Robust and Data-Efficient 3D Perception

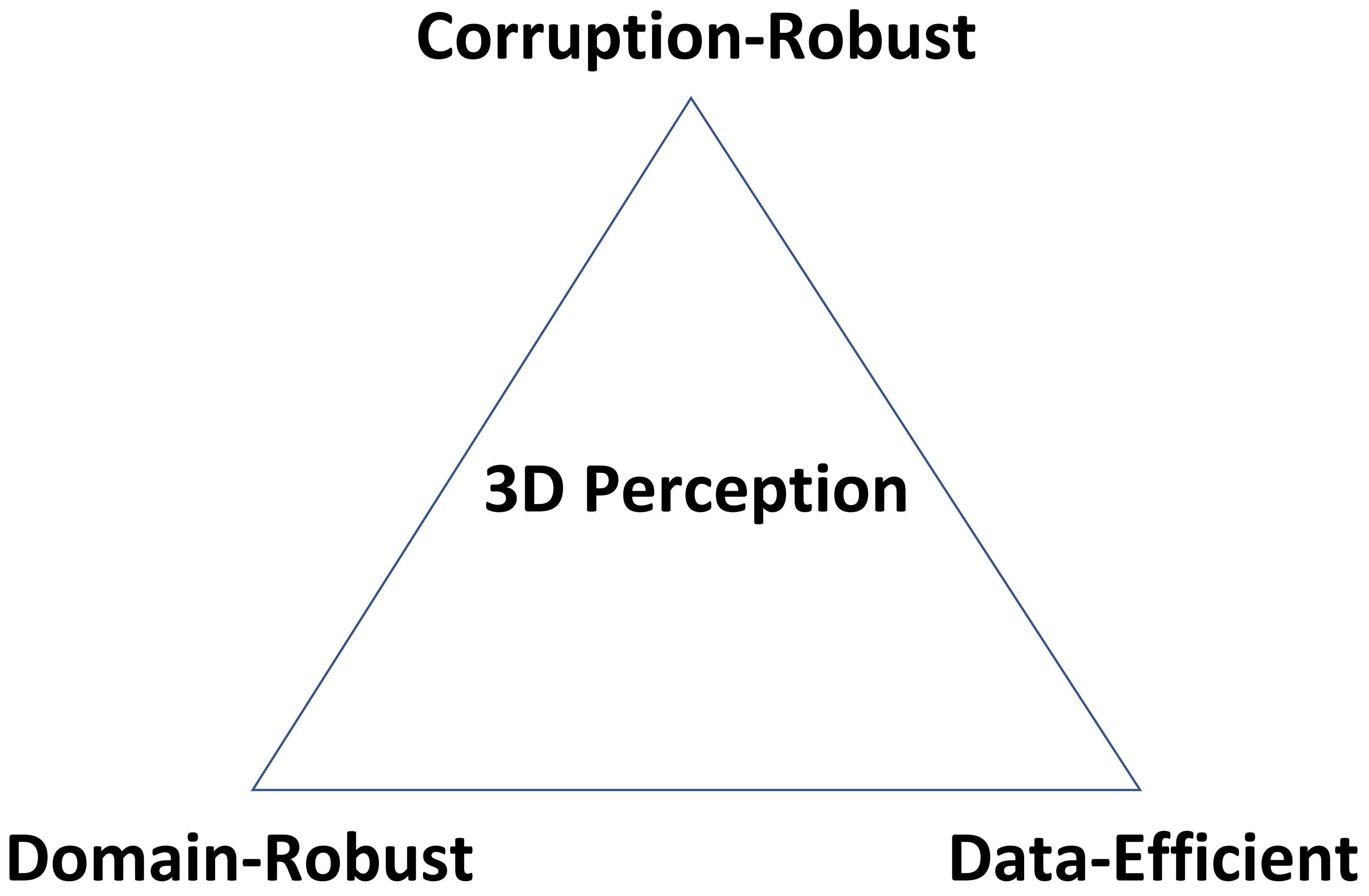
鲁棒高效的三维感知

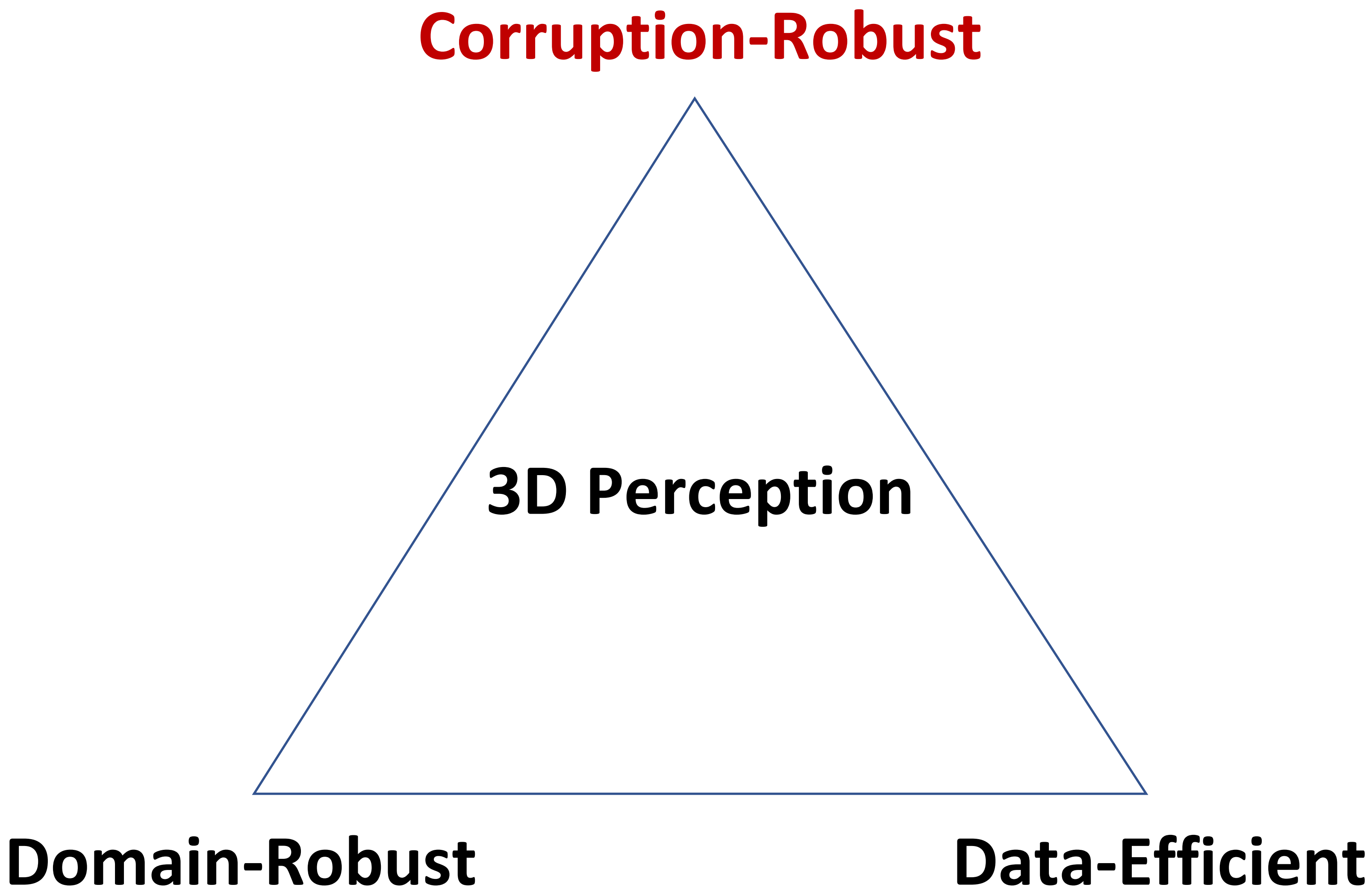
Ziwei Liu　刘子纬

Nanyang Technological University

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE | S-LAB FOR ADVANCED INTELLIGENCE


150m
50
600m
12:56


Corruption-Robust
3D Perception
Domain-Robust
Data-Efficient

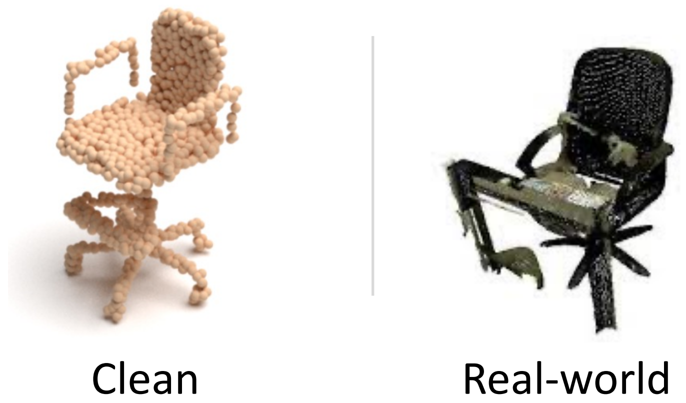

Corruption-Robust
3D Perception
Domain-Robust
Data-Efficient

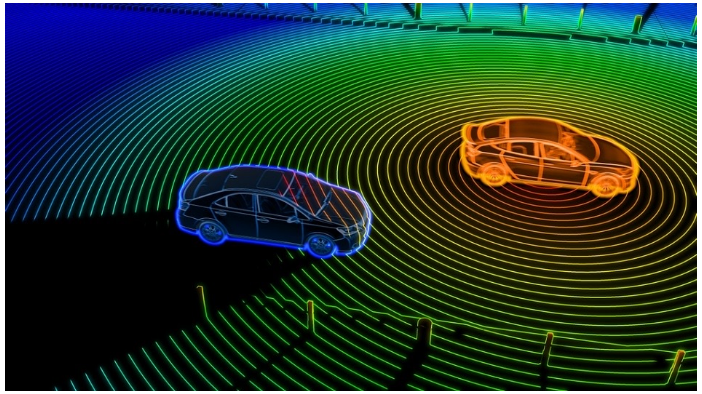
PointCloud-C: Benchmarking and Analyzing Point Cloud Perception Robustness under Corruptions

Jiawei Ren*, Lingdong Kong*, Liang Pan, and Ziwei Liu

Robustness is Crucial in Point Cloud

- Point clouds are used in **safety-critical** applications but often suffer from severe **OOD corruptions**.



