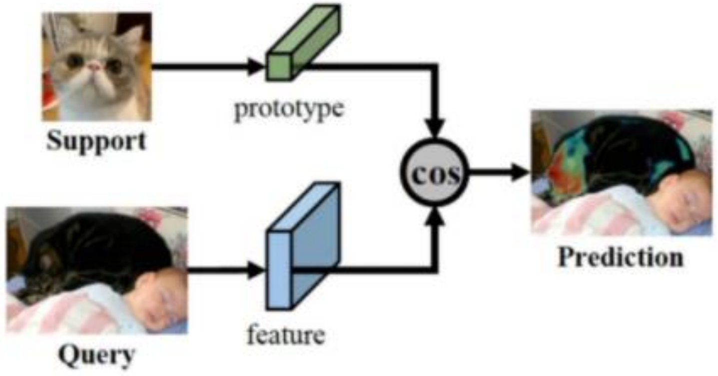
Self-Support Few-Shot Semantic Segmentation，ECCV 2022

自支持匹配的思想 用于FSS

Motivation

现有的few-shot方法大多基于查询-支持匹配的框架。
支持的类是少量的，覆盖范围有限 such as：
　即：S和Q中的同类物体是有可能存在很大的外观差异性。

提出了自支持匹配的策略来解决问题



Motivation

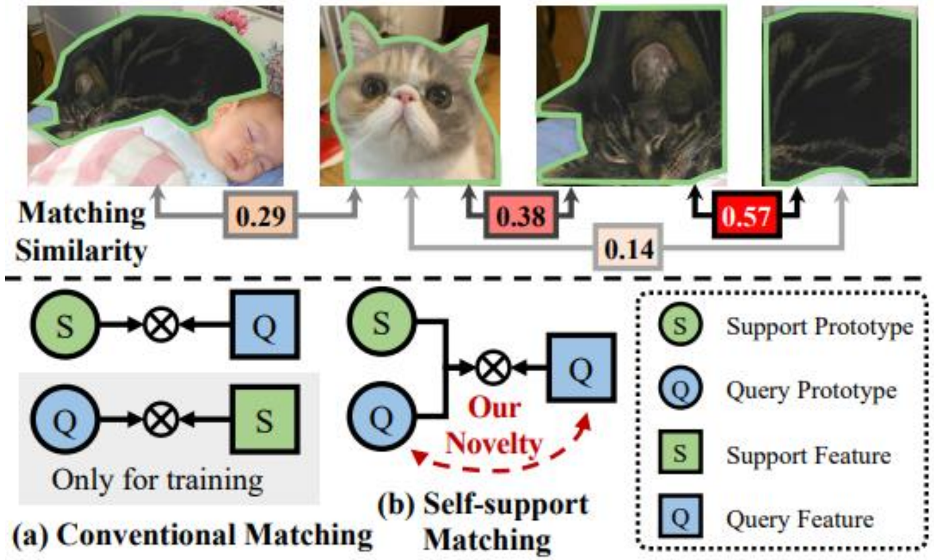
之前的方法:

