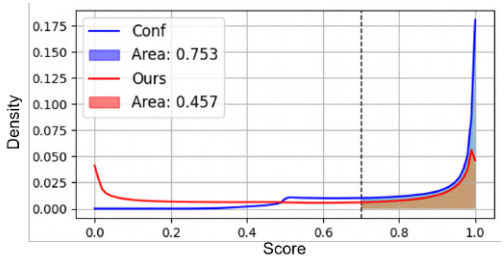
Supplementary Material: Pruning-Guided Curriculum Learning for Semi-Supervised Semantic Segmentation

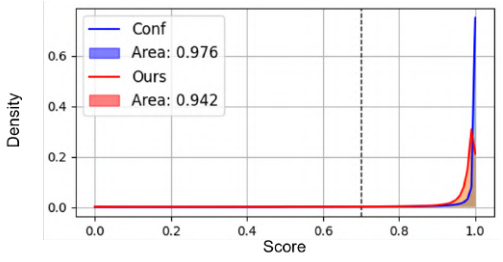
Heejo Kong[1] Gun-Hee Lee[2] Suneung Kim[3] Seong-Whan Lee[3†]
[1]Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea
[2]Department of Computer Science and Engineering, Korea University, Seoul, South Korea
[3]Department of Artificial Intelligence, Korea University, Seoul, South Korea
{hj_kong, gunhlee, se_kim, sw.lee}@korea.ac.kr

1. Detailed Figures of PDF

Fig. 1 shows probability density functions (PDFs) of the confidence (Conf.) and our proposed scores. It is based on the pixel-wise softmax probability of the network trained for 10 epochs (out of total of 80 epochs) on PASCAL VOC 2012 with the 1/8 partition protocol. 'False' and 'True' represent the cases in which the predicted pseudo-labels are incorrect or correct respectively. 'Area' indicates the ratio of pseudo-labels that are reflected in training, i.e., their scores exceed a predefined threshold (0.7). As shown in this figure, our method effectively refines the score of confident samples for the false case, and significantly reduces the area from 0.753 to 0.457. On the other hand, the confidence score does not change much for the true case, and the area of the true case hardly changes from 0.976 to 0.942.



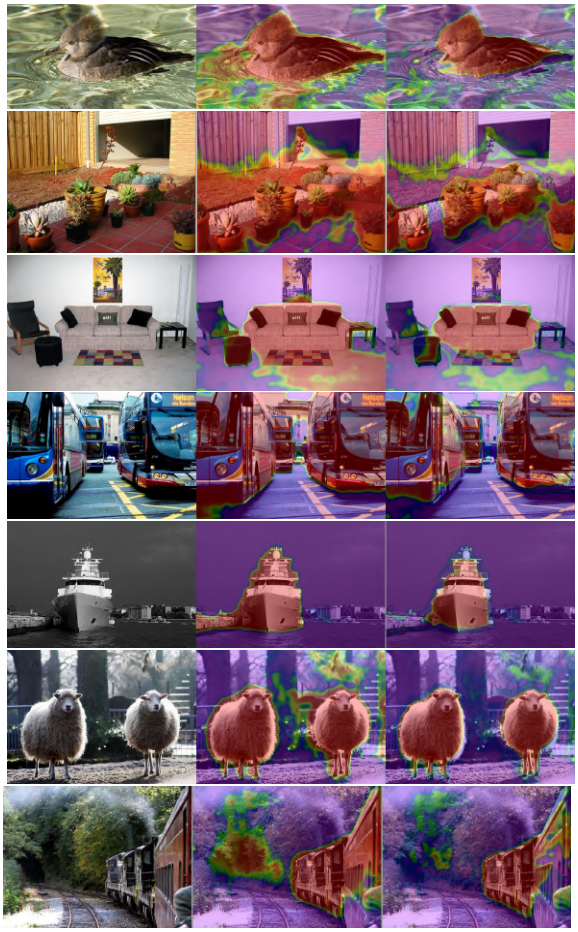
(a) False



(b) True

Figure 1. Probability density functions (PDFs) comparison between the confidence (Conf.) and our proposed scores.

2. Qualitative Results



Input Confidence Ours

Figure 2. Heatmap visualizations of pixel-wise confidence scores for foreground objects. The confidence scores are estimated by the trained network for 10 epochs (out of a total of 80 epochs) on PASCAL VOC 2012 [3].