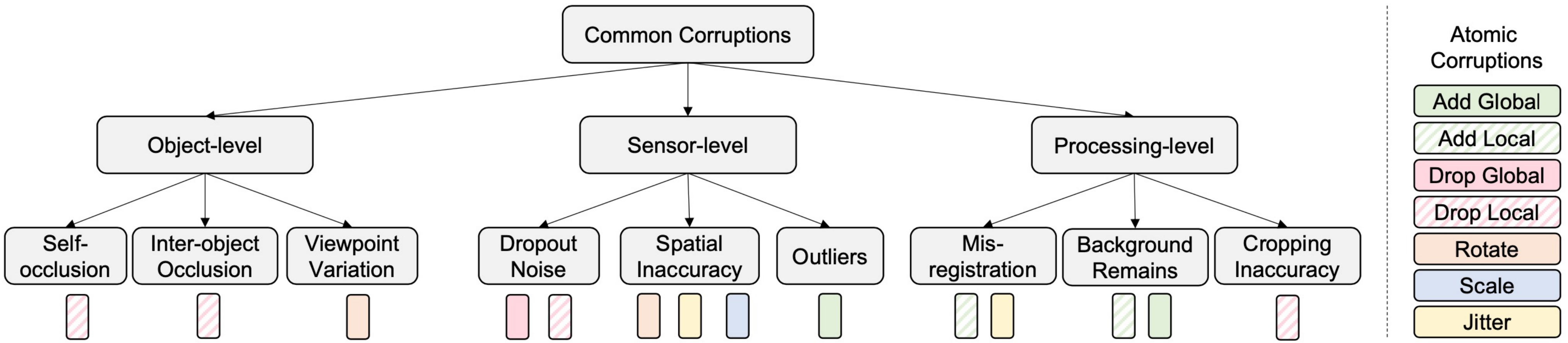


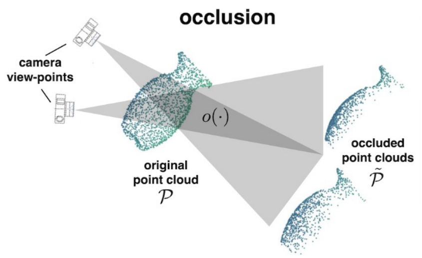
Corruptions are severe and OOD
e.g., occlusions, sensory noise

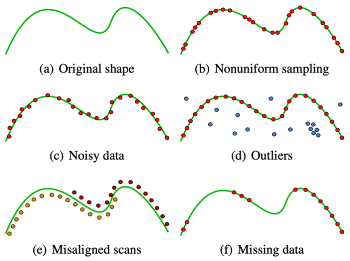
Applications are safety-critical
e.g., autonomous driving

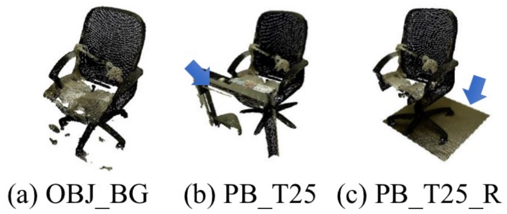
3D Sensory Data with Distribution Shifts

- **Corruptions Taxonomy**: We break down common corruptions into detailed corruption sources, and further simplify them into a combination of atomic corruptions.



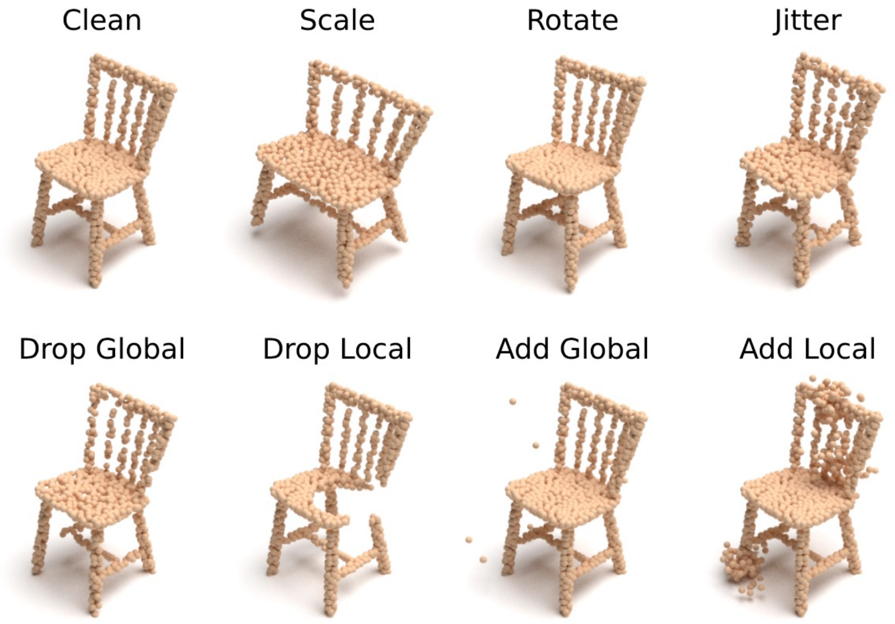




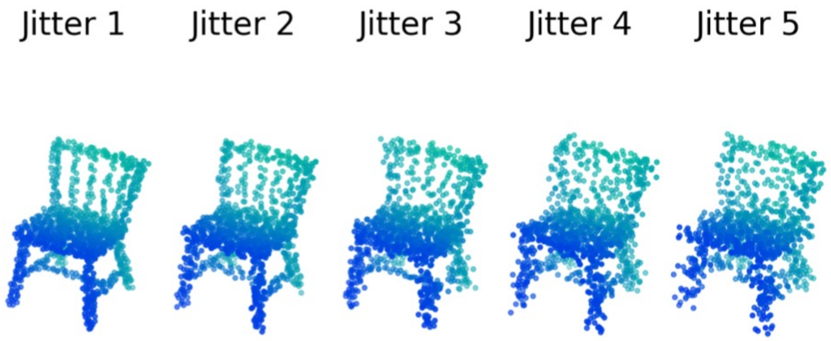


Comprehensive Benchmarking Suite

ModelNet-C: ModelNet40 is one of the most used benchmarks. We corrupt the ModelNet40 testset using the atomic corruptions with varying severities.



Atomic Corruptions



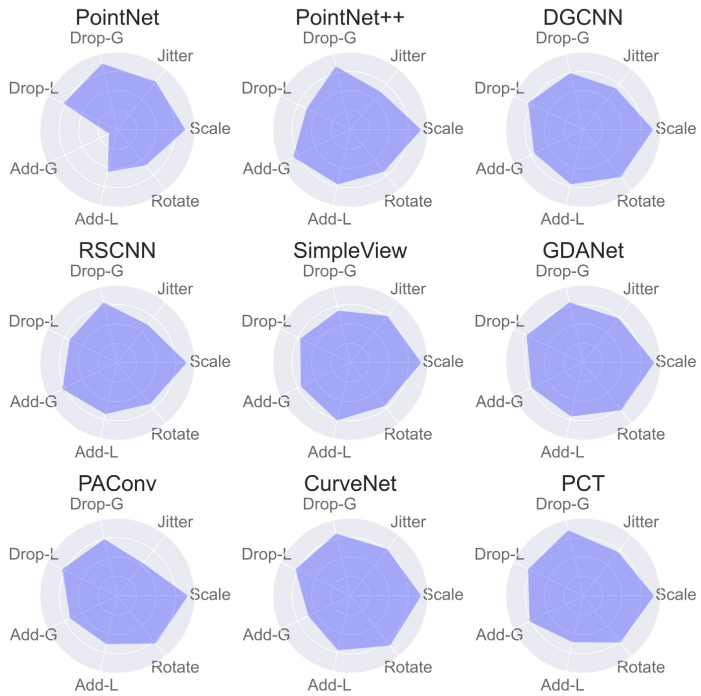
Different Severities

Evaluation Protocol

Evaluation Metrics: Inspired by the ImageNet-C, we use mean CE (mCE), as the primary metric. Compared to the commonly used Overall Accuracy (OA), mCE shows average performance under all types of corruptions.

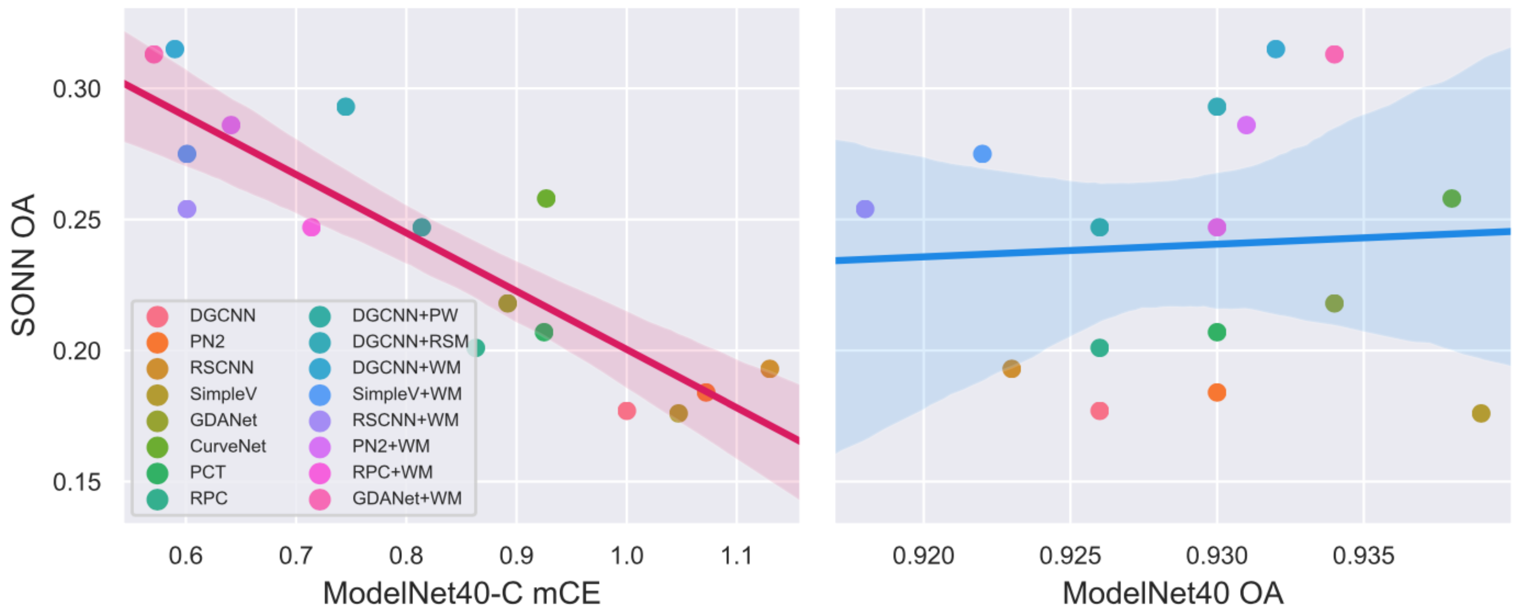
$$\mathrm{CE}_i = \frac{\sum_{l=1}^{5}(1-\mathrm{OA}_{i,l})}{\sum_{l=1}^{5}(1-\mathrm{OA}_{i,l}^{\mathrm{DGCNN}})},$$

