Novel Class Discovery in Semantic Segmentation
(Supplementary Material)

Yuyang Zhao[1] Zhun Zhong[2] Nicu Sebe[2] Gim Hee Lee[1]
[1] Department of Computer Science, National University of Singapore
[2] Department of Information Engineering and Computer Science, University of Trento
Project: https://ncdss.github.io

A. Performance of All Classes

Despite the focus on novel classes as previous works [1, 6], our NCDSS setting still maintains the ability of segmenting base classes. As shown in Tab. 1, our EUMS achieves nearly 70% base mIoU on Fold0 and Fold3, and more than 60% base mIoU on Fold1 and Fold2. However, our performance is not as competitive as the base model of stage one despite the use of fully-annotated base data. There are two reasons for this limitation. First, the model is required to segment all of base and novel classes together, increasing the task difficulty. Second, there exist unlabeled base classes in the novel images and the generated pseudo-labels of the base classes are also not completely accurate. Thus, bad cases are introduced into the base classes. The above two factors limit the base class performance. How to maintain the high base performance while discovering novel classes in semantic segmentation deserves further explorations in the future.

Fold	mIoU Base	Novel	All
PASCAL-5^0	69.28	69.79	69.40
PASCAL-5^1	66.95	60.11	65.32
PASCAL-5^2	62.87	56.28	61.30
PASCAL-5^3	69.83	50.18	65.15

Table 1. Performance of all classes.

B. Comparison with Related Settings

We further compare several methods under related settings in Tab. 2 on PASCAL-5^i. Our method clearly outperforms the unsupervised learning method on all the folds. Interestingly, our method performs higher on Fold0 when compared with the methods using image/pixel-level labels. However, these weak&few-shot methods generally give better results on the other folds. Please note that directly comparing our method with these weak&few-shot methods is not exactly fair.

Method	Setting	Fold0	Fold1	Fold2	Fold3
PFENet [4]	1-Shot	61.7	69.5	55.4	56.3
ASGNet [2]	1-Shot	58.8	67.9	56.8	53.7
PFENet [4]	5-Shot	63.1	70.7	55.8	57.9
ASGNet [2]	5-Shot	63.7	70.6	64.2	57.4
CAM+RETAB [7]	Weak-Shot	69.2	**76.1**	72.0	58.5
SEAM+RETAB [7]	Weak-Shot	65.4	74.5	**73.0**	**58.9**
MaskContrast [5]	Unsupervised	55.3	38.9	35.6	37.0
EUMS	NCDSS	**69.8**	60.1	56.3	50.2

Table 2. Comparison with peer methods.

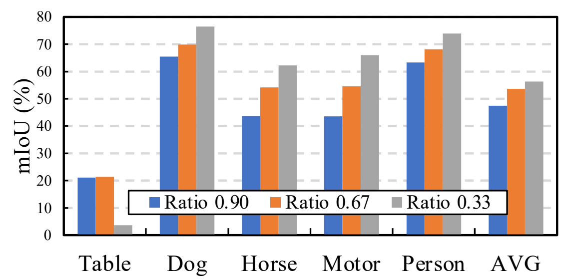

Figure 1. Accuracy (mIoU) of clustering pseudo-labels with different easy split ratio λ in Fold2.

C. Further Explanation on Easy Split Ratio

In our method, we set the easy split ratio λ as a hyper-parameter. We study its impact on PASCAL-5^i and observe two phenomena. First, more incorrect labels are included when λ is too large. Second, hard classes will be largely ignored when λ is too small. We show the accuracy of pseudo-labels with different λ in Fig. 1. The average accuracy is poor when λ is 0.90, while the dinning table is almost ignored when λ is 0.33. This motivates us to select the easy split ratio and $\lambda = 0.67$ is the best choice.

D. Limitation

Semantic-relevant knowledge between base and novel classes is required for novel class discovery. For example, "potted plant" in Fold3 is a semantically different class from the base classes. The mIoU of this class is only 34.5%,

