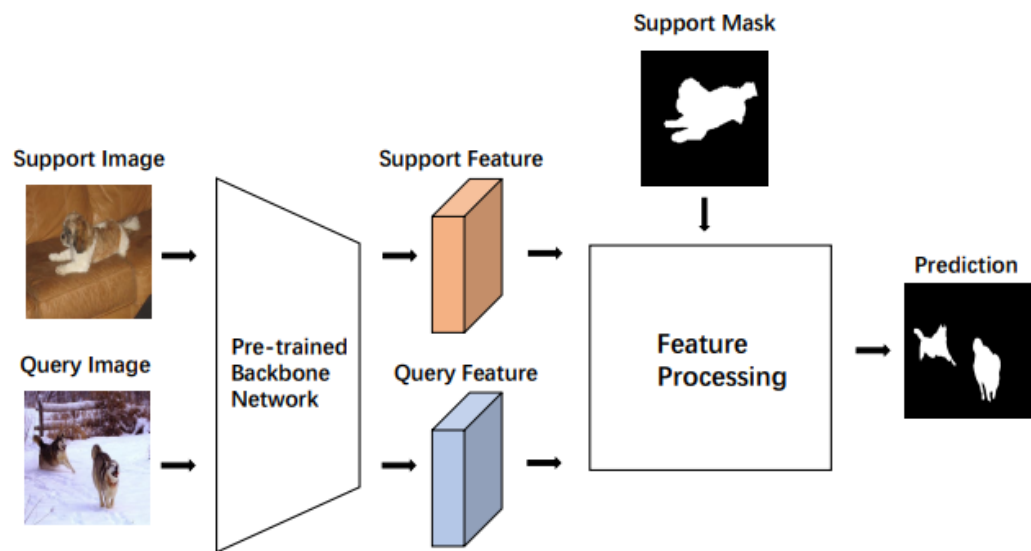


图1. 小样本语义分割的传统架构

问题背景

小样本语义分割所面临的问题：

①高层特征的滥用

block2	block3	block4	meanIoU
✓			46.6
	✓		50.8
		✓	48.4
✓	✓		51.2
	✓	✓	49.2
✓		✓	49.6
✓	✓	✓	49.5

