

CGSA: Class-Guided Slot-Aware Adaptation for Source-Free Object Detection

Boyang Dai, Zeng Fan, Zihao Qi, Meng Lou and Yizhou Yu



香港大學
THE UNIVERSITY OF HONG KONG



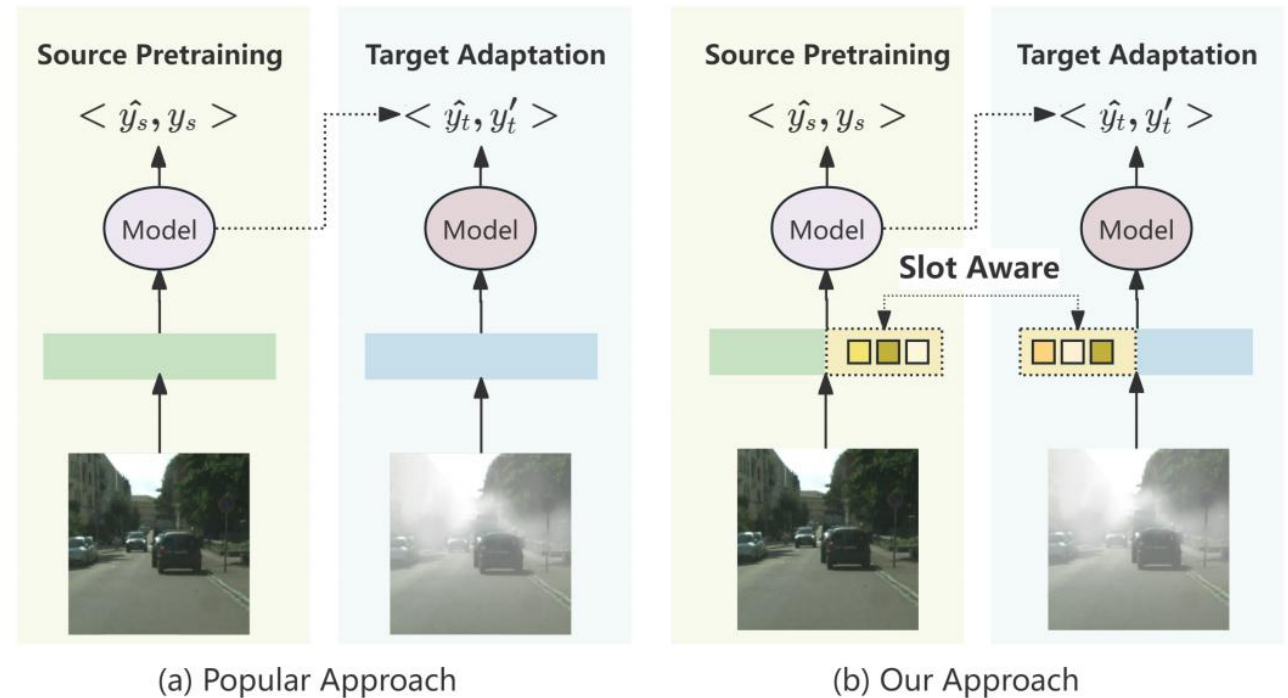
NUS
National University
of Singapore



