Supplementary Material

1. Impact of loss

Table 1. The impact of different losses on COCO-20^i to PASCAL-5^i 5-shot experiments.

COCO-20^i to PASCAL-5^i						
L_{orth}	L_{cont}	split-0	split-1	split-2	split-3	mean
		56.3	57.5	65.4	63.2	60.6
✓		57.6	59.1	67.0	66.4	62.5
	✓	58.4	60.5	69.5	70.1	64.6
✓	✓	57.4	62.2	68.0	74.8	65.6

In Table 1, we analyzed the impact of L_{orth} and L_{cont} to the final result. We observed that both L_{orth} and L_{cont} could improve the segmentation effects from rows 2 and 3, respectively. The advance of L_{orth} achieves 1.9%, which proves the irrelevance among meta-memory can help the model to capture more abundant style information. L_{cont} is also valuable to segmentation, which improves 4% effects. Two losses help the model collect different stylistic features and achieve consistent high-level representation.

2. 5-shot Experiments

In Table 2, we compare our method with other few-shot segmentation network on the 5-shot experiments from COCO-20^i to FSS-1000. Our model achieve 11.3% improve of ASGNet [2], and the performance of HSNet [4] is also stunning, but our model still have 0.6% lead.

Table 3 and Table 4 show the 5-shot transfer effect from COCO-20^i and PASCAL-5^i to SUIM dataset, the domain gap is significantly present in SUIM and source domains. As shown in Table 3, our method achieve better results on all splits, it outperforms the ASGNet [2] of 2.1%.

The segmentation results in Table 4, with SCL [7] achieving 30.8% mIoU compare 26.5% in Table 3, may caused by the gentler domain gap between PASCAL-5^i and SUIM. Our method provides 43.0% mIoU results for PASCAL-5^i to SUIM tasks, which outperforms SCL [7] by 12.2%.

3. Intra-domain segmentation

We also compare our method with state-of-the-art methods on the intra-domain few-shot segmentation. In Table 5, we provide 1-shot and 5-shot semantic segmentation results

Table 2. Cross-domain 5-shot semantic segmentation results on COCO-20^i to FSS-1000 task.

COCO-20^i to FSS-1000						
Backbone	Method	split-0	split-1	split-2	split-3	mean
ResNet50	ASGNet [2] (CVPR21)	72.5	74.6	70.2	71.2	72.1
	HSNet [4] (ICCV21)	81.8	82.3	**82.4**	**84.9**	82.8
	SCL [7] (CVPR21)	82.5	78.3	77.3	74	78.0
	Ours	**84.4**	**83.5**	81.2	84.5	**83.4**

Table 3. Cross-domain 5-shot semantic segmentation results on COCO-20^i to SUIM task.

COCO-20^i to SUIM						
Backbone	Method	split-0	split-1	split-2	split-3	mean
ResNet50	ASGNet [2] (CVPR21)	41.1	42.5	41.4	42.2	41.8
	HSNet [4] (ICCV21)	41.1	42.7	42.7	41.1	41.9
	SCL [7] (CVPR21)	28.1	27.7	23.3	27	26.5
	Ours	**41.2**	**44.5**	**45.9**	**44.1**	**43.9**

Table 4. Cross-domain 5-shot semantic segmentation results on PASCAL-5^i to SUIM task.

PASCAL-5^i to SUIM						
Backbone	Method	split-0	split-1	split-2	split-3	mean
ResNet50	ASGNet [2] (CVPR21)	**43.2**	41.3	39.5	43.2	41.8
	HSNet [4] (ICCV21)	35.3	34.1	34	32.5	33.9
	SCL [7] (CVPR21)	31.3	29.6	31	31.5	30.8
	Ours	42.9	**43.1**	**42.5**	**43.6**	**43.0**

compared with previous methods, where the train and test dataset are sampled from COCO-20^i. We can observe that our model has achieved the best performance in the 1-shot segmentation task, which is 1.4% higher than HSNet [4]. In the 5-shot experiment, the proposed method is also close to the state-of-art approaches, which further indicates that the improvement of our method comes from the reducing domain gap.

4. Qualitative results

In the Figure 1 left part, we visualize the effect of COCO-20^i to PASCAL-5^i. As the figure shows, our network can complete the segmentation task more finely than other methods. For example, in the second row, our network avoids the dilemma of segmenting humans as the bottle category. In the last Figure 1 row, our network can segment more instances.

