Mixup-CAM: Weakly-supervised Semantic Segmentation via Uncertainty Regularization

Yu-Ting Chang[1]
ychang39@ucmerced.edu
Qiaosong Wang[2]
qiaowang@ebay.com
Wei-Chih Hung[2]
whung8@ucmerced.edu
Robinson Piramuthu[2]
robinpir@amazon.com
Yi-Hsuan Tsai[3]
wasidennis@gmail.com
Ming-Hsuan Yang[14]
mhyang@ucmerced.edu

[1] UC Merced
[2] eBay Inc.
[3] NEC Labs America
[4] Google Research

1 Overview

In this supplementary document, we provide additional visualizations results. More qualitative results are presented in Section 2, including initial response maps, semantic segmentation results, and few failure cases.

2 Qualitative Comparisons

Figure 1 presents more initial response maps generated by the CAM [1] method and ours. A number of qualitative examples of our final semantic segmentation results are presented in Figure 2. In addition, we show some failure segmentation cases in Figure 3. There are three main issues that would affect the quality of segments: 1) the incompleteness on detailed parts, 2) the ambiguity on object boundaries, and 3) the noise on background. The first two issues are the common problem of the WSSS task. The third issue could be raised by the increased uncertainty from mixup augmentations, such that the expanded initial response could attend to the non-object region. Although there are some failure examples, our approach generally produces high-quality semantic segmentation results.

References

[1] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *CVPR*, 2016. 1, 2

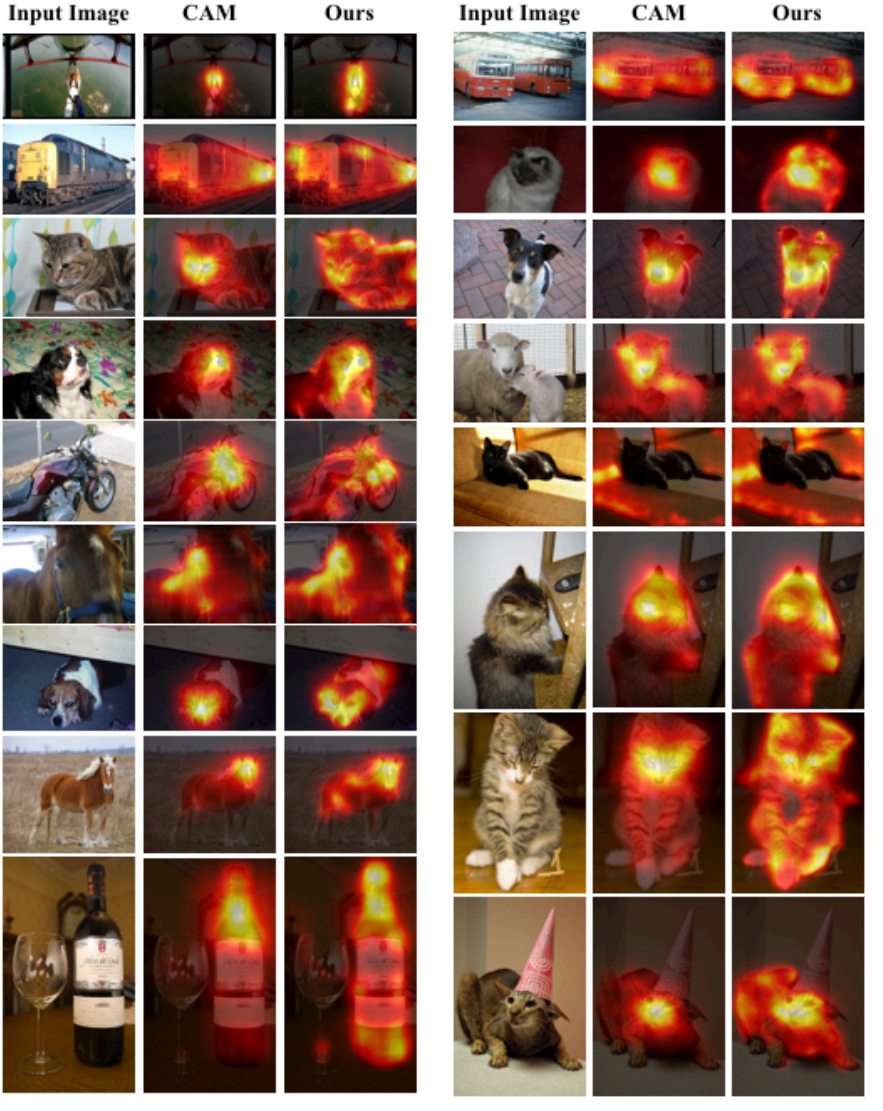

Figure 1: Qualitative comparison of the initial response map with [1] on the PASCAL VOC 2012 val images.