Figure 3. Qualitative results for comparison the pseudo-labeling quality between confidence thresholding and our PGCL on PASCAL VOC 2012 [3] validation set. All predictions are estimated by the trained network for 10 epochs (out of a total of 80 epochs) with a 1/8 split. The white region indicates pixels that are not assigned as pseudo-labels, i.e., their scores are lower than the predefined threshold (0.7).

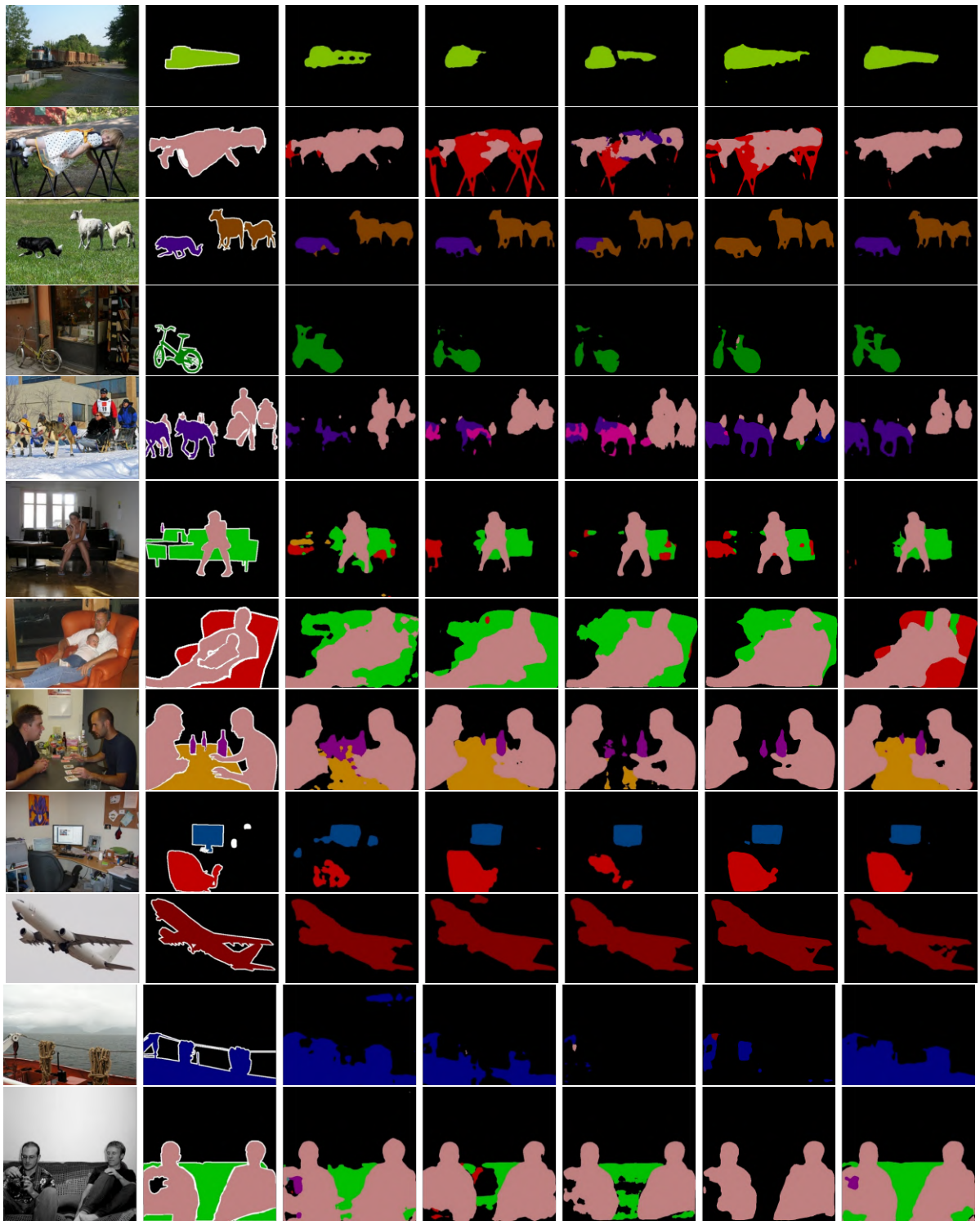
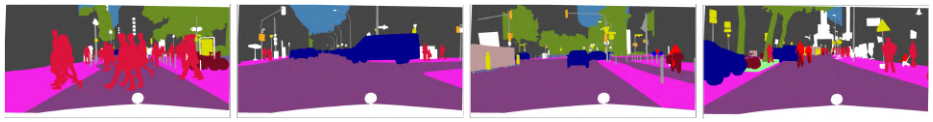
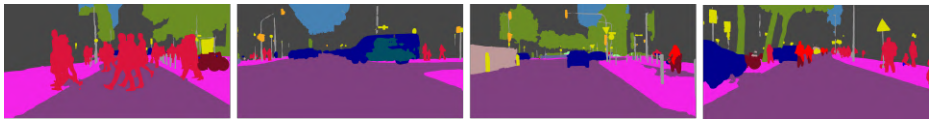