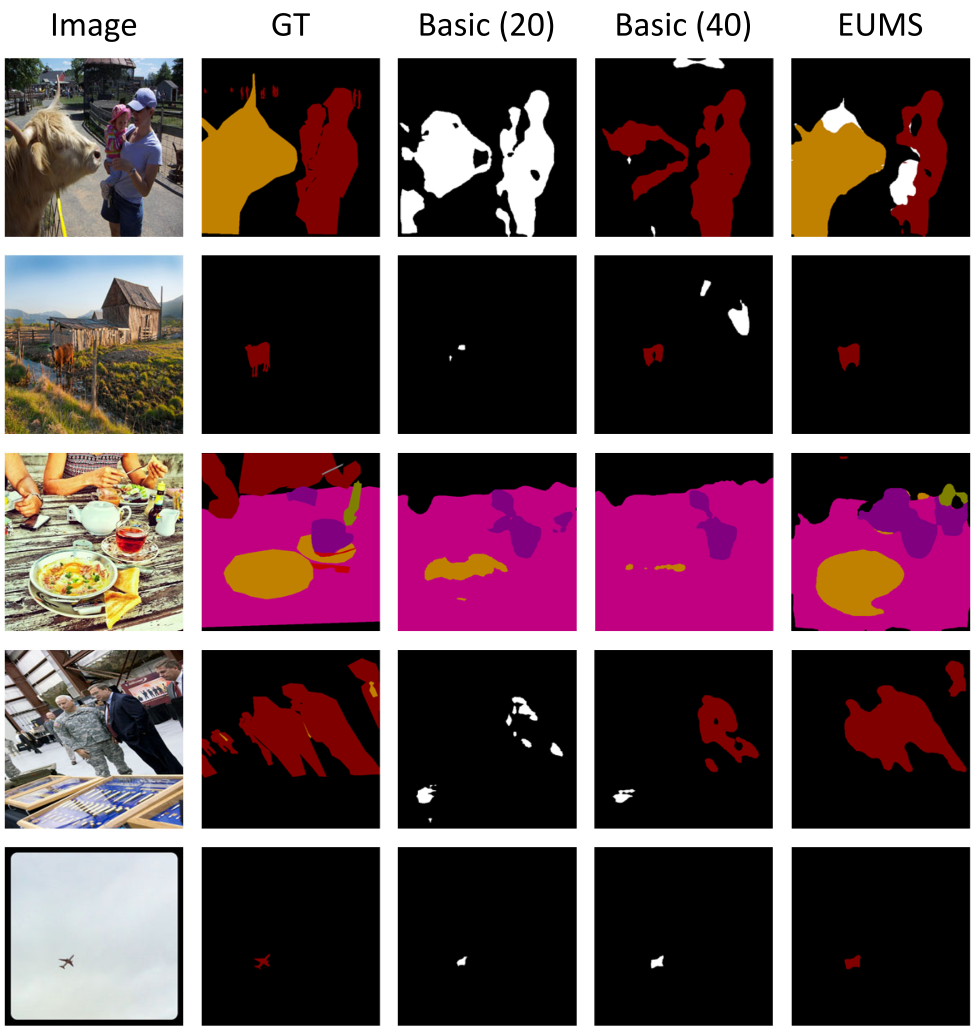

Figure 2. Qualitative comparison of segmentation results in MS COCO 2014 validation set. “Basic (20)” and “Basic (40)” denote the basic framework with 20 and 40 clusters.

which is much less than the other novel classes in Fold3.

E. Visualization

We provide the qualitative comparison on COCO-20^i [3] in Fig. 2. Our EUMS can well handle the circumstances that multiple classes exist in one image (the first and third examples) and the cases that the object is tiny and hard to segment (the second and fourth examples).

References

[1] Kai Han, Sylvestre-Alvise Rebuffi, Sebastien Ehrhardt, Andrea Vedaldi, and Andrew Zisserman. Automatically discovering and learning new visual categories with ranking statistics. In *ICLR*, 2020. 1

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

[3] Khoi Nguyen and Sinisa Todorovic. Feature weighting and boosting for few-shot segmentation. In *ICCV*, 2019. 2

[4] Zhuotao Tian, Hengshuang Zhao, Michelle Shu, Zhicheng Yang, Ruiyu Li, and Jiaya Jia. Prior guided feature enrichment network for few-shot segmentation. *IEEE TPAMI*, 2020. 1

[5] Wouter Van Gansbeke, Simon Vandenhende, Stamatios Georgoulis, and Luc Van Gool. Unsupervised semantic segmentation by contrasting object mask proposals. In *ICCV*, 2021. 1

[6] Zhun Zhong, Enrico Fini, Subhankar Roy, Zhiming Luo, Elisa Ricci, and Nicu Sebe. Neighborhood contrastive learning for novel class discovery. In *CVPR*, 2021. 1

[7] Siyuan Zhou, Li Niu, Jianlou Si, Chen Qian, and Liqing Zhang. Weak-shot semantic segmentation by transferring semantic affinity and boundary. *arXiv preprint arXiv:2110.01519*, 2021. 1