图2. CANet中的消融实验

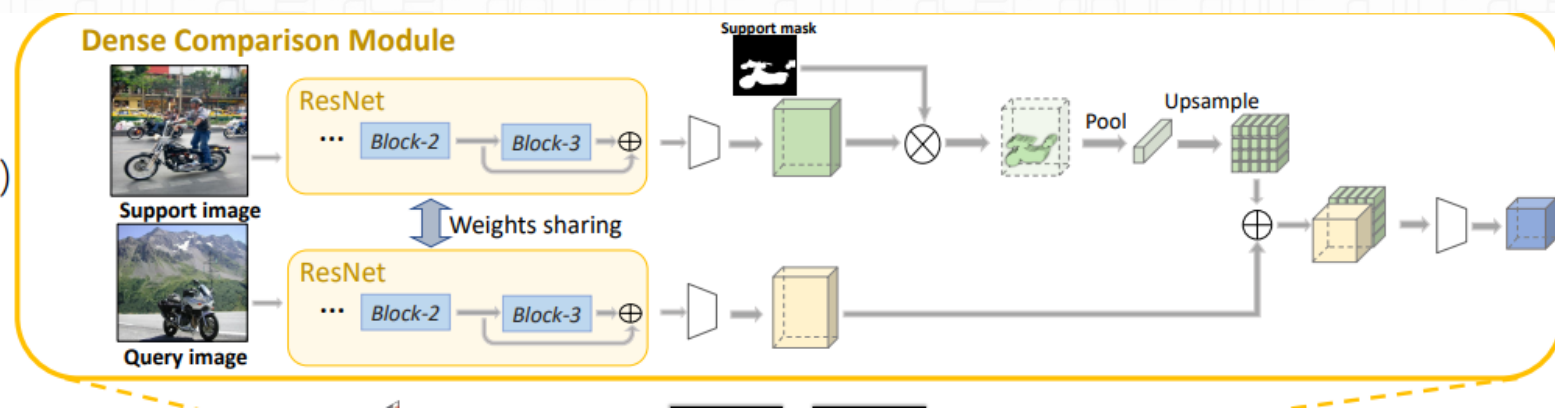


图3. CANet中的DCM模块

高层特征中包含的语义信息比中间层特征更加的class-specific, 所以高层特征有可能对未见类的泛化能力产生影响

高层特征直接提供语义信息, 会导致对训练类别的偏爱。

问题背景

小样本语义分割所面临的问题：

②支持集和查询集的目标空间不一致

查询图像中的目标可能比支持样本大得多或者小得多。

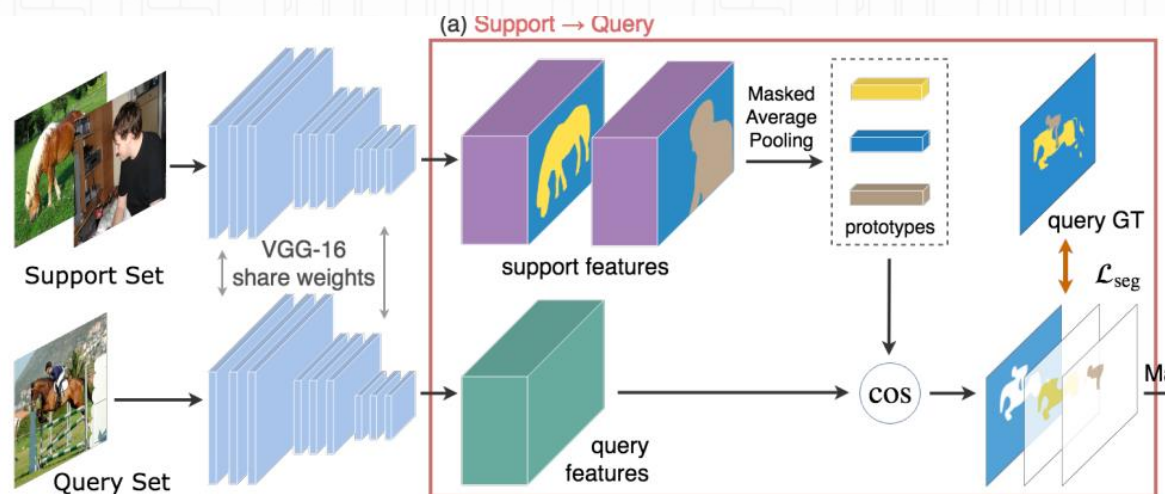
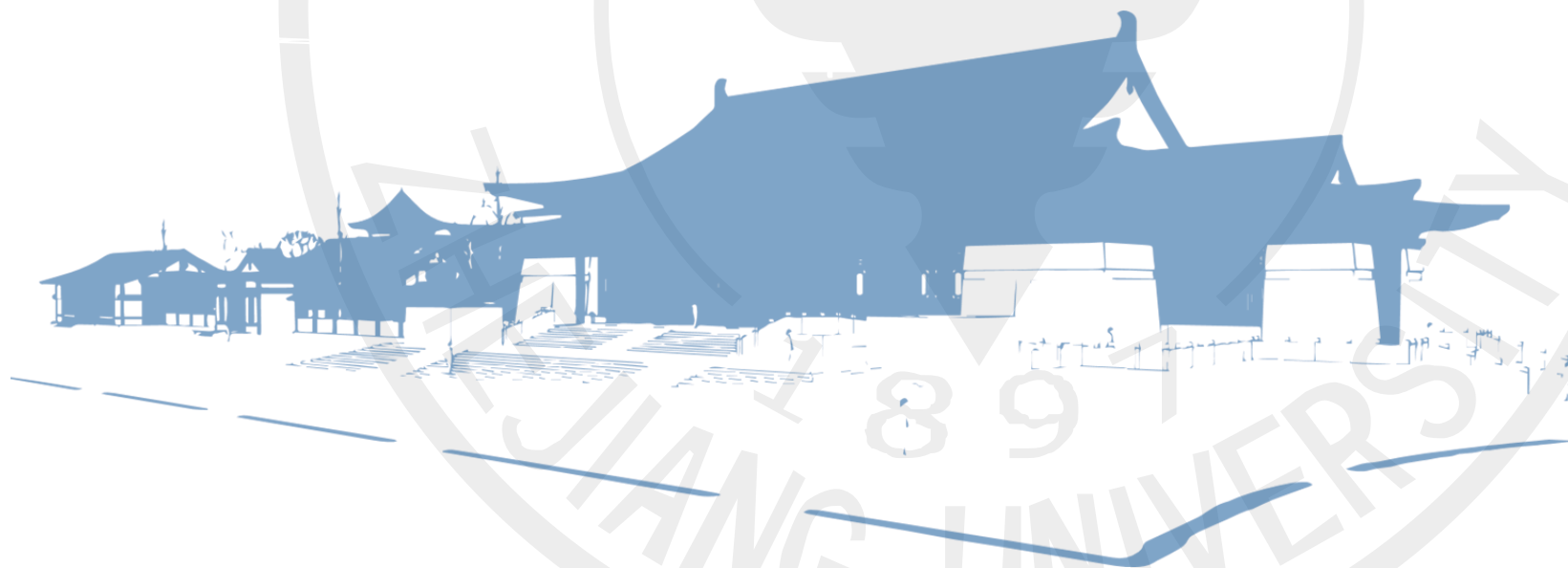


图4. PANet中所采用的全局平均池化方法来提原型

02

论文创新点

Prior Mask生成和FEM模块



论文创新点

针对**高层特征误用导致泛化损失**的问题：先验掩码

“既要，又要”：既要使用高层特征，又不希望它对模型的泛化能力产生影响。

- 利用查询和支持图像的高层特征来生成模型的先验值
- 高层特征从预训练的CNN中得到，生成先验的过程无需训练，模型不会丧失对未见类的泛化能力。
- 提高了预测精度，保持了高泛化性。