$$\mathrm{mCE} = \frac{1}{N}\sum_{i=1}^{N}\mathrm{CE}_i$$



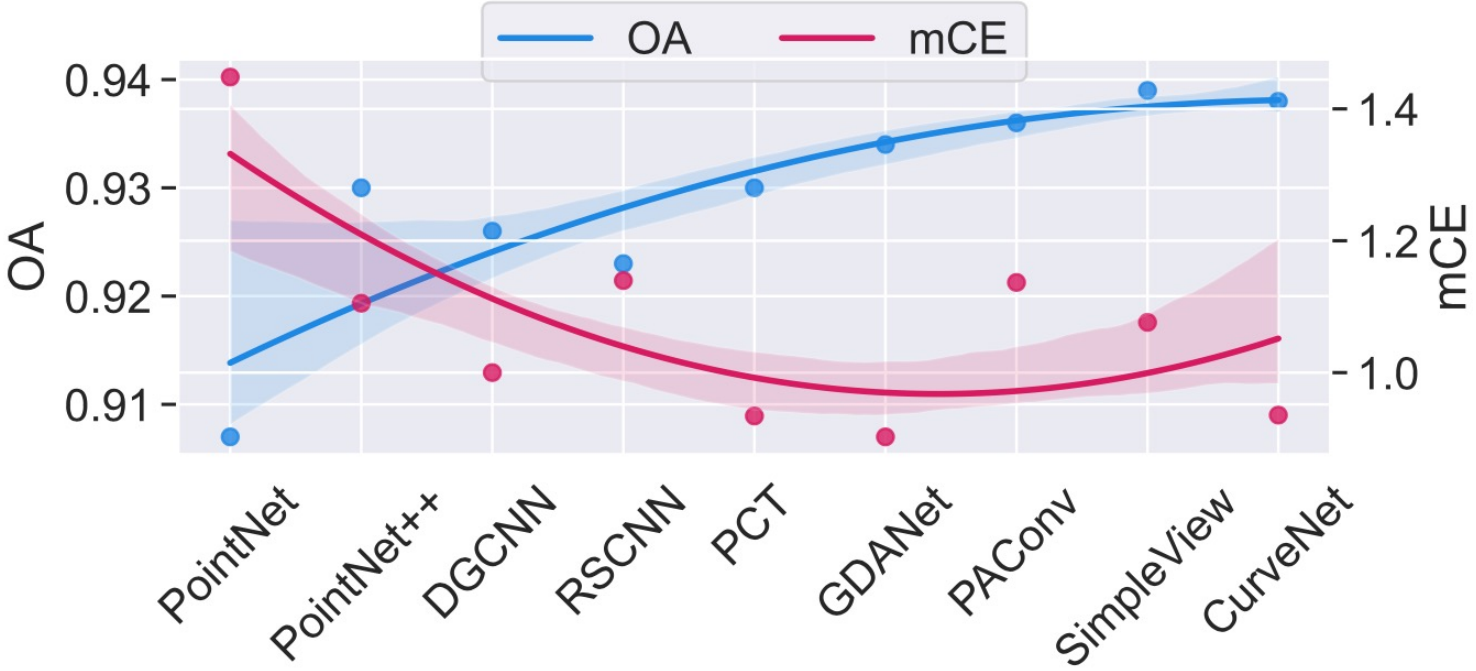
Indicative of real-world robustness?

- **Yes.** We observe that ModelNet-C mCE strongly correlates to ScanObjectNN (SONN) OA. In comparison, ModelNet40 OA has nearly no correlation to SONN OA.



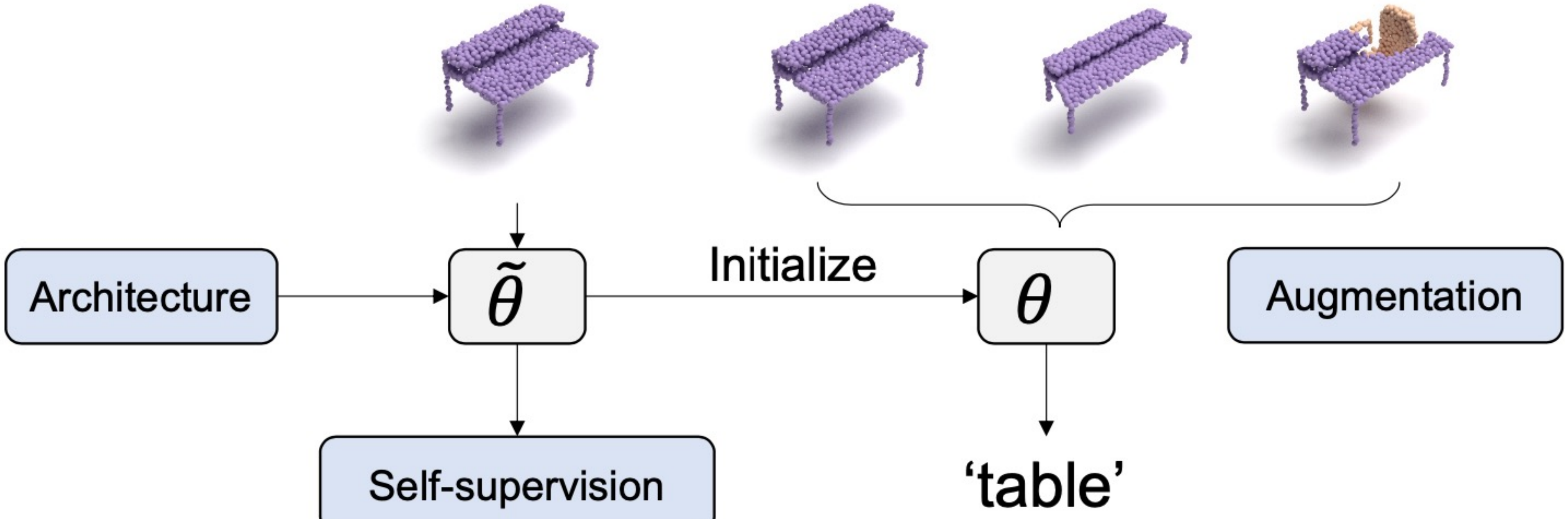
Point cloud classifier getting more robust?

- **No.** Although the accuracy on ModelNet40 gradually saturates, the robustness is at the risk of getting worse, due to the lack of a standard test suite.



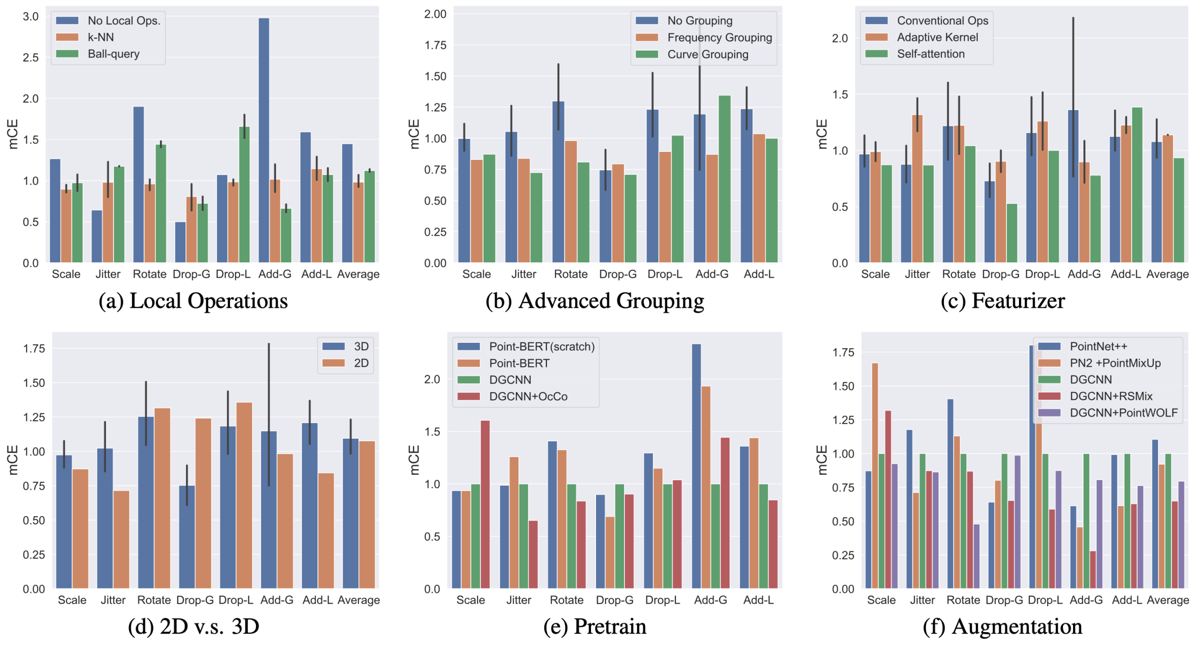
What makes a robust point cloud classifier?

- **Three main components**: 1) architecture design, 2) self- supervised pretraining 3) augmentation methods.



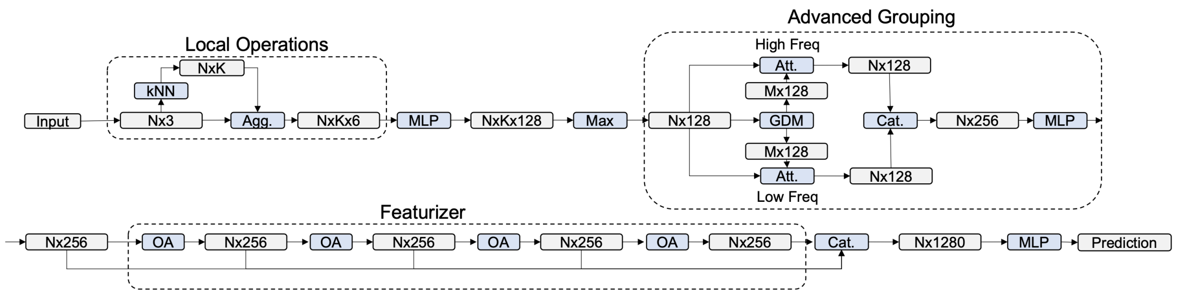
What makes a robust point cloud classifier?

- We conduct a comprehensive analysis and observe:
  - Proper architecture designs can improve robustness, e.g., advanced grouping and self-attention.
  - Pretrain signals can be transferred, benefiting robustness under specific corruptions.
  - Mixing and deformation augmentations can bring significant improvements to model robustness.

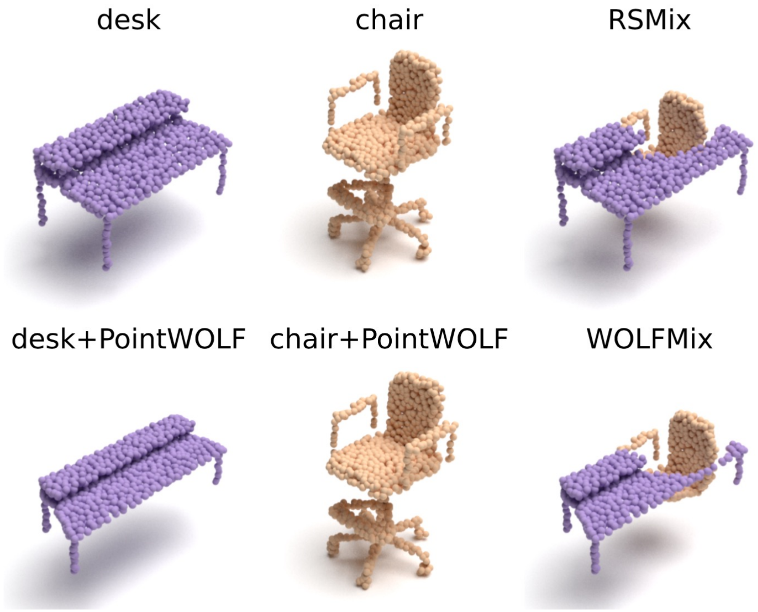


Enhancing Robustness in Point Cloud

- For verification, we propose a new architecture and a new augmentation technique strictly following our empirical findings.
- They *outperform* existing methods.