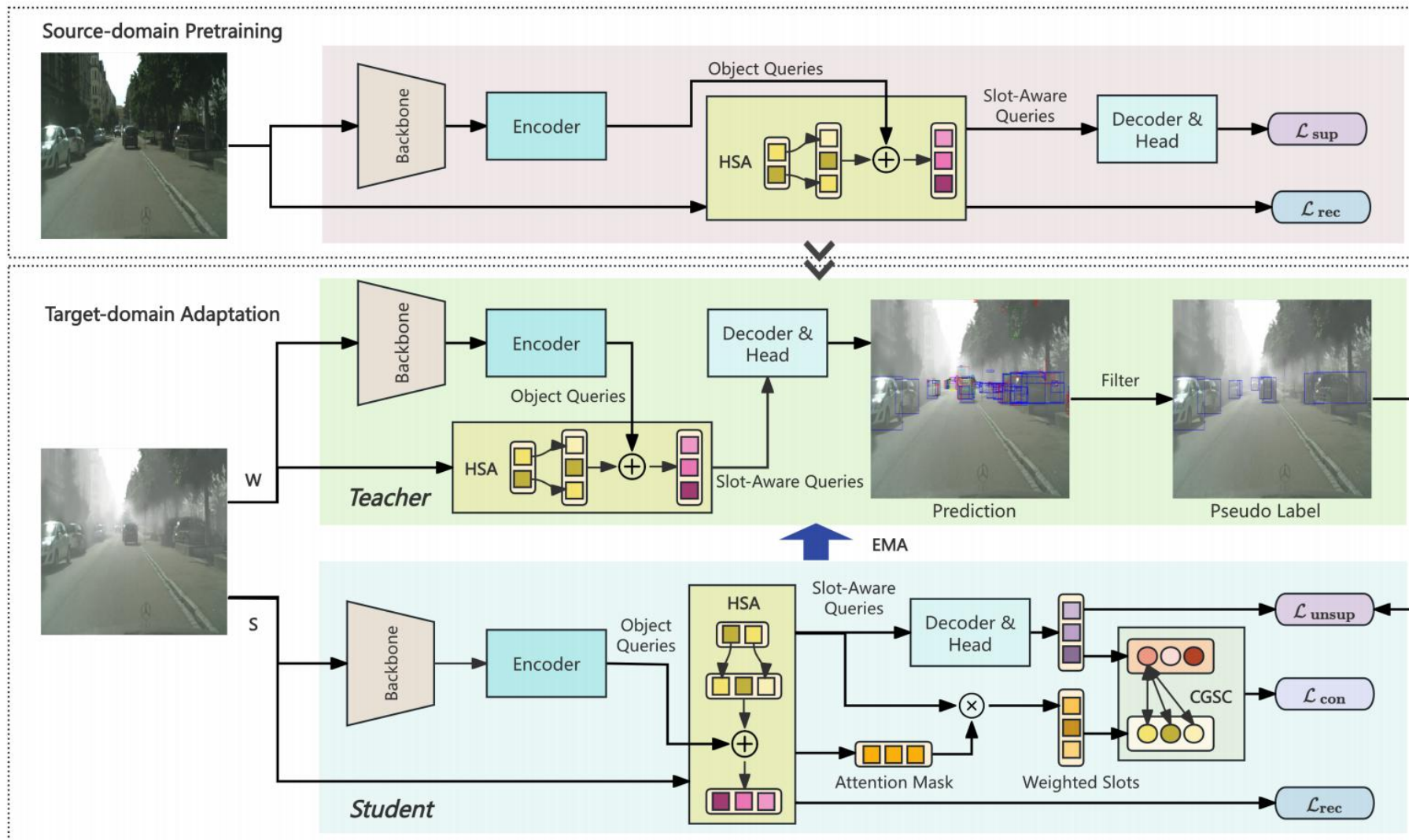
Motivation

Why SF-DAOD is hard:

- Source-Free data: can't explicitly align source/target.
- Unlabeled target domain: only rely on pseudo labels.
- Existing methods emphasize label filtering, weakening object structure considerations.