自适应学习 注意力机制 或是更好的训练方法

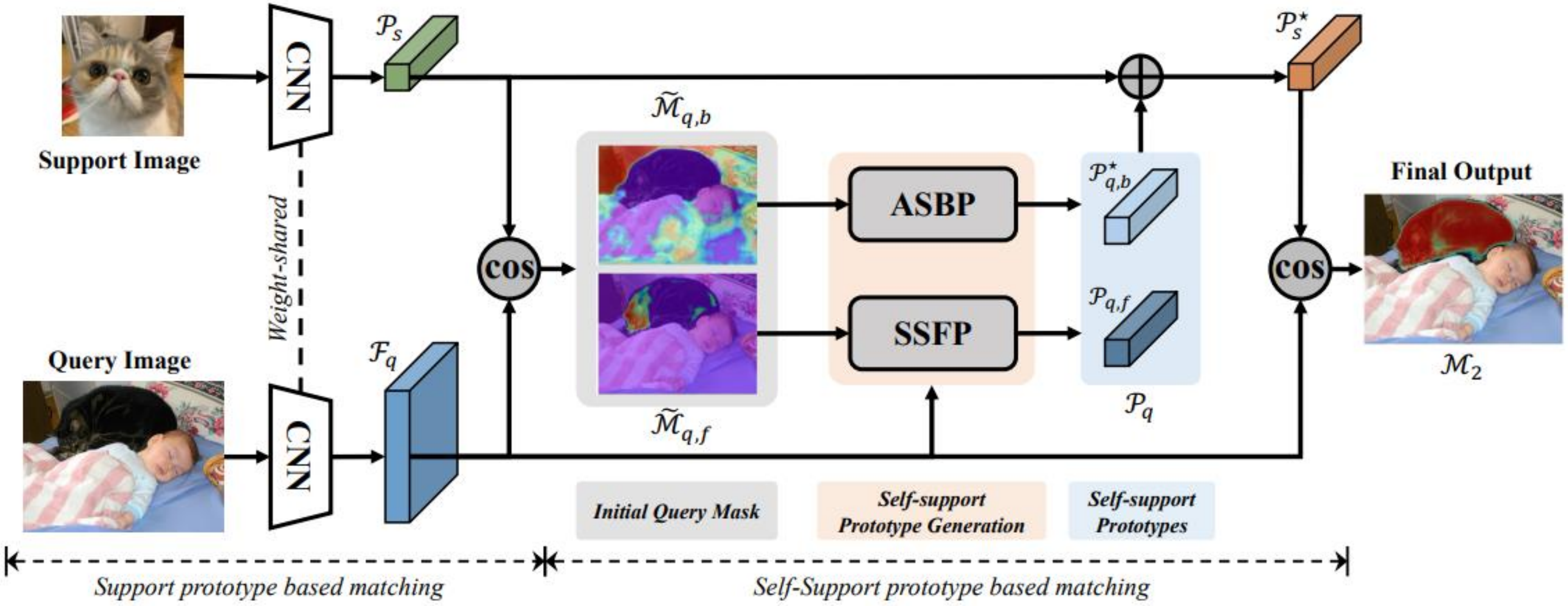
但都没有很好的解决上述问题，因为都局限在用少量的support去分割无穷的query

Solution

想法是利用查询特征原型P_{q}去匹配查询图像特征F_{q}。P_{q}是从query prediction mask 提取对应的查询特征F_{q}得到的。 query prediction mask 是用传统的匹配算法生成的。



Framework



Support Image
CNN
$\mathcal{P}_s$
$\mathcal{P}_s^\star$
$\widetilde{\mathcal{M}}_{q,b}$
Weight-shared
cos
ASBP
SSFP
$\mathcal{P}_{q,b}^\star$
$\mathcal{P}_{q,f}$
$\mathcal{P}_q$
Final Output
$\mathcal{M}_2$
Query Image
$\mathcal{F}_q$
$\widetilde{\mathcal{M}}_{q,f}$
Initial Query Mask
Self-support Prototype Generation
Self-support Prototypes
Support prototype based matching
Self-Support prototype based matching