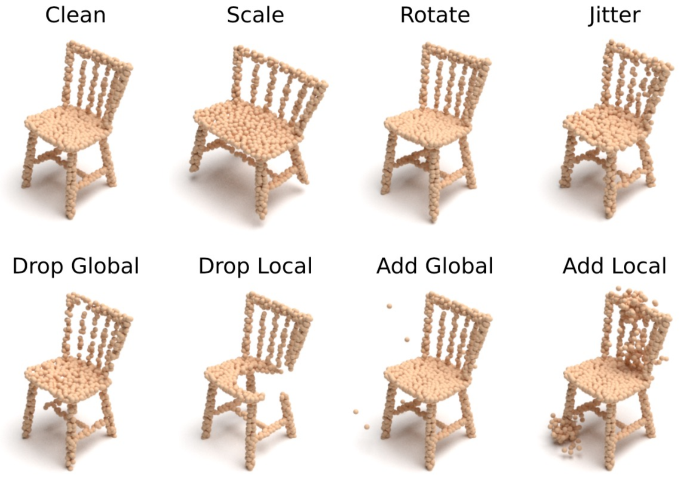
Our proposed architecture *RPC*



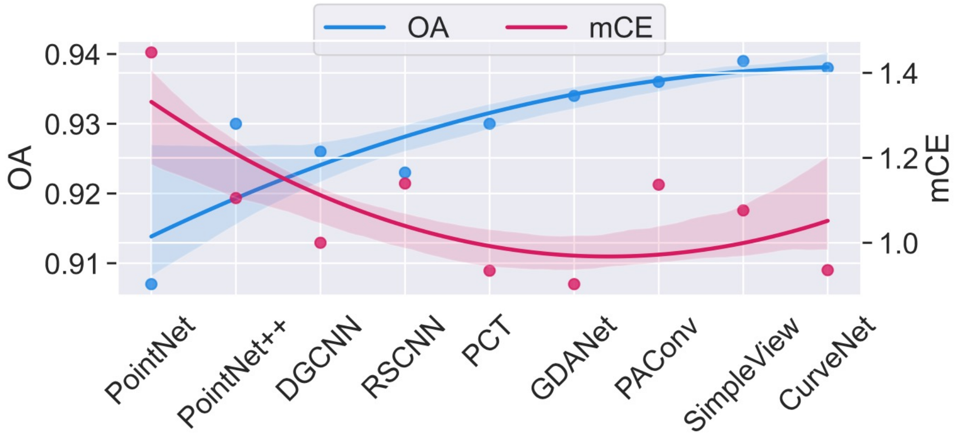
Our proposed augmentation *WolfMix*

Conclusion

- The SoTA methods for point cloud classification on clean data are becoming **less robust** to random real-world corruptions.
- We highly encourage future research to **focus on classification robustness** so as to benefit real applications.



ModelNet-C



Benchmark Result

Code, Models & Dataset

Released at https://github.com/ldkong1205/PointCloud-C

PointCloud-C

Benchmarking and Analyzing Point Cloud Perception Robustness under Corruptions

Jiawei Ren, Lingdong Kong, Liang Pan, Ziwei Liu
S-Lab, Nanyang Technological University

Paper | Project | Demo | 中译版

About

PointCloud-C is the very first test-suite for **point cloud perception robustness analysis under corruptions**. It includes two sets: ModelNet-C (ICML'22) for point cloud **classification** and ShapeNet-C (arXiv'22) for **part segmentation**.

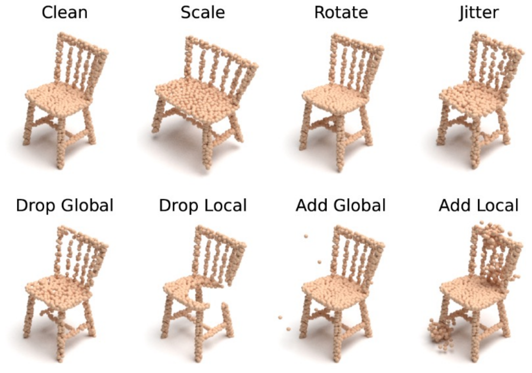
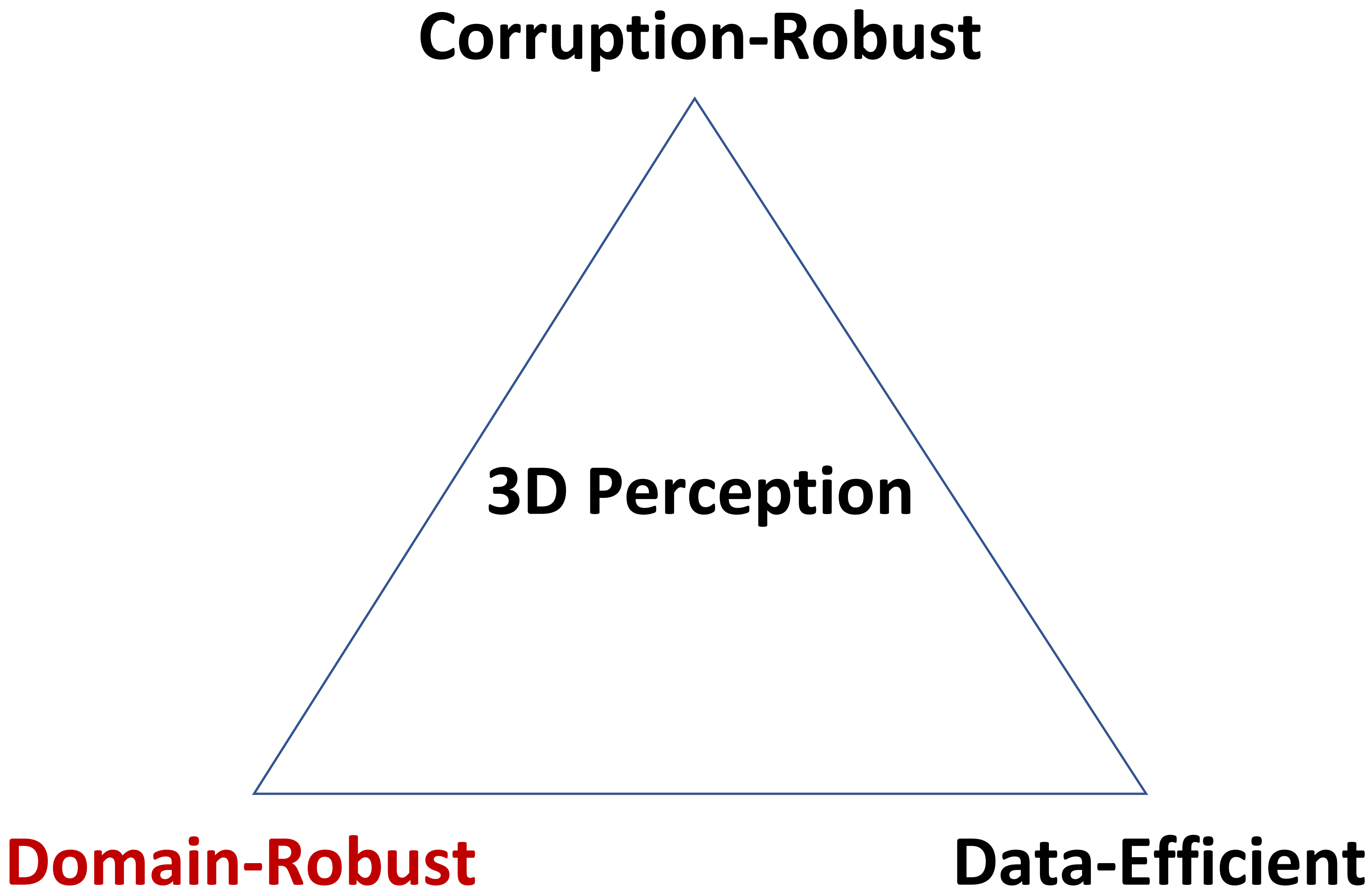


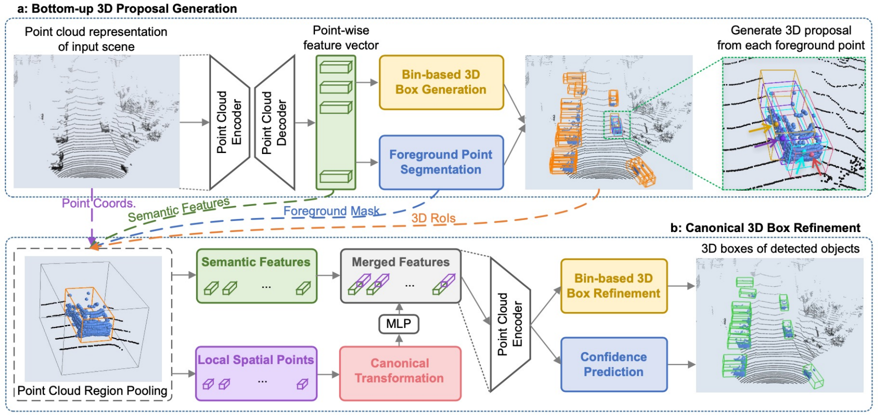
Fig. Examples of point cloud corruptions in PointCloud-C.


Corruption-Robust
3D Perception
Domain-Robust
Data-Efficient

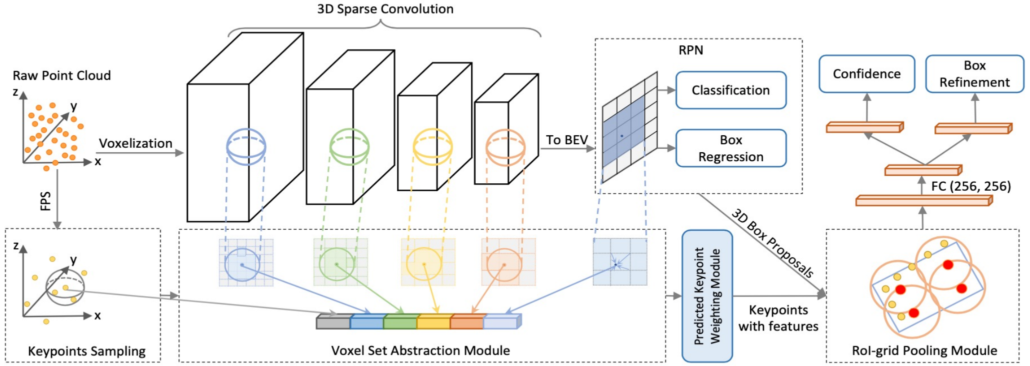
Unsupervised Domain Adaptive 3D Detection with Multi-Level Consistency

Zhipeng Luo[1,2*] Zhongang Cai[1,2,3*] Changqing Zhou[2,4*] Gongjie Zhang[4] Haiyu Zhao[2,3]
Shuai Yi[2,3] Shijian Lu[4†] Hongsheng Li[5] Shanghang Zhang Ziwei Liu[1]
[1] S-Lab, Nanyang Technological University [2] Sensetime Research [3] Shanghai AI Laboratory
[4] Nanyang Technological University [5] Chinese University of Hong Kong