Overview

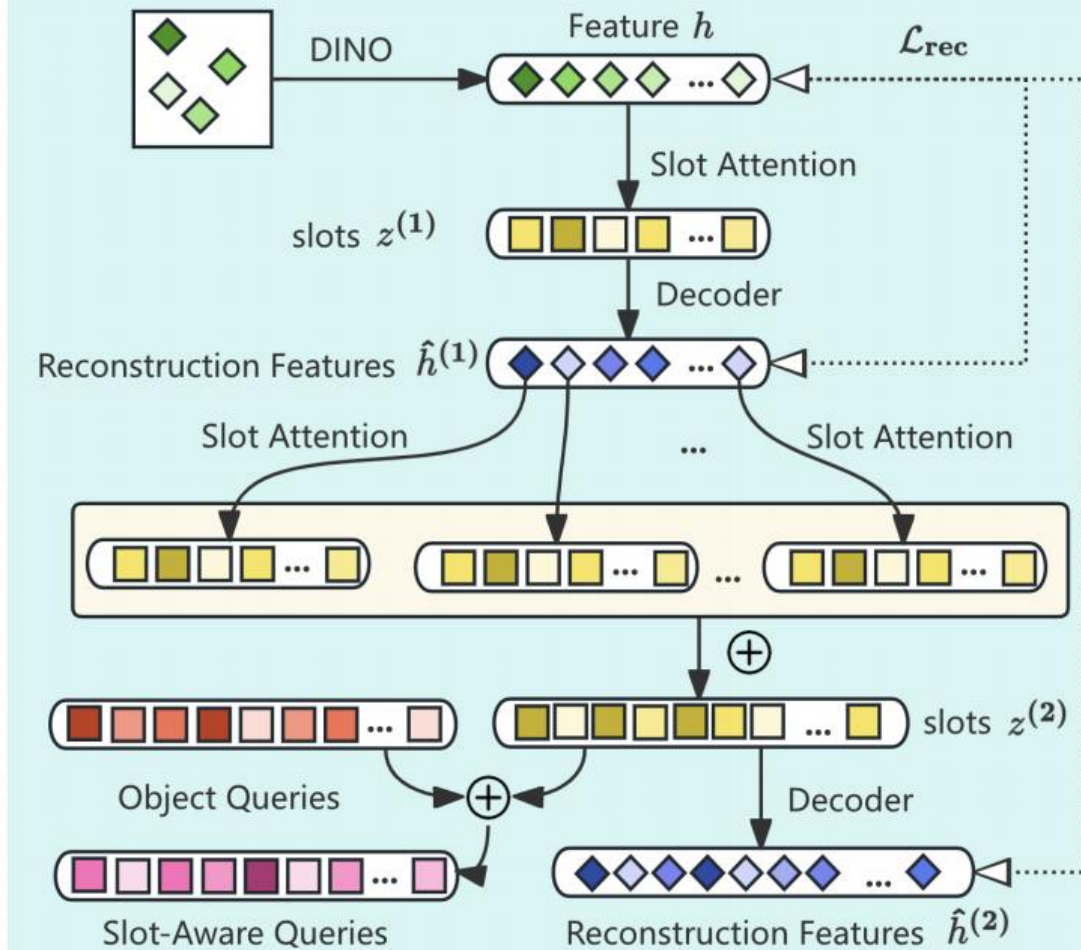


Hierarchical Slot Awareness

Key idea:

- Coarse-to-fine slot decomposition.
- Reconstruction supervision for stable slot discovery.
- Slot-aware queries = object queries + projected slots.

