Supplementary Material: Class-wise Dynamic Graph Convolution for Semantic Segmentation

Hanzhe Hu[1[0000-0003-2799-2655]*], Deyi Ji[2[0000-0001-7561-9789]], Weihao Gan[2], Shuai Bai[3], Wei Wu[2], and Junjie Yan[2]

[1] Peking University, Beijing, China
[2] SenseTime Group Limited, Beijing, China
[3] Beijing University of Posts and Telecommunications, Beijing, China
huhz@pku.edu.cn, {jideyi,ganweihao,wuwei, yanjunjie}@sensetime.com, baishuai@bupt.edu.cn

In this supplementary material, we provide additional experimental and visualization results. In section 1, we present the impact on boundary accuracy of our proposed CDGCNet. And in section 2, class-wise feature maps are presented. Then some qualitative results on evaluated datasets are shown in section 3.

1 Impact on boundary accuracy

To validate the effectiveness of dynamic sampling method in mining hard samples which occur in boundary areas, we design an experiment to show that our method improves boundary precision. Following [4,5], we adopt a standard trimap approach where we compute the classification accuracy within a band (called trimap) of varying width around boundaries of segments(shown in Fig. 2). As shown in Fig. 1, we use deeplab v3 as baseline and compute the boundary accuracy gain over baseline of CDGC module with two graph construction methods: similarity graph and dynamic sampling. The results indicate that dynamic sampling graph construction method effectively improves the vicinity of boundaries compared with baseline method and CDGC module with simple similarity graph, hence preserving more details.

2 Visualizations of Class-wise Features

As shown in Fig. 3, we provide additional visualizations of class-wise features before and after graph convolution evaluated on Cityscapes[2] validation set. It can be seen that our proposed CDGC module can effectively deal with hard samples, hence obtaining a better segmentation results in a class-wise manner. Concisely, ambiguity such as hard positive and hard negative samples is well taken care of. Moreover, CDGC module can significantly increase the confidence of classification on indistinguishable pixels.

* This work is done when Hanzhe Hu is an intern at SenseTime Group Limited.

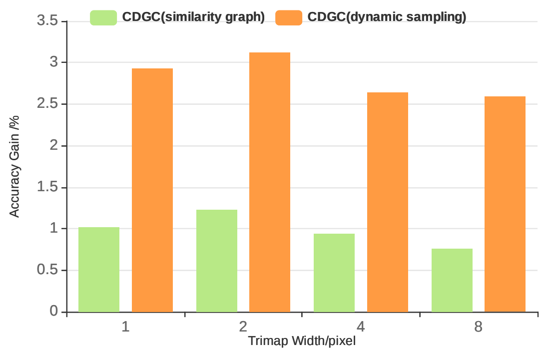

Fig. 1. Boundary accuracy gain over baseline for two graph construction methods on Cityscapes validation set.

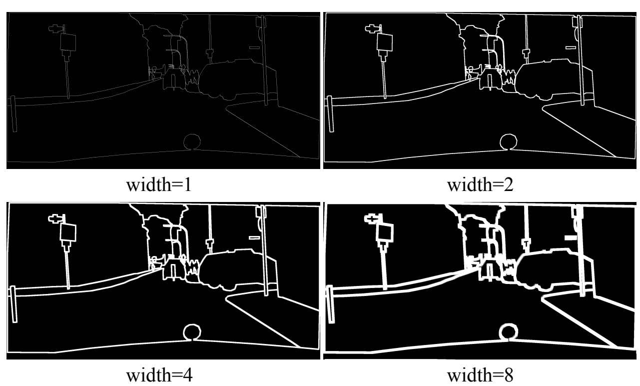

Fig. 2. Trimaps used for boundary accuracy evaluation with different band width from Cityscapes validation set