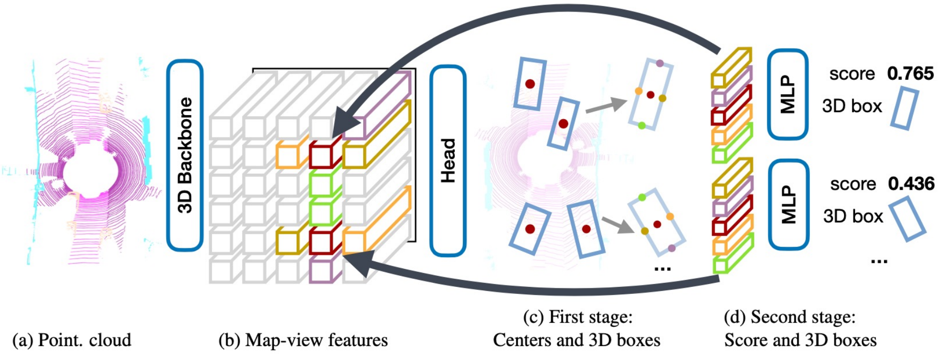
3D Object Detection



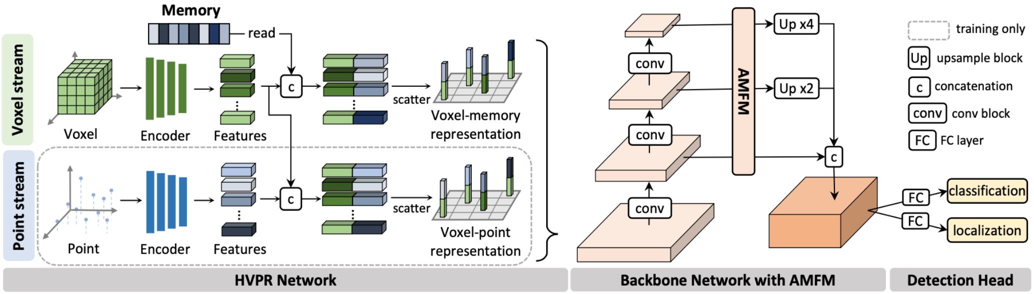
[PointRCNN CVPR2019]



[PVRCNN CVPR2020]

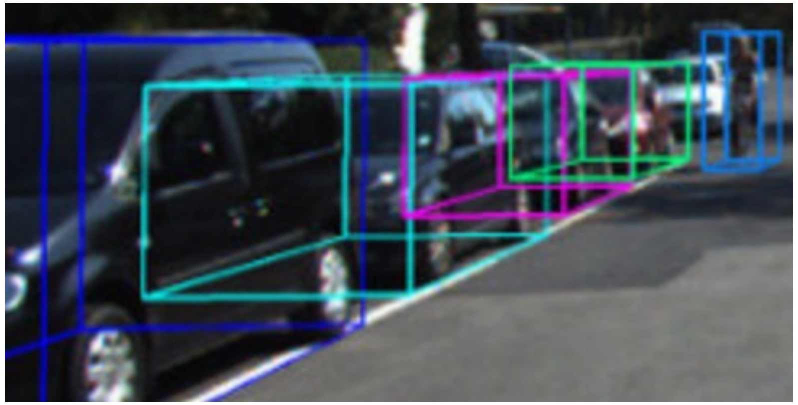


[CenterPoint CVPR2021]

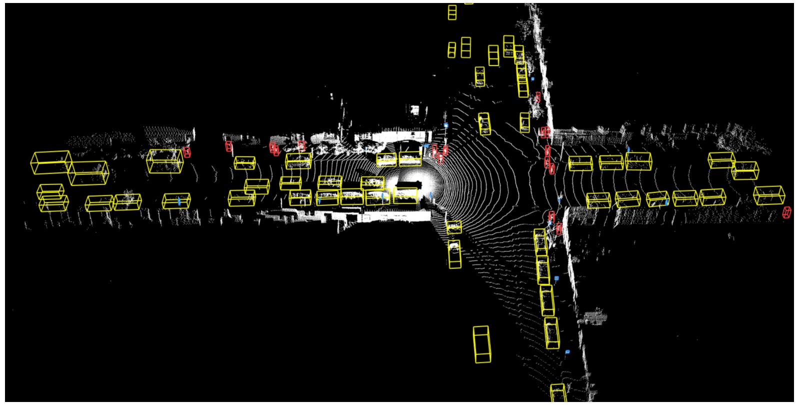


[HVPR CVPR2021]

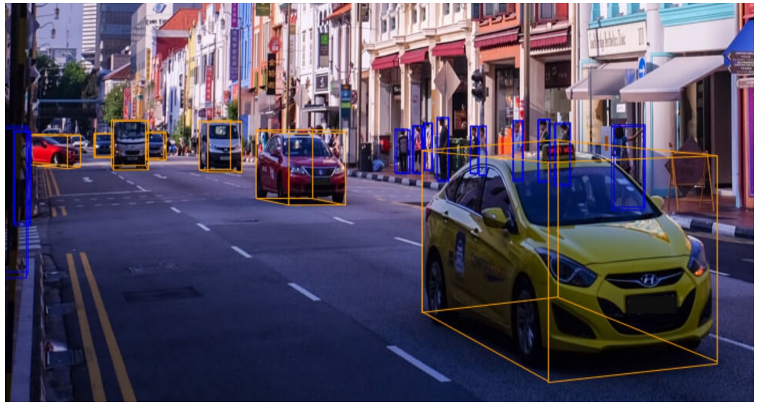
3D Object Detection Datasets



[KITTI Dataset]

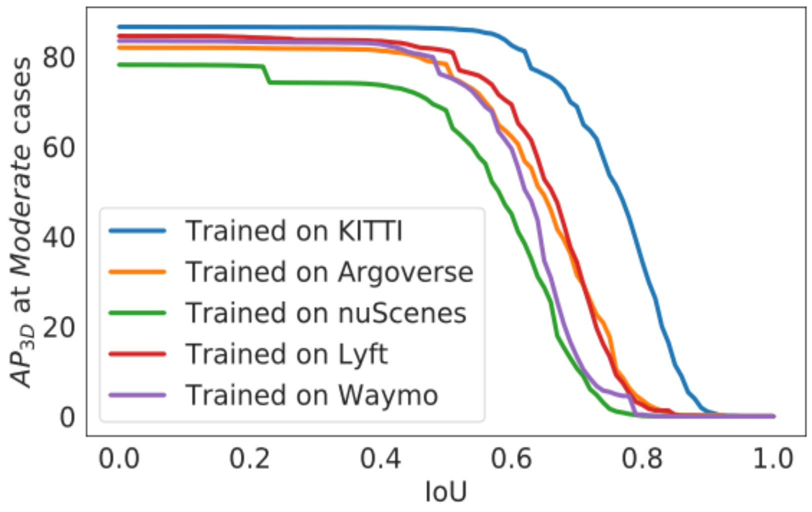


[Waymo Open Dataset]

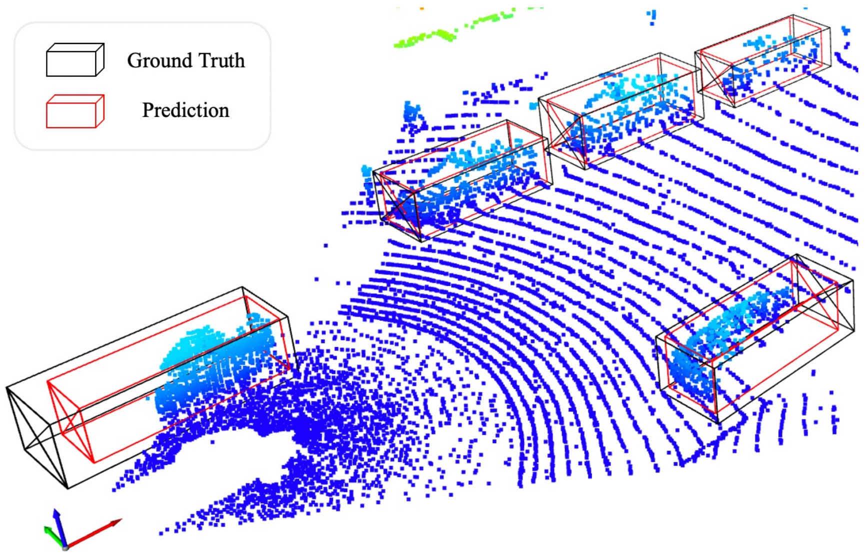


[nuScenes Dataset]

How Do Models Generalize Across Domains?



Evaluation performance on KITTI for models trained on different domains [1]

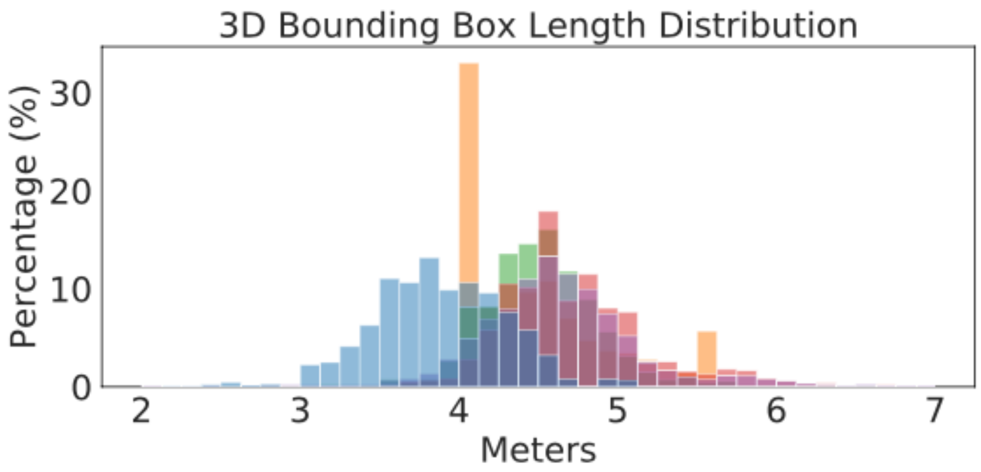


Inaccurate cross-domain predictions

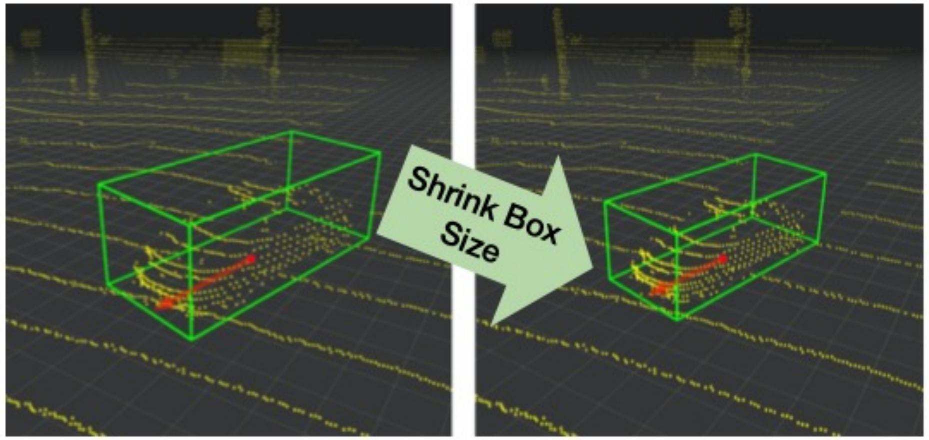
- Performance drops dramatically across domains
- Largely due to scale mismatch

[1] Wang, Yan, et al. "Train in germany, test in the usa: Making 3d object detectors generalize." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

How Do Models Generalize Across Domains?



