Novel Scenes & Classes: Towards Adaptive Open-set Object Detection

Wuyang Li[1], Xiaoqing Guo[1,2] and Yixuan Yuan[1,3] ✉

[1]City University of Hong Kong, [2]University of Oxford, [3]The Chinese University of Hong Kong

CODE

LAB

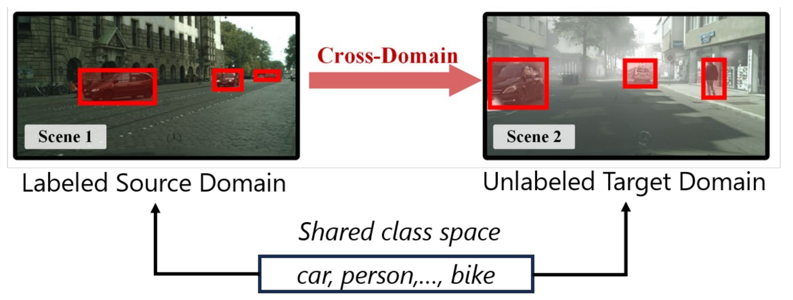
ICCV23 PARIS

Background

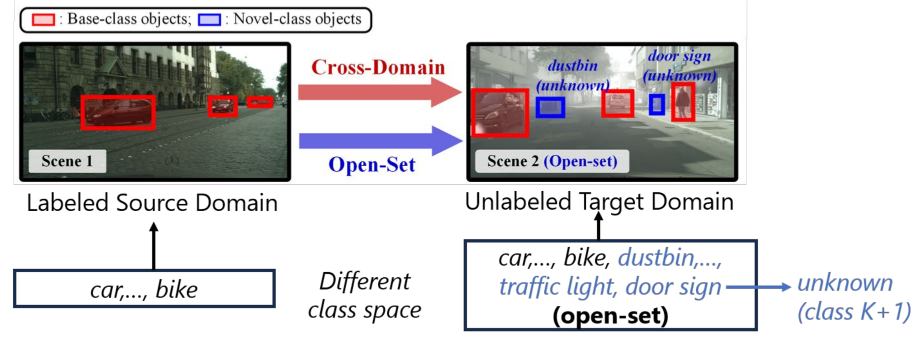
Adaptive Open-set Object Detection (AOOD)

AOOD aims transfer an object detector from a labeled source domain to an unlabeled target domain with open-set classes, different from existing DAOD with shared class definition.

Domain Adaptive Object Detection (DAOD)

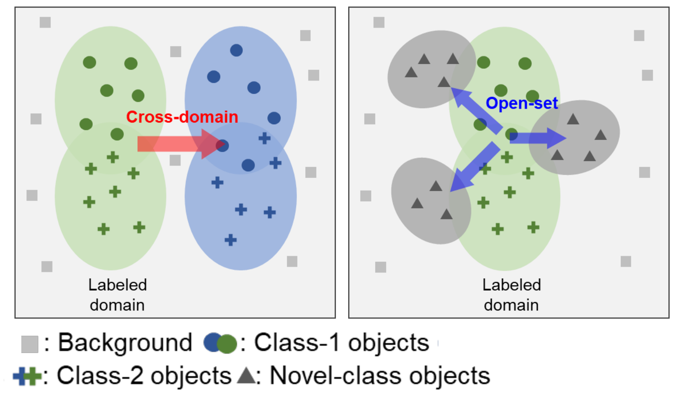


Adaptive Open-set Object Detection (AOOD)