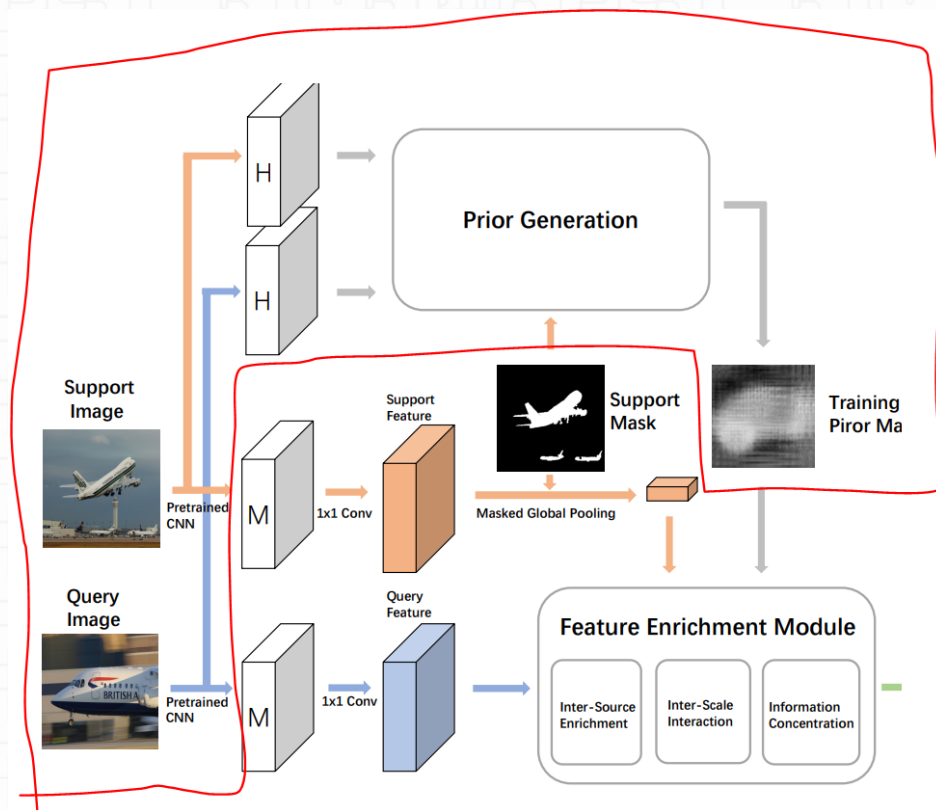


图5. 红色方框部分为文中提出的先验掩码生成模块

论文创新点

针对**支持和查询样本空间不一致**的问题：FEM方法

- 水平的交互查询特征和支持特征和先验掩码
- 垂直利用层次关系，从精细特征中提取必要的信息丰富粗特征图
- 水平和垂直优化之后，收集不同尺度的特征，形成新的查询特征。