Object scale distributions of different datasets [1]

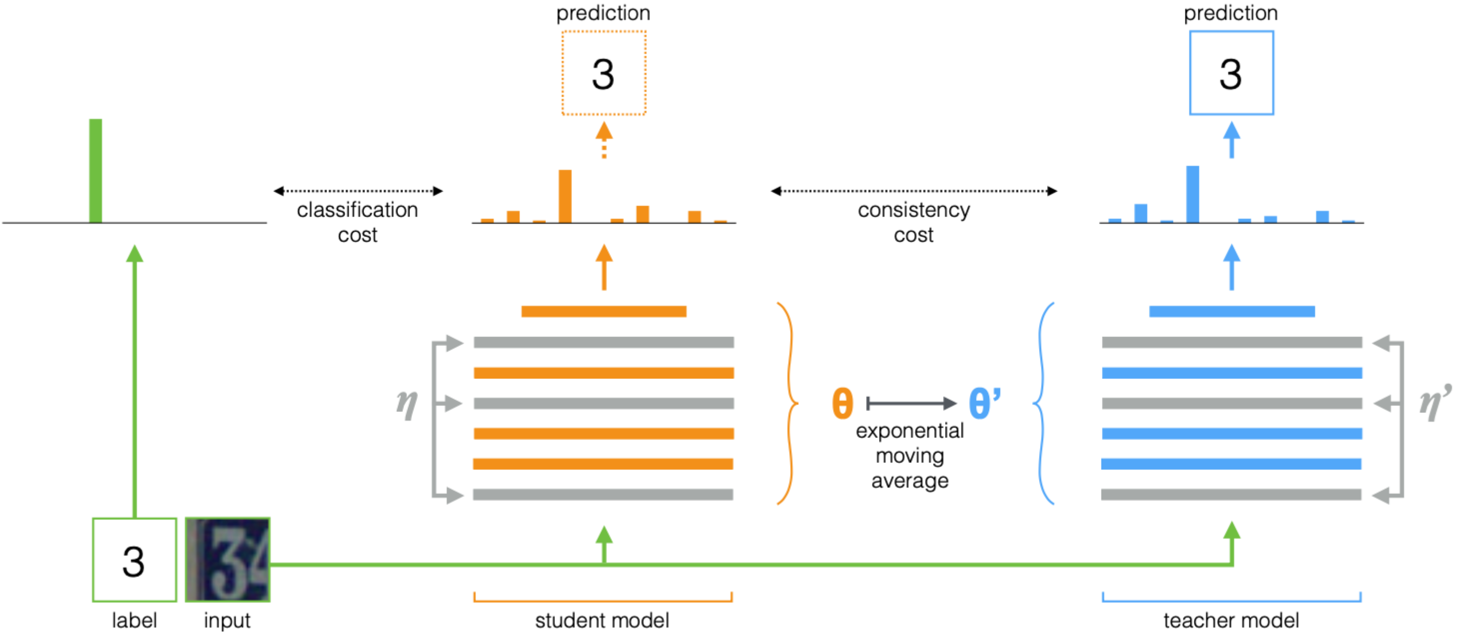


Scale rectification based on statistical information [1]

- Early study [1] mitigates scale mismatch based on statistical information
- Such information is not always available

[1] Wang, Yan, et al. "Train in germany, test in the usa: Making 3d object detectors generalize." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

The Mean Teacher Paradigm

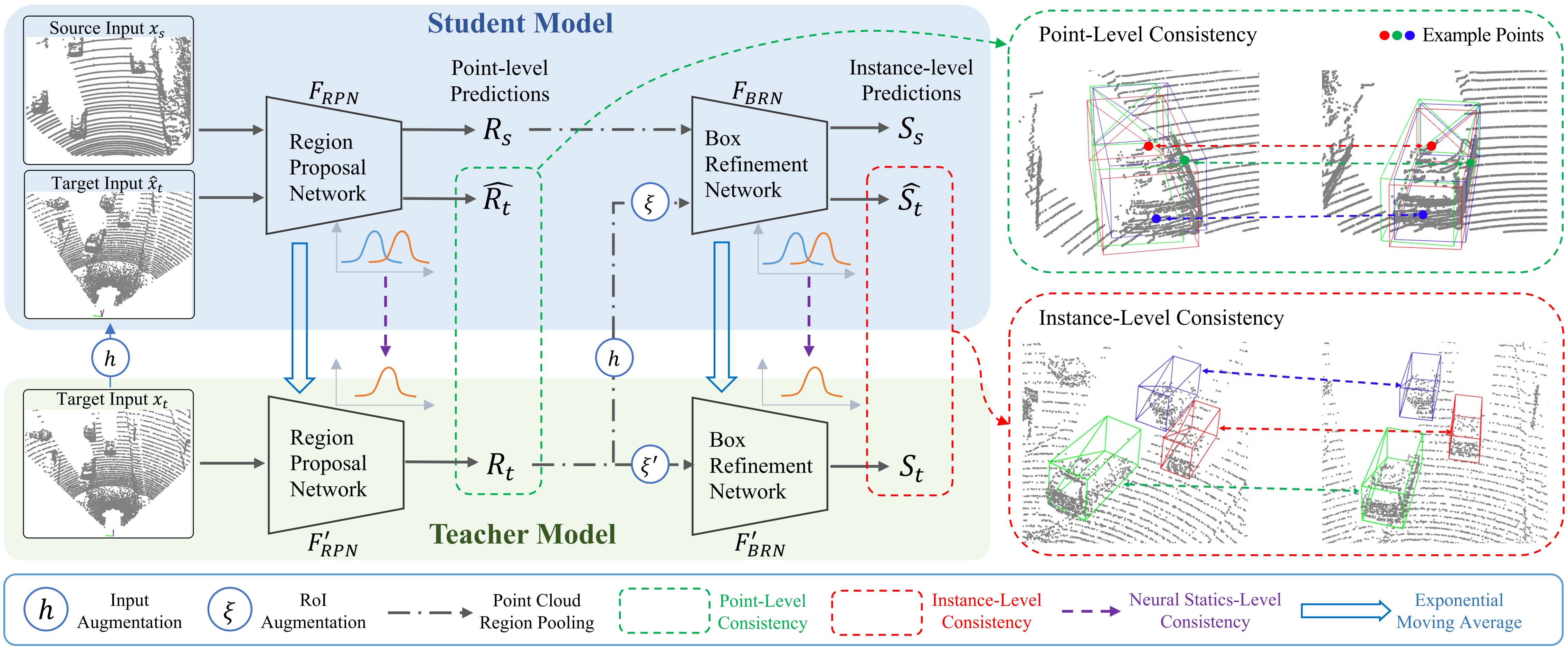


- Widely used for semi-supervised learning, self-supervised learning, domain adaptation
- Teacher model obtained from exponential average of student model

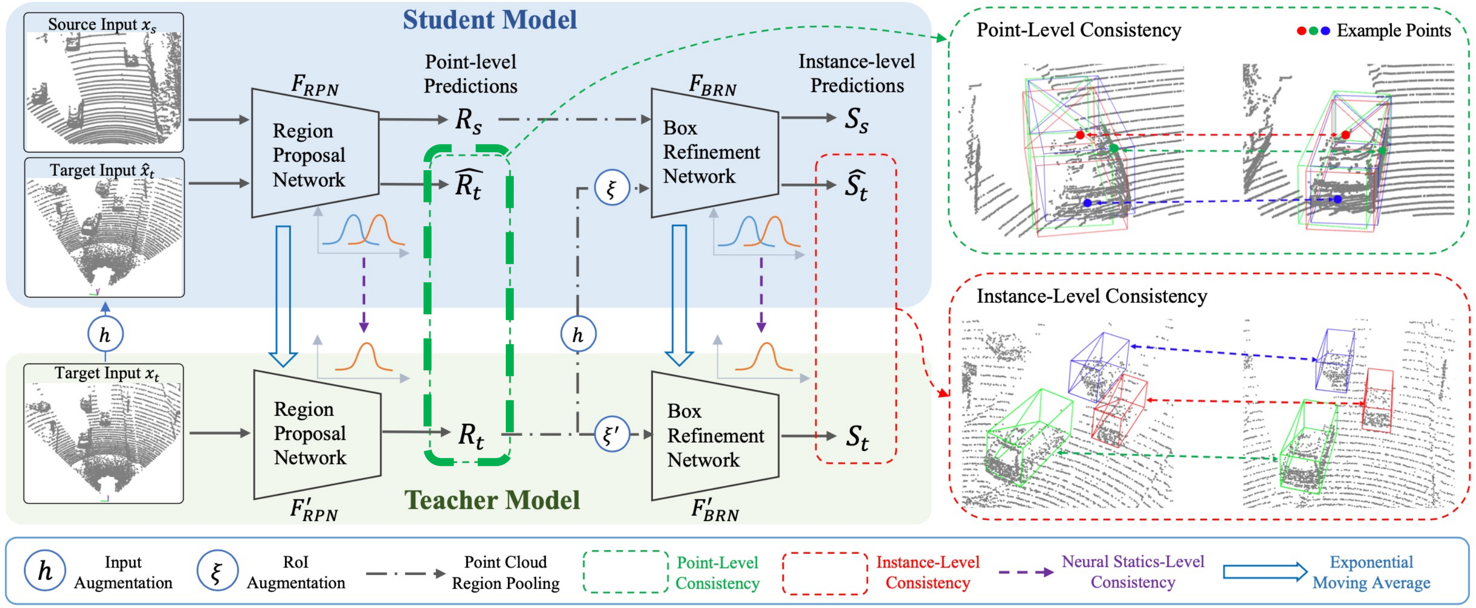
$$\theta' = m\theta' + (1-m)\theta$$

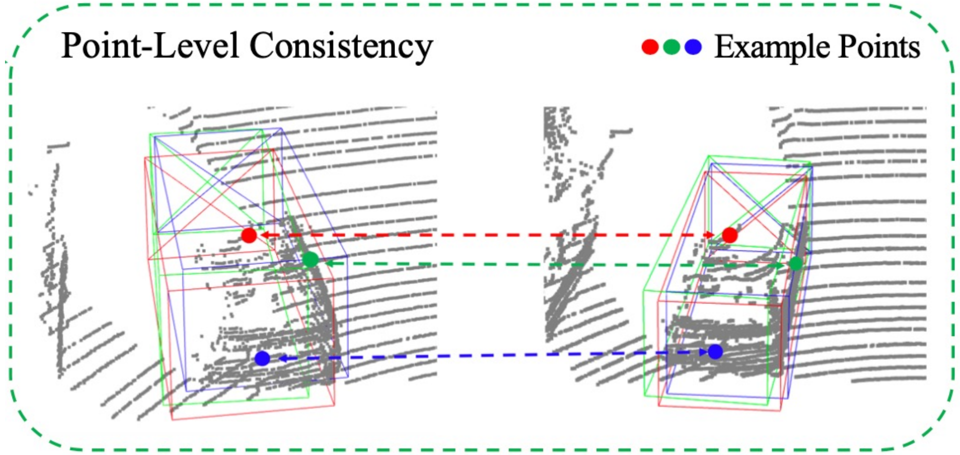
- Trained with consistency loss between student and teacher predictions

Multi-Level Consistency Network (MLC-Net)