(a) Hierarchical Slot Awareness

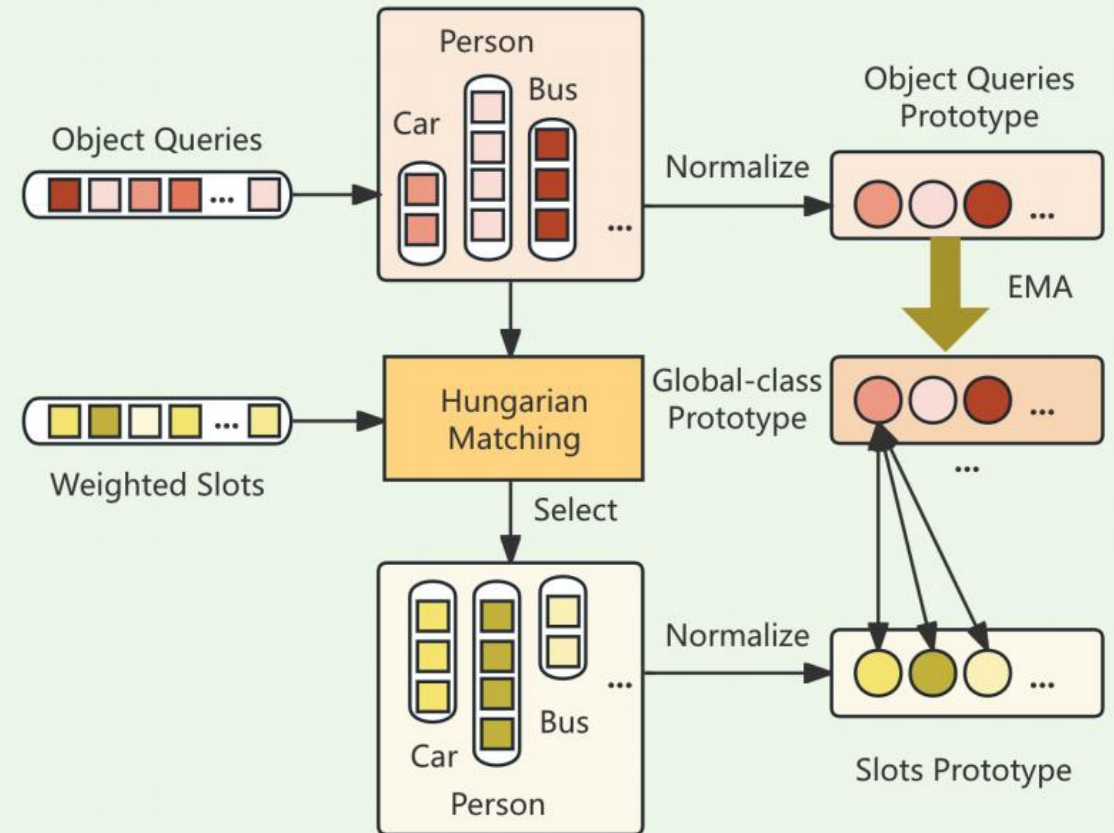


Class-Guided Slot Contrast

Key idea:

- Maintain global class prototypes from decoder queries.
- Construct weighted slots using attention masks.
- Assign pseudo-classes to slots using Hungarian matching.
- Perform prototype-slot contrastive learning.

(b) Class-Guided Slot Contrast



Main Result



➤ Adaptation from Small-Scale to Large-Scale Dataset: Cityscapes → BDD100K

Method	SF	Base	Person	Rider	Car	Truck	Bus	Mcycle	Bicycle	mAP
Oracle	-	DETR	68.2	53.2	84.0	67.7	67.9	54.1	54.7	64.3
DA-Faster(Chen et al., 2018)	✗	FRCNN	28.9	27.4	44.2	19.1	18.0	14.2	22.4	24.9
SFA(Wang et al., 2021)	✗	DETR	40.2	27.6	57.5	19.1	23.4	15.4	19.2	28.9
AQT(Huang et al., 2022)	✗	DETR	38.2	33.0	58.4	17.3	18.4	16.9	23.5	29.4
MRT(Zhao et al., 2023)	✗	DETR	48.4	30.9	63.7	24.7	25.5	20.2	22.6	33.7
DATR(Chen et al., 2025)	✗	DETR	58.5	42.8	73.4	26.9	39.9	24.2	37.3	43.3
SFOD(Li et al., 2021)	✓	FRCNN	31.0	32.4	48.8	20.4	21.3	15.0	24.3	27.6
SFOD-M(Li et al., 2021)	✓	FRCNN	32.4	32.6	50.4	20.6	23.4	18.9	25.0	29.0
PETS(Liu et al., 2023)	✓	FRCNN	42.6	34.5	62.4	19.3	16.9	17.0	26.3	31.3
A ² SFOD(Chu et al., 2023)	✓	FRCNN	33.2	36.3	50.2	26.6	24.4	22.5	28.2	31.6
TITAN(Ashraf & Bashir, 2025)	✓	DETR	49.9	35.6	65.7	24.6	35.9	31.5	29.2	38.3
CGSA(Ours)	✓	DETR	60.0	47.6	75.0	48.4	53.1	43.5	43.4	53.0