Motivation

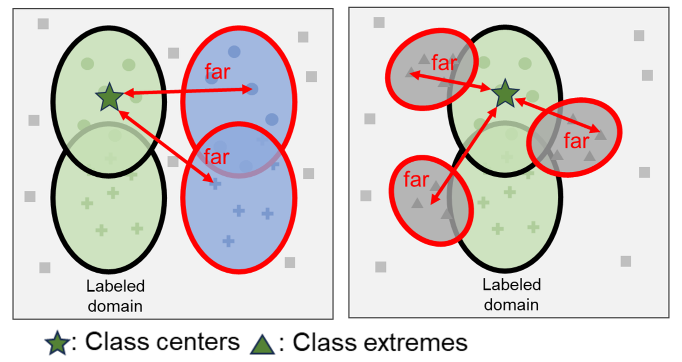
Problem Analysis



AOOD needs to address

- cross-domain challenge
- open-set challenge

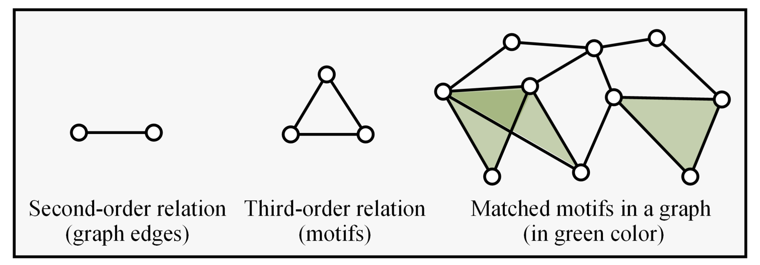
Challenges



What limits AOOD?

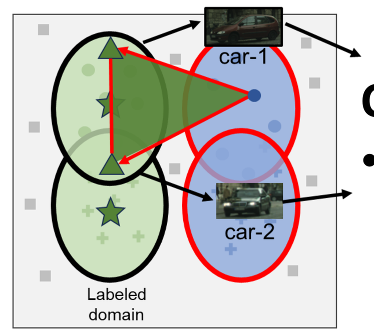
- Cross-domain/open-set objects are both out of labeled distribution, leading to the failure of existing low-order solutions, e.g., selecting objects with low confidence.

Preliminaries



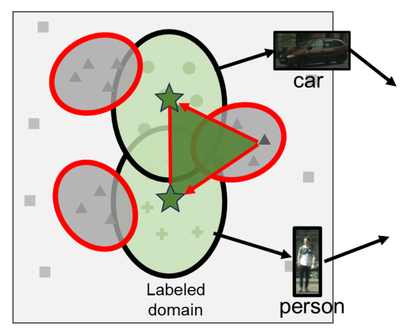
Motif is a statistically significant subgraph with high-order pattern.

The Proposed High-order Solutions



Cross-domain Solution

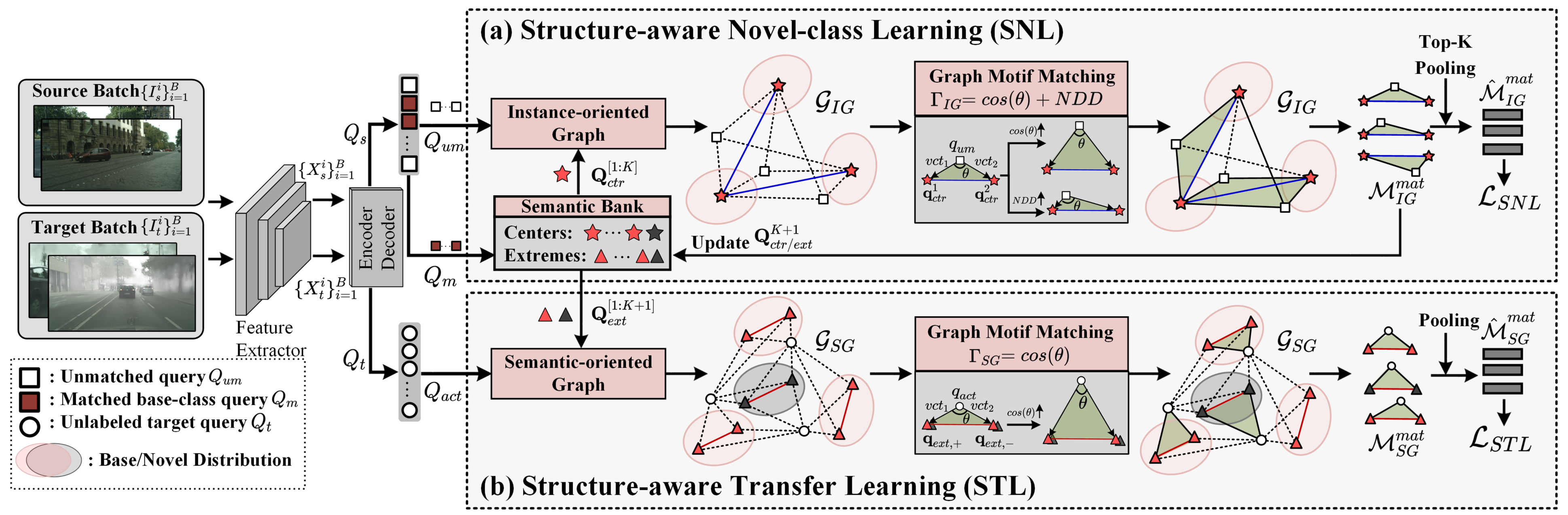
- Exploring within-class diversity for each class



