Multiscale Attention-Based Prototypical Network For Few-Shot Semantic Segmentation

Yifei Zhang[1], Désiré Sidibé[2], Olivier Morel[1], Fabrice Meriaudeau[1]*

[1]ERL VIBOT CNRS 6000, ImViA, Université Bourgogne Franche Comté, 71200, Le Creusot, France
[2]Université Paris-Saclay, Univ Evry, IBISC, 91020, Evry, France
**Yifei.Zhang@u-bourgogne.fr*

ICPR 2020, Milan, Italy

Summary

Introduction and motivation

Few-shot semantic segmentation aims at generalizing the segmentation ability of the model to new categories given only a few labeled samples.

- Rich contextual information of labeled support images
- The integration of multiple similarity-guided probability maps by attention mechanism

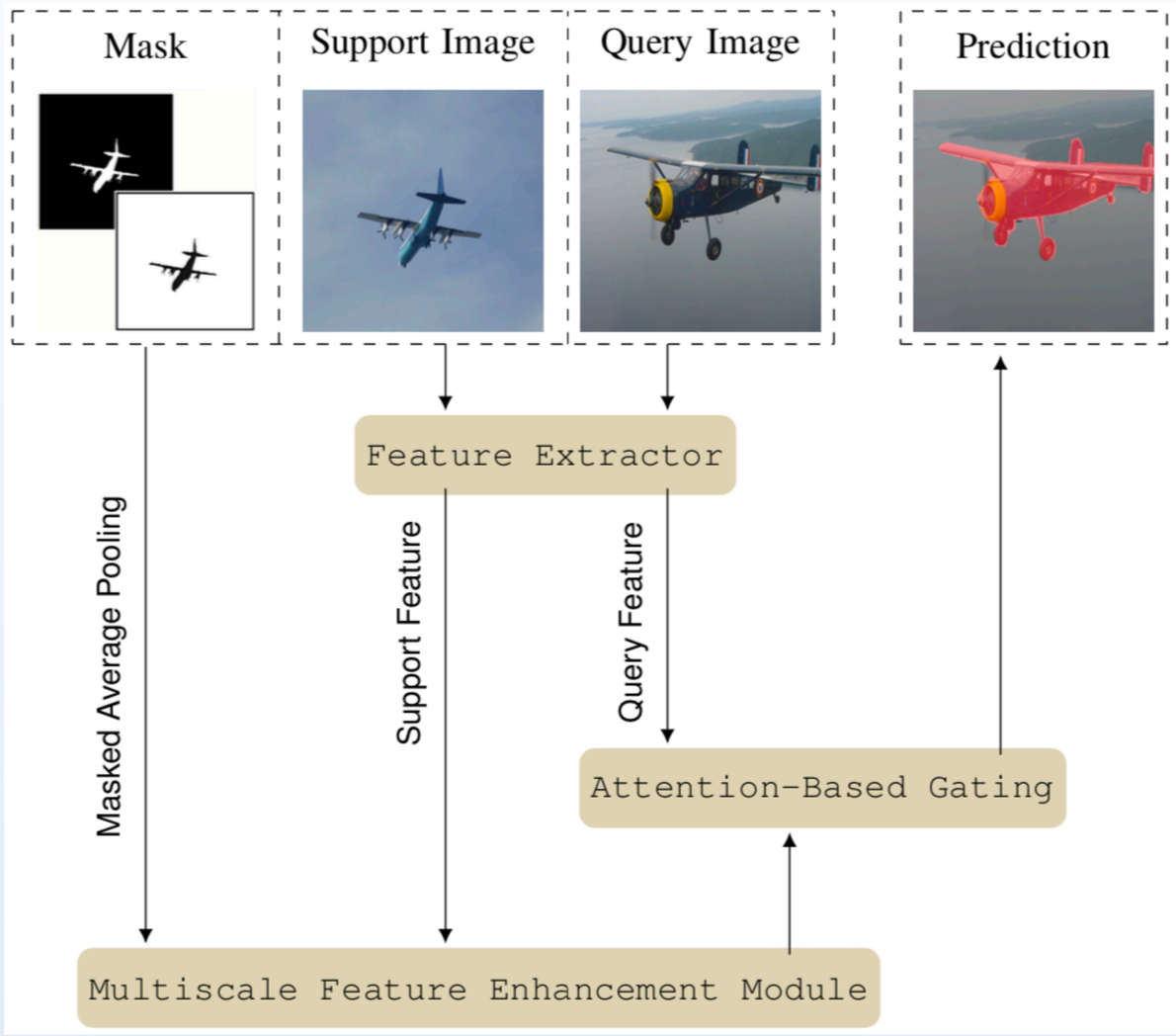

Figure 1: An overview of the proposed method (MAPnet).