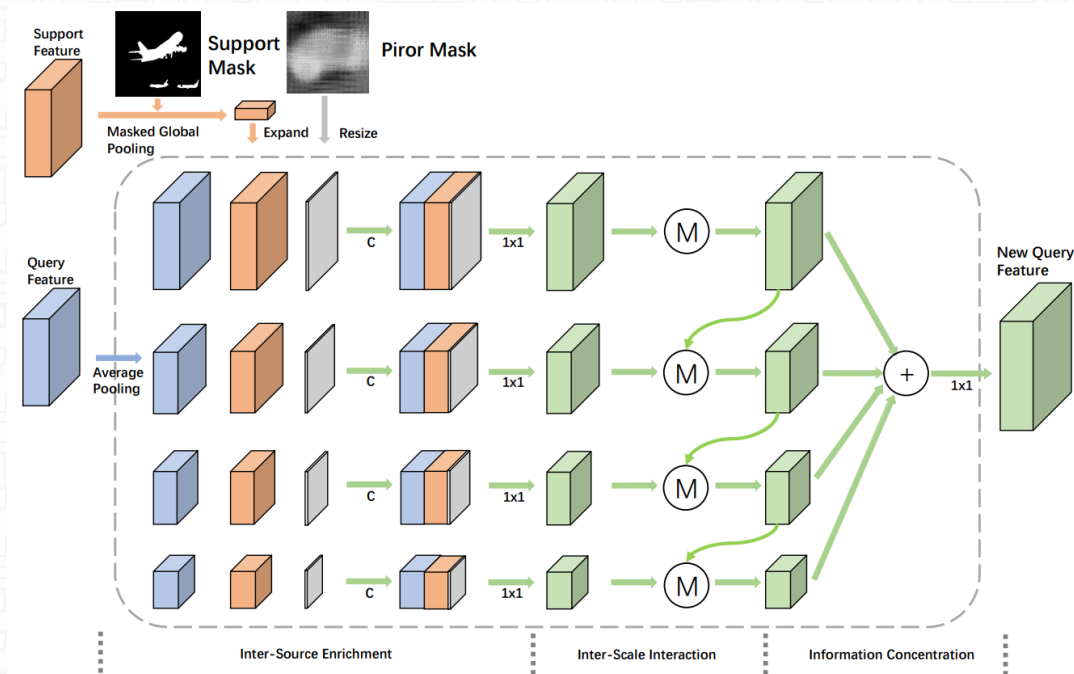


图6. FEM特征富集模块

03

网络框架



网络框架

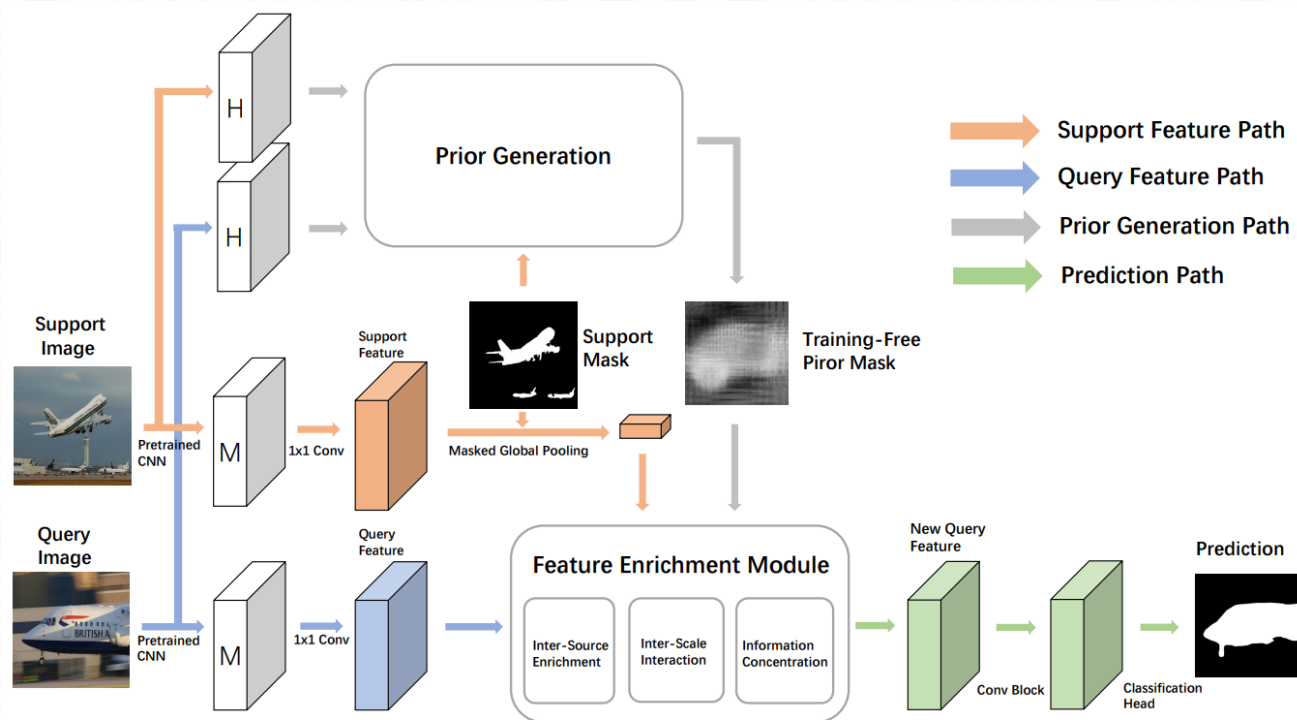


图7. PFENet网络框架

网络框架

Prior Mask生成过程:

X_Q 和 X_S 是Q和S的高层特征, Y_Q 是要生成的先验掩码。

$$\cos(x_q, x_s) = \frac{x_q^T \cdot x_s}{\|x_q\| \|x_s\|} \quad q, s \in \{1, 2, \dots, hw\}$$