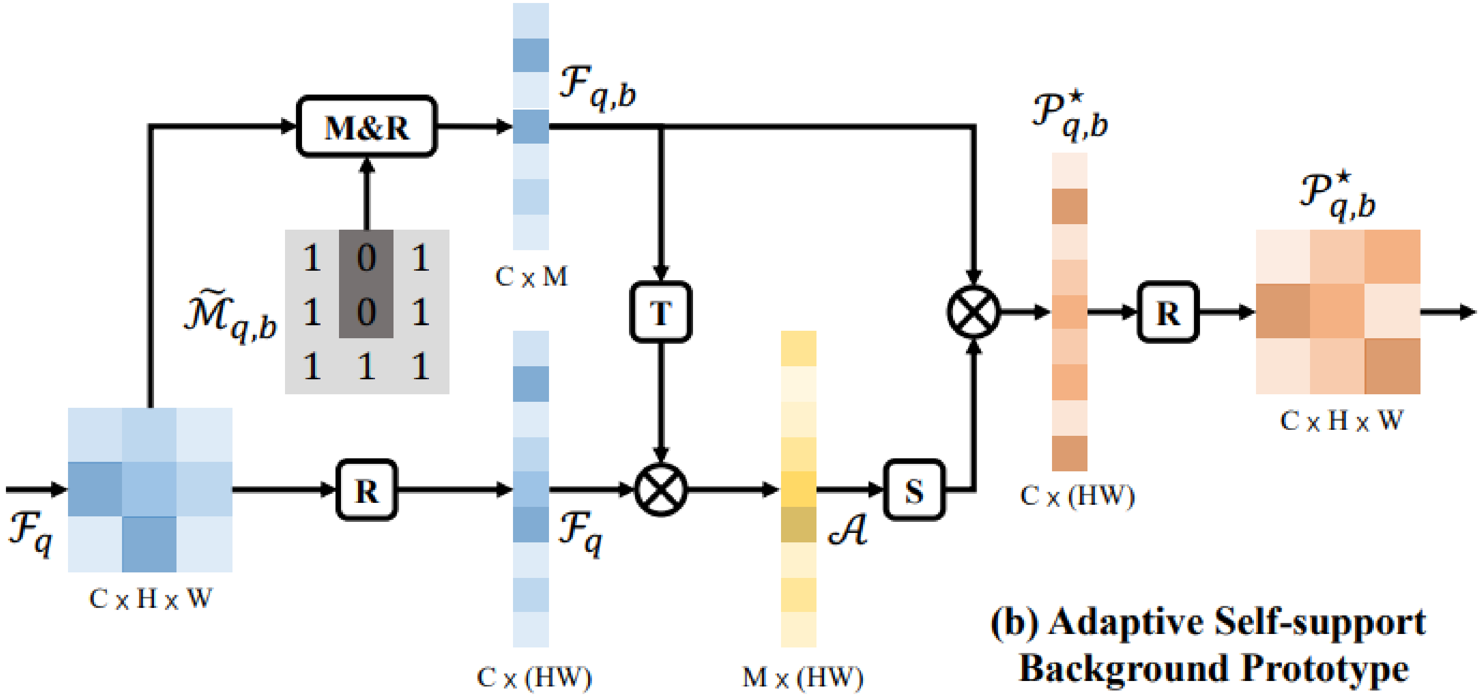
Framework

$$\mathcal{M}_1 = \text{softmax}(\text{cosine}(\mathcal{P}_s, \mathcal{F}_q)),$$

where $\widetilde{\mathcal{M}}_q = \mathbb{1}(\mathcal{M}_1 > \tau)$, and $\mathcal{M}_1$ is the estimated query mask generated by Equation 2, $\mathbb{1}$ is the indicator function. The mask threshold τ is used to control the query feature sampling scope which is set as $\{\tau_{fg} = 0.7, \tau_{bg} = 0.6\}$ for foreground and background query masks respectively. The estimated self-support prototype $\mathcal{P}_q = \{\mathcal{P}_{q,f}, \mathcal{P}_{q,b}\}$ will be utilized to match query features.