Summary

Proposed model

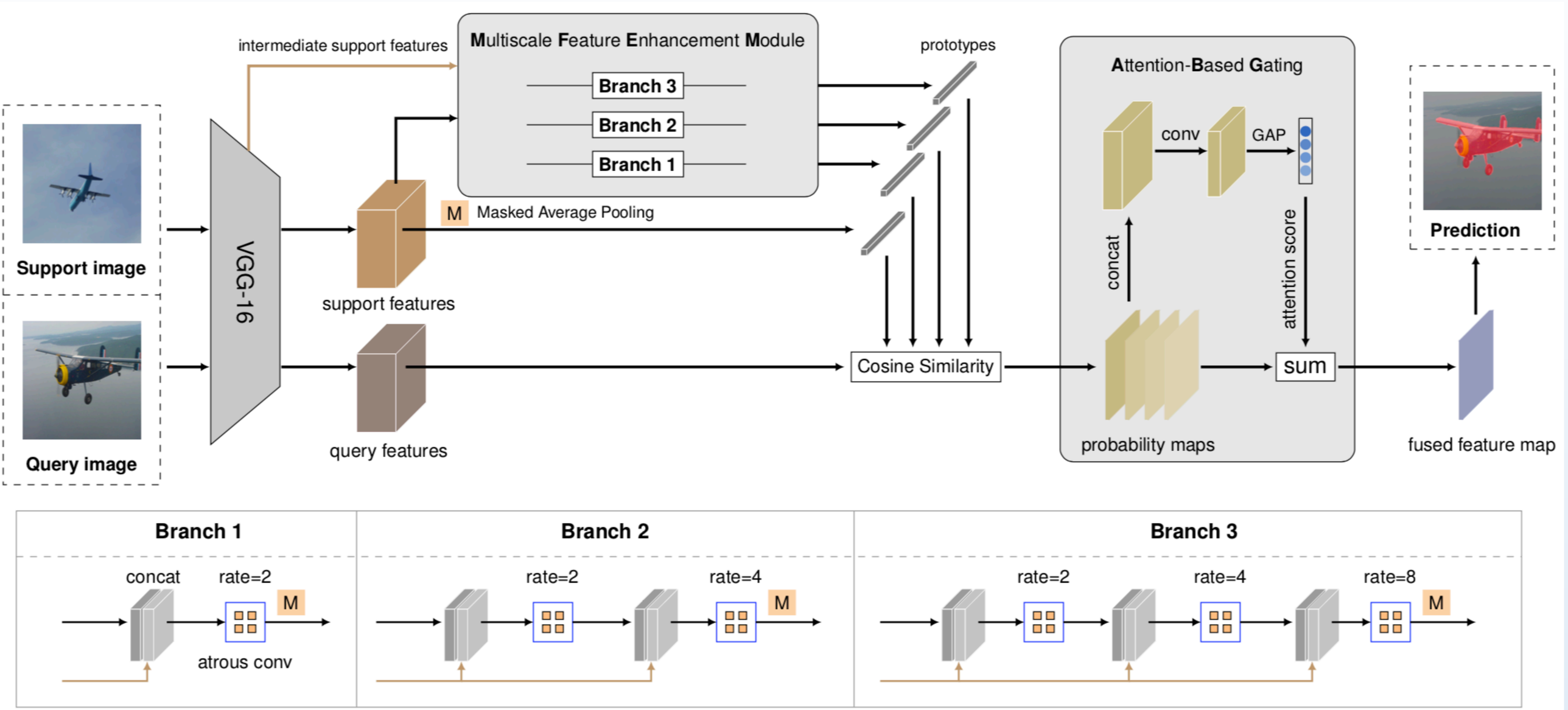

Figure 2: Illustration of the proposed method (MAPnet)