Point-Level Consistency





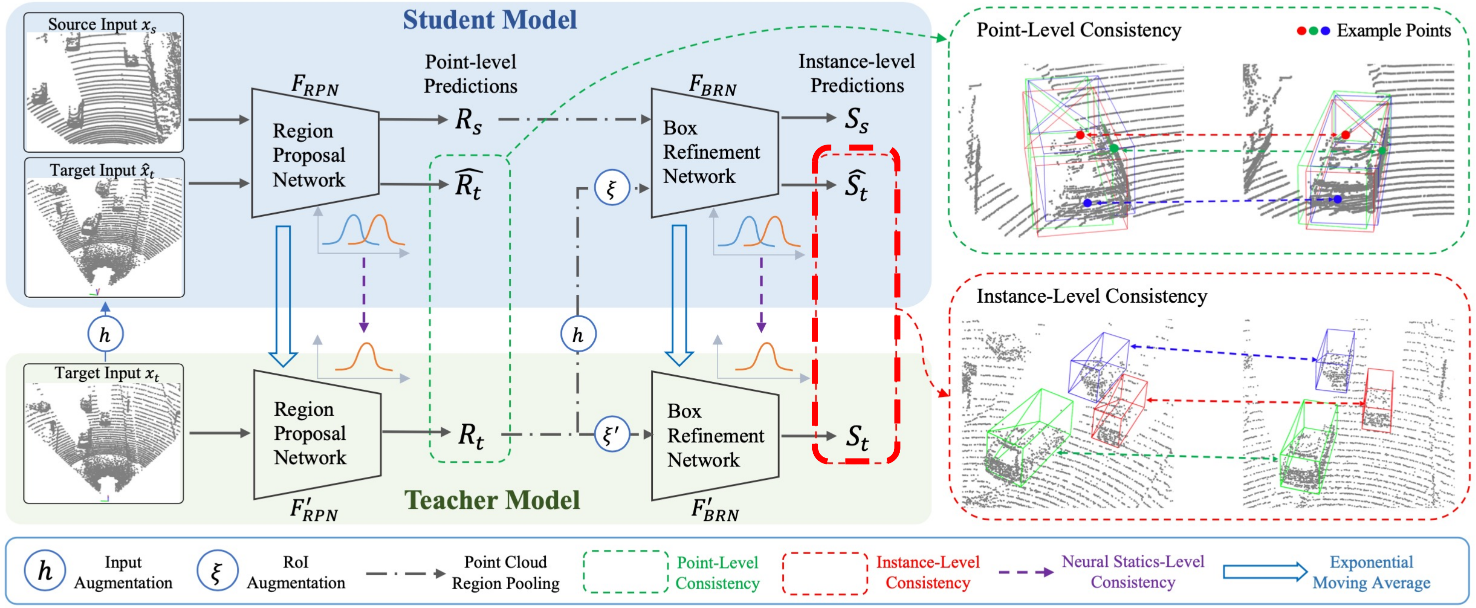
- Point correspondence remains after input augmentation
- Classification consistency between each pair of points

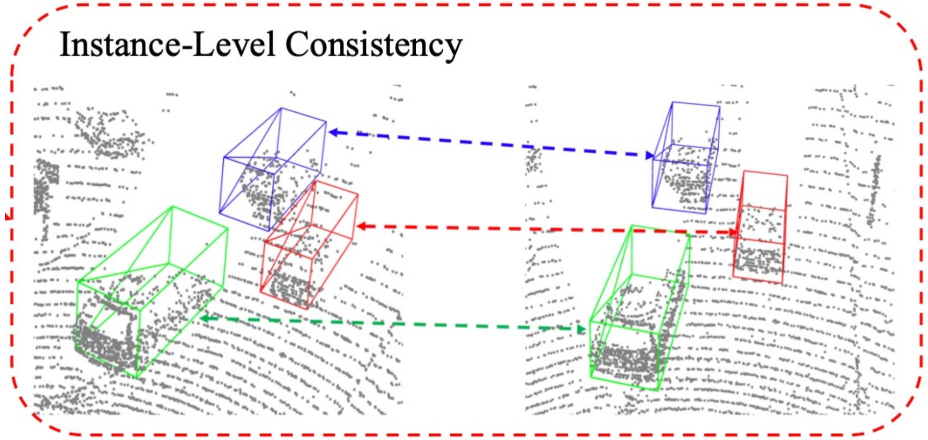
$$L_{pt,cls} = \frac{1}{|x_t|}\sum D_{KL}(\hat{R}_t^c || R_t^c)$$

- Box consistency for points belonging to the foreground

$$L_{pt,box} = \frac{1}{|\mathbb{P}_{pos}|}\sum_{p^i \in \mathbb{P}_{pos}} d(\hat{R}_t^{c(i)}, h(R_t^{c(i)}))$$

Instance-Level Consistency



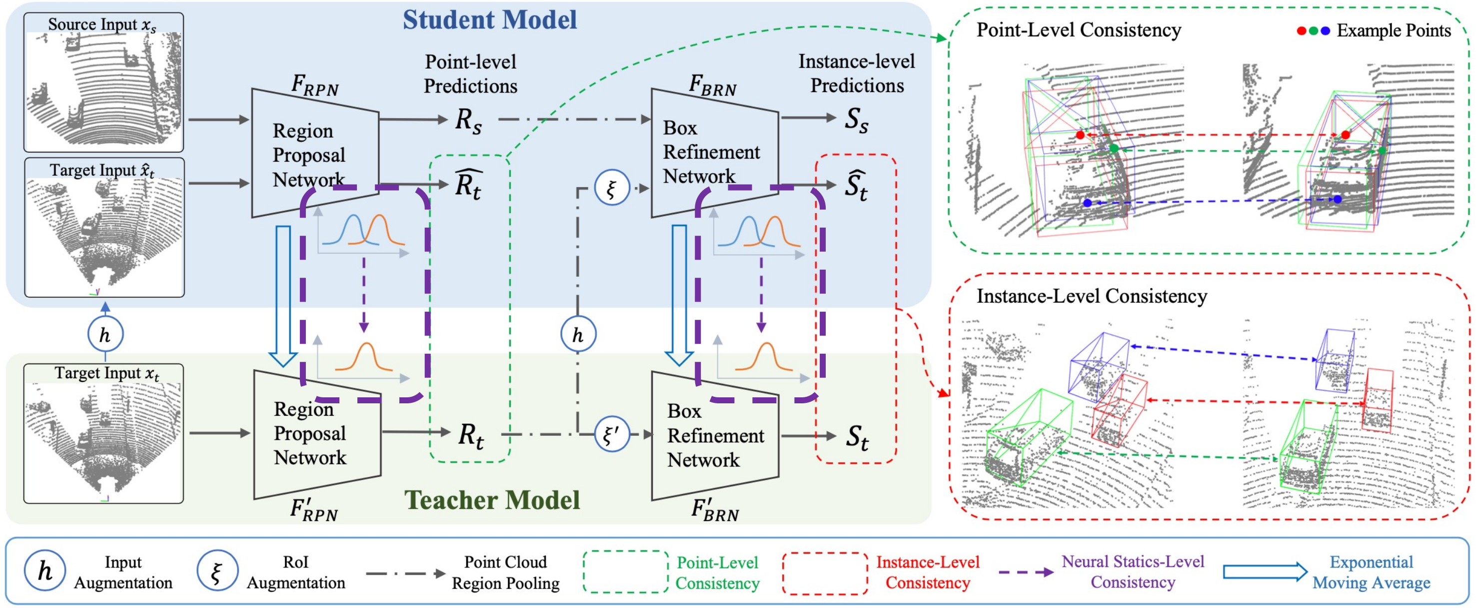


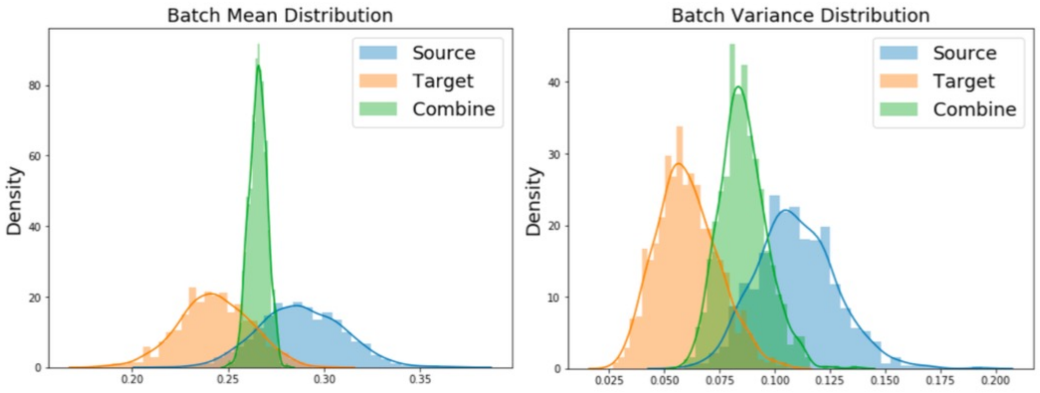
- Correspondence breaks due to proposal sampling
- Map teacher proposals to student to establish correspondence
- Apply RoI augmentation to avoid trivial solution
- Compute instance-level consistency losses

$$L_{ins,cls} = \frac{1}{|G_t|} \sum D_{KL}(\hat{S}_t^c || S_t^c)$$

$$L_{ins,box} = \frac{1}{|\mathbb{S}_{pos}|} \sum_{S_t^{(i)} \in \mathbb{S}_{pos}} d(\hat{S}_t^{b(i)}, S_t^{b(i)})$$

Neural Statistics-Level Consistency





Neural Statistical mismatch

- Significant mismatch exists in layer statistics between source and target domain
- Lead to suboptimal training behaviors
- Apply running statistics of student model to teacher model during training

$$\mu' = (1-\alpha)\mu' + \alpha\mu$$
$$\sigma' = (1-\alpha)\sigma' + \alpha\sigma$$

Experimental Results

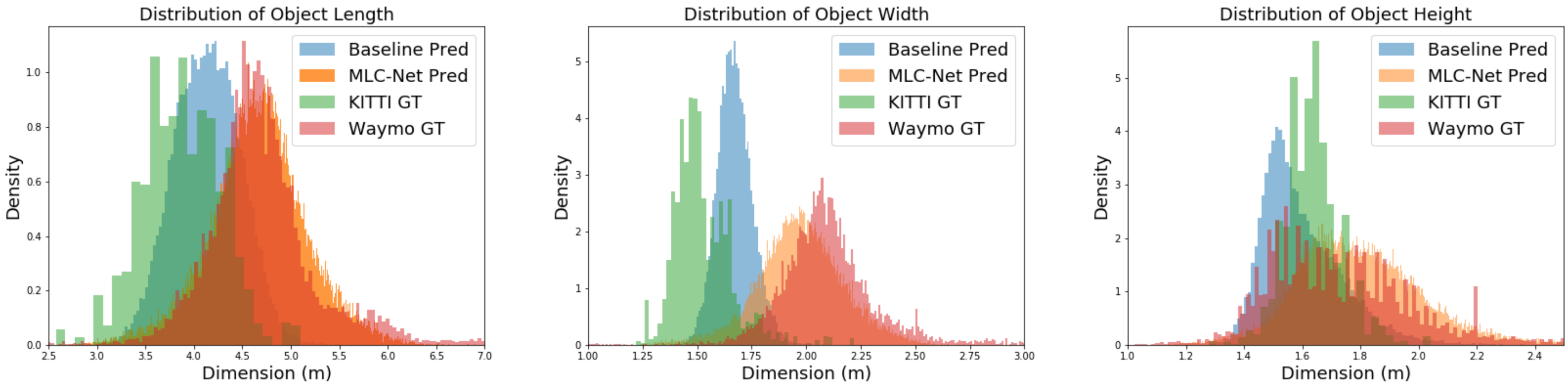
KITTI → Waymo				
Methods	AP/L1	APH/L1	AP/L2	APH/L2
Direct Transfer	9.17	8.99	7.94	7.78
Wide-Range Aug	18.61	18.18	16.77	16.40
DA-Faster	6.96	6.87	6.42	6.33
Output Transform	26.48	25.84	23.85	23.29
Statistical Norm	30.69	30.06	27.23	26.67
Ours	**38.21**	**37.74**	**34.46**	**34.04**