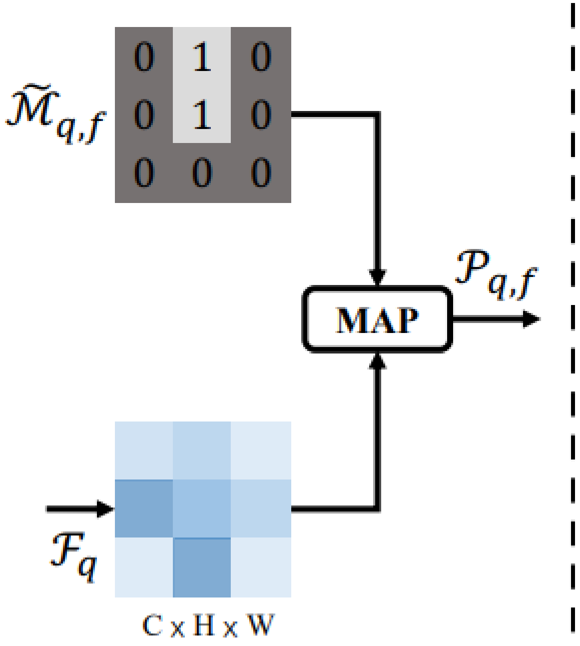
Framework：ASBP

自适应地为每个查询像素生成自支持背景原型。



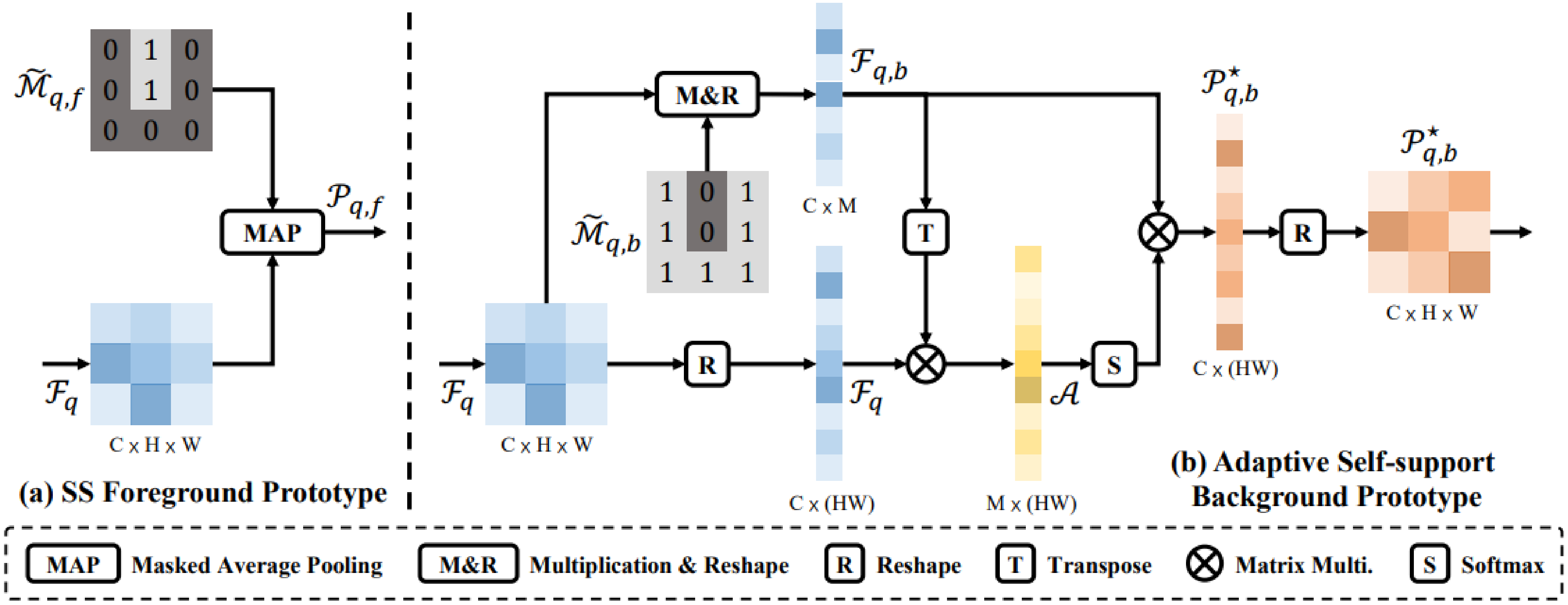
(b) Adaptive Self-support Background Prototype

