Supplementary Material for

Feature-level and Spatial-level Activation Expansion for Weakly-Supervised Semantic Segmentation

A. Details of Non-geometric Transforms

Our FSAE framework uses strong non-geometric image transform for consistency regularization. In Tab. 1, candidates of transformation operations are listed. For each image, three out of nine RandAugment [1] transforms are randomly selected. By imposing consistency between outputs of differently perturbed input images, the model generates consistent CAM activation across wide range of object region.

Table 1. Type of used non-geometric transforms of RandAugment.

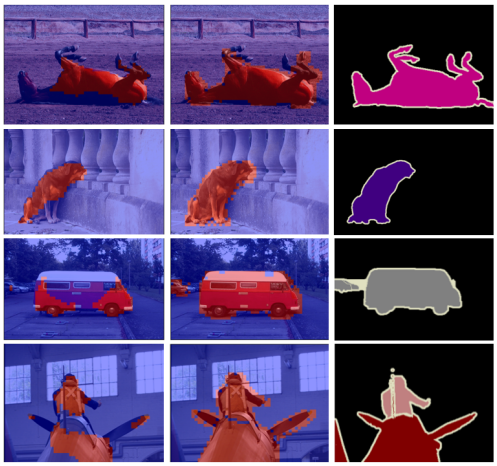
Index	Transforms
1	*Identity*
2	*AutoContrast*
3	*Equalize*
4	*Solarize*
5	*Color*
6	*Posterize*
7	*Contrast*
8	*Brightness*
9	*Sharpness*

B. Comparison of High-confidence Pseudo-label Area

In Fig. 1, we provide visualization examples of high-confidence pseudo-label area of base model with and without our FSAE. In our method, we use the area for the explicit pixel-wise supervision. It can be seen that our method results in more broad and accurate pseudo-label area compared to the base model. Consequently, our method can produce more high-quality CAMs as shown in experiment section.

C. More Qualitative Results

We provide more qualitative segmentation results of our proposed FSAE framework on PASCAL VOC 2012 and MS COCO 2014 datasets as shown in Fig. 2 and 3. Our proposed method produces more accurate segmentation results in various scenes.



(a) (b) (c)

Figure 1. Visualization examples of high-confidence pseudo-label area. Our method produces accurate pseudo-label for explicit supervision. (a) PPCw/EPS, (b) Ours, and (c) GT.

References

[1] Ekin D Cubuk, Barret Zoph, Jonathon Shlens, and Quoc V Le. Randaugment: Practical automated data augmentation with a reduced search space. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pages 702–703, 2020. 1