

Supplementary Material for Paper: Semi-Supervised Object Detection via Multi-instance Alignment with Global Class Prototypes

Aoxue Li Peng Yuan Zhenguo Li
Huawei Noah's Ark Lab, China

lax@pku.edu.cn, yuanpeng126@huawei.com, Li.Zhenguo@huawei.com

1. Discussion of Failure Cases

Figure 1 provides examples of failure cases of the proposed approach on the PASCAL VOC dataset. As shown in this figure, the failures of our approach can be grouped into two categories. The first failure case is object in complex backgrounds. As shown in the first row of Figure 1, our approach misdetects a house as a boat when it is occluded by a train or misrecognizes a trash can as a person when it is squeezed between two cupboards. This failure might be caused by the fact: The learned SSOD model infer the potential labels without considering context objects. For example, if a SSOD model had taken the train into account as a context, it wouldn't infer the candidate proposal as a boat because a train and a boat are hardly occurred in the same picture. In the future work, we will focus on addressing the

challenging issues for better detection performance.

The second failure case is object in rare views. As shown in the second row of Figure 1, our approach failed to detect tv monitors that takes pictures from a circular window, or misclassified a person as a cat when she made up with feathers. This problem is also called corner case, which is a challenging issue suffered by most learning-based models. To address this issue, case-specific data augmentation or additional knowledge of experts might be potential solutions. We will further explore this problem in the future work.

Since our model still performs limited robustness in detecting objects in rare views or complex contexts. Therefore, we should be really careful when using it in the scenarios where failures will cause severe consequences.

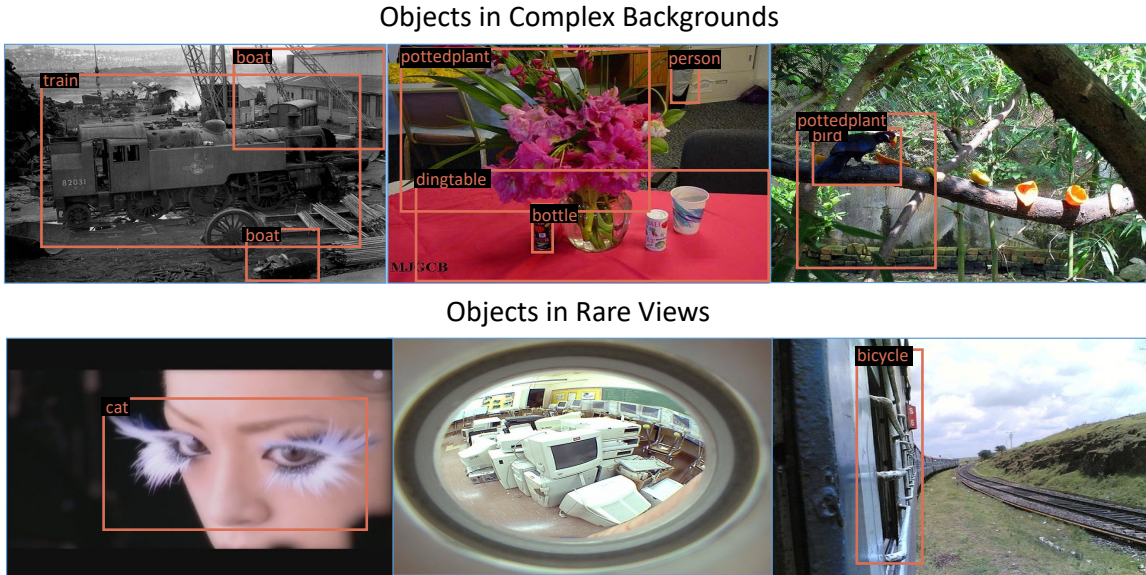


Figure 1. Examples of failure cases obtained by our approach on the PASCAL VOC dataset. We illustrate two types of failure cases. The first row provides failure objects in complex backgrounds and the second row provides failure objects in rare views.