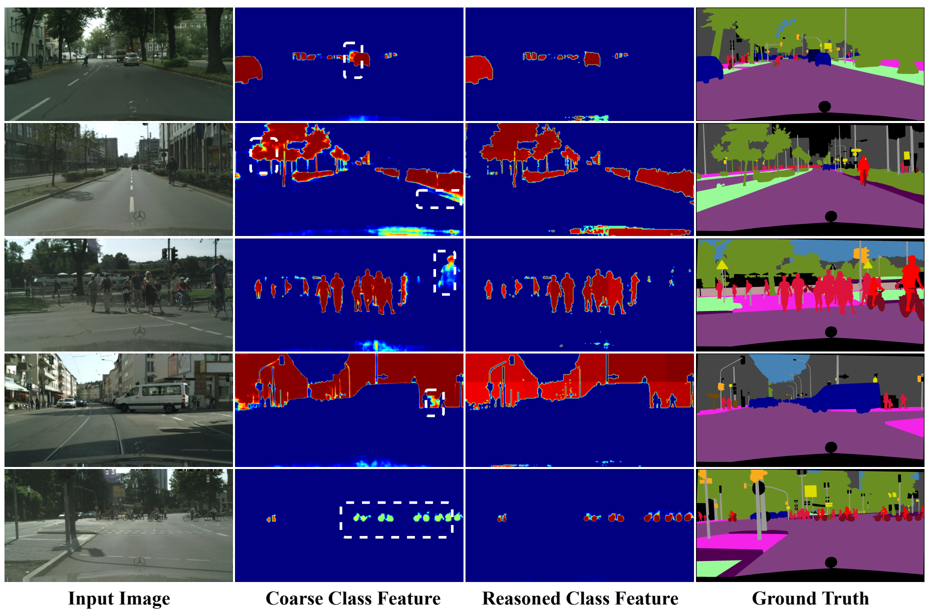

Fig. 3. Visualizations of class-wise features before and after graph convolution on Cityscapes validation set. From left to right: input image, class-wise feature before CDGC module, class-wise feature after CDGC module, ground truth. From top to bottom, the visualized category is car, vegetation, person, building and bicycle.

3 Qualitative Results of CDGCNet

We provide visualization results on PASCAL VOC 2012[3] validation set and COCO Stuff[1] validation set,shown in Fig. 4 and Fig. 5. It can be seen that our proposed CDGC module can effectively improve the segmentation results on these evaluated datasets.

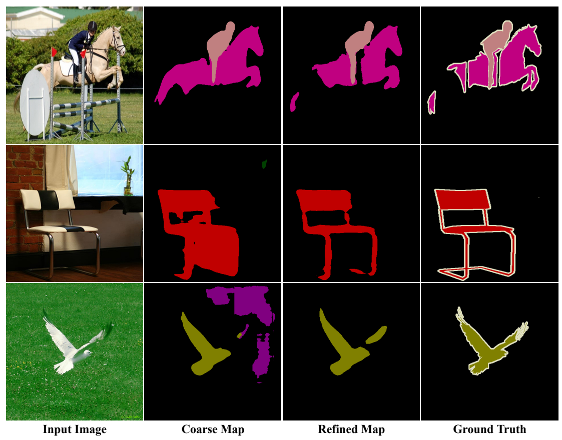

Fig. 4. Qualitative results on PASCAL VOC 2012 validation set.

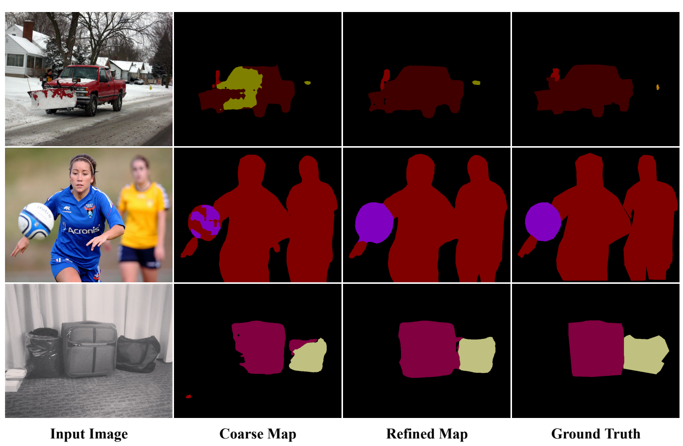

Fig. 5. Qualitative Results on COCO Stuff validation set

References

1. Caesar, H., Uijlings, J., Ferrari, V.: Coco-stuff: Thing and stuff classes in context. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1209–1218 (2018)
2. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B.: The cityscapes dataset for semantic urban scene understanding. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3213–3223 (2016)
3. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International journal of computer vision **88**(2), 303–338 (2010)
4. Kohli, P., Torr, P.H., et al.: Robust higher order potentials for enforcing label consistency. International Journal of Computer Vision **82**(3), 302–324 (2009)
5. Marin, D., He, Z., Vajda, P., Chatterjee, P., Tsai, S., Yang, F., Boykov, Y.: Efficient segmentation: Learning downsampling near semantic boundaries. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2131–2141 (2019)