Open-set Solution

- Exploring object-level property shared among different classes

Methodology



Structure-aware Novel-class Learning (SNL)

- **Establish a semantic bank saving class centers and extremes**

$$\mathbf{Q}^k_{ctr} \leftarrow \alpha f_{\text{mean}}(Q^k_m) + (1-\alpha)\mathbf{Q}^k_{ctr}$$
$$\mathbf{Q}^k_{std} \leftarrow \alpha f_{\text{std}}(Q^k_m) + (1-\alpha)\mathbf{Q}^k_{std}$$

- **Build up the instance-oriented graph**

  Link class centers with the farthest counterpart

  Link unmatched object queries with class centers

- **Select motifs via a high-order metric**

$$M^{mat}_{IG} = \text{argmin}_{\Gamma_{IG}} \mathcal{M}_{IG} \qquad \Gamma_{IG} := cos(\theta) + NDD$$

where

$$cos(\theta) := \frac{vct_1 \cdot vct_2}{||vct_1||_2 \cdot ||vct_2||_2} \qquad NDD := \left|\frac{||vct_1||_2 - ||vct_2||_2}{||\mathbf{q}^1_{ctr} - \mathbf{q}^2_{ctr}||_2}\right|$$

- **Optimize with selected motifs**

$$\mathcal{L}_{SNL} = -\frac{1}{\mathrm{K}}\sum_{i=1}^{\mathrm{K}} log(\boxed{p(f_{cls}(\hat{M}^{mat}_{IG,i}) = K+1|\hat{M}^{mat}_{IG,i})})$$

Unknown posterior

Structure-aware Transfer Learning (STL)

- **Build up the semantic-oriented graph**

  Link class extreme pairs in the same class

  Link activated object queries with class extremes

- **Select motifs via a high-order metric**

$$M^{mat}_{SG} = \text{argmin}_{\Gamma_{SG}} \mathcal{M}_{SG} \qquad \Gamma_{SG} := cos(\theta)$$

where

$$cos(\theta) := \frac{vct_1 \cdot vct_2}{||vct_1||_2 \cdot ||vct_2||_2}$$

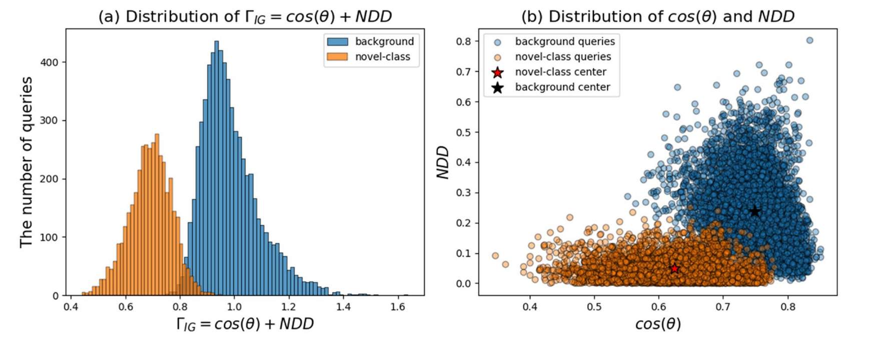
- **Optimize with selected motifs**

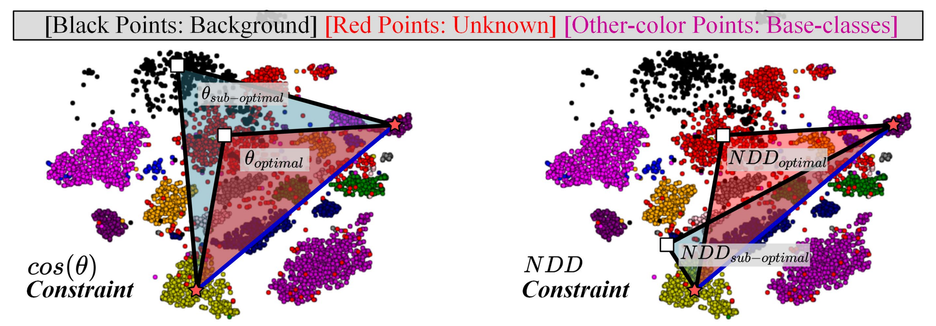
$$\mathcal{L}_{STL} = -\frac{1}{|\mathcal{M}^{mat}_{SG}|}\sum_{i=1}^{|\mathcal{M}^{mat}_{SG}|} \boxed{\tilde{y}} log(p(f_{cls}(\hat{M}^{mat}_{SG,i})|\hat{M}^{mat}_{SG,i}))$$