Summary

Experiments

Methods	1-shot					5-shot				
	Pascal-5^0	Pascal-5^1	Pascal-5^2	Pascal-5^3	Mean	Pascal-5^0	Pascal-5^1	Pascal-5^2	Pascal-5^3	Mean
1-NN	25.3	44.9	41.7	18.4	32.6	34.5	53.0	46.9	25.6	40.0
LogReg	26.9	42.9	37.1	18.4	31.4	35.9	51.6	44.5	25.6	39.3
OSLSM [23]	33.6	55.3	40.9	33.5	40.8	35.9	58.1	42.7	39.1	43.9
co-FCN [24]	36.7	50.6	44.9	32.4	41.1	37.5	50.0	44.1	33.9	41.4
SG-One [29]	40.2	**58.4**	48.4	38.4	46.3	41.9	58.6	48.6	39.4	47.1
PANet [30]	42.3	58.0	**51.1**	41.2	48.1	**51.8**	64.6	**59.8**	46.5	55.7
MAPnet	**42.9**	58.3	48.8	**42.6**	**48.2**	51.6	**65.1**	58.4	**48.8**	**56.0**

Table 1: Results of 1-way 1-shot and 1-way 5-shot semantic segmentation on PASCAL-5^i using mean-IoU(%) metric.

Mehtods	1-shot	5-shot	Δ
co-FCN [24]	60.1	60.2	0.1
OSLSM [23]	61.3	61.5	0.2
MDL [26]	63.2	63.7	0.5
PL+SEG [25]	61.2	62.3	1.1
AMP-2+FT [27]	62.2	63.8	1.6
SG-One [29]	63.1	65.9	2.8
PANet [30]	66.5	70.7	4.2
MAPnet	**66.7**	**71.8**	**5.1**

Table 2: Results of 1-way 1-shot and 1-way 5-shot semantic segmentation using binary-IoU(%) metric. Δ denotes the difference between 1-shot and 5-shot.

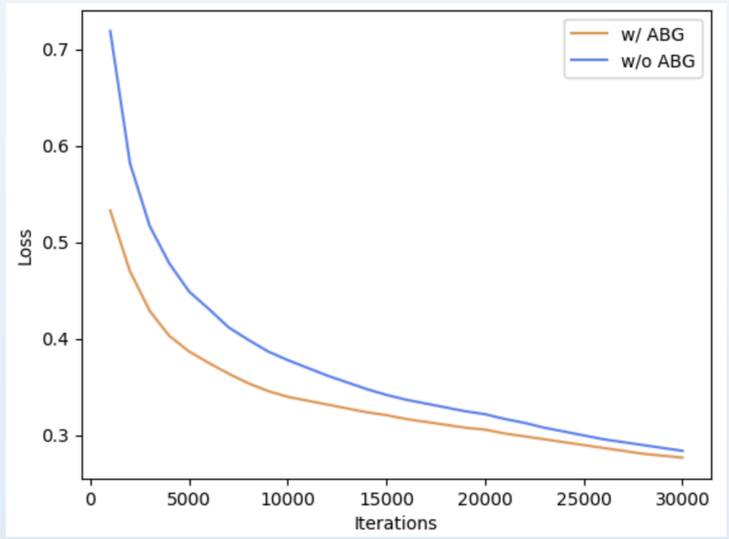

Figure 3: Training loss of models with and without attention-based gating (ABG) for 1-way 1-shot segmentation.

Experiments

PASCAL-5^i dataset

Dataset	Test classes
Pascal-5^0	aeroplane, bicycle, bird, boat, bottle
Pascal-5^1	bus, car, cat, chair, cow
Pascal-5^2	diningtable, dog, horse, motorbike, person
Pascal-5^3	potted plant, sheep, sofa, train, tv/monitor

