Supplementary Material: Seminar Learning for Click-Level Weakly Supervised Semantic Segmentation

Hongjun Chen[1], Jinbao Wang[1], Hong Cai Chen[1], Xiantong Zhen[2],
Feng Zheng[1*], Rongrong Ji[3], Ling Shao[4]
[1] Southern University of Science and Technology [2] University of Amsterdam
[3] Xiamen University [4] Inception Institute of Artificial Intelligence

1. Seminar learning in Scribble-Level Supervision

Methods	mIOU(%)
Regularized Loss [2]	75.18
Ours	76.17
Full Supervision	78.59

Table 1. Performance comparison on the Pascal VOC 2012 validation set. All results are evaluated on DeeplabV3+ and ResNet101.

Considering that scribble-level annotations can be treated as the extension of click-level annotations, we also apply the seminar learning to the scribble-level supervised semantic segmentation. We use the scribble annotation set first proposed in [1] to train our model. As shown in Table 1, we use 'regularized loss' method [2] as our baseline and gain 75.18% mIOU under our experimental settings. After applying the seminar learning, we gain a competitive result of 76.17% (0.99% higher). The results validate that our method has the potential to improve the performance of weak supervision at various levels.

Notably, since scribble-level labels are able to cross major regions of an object and carry stronger information, it is unnecessary to expand the learning range of the network by building a teacher-student network. However, uncertainties are still inevitable in such a setting as insufficient weak annotations covering the full objects. Therefore, we simplify our seminar learning pipeline for scribble-level supervision, excluding the teacher-student modules but retaining the two-stage training of the student-student module. specifically, the ancillary model is trained with the $L^*_{scribble}$ loss, which is defined as

$$L^*_{scribble} = L_{pCE} + \lambda_{CRF} L_{CRF}, \tag{1}$$

where L_{CRF} is also an important regularization term for improving segmentation performance of scribble-level supervised semantic segmentation [2]. The overall $L_{scribble}$ loss for training the primary model is defined as

$$L_{scribble} = L^*_{scribble} + \lambda_{pseudo} L_{pseudo}. \tag{2}$$

2. Effect of Label Size

Click Size	Only Use L_{pCE} mIOU(%)	Seminar Learning mIOU(%)
1×1	54.70	72.51
3×3	54.87	72.35
5×5	55.85	72.71

Table 2. The effect of enlarging size of the click-level labels. When the size is more than 1×1, click-level labels extend to their neighborhood.

As done by many existing approaches, we evaluate the sizes of click-level labels from 1×1 to 5×5 under our experimental setting. Table 2 conveys that increasing label size can slightly improve the segmentation performance. However, our seminar learning method gains the similar results in all sizes. It shows that our method is hardly affected by click-level label size.

3. More Visualizations

More examples of semantic segmentation in click-level weakly supervised semantic segmentation are shown in Fig. 1. We compare segmentation results among different combinations of losses, which shows the contribution of each loss function in our method.

References

[1] Di Lin, Jifeng Dai, Jiaya Jia, Kaiming He, and Jian Sun. Scribblesup: Scribble-supervised convolutional networks for semantic segmentation. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. 1

[2] Meng Tang, Federico Perazzi, Abdelaziz Djelouah, Ismail Ben Ayed, Christopher Schroers, and Yuri Boykov. On regularized losses for weakly-supervised cnn segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 507–522, 2018. 1

*Corresponding author: Feng Zheng (Email: f.zheng@ieee.org). This work is supported by the National Natural Science Foundation of China under Grant No. 61972188.

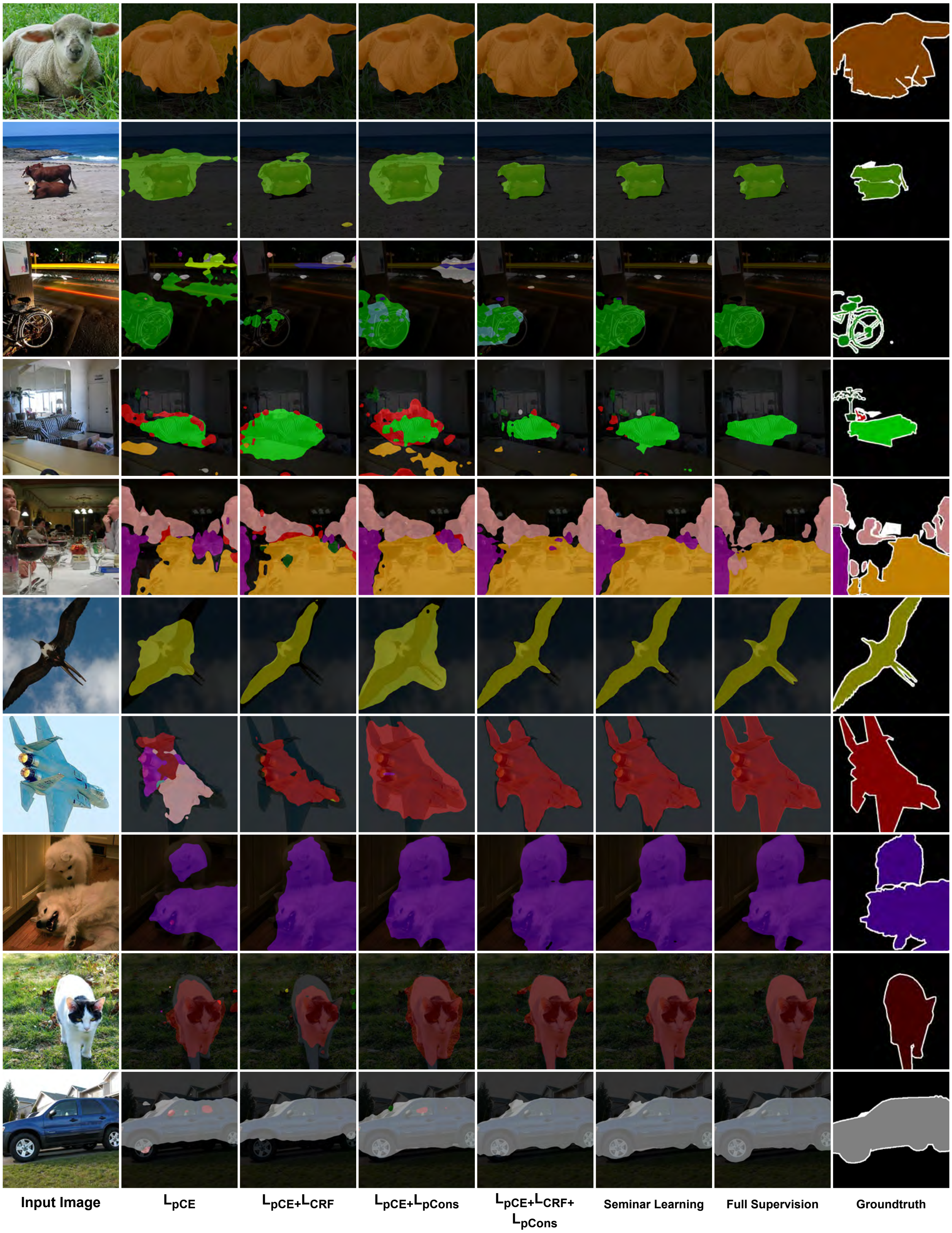

Figure 1. Qualitative results in various combinations of losses.