Table 5. Test the split effect on the original COCO

Backbone	Method	1-shot					5-shot				
		split-0	split-1	split-2	split-3	mean	split-0	split-1	split-2	split-3	mean
ResNet50	RPMMs [6] (ECCV20)	29.5	36.8	29.0	27.0	30.6	33.8	42.0	33.0	33.3	35.5
	CWT [3] (ICCV21)	30.3	36.6	30.5	32.2	32.4	38.5	46.7	39.4	43.2	42.0
	PFENet [5] (TPAMI)	36.5	38.6	34.5	33.8	35.8	36.5	43.3	37.8	38.4	39.0
	RePRI [1] (CVPR21)	31.2	38.1	33.3	33.0	34.0	38.5	46.2	40.0	43.6	42.1
	ASGNet [2] (CVPR21)	-	-	-	-	34.5	-	-	-	-	42.5
	HSNet [4] (ICCV21)	36.3	**43.1**	38.7	38.7	39.2	43.3	**51.3**	**48.2**	**45.0**	**46.9**
	Ours	**38.9**	40.1	**44.0**	**39.5**	**40.6**	**44.7**	49.8	47.9	44.4	46.7

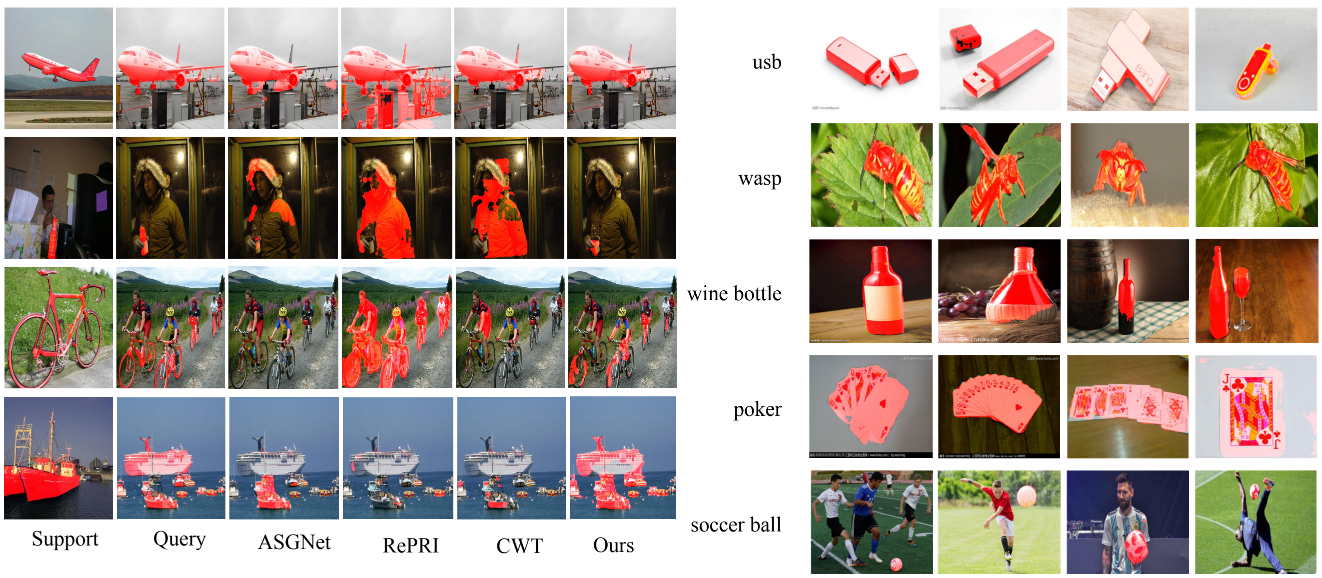

Figure 1. left part shows the 1-shot results on COCO to PASCAL with ResNet50. Right part shows the 1-shot results on COCO to FSS-1000 with ResNet50.

In the Figure 1 right part, we visualize the segmentation effect on FSS-1000. There are many daily objects in FSS-1000, and the data distribution significantly differs from COCO datasets. The figure shows that even facing small objects, our network can still achieve good segmentation results.

References

[1] Malik Boudiaf, Hoel Kervadec, Imtiaz Masud Ziko, Pablo Piantanida, Ismail Ben Ayed, and José Dolz. Few-shot segmentation without meta-learning: A good transductive inference is all you need? In *Proc. CVPR*, pages 13974–13983, 2021. 2

[2] Gen Li, V. Jampani, Laura Sevilla-Lara, Deqing Sun, Jonghyun Kim, and Joongkyu Kim. Adaptive prototype learning and allocation for few-shot segmentation. In *Proc. CVPR*, pages 8330–8339, 2021. 1, 2

[3] Zhihe lu, Sen He, Xiatian Zhu, Li Zhang, Yi-Zhe Song, and Tao Xiang. Simpler is better: Few-shot semantic segmentation with classifier weight transformer. In *Proc. ICCV*, volume abs/2108.03032, 2021. 2

[4] Juhong Min, Dahyun Kang, and Minsu Cho. Hypercorrelation squeeze for few-shot segmentation. In *Proc. ICCV*, volume abs/2104.01538, 2021. 1, 2

[5] Zhuotao Tian, Hengshuang Zhao, Michelle Shu, Zhicheng Yang, Ruiyu Li, and Jiaya Jia. Prior guided feature enrichment network for few-shot segmentation. *IEEE transactions on pattern analysis and machine intelligence*, PP, 2020. 2

[6] Boyu Yang, Chang Liu, Bohao Li, Jianbin Jiao, and Qixiang Ye. Prototype mixture models for few-shot semantic segmentation. In *Proc. ECCV*, volume abs/2008.03898, 2020. 2

[7] Bingfeng Zhang, Jimin Xiao, and Terry Qin. Self-guided and cross-guided learning for few-shot segmentation. In *Proc. CVPR*, pages 8308–8317, 2021. 1