Main Result

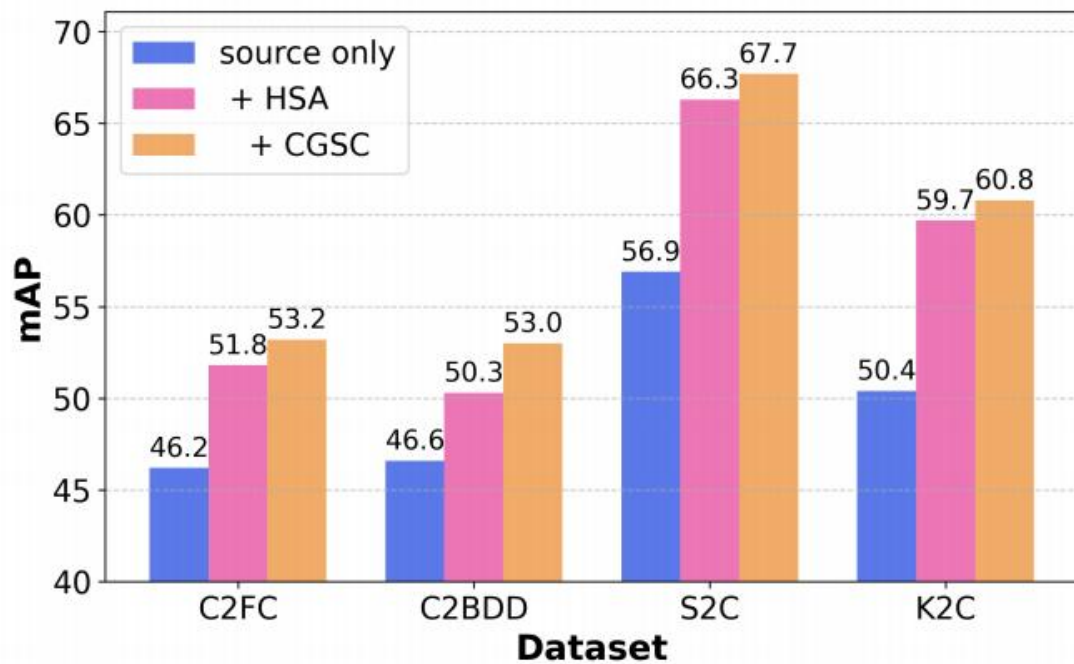


➤ Adaptation from Normal to Foggy Weather: Cityscapes → Foggy-Cityscapes

Method	SF	Base	Person	Rider	Car	Truck	Bus	Train	Mcycle	Bicycle	mAP
Oracle	-	DETR	55.6	55.8	75.2	44.8	63.5	56.1	45.6	52.2	56.1
DA-Faster(Chen et al., 2018)	✗	FRCNN	29.2	40.4	43.4	19.7	38.3	28.5	23.7	32.7	32.0
SFA(Wang et al., 2021)	✗	DETR	46.5	48.6	62.6	25.1	46.2	29.4	28.3	44.0	41.3
AQT(Huang et al., 2022)	✗	DETR	49.3	52.3	64.4	27.7	53.7	46.5	36.0	46.4	47.1
AT(Li et al., 2022b)	✗	FRCNN	43.7	54.1	62.3	31.9	54.4	49.3	35.2	47.9	47.4
MRT(Zhao et al., 2023)	✗	DETR	52.8	51.7	68.7	35.9	58.1	54.5	41.0	47.1	51.2
CAT(Kennerley et al., 2024)	✗	FRCNN	44.6	57.1	63.7	40.8	66.0	49.7	44.9	53.0	52.5
DATR(Chen et al., 2025)	✗	DETR	61.6	60.4	74.3	35.7	60.3	35.4	43.6	55.9	53.4
SFOD(Li et al., 2021)	✓	FRCNN	21.7	44.0	40.4	32.6	11.8	25.3	34.5	34.3	30.6
SFOD-M(Li et al., 2021)	✓	FRCNN	25.5	44.5	40.7	33.2	22.2	28.4	34.1	39.0	33.5
LODS(Li et al., 2022a)	✓	FRCNN	34.0	45.7	48.8	27.3	39.7	19.6	33.2	37.8	35.8
A ² SFOD(Chu et al., 2023)	✓	FRCNN	32.3	44.1	44.6	28.1	34.3	29.0	31.8	38.9	35.4
IRG(VS et al., 2023)	✓	FRCNN	37.4	45.2	51.9	24.4	39.6	25.2	31.5	41.6	37.1
PETS(Liu et al., 2023)	✓	FRCNN	46.1	52.8	63.4	21.8	46.7	25.5	37.4	48.4	40.3
LPLD(Yoon et al., 2024)	✓	FRCNN	39.7	49.1	56.6	29.6	46.3	26.4	36.1	43.6	40.9
SF-UT(Hao et al., 2024)	✓	FRCNN	40.9	48.0	58.9	29.6	51.9	50.2	36.2	44.1	45.0
TITAN(Ashraf & Bashir, 2025)	✓	DETR	52.8	51.7	68.0	43.2	65.5	41.8	46.0	48.7	52.2
CGSA(Ours)	✓	DETR	49.6	53.1	68.6	43.7	62.1	54.3	44.5	49.3	53.2