$$C_q = \max_{s \in \{1, 2, \dots, hw\}} (\cos x_q, x_s)$$

$$C_Q = [C_1, C_2, \dots, C_{hw}] \in R^{hw \times 1}$$

Reshape $C_Q^{hw \times 1}$ to $Y_Q^{h \times w \times 1}$

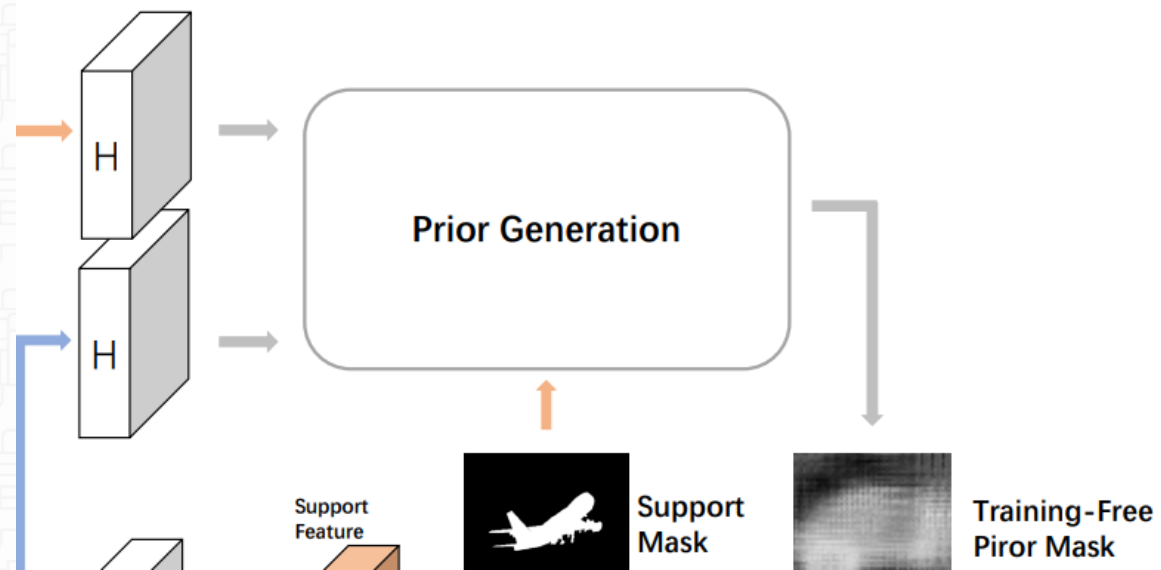


图8. 先验掩码生成示意图

网络框架

FEM: Inter-Source Enrichment

$B = [B^1, B^2, \dots, B^n]$ 表示平均池化的不同大小

$$X_Q^{FEM} = [X_Q^1, X_Q^2, \dots, X_Q^n], X_Q^i \in R^{B^i \times B^i \times C}$$

$X_S \in R^{1 \times 1 \times C}$ expand to

$$X_S^{FEM} = [X_S^1, X_S^2, \dots, X_S^n], X_S^i \in R^{B^i \times B^i \times C}$$

$Y_Q \in R^{h \times w \times 1}$ resize to

$$Y_Q^{FEM} = [Y_Q^1, Y_Q^2, \dots, Y_Q^n], Y_Q^i \in R^{B^i \times B^i \times C}$$

$$X_{Q,m}^i = F_{1 \times 1}(X_Q^i \oplus X_S^i \oplus Y_Q^i) \in R^{B^i \times B^i \times C}$$