Pseudo labels

Experiments

Method	Set	num. novel-class: 3				num. novel-class: 4				num. novel-class: 5			
		$mAP_b\uparrow$	$AR_n\uparrow$	WI↓	AOSE↓	$mAP_b\uparrow$	$AR_n\uparrow$	WI↓	AOSE↓	$mAP_b\uparrow$	$AR_n\uparrow$	WI↓	AOSE↓
DDETR [60] $_{ICLR'21}$	het-sem	47.52	0.00	0.341	459	45.24	0.00	0.506	1028	42.38	0.00	0.659	1968
PROSER [59] $_{CVPR'21}$	het-sem	46.92	1.80	0.271	218	44.19	2.02	0.415	531	41.99	2.00	0.584	1127
OpenDet [19] $_{CVPR'22}$	het-sem	47.04	1.92	0.269	221	45.71	1.89	0.499	511	42.09	1.70	0.579	922
OW-DETR [18] $_{CVPR'22}$	het-sem	43.31	1.84	0.432	192	42.52	2.10	0.619	451	39.92	1.98	0.684	814
SOMA (ours)		**50.87**	**3.78**	**0.268**	**139**	**48.06**	**4.41**	**0.412**	**340**	**45.55**	**4.08**	**0.526**	**649**
DDETR [60] $_{ICLR'21}$	hom-sem	44.62	0.00	1.860	2937	43.55	0.00	2.000	3565	40.18	0.00	2.462	6770
PROSER [59] $_{CVPR'21}$	hom-sem	43.15	4.59	1.842	2146	43.31	4.99	2.018	2641	39.99	5.99	2.563	4963
OpenDet [19] $_{CVPR'22}$	hom-sem	45.51	5.28	1.336	1458	44.02	5.67	1.653	1798	40.87	6.58	2.303	3416
OW-DETR [18] $_{CVPR'22}$	hom-sem	43.22	3.15	1.355	1076	42.83	3.46	1.593	1320	39.45	4.38	2.384	3399
SOMA (ours)		**48.67**	**6.96**	**1.257**	**915**	**47.02**	**7.42**	**1.527**	**1232**	**43.37**	**8.42**	**2.281**	**2886**
DDETR [60] $_{ICLR'21}$	freq-dec	56.99	0.00	0.579	1240	55.02	0.00	0.835	2136	53.89	0.00	0.93	2625
PROSER [59] $_{CVPR'21}$	freq-dec	55.70	6.68	0.589	**536**	54.51	7.88	0.780	952	53.43	8.22	0.943	**1072**
OpenDet [19] $_{CVPR'22}$	freq-dec	57.28	9.35	0.519	720	54.89	10.59	0.781	1251	53.51	10.37	0.839	1470
OW-DETR [18] $_{CVPR'22}$	freq-dec	56.63	6.61	0.585	698	55.45	7.90	0.745	**930**	53.60	7.90	0.807	1105
SOMA (ours)		**59.18**	**11.41**	**0.507**	669	**56.85**	**12.47**	**0.723**	1140	**55.63**	**12.36**	**0.759**	1315
DDETR [60] $_{ICLR'21}$	freq-inc	44.72	0.00	2.862	2859	43.91	0.00	3.270	4907	41.12	0.00	3.609	8291
PROSER [59] $_{CVPR'21}$	freq-inc	44.23	2.94	2.881	1090	42.47	2.98	2.745	1866	39.11	3.01	3.119	3242
OpenDet [19] $_{CVPR'22}$	freq-inc	44.85	3.23	2.579	1700	42.92	3.30	2.741	2835	40.34	3.44	2.970	4965
OW-DETR [18] $_{CVPR'22}$	freq-inc	43.92	3.85	2.032	1377	43.01	3.99	2.219	1891	40.21	2.98	2.184	2293
SOMA (ours)		**46.62**	**8.32**	**1.452**	**733**	**47.30**	**8.43**	**1.566**	**1166**	**44.45**	**7.95**	**1.792**	**1974**





Conclusion

- A more practical problem definition, Adaptive Open-set Object Detection (AOOD) by considering both novel classes and novel scenes.
- Exploring the low-order limitation in the AOOD scenario.
- Solving AOOD with a unified motif-based framework with high-order evidence, including the structure-aware novel-class learning and transfer learning.