Ablation Study

➤ Components Ablation

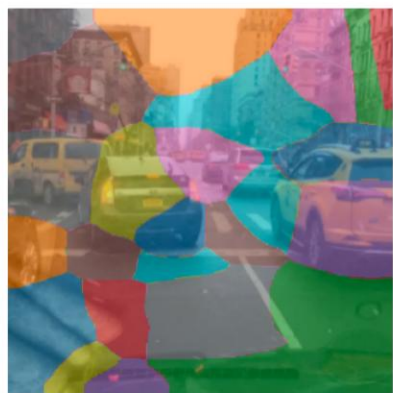


➤ Hyperparameter Sensitivity

Depth	Number	Total Slot Numbers	mAP
2	2	4	50.9
2	4	16	52.9
2	5	25	53.2
2	6	36	52.6
2	8	64	52.5
2	10	100	51.2
3	2	8	50.8
3	3	27	50.3
3	4	64	48.7

Visualization Analysis

➤ Slots:



epoch 5



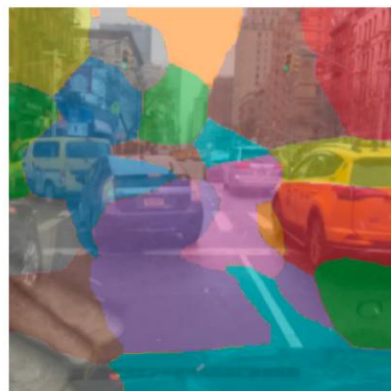
epoch 20



epoch 40



epoch 60

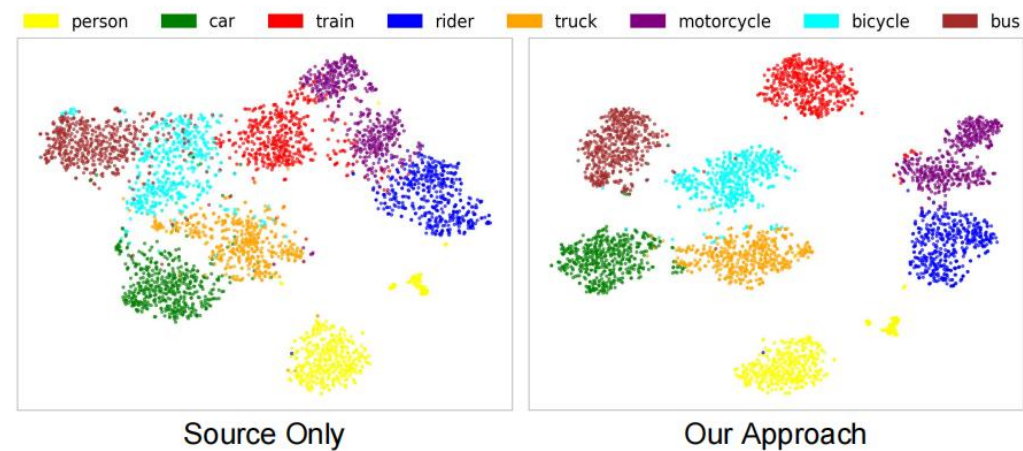


epoch 80



epoch 100

➤ Object Queries:



Conclusion



Contribution:

- First OCL-based framework for SF-DAOD. HSA supplies structural priors; CGSC provides semantic guidance.
- Consistent SOTA across diverse shifts.
- Object-centric priors are a promising direction for privacy-preserving transfer.

Future Work:

- Extend to any framework, not just limited to DETR-based detectors.
- Explore more vision tasks other than object detection.

Thank You !

More details are available in our paper and code.

<https://openreview.net/forum?id=yiF4Jo38AP>

<https://github.com/Michael-McQueen/CGSA>