Towards Unsupervised Object Detection From LiDAR Point Clouds

waabi

Lunjun Zhang, Anqi Joyce Yang, Yuwen Xiong, Sergio Casas, Bin Yang, Mengye Ren, Raquel Urtasun

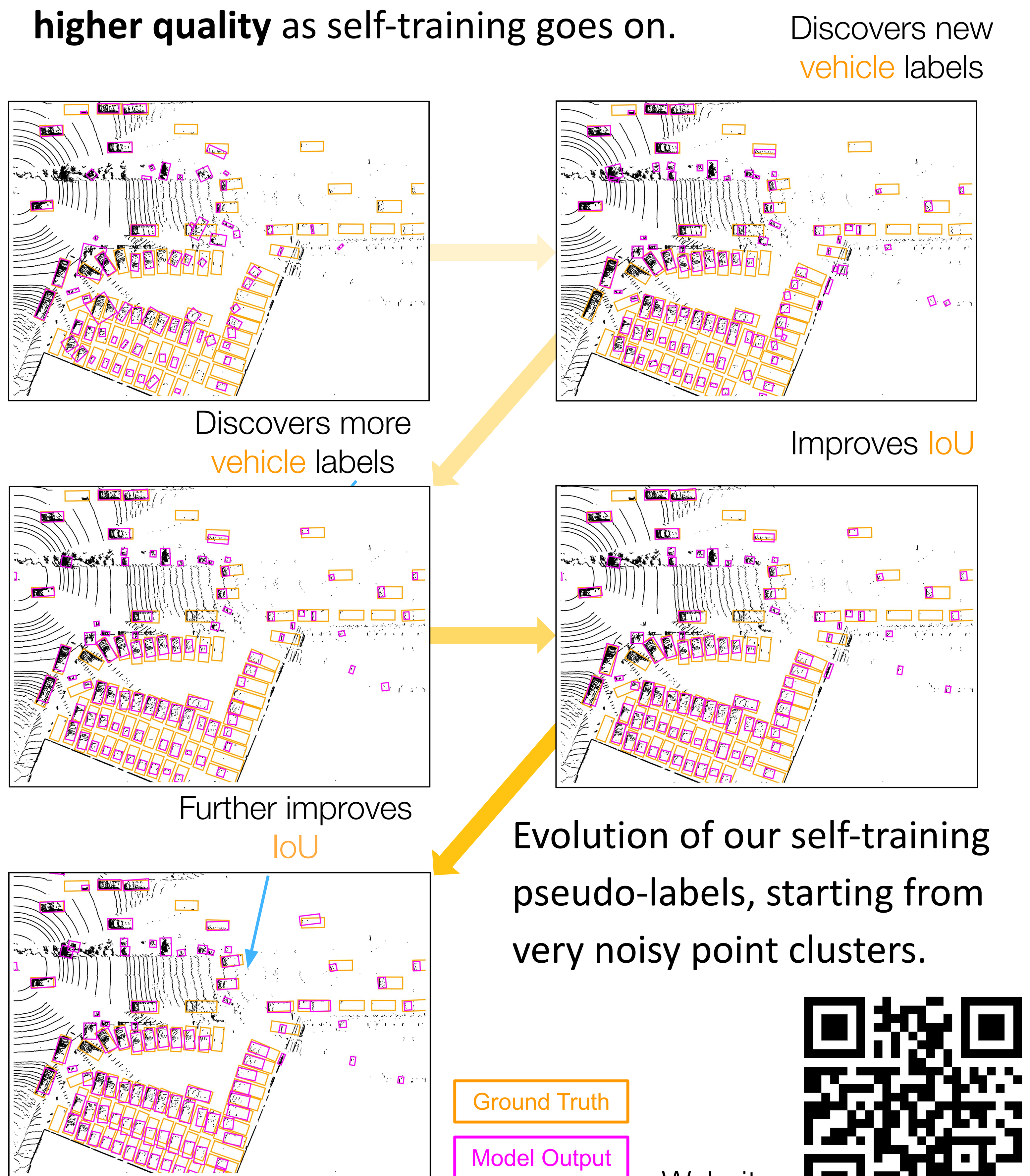
UNIVERSITY OF TORONTO

Why Unsupervised Object Detection?

- Obtaining human annotations is costly and tedious.
- Most existing visual data is **unlabeled**.
- Humans and animals learn to perceive objects without explicit labels at all.
- In this paper, we study **unsupervised object detection** from **LiDAR point clouds** in **real-world self-driving scenes**.