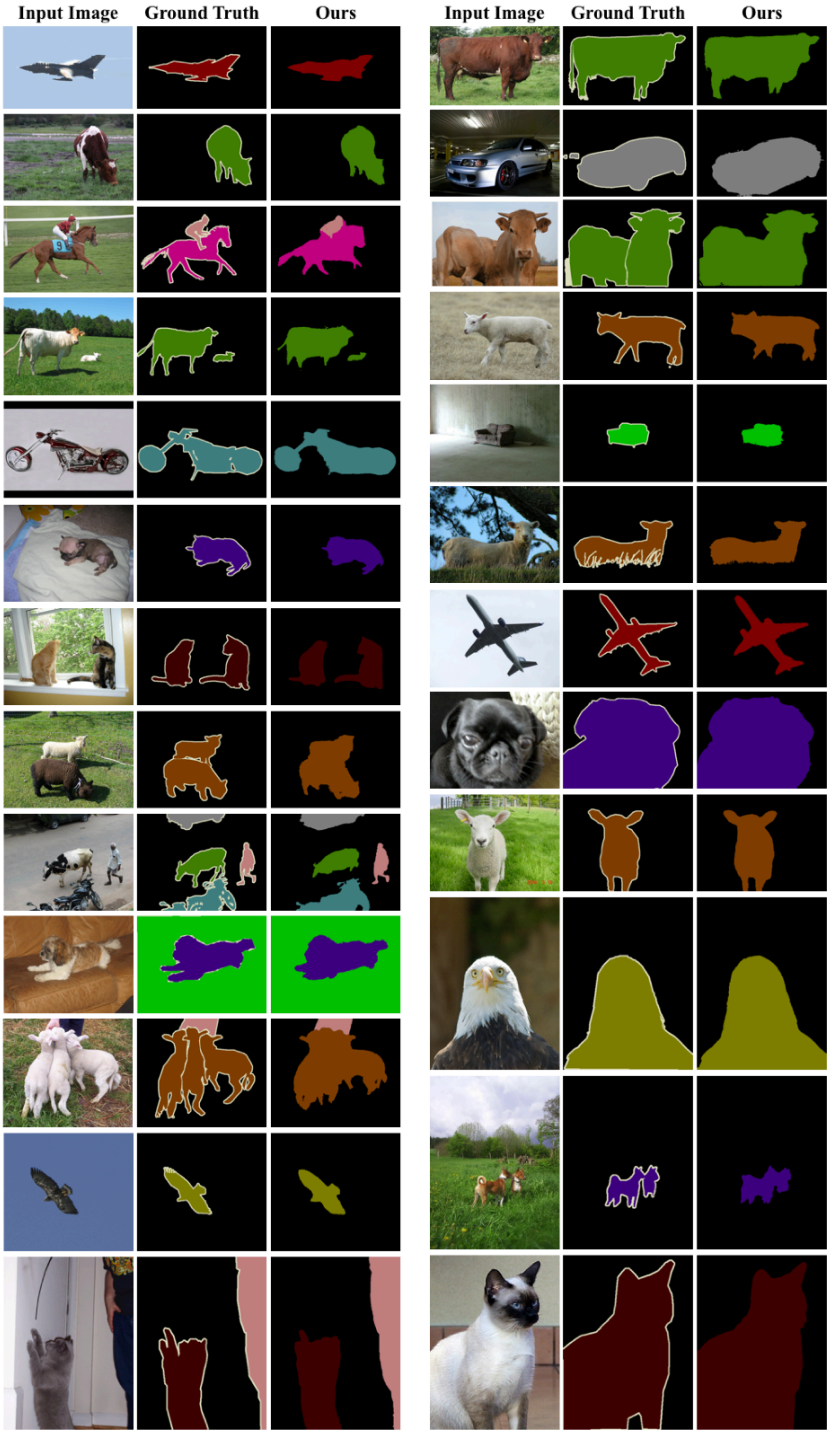

Figure 2: Semantic segmentation results on the PASCAL VOC 2012 val images.

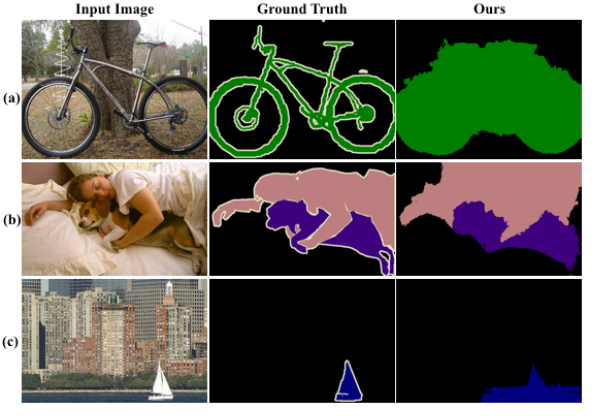

Figure 3: Failure semantic segmentation examples. (a) Missing detailed parts. Details of the bicycle are missing in the segment. (b) The ambiguity on object boundary. There are errors on the boundary region between two objects. (c) The background noise.