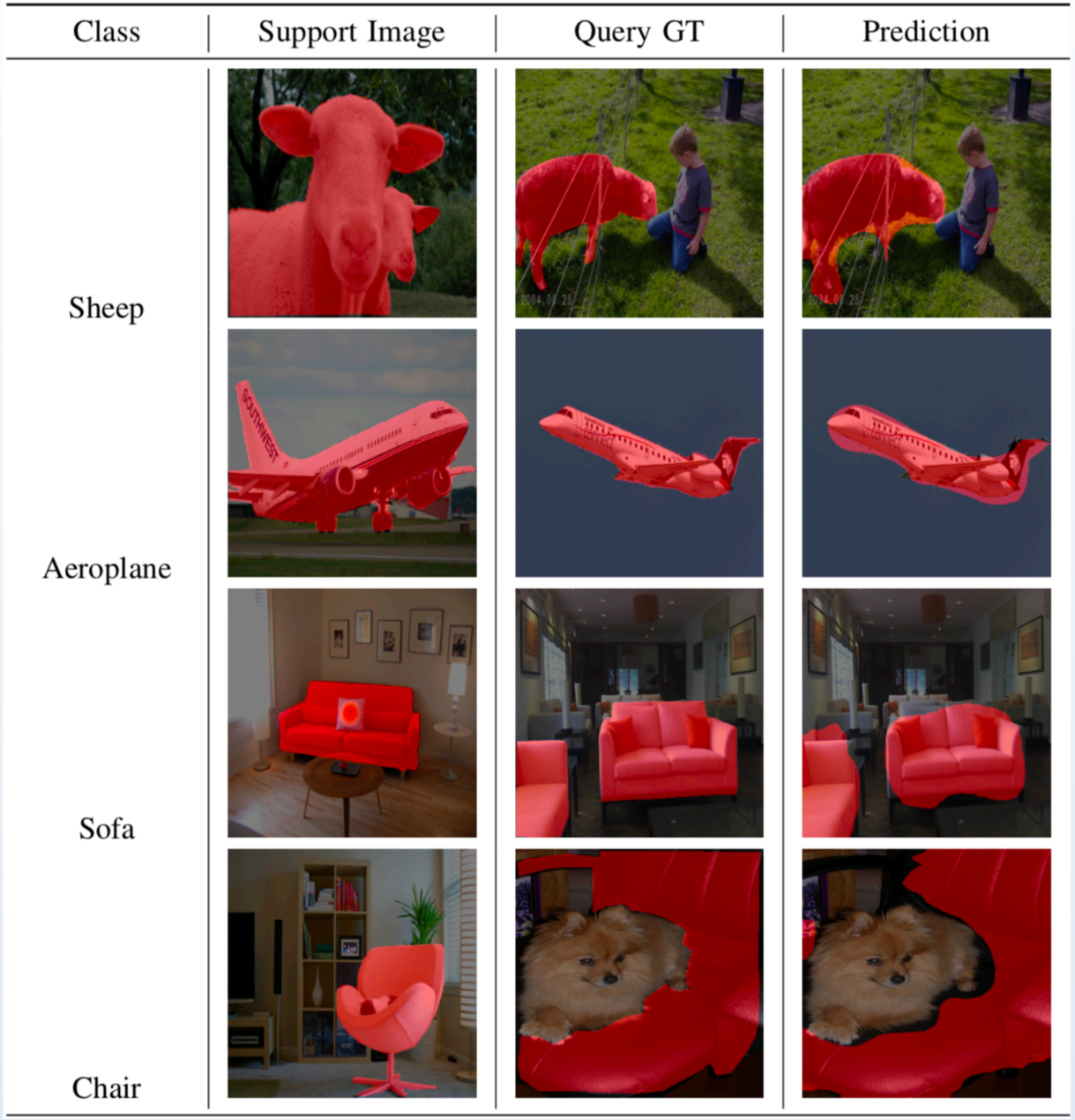

Figure 4: Qualitative results of our method for 1-way 1-shot on PASCAL-5^i.

Experiments

Test with weak annotations

Methods	1-shot			5-shot		
	Dense	Scribble	Bbox	Dense	Scribble	Bbox
PANet [30]	48.1	44.8	45.1	55.7	54.6	52.8
MAPnet	48.2	44.1	45.7	56.0	53.5	53.7

Table 3: Evaluation results of using different types of annotations in mean-IoU(%) metric.