Intuition behind OYSTER

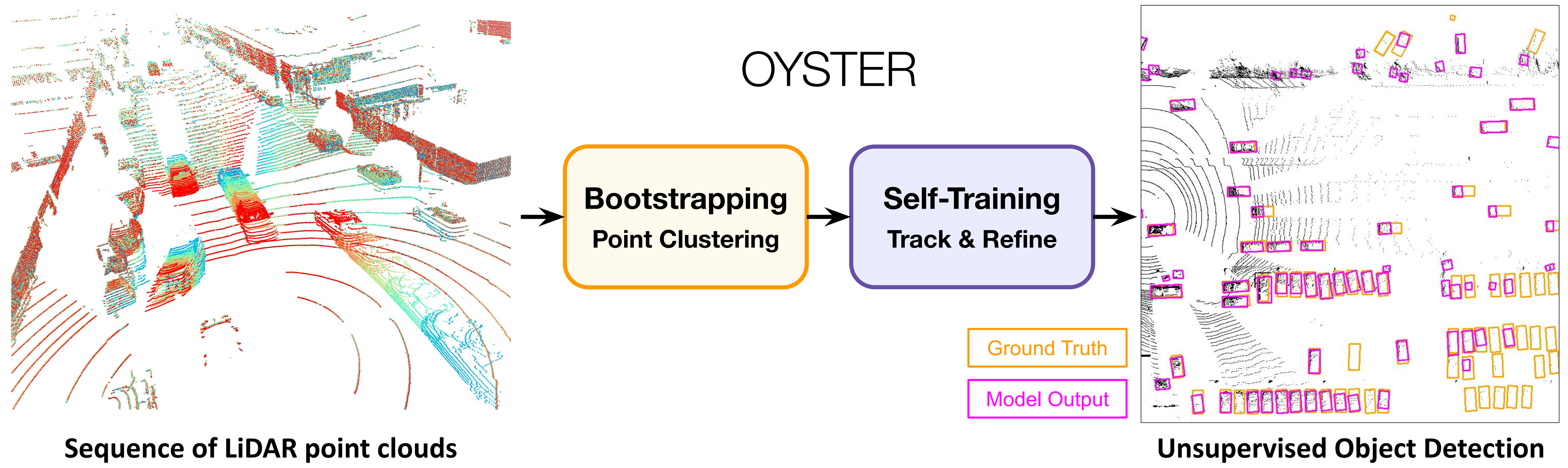
- Construct an **object discovery loop** where the detector is **iteratively re-trained** on **pseudo-labels of increasingly higher quality** as self-training goes on.



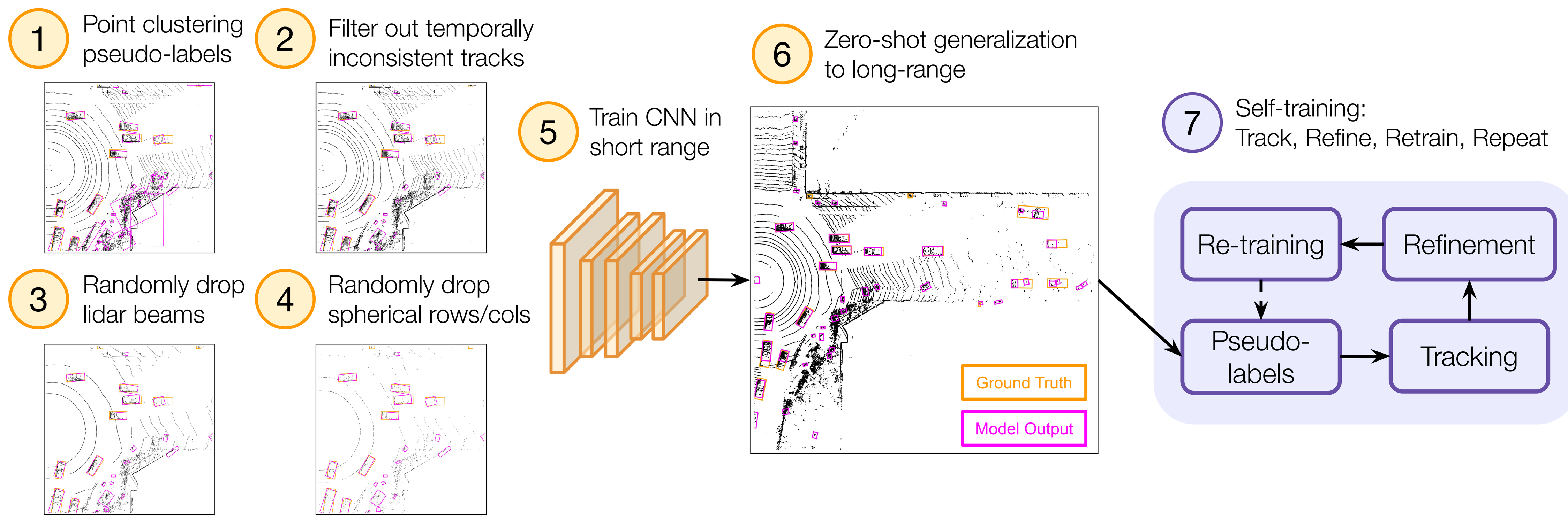
Evolution of our self-training pseudo-labels, starting from very noisy point clusters.