Figure 4. Qualitative results for comparison PGCL to previous state-of-the-art methods on PASCAL VOC 2012 [3] validation set. "Baseline" stands for the results of supervised training on the labeled dataset only. For a fair comparison, all models are trained with a 1/8 split.

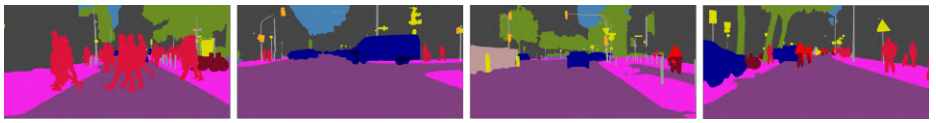
(a) Image



(b) Ground truth



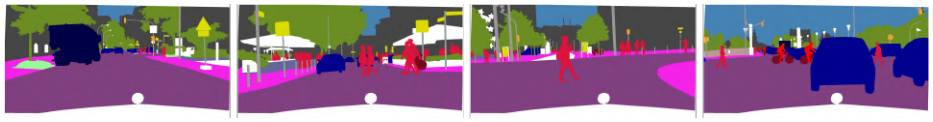
(c) 1/8 split



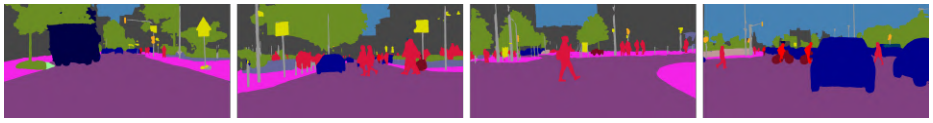
(d) 1/4 split

(e) 1/2 split

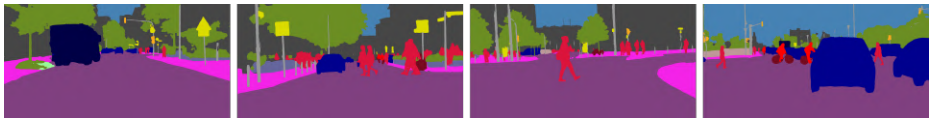
(a) Image



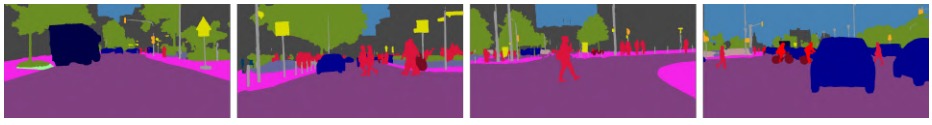
(b) Ground truth



(c) 1/8 split



(d) 1/4 split



(e) 1/2 split

Figure 5. Qualitative results on Cityscapes [2] validation set in various proportions of labeled data to unlabeled data.

References

[1] Xiaokang Chen, Yuhui Yuan, Gang Zeng, and Jingdong Wang. Semi-supervised semantic segmentation with cross pseudo supervision. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[2] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[3] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision (IJCV)*, 2010.

[4] Donghyeon Kwon and Suha Kwak. Semi-supervised semantic segmentation with error localization network. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[5] Xin Lai, Zhuotao Tian, Li Jiang, Shu Liu, Hengshuang Zhao, Liwei Wang, and Jiaya Jia. Semi-supervised semantic segmentation with directional context-aware consistency. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.