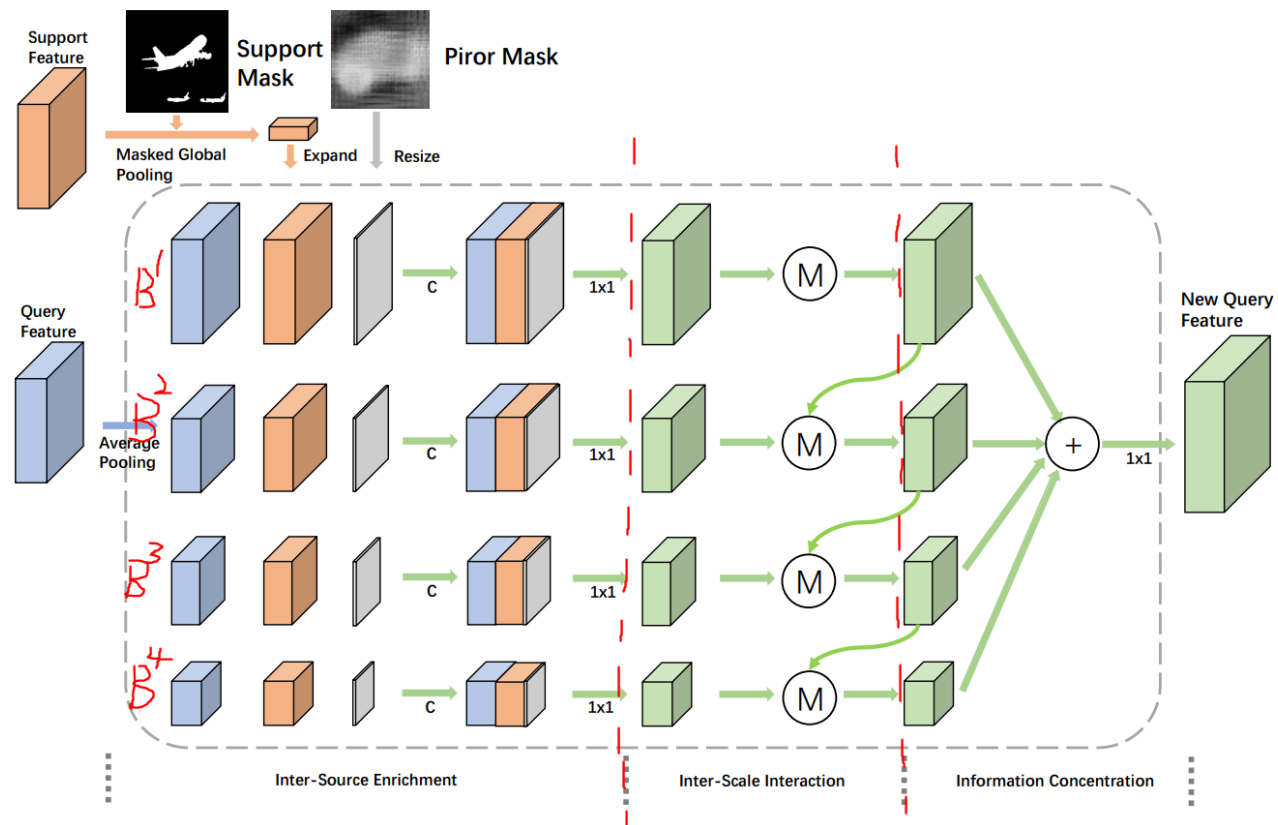


图8. FEM模块示意图

网络框架

FEM: Inter-Scale Interaction

$$X_{Q,new}^i = M(X_{Q,m}^i, X_{Q,m}^{i-1}) \quad i > 1$$

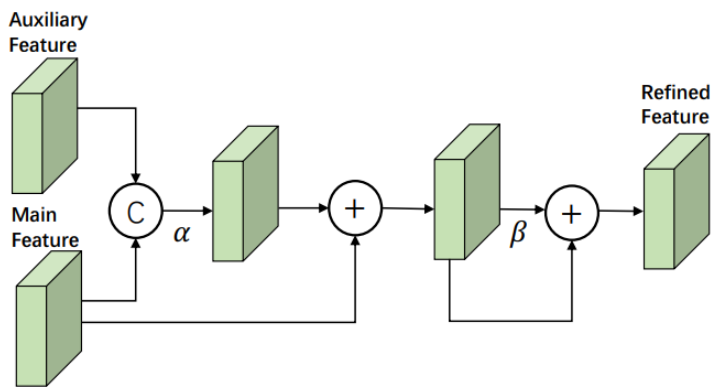


图9. FEM模块中的M示意图

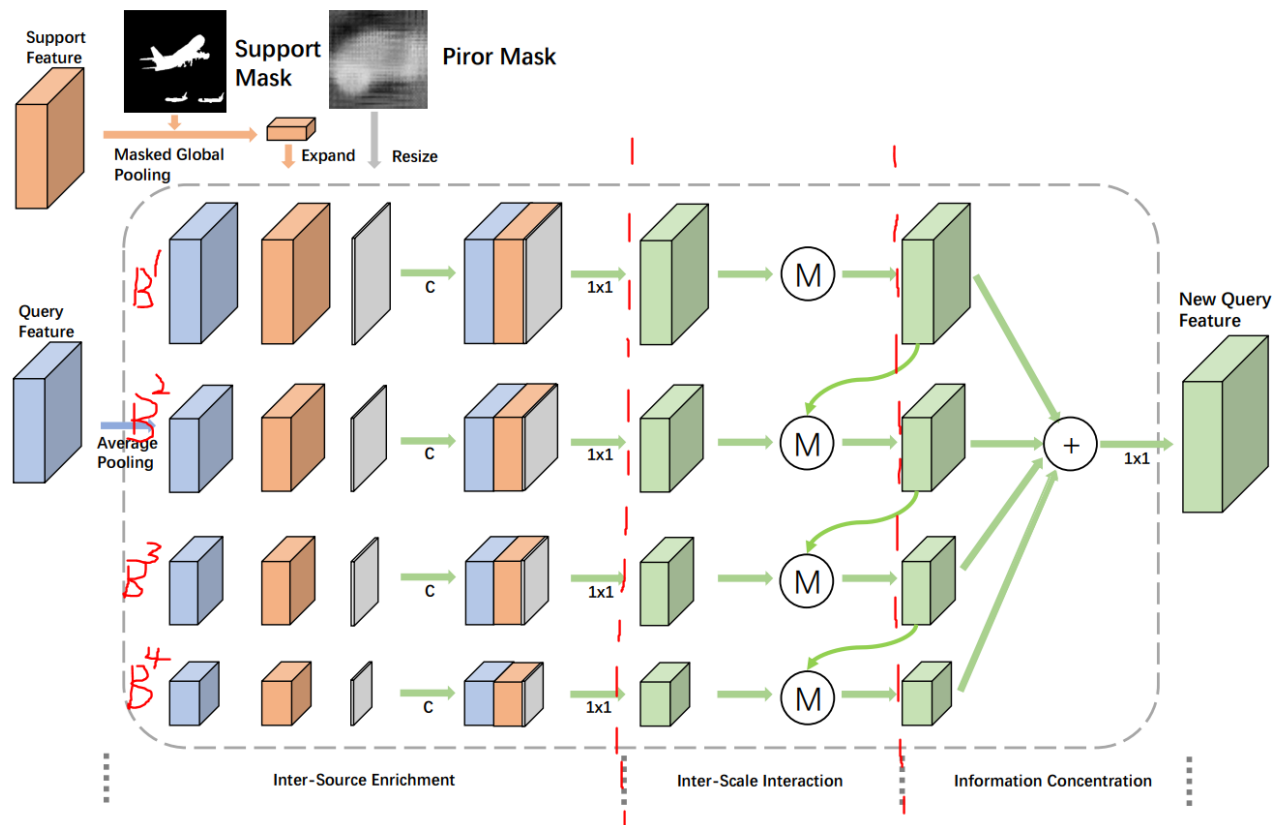


图8. FEM模块示意图

网络框架

FEM: Inter-Scale Interaction

对 $X_{Q,new}^i$ 进行插值, 得到 $X_{Q,new}^i \in R^{h \times w \times c}$

最后生成:

$$X_{Q,new} = F_{1 \times 1}(X_{Q,new}^1 \oplus X_{Q,new}^2 \oplus \dots \oplus X_{Q,new}^n)$$