Website:

Object Discovery via Spatio-Temporal Refinement

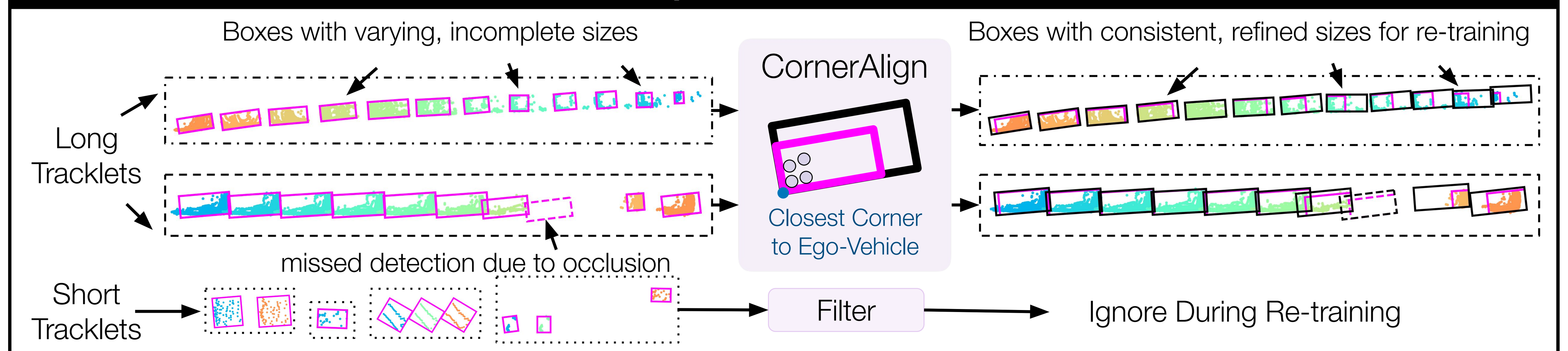


OYSTER: Bootstrapping + Self-Training



- Near-range point clustering provides initial pseudo-labels. We utilize translation equivariance of CNNs and specialized data augmentations to auto-label long range.

Track, Refine, Retrain, Repeat



- We use an unsupervised offline **tracker** to find object tracks of various lengths, discard short tracks, **refine** long tracks, **re-train** on refined pseudo-labels, **repeat**.
- Refinement uses temporal consistency of object tracks as a self-supervision signal: objects tend to stay the same 3D size and move in a temporally consistent way.

Results

	AP @ IoU			Recall @ IoU		
	0.3	0.5	0.7	0.3	0.5	0.7
DBSCAN [17]	3.5	1.1	0.3	28.6	15.9	8.3
DBSCAN + init-train	21.6	12.0	6.4	42.0	26.0	13.8
DBSCAN + self-train [66]	20.1	12.3	6.3	42.2	26.1	13.5
PP score [66]	6.4	2.0	0.4	32.8	18.7	8.8
PP score + init-train [66]	28.4	14.0	4.9	47.4	26.6	12.5
MODEST [66] (1 traversal)	22.8	7.5	2.8	49.7	28.9	14.9
Ours	**43.5**	**29.5**	**18.1**	**62.8**	**44.8**	**28.1**

PandaSet results

	AP @ IoU			Recall @ IoU		
	0.3	0.5	0.7	0.3	0.5	0.7
DBSCAN [17]	2.1	1.0	0.4	26.4	17.9	10.9
DBSCAN + init-train	15.5	10.7	5.7	29.7	19.3	10.1
DBSCAN + self-train [66]	12.5	9.1	5.5	30.1	19.6	10.3
PP score [66]	5.5	2.5	0.9	33.7	22.8	13.8
PP score + init-train [66]	24.4	15.8	7.4	43.9	27.4	14.2
MODEST [66] (1 traversal)	20.6	9.9	2.7	47.0	28.6	13.0
Ours	**35.4**	**24.5**	**12.9**	**55.1**	**37.5**	**21.4**

Argoverse2 Sensor results