SFUOD: Source-Free Unknown Object Detection

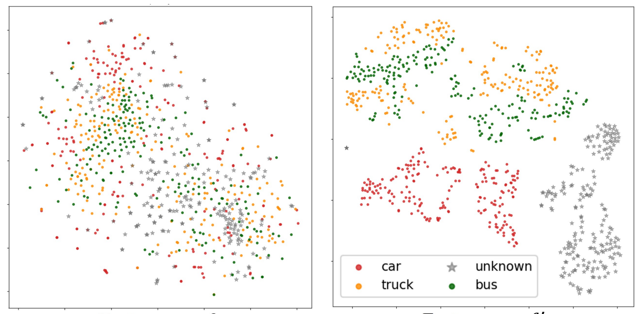
Supplementary Material

top-k selection	Cityscapes $\rightarrow$ Foggy Cityscapes		
	Known mAP	U-Recall	H-Score
Baseline	22.97	3.60	6.22
$k = 0$	29.07	7.52	11.95
$k = 25$	31.84	7.23	11.78
k = 50	**32.32**	**10.59**	**15.95**
$k = 100$	27.95	8.27	12.76

S.Table 1. Ablation study of selecting top-k activated features.

top-r Recon	Cityscapes $\rightarrow$ Foggy Cityscapes		
	Known mAP	U-Recall	H-Score
r = 5	**32.32**	**10.59**	**15.95**
$r = 10$	30.77	8.47	13.28
$r = 20$	30.78	8.09	12.81
$r = 30$	31.63	7.72	12.41

S.Table 2. Ablation study of the reconstruction on top-r



Feature map f_a Feature map f'_a

S.Figure 1. Visualization of the feature map.

Threshold ϵ	Cityscapes $\rightarrow$ Foggy Cityscapes		
	Known mAP	U-Recall	H-Score
$\epsilon = 0.1$	29.51	7.88	12.44
$\boldsymbol{\epsilon = 0.3}$	32.32	**10.59**	**15.95**
$\epsilon = 0.5$	32.81	9.23	14.41
$\epsilon = 0.7$	33.45	7.36	12.07
$\epsilon = 0.9$	**34.55**	7.11	11.79

S.Table 3. Ablation study of the Unknown threshold ϵ

Unknown Pseudo Label	Cityscapes $\rightarrow$ Foggy Cityscapes		
	Known mAP	U-Recall	H-Score
OpenDet [9]	42.09	1.79	3.43
OW-DETR [8]	39.92	1.98	3.77
SOMA [17]	**45.55**	4.08	7.49
CollaPAUL (Ours)	32.32	**10.59**	**15.95**

S.Table 4. Comparison of the performance on open-set methods.

A. Further Ablation Study

Ablation Study of the Target-dependent Knowledge. Given the feature maps f from the backbone, we select the top-k activated features f_a and apply truncated reconstruction with top-r decomposed components via SVD to extract latent target knowledge shown in S.Figure1. To evaluate the effectiveness of top-k selection and top-r decomposition, we conducted ablation studies. S.Table 1 presents the results for selecting top-k features from the feature maps. K=0 indicates truncated decomposition applied to all feature maps without selection. We found that K=50 achieved the best performance across known mAP, U-Recall, and H-Score. S.Table 2 compares performance variations with different top-r values for truncated decomposition. Using 5 components (r=5) for the truncated reconstruction achieved a promising performance across all the metrics.

Ablation Study on the Threshold. We further analyzed the effectiveness of the unknown threshold in generating the confidence mask for principal axis-based unknown labeling. As shown in S.Table 3, we evaluated the threshold ϵ within the range 0.1 to 0.9. Our findings indicate that setting ϵ = 0.3 achieves strong performance in U-Recall and H-Score. Notably, we also observed that increasing ϵ tends to improve known mAP but leads to a decline in U-Recall.

B. Further Analysis

Comparison with other Open-set Methods. To validate the effectiveness of the proposed CollaPAUL in detecting unknown objects as well as identifying known objects, we compared CollaPAUL with other open recognition methods: OpenDet[9], OW-DETR[8], and SOMA [17]. OpenDet is designed to detect unknown objects during the training of annotated known objects. OW-DETR continuously learns annotated known objects and detects unannotated objects as unknowns. SOMA transfers knowledge for domain adaptation, enabling the detection of unknown objects while conveying knowledge of known objects from the source to the target domain. However, these methods struggle to be applied to the proposed SFUOD setting because OpenDet assumes supervised learning, OW-DETR requires continuous learning in a single domain, and SOMA, which assumes domain-adaptive learning, depends on the source domain dataset for knowledge transfer.

Thus, we compared our performance at the proposed SFUOD scenario with these methods at the Adaptive Open-Set Object Detection (AOOD) scenario, where the detector can access both the labeled source dataset and the unlabeled target dataset during training. As shown in S.Table 4, CollaPAUL achieved promising performance in both U-Recall and H-Score. Notably, unlike other methods that use labeled source domain datasets during training, CollaPAUL, which relies solely on unlabeled target domain datasets, achieved comparable performance in known mAP. These results demonstrate that CollaPAUL effectively detects unknown objects while maintaining strong performance compared to existing methods.