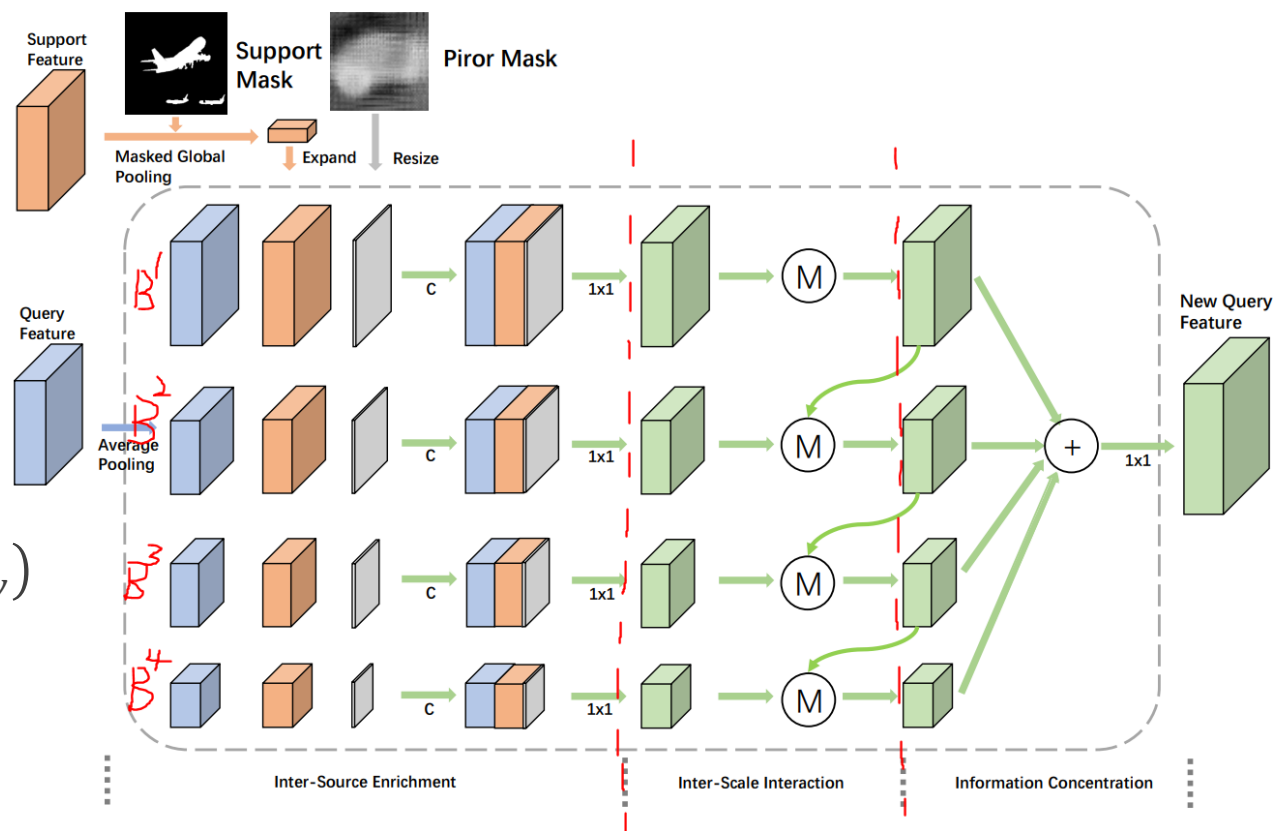


图8. FEM模块示意图

04

实验分析

基于PASCAL-5i数据集



实验分析

消融实验：FEM与其他方法对比

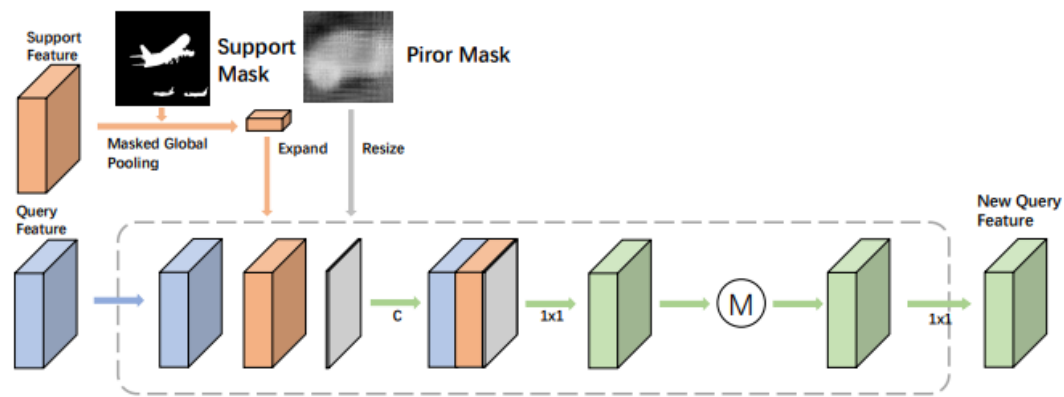


图10. baseline的结构

Methods	1-Shot					5-Shot				
	Fold-0	Fold-1	Fold-2	Fold-3	Mean	Fold-0	Fold-1	Fold-2	Fold-3	Mean
{60} (Baseline)	54.3	67.3	53.3	50.4	<u>56.3</u>	57.1	68.0	53.8	52.9	58.0
{60} + PPM [60]	55.4	68.4	53.2	51.4	<u>57.1</u>	58.3	68.9	53.5	50.8	<u>57.9</u>
{60} + ASPP [4]	57.6	68.4	52.8	49.0	<u>56.9</u>	59.5	69.3	52.6	50.7	<u>58.0</u>
{60, 6, 3, 2, 1}	58.8	68.0	54.1	51.2	58.0	59.8	68.4	53.8	52.1	58.5
{60, 30}	55.3	67.8	54.7	51.2	57.3	58.4	68.7	54.5	53.1	58.7
{60, 30, 15}	56.6	68.0	54.6	52.9	58.0	59.0	68.7	55.0	54.0	59.2
{60, 30, 15, 8}	59.4	68.9	54.7	53.6	<u>59.2</u>	61.5	69.5	55.4	55.3	<u>60.4</u>
{60, 30, 15, 8, 4}	58.7	68.5	54.1	54.5	58.9	60.3	69.3	54.9	56.4	60.2
{60, 30, 15, 8}-WO	57.9	67.4	53.7	53.6	58.2	60.5	68.0	54.2	53.8	59.1