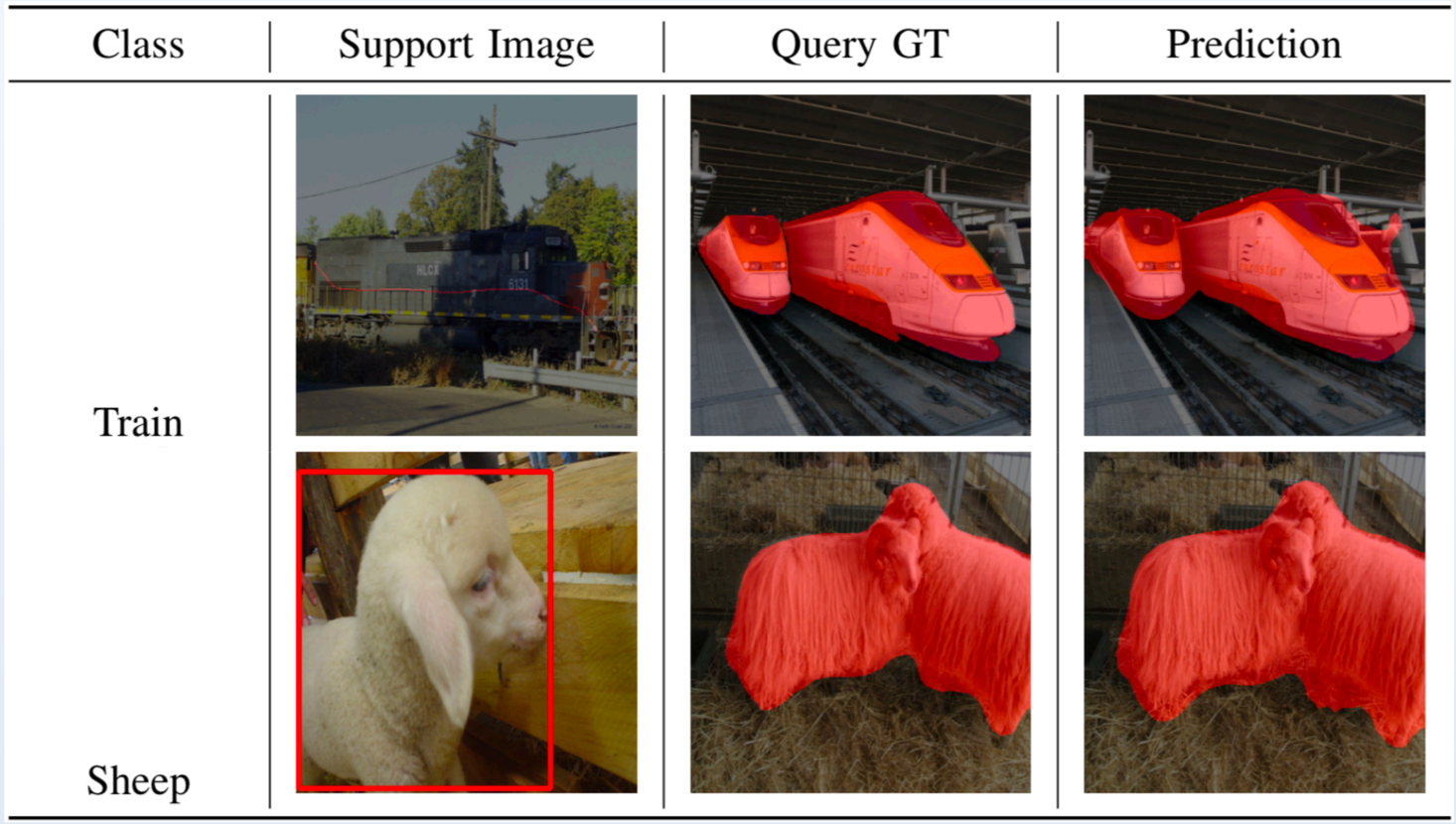

Figure 6: Qualitative results of our model using scribble and bounding box annotations for 1-way 5-shot setting.

Summary

Conclusion and future work

Advantages:

- The proposed method provides effective semantic guidance on the query feature and adaptive information integration for an optimal pixel-wise prediction.
- Experiments on PASCAL-5^i dataset show that our method achieves a comparable accuracy with the state-of-the-art and faster convergence.

Disadvantages:

- The distinction of the objects with similar characteristics, especially when these objects are placed in an overlapping manner.
- Limited capacity to recognize irregular objects and their boundary delineation.

Future work:

- The integration of multimodal information in few-shot semantic segmentation

Thank you!