

Variational Prototype Inference for Few-Shot Semantic Segmentation

SUPPLEMENTARY MATERIALS

Haochen Wang^{1*}, Yandan Yang^{1*}, Xianbin Cao^{1,2,3†}, Xiantong Zhen^{4,5}, Cees Snoek⁴, Ling Shao⁵

¹School of Electronic and Information Engineering, Beihang University, Beijing, China

²Key Laboratory of Advanced Technology of Near Space Information System,
Ministry of Industry and Information Technology of China

³Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, China

⁴University of Amsterdam, Amsterdam, Netherlands

⁵Inception Institute of Artificial Intelligence, Abu Dhabi, UAE

{haochenwang, yangyandan, xbcao}@buaa.edu.cn,

x.zhen@uva.nl, cgmsnoek@uva.nl, ling.shao@ieee.org

A. Visualization of results on FSS-1000

The qualitative segmentation results on the FSS-1000 dataset in Sec.4.2.3 are illustrated in Fig. 1. VPI produces accurate segmentation close to the ground truth.

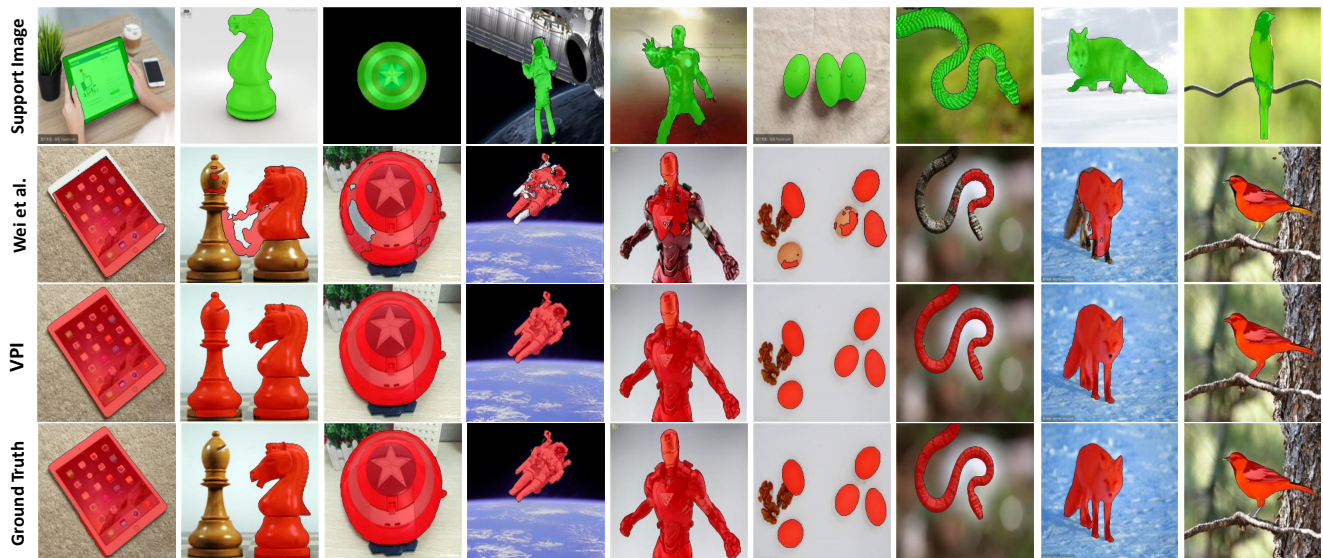


Figure 1. Qualitative visualizations of one-shot results on FSS-1000. Our VPI can accurately produce the segmentation maps for objects by predicting almost all the pixels on them, while Wei et al. [?] fails in some cases by missing some foreground pixels.

B. Failure Cases

To gain further insights into the proposed method, we show some failure cases on PASCAL-5ⁱ in Fig 2. In the first case (top left), our method fails to predict detailed information of the bicycle though it can roughly capture the object from the

*These authors contribute equally.

†Corresponding Author.

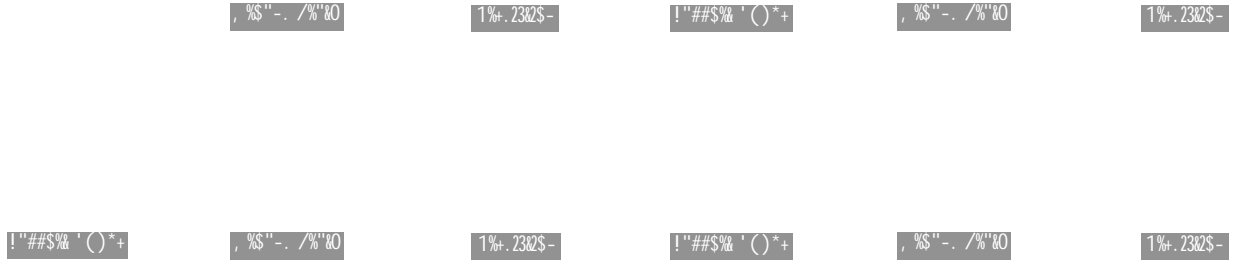


Figure 2. Several failure cases on PASCAL-5^f. In these failure cases, query images demonstrate considerable visual differences from the corresponding support images, which still poses challenges for accurate segmentation.

complex background. This is mainly due to the complex line structure of the bicycle and the variation in size of objects between query and support images. In the third case (bottom left), the annotated object in the support image is largely occluded, which offers only weak support for segmentation. As a result, our method fails to distinguish the foreground bus from the surrounding cars.