图11. 在PASCAL-5i上消融实验, backboe为ResNet-50。

实验分析

消融实验：Prior Mask的生成方式对比

Methods	1-Shot					5-Shot				
	Fold-0	Fold-1	Fold-2	Fold-3	Mean	Fold-0	Fold-1	Fold-2	Fold-3	Mean
Baseline	49.4	64.6	53.3	46.0	53.3	51.5	65.5	52.5	47.0	54.1
Baseline + Prior _{LM}	50.3	54.5	53.0	46.2	53.5	51.9	65.7	52.9	47.2	54.4
Baseline + Prior _{LH}	37.8	60.8	53.5	43.4	48.9	42.5	64.2	57.8	47.6	53.0
Baseline + Prior _{FM}	51.2	64.4	53.9	45.7	53.8	52.8	65.1	53.2	47.5	54.7
Baseline + Prior _{FH}	53.5	65.6	53.6	48.8	55.4	55.7	66.4	53.8	49.8	56.4
Baseline + Prior-A _{FH}	52.2	65.4	54.5	48.5	55.1	54.8	66.0	54.3	50.2	56.3
Baseline + Prior-P _{FH}	52.4	65.8	53.1	47.6	54.7	54.9	67.0	53.5	48.8	56.1
Baseline + Prior-FW _{LM}	50.6	64.9	52.4	42.9	52.7	53.4	65.5	51.7	43.2	53.5
Baseline + Prior-FW _{LH}	37.5	60.3	54.8	43.9	49.1	44.2	62.8	58.5	47.0	53.1
Baseline + Prior-FW _{FM}	50.6	64.7	54.4	47.0	54.2	52.5	65.4	53.7	47.8	54.9
Baseline + Prior-FW _{FH}	51.0	65.1	53.9	48.8	54.7	52.7	66.1	53.8	50.4	55.8

图12. PASCAL-5i上的结果，backbone为VGG16.

实验分析

TABLE 1
Class mIoU results on four folds of PASCAL-5ⁱ. **Params**: number of learnable parameters.

Methods	1-Shot					5-Shot					Params
	Fold-0	Fold-1	Fold-2	Fold-3	Mean	Fold-0	Fold-1	Fold-2	Fold-3	Mean	
VGG-16 Backbone											
OSLSM ₂₀₁₇ [33]	33.6	55.3	40.9	33.5	40.8	35.9	58.1	42.7	39.1	44.0	276.7M
co-FCN ₂₀₁₈ [29]	36.7	50.6	44.9	32.4	41.1	37.5	50.0	44.1	33.9	41.4	34.2M
SG-One ₂₀₁₈ [58]	40.2	58.4	48.4	38.4	46.3	41.9	58.6	48.6	39.4	47.1	19.0M
AMP ₂₀₁₉ [35]	41.9	50.2	46.7	34.7	43.4	41.8	55.5	50.3	39.9	46.9	34.7M
PANet ₂₀₁₉ [45]	42.3	58.0	51.1	41.2	48.1	51.8	64.6	59.8	46.5	55.7	14.7M
FWBF ₂₀₁₉ [28]	47.0	59.6	52.6	48.3	51.9	50.9	62.9	56.5	50.1	55.1	-
Ours	56.9	68.2	54.4	52.4	<u>58.0</u>	59.0	69.1	54.8	52.9	<u>59.0</u>	10.4M
ResNet-50 Backbone											
CANet ₂₀₁₉ [54]	52.5	65.9	51.3	51.9	55.4	55.5	67.8	51.9	53.2	57.1	19.0M
PGNet ₂₀₁₉ [53]	56.0	66.9	50.6	50.4	56.0	54.9	67.4	51.8	53.0	56.8	17.2M
Ours	61.7	69.5	55.4	56.3	<u>60.8</u>	63.1	70.7	55.8	57.9	<u>61.9</u>	10.8M
ResNet-101 Backbone											
FWBF ₂₀₁₉ [28]	51.3	64.5	56.7	52.2	56.2	54.8	67.4	62.2	55.3	59.9	-
Ours	60.5	69.4	54.4	55.9	<u>60.1</u>	62.8	70.4	54.9	57.6	<u>61.4</u>	10.8M

图13. 在PASCAL-5ⁱ数据集上的结果

05

总结



总结

PFENet通过无训练的先验掩码生成方法，不仅保留了泛化能力，也提高了模型性能。达到了“既要、又要的效果”。

通过特征丰富模块增强了查询特征，克服空间不一致性。达到了新的SOTA。



THANKS!