Waymo → KITTI			
Methods	Easy	Moderate	Hard
Direct Transfer	20.22	21.43	20.49
Wide-Range Aug	30.23	31.49	32.85
DA-Faster	4.42	5.55	5.53
Output Transform	39.78	37.82	39.55
Statistical Norm	61.93	58.07	**58.44**
Ours	**69.35**	**59.44**	56.29

KITTI → nuScenes				
Methods	ATE	ASE	AOE	AP^{3D}
Direct Transfer	0.207	0.248	0.212	13.01
Wide-Range Aug	0.200	0.228	0.211	16.01
DA-Faster	0.247	0.253	0.292	10.77
Output Transform	0.207	0.220	0.212	14.67
Statistical Norm	0.227	**0.168**	0.368	23.15
Ours	**0.197**	0.179	**0.197**	**23.47**

nuScenes → KITTI			
Methods	Easy	Moderate	Hard
Direct Transfer	49.13	39.56	35.51
Wide-Range Aug	58.71	45.37	43.03
DA-Faster	52.25	40.62	35.90
Output Transform	23.13	27.26	29.10
Statistical Norm	44.81	45.15	47.60
Ours	**71.26**	**55.42**	**48.99**

Effectiveness of scale distribution rectification



Scale distribution comparison of models trained on KITTI and tested on Waymo

Ablation Studies

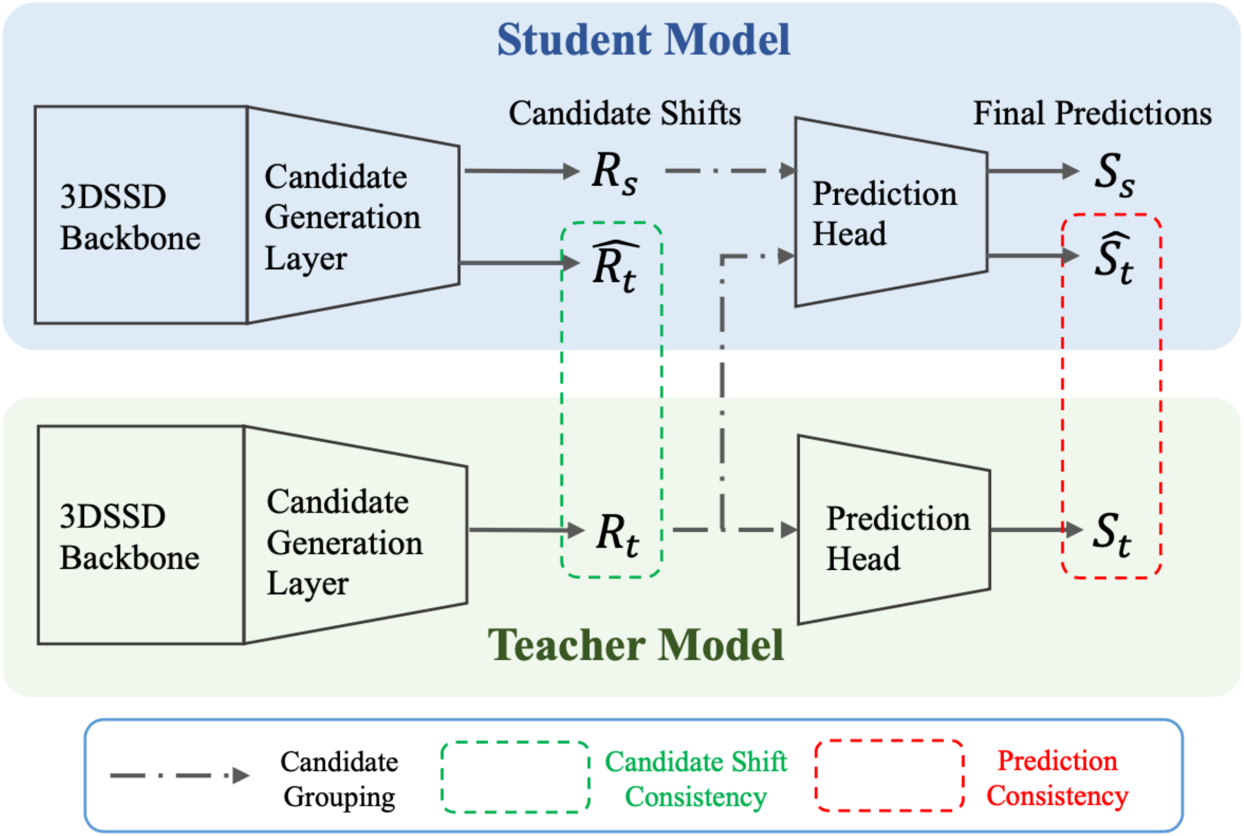
$L_{pt,cls}$	$L_{pt,box}$	$L_{ins,cls}$	$L_{ins,box}$	AP/L1	APH/L1	AP/L2	APH/L2
				18.61	18.18	16.77	16.40
✓				20.34	19.91	18.07	17.70
	✓			30.34	29.69	27.08	26.49
✓	✓			31.00	30.39	27.64	27.09
		✓		21.12	20.87	18.79	18.57
			✓	33.21	32.44	29.95	29.26
		✓	✓	34.95	34.53	31.43	31.05
✓	✓	✓	✓	**38.21**	**37.74**	**34.46**	**34.04**

Effect of point-level and instance-level consistency losses

Setting	AP/L1	APH/L1	AP/L2	APH/L2
Disabled	2.79	2.74	2.54	2.49
Separate	29.88	29.45	26.85	26.48
Enabled	**38.21**	**37.74**	**34.46**	**34.04**

Effect of neural statistics-level consistency. (Separate: batch norm performed for each domain individually.)

Extension to other detection models



Extension to the one-stage 3DSSD detector.

KITTI → Waymo				
Methods	AP/L1	APH/L1	AP/L2	APH/L2
Direct Transfer	0.0329	0.0326	0.0278	0.0275
Wide-Range Aug	0.1667	0.1648	0.1473	0.1456
OT [9]	0.2456	0.2423	0.2270	0.2239
SN [9]	0.2595	0.2561	0.2400	0.2367
Ours	**0.2987**	**0.2927**	**0.2680**	**0.2627**

Waymo → KITTI			
Methods	Easy	Moderate	Hard
Direct Transfer	6.3059	6.4088	6.2498
Wide-Range Aug	37.8330	35.3351	34.0547
OT [9]	45.4223	40.4968	41.0412
SN [9]	47.8134	45.9175	46.4571
Ours	**56.8611**	**48.7393**	**48.3180**

Qualitative Results

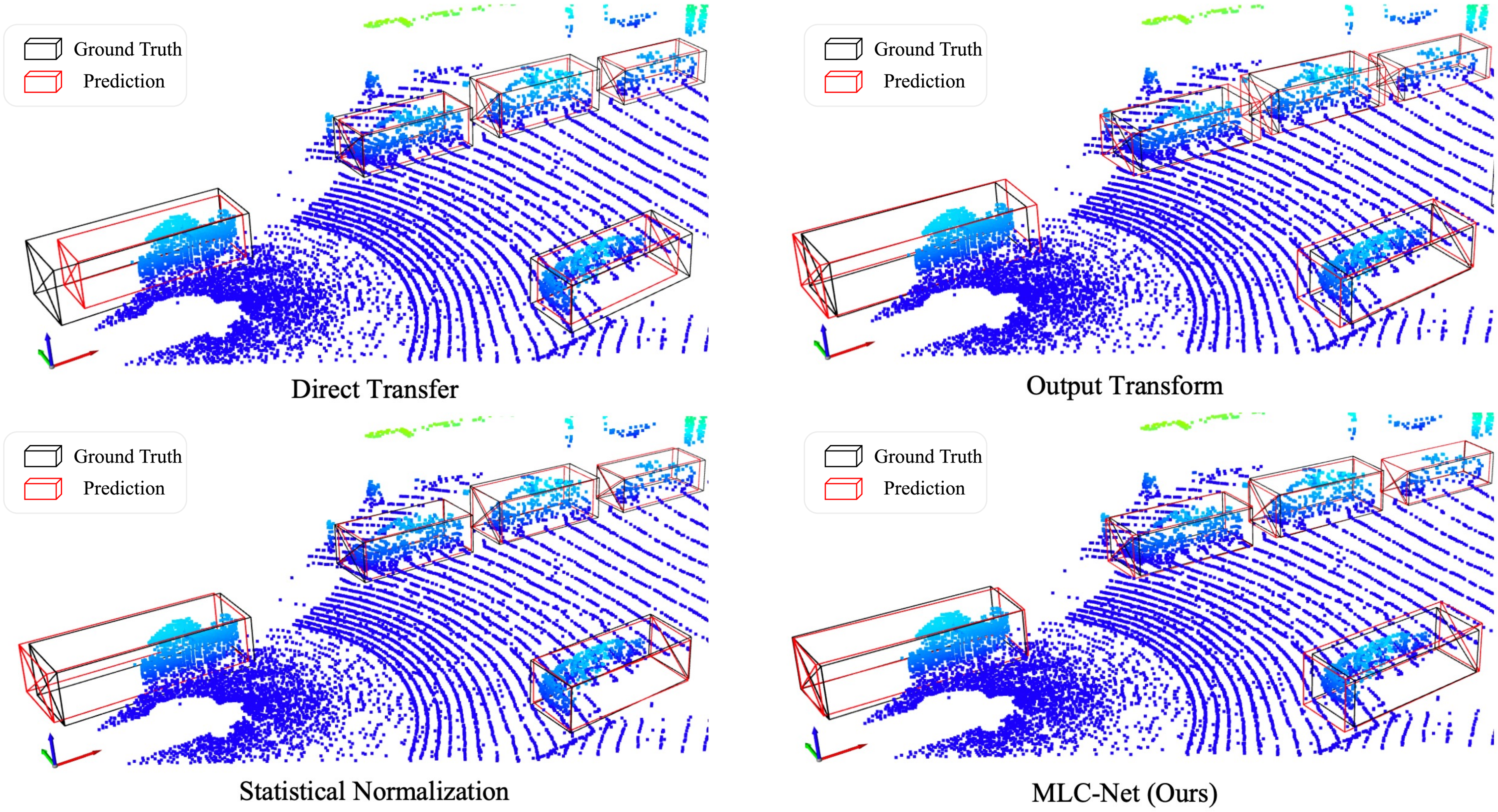


Figure 2: Qualitative results on Waymo validation dataset for KITTI to Waymo transfer.

Qualitative Results