Framework：SSFP



(a) SS Foreground Prototype

Framework



$\widetilde{M}_{q,f}$
0 1 0
0 1 0
0 0 0
MAP
$\mathcal{P}_{q,f}$
$\mathcal{F}_q$
C x H x W
(a) SS Foreground Prototype
M&R
$\mathcal{F}_{q,b}$
$\widetilde{M}_{q,b}$
1 0 1
1 0 1
1 1 1
C x M
T
$\mathcal{F}_q$
C x H x W
R
$\mathcal{F}_q$
C x (HW)
$\mathcal{A}$
M x (HW)
S
$\mathcal{P}^{\star}_{q,b}$
C x (HW)
R
$\mathcal{P}^{\star}_{q,b}$
C x H x W
(b) Adaptive Self-support Background Prototype
MAP Masked Average Pooling
M&R Multiplication & Reshape
R Reshape
T Transpose
Matrix Multi.
S Softmax

Results

Table 3. Quantitative comparison results on PASCAL-5^i dataset. The **best** and <u>second best</u> results are highlighted with **bold** and <u>underline</u>, respectively.

Method	Backbone	1-shot fold0	1-shot fold1	1-shot fold2	1-shot fold3	1-shot **Mean**	5-shot fold0	5-shot fold1	5-shot fold2	5-shot fold3	5-shot **Mean**	Params
PANet [75]	Res-50	44.0	57.5	50.8	44.0	49.1	55.3	67.2	61.3	53.2	59.3	<u>23.5 M</u>
PPNet [56]	Res-50	48.6	60.6	55.7	46.5	52.8	58.9	68.3	66.8	58.0	63.0	31.5 M
PFENet [73]	Res-50	<u>61.7</u>	69.5	55.4	<u>56.3</u>	60.8	63.1	70.7	55.8	57.9	61.9	34.3 M
CWT [59]	Res-50	56.3	62.0	59.9	47.2	56.4	61.3	68.5	68.5	56.6	63.7	-
HSNet [61]	Res-50	**64.3**	<u>70.7</u>	60.3	**60.5**	**64.0**	**70.3**	**73.2**	67.4	**67.1**	**69.5**	26.1 M
MLC [82]	Res-50	59.2	**71.2**	<u>65.6</u>	52.5	<u>62.1</u>	63.5	71.6	71.2	58.1	66.1	**8.7 M**
SSP (Ours)	Res-50	61.4	67.2	65.4	49.7	60.9	<u>68.0</u>	72.0	<u>74.8</u>	60.2	68.8	**8.7 M**
SSP_{refine}	Res-50	60.5	67.8	**66.4**	51.0	61.4	67.5	<u>72.3</u>	**75.2**	<u>62.1</u>	<u>69.3</u>	**8.7 M**
FWB [62]	Res-101	51.3	64.5	56.7	52.2	56.2	54.8	67.4	62.2	55.3	59.9	<u>43.0 M</u>
PPNet [56]	Res-101	52.7	62.8	57.4	47.7	55.2	60.3	70.0	69.4	60.7	65.1	50.5 M
PFENet [73]	Res-101	60.5	69.4	54.4	55.9	60.1	62.8	70.4	54.9	57.6	61.4	53.4 M
CWT [59]	Res-101	56.9	65.2	61.2	48.8	58.0	62.6	70.2	68.8	57.2	64.7	-
HSNet [61]	Res-101	**67.3**	**72.3**	62.0	**63.1**	**66.2**	**71.8**	74.4	67.0	**68.3**	70.4	45.2 M
MLC [82]	Res-101	60.8	<u>71.3</u>	61.5	<u>56.9</u>	62.6	65.8	74.9	71.4	63.1	68.8	**27.7 M**
SSP (Ours)	Res-101	<u>63.7</u>	70.1	<u>66.7</u>	55.4	64.0	70.3	<u>76.3</u>	<u>77.8</u>	65.5	<u>72.5</u>	**27.7 M**
SSP_{refine}	Res-101	63.2	70.4	**68.5**	56.3	<u>64.6</u>	<u>70.5</u>	**76.4**	**79.0**	<u>66.4</u>	**73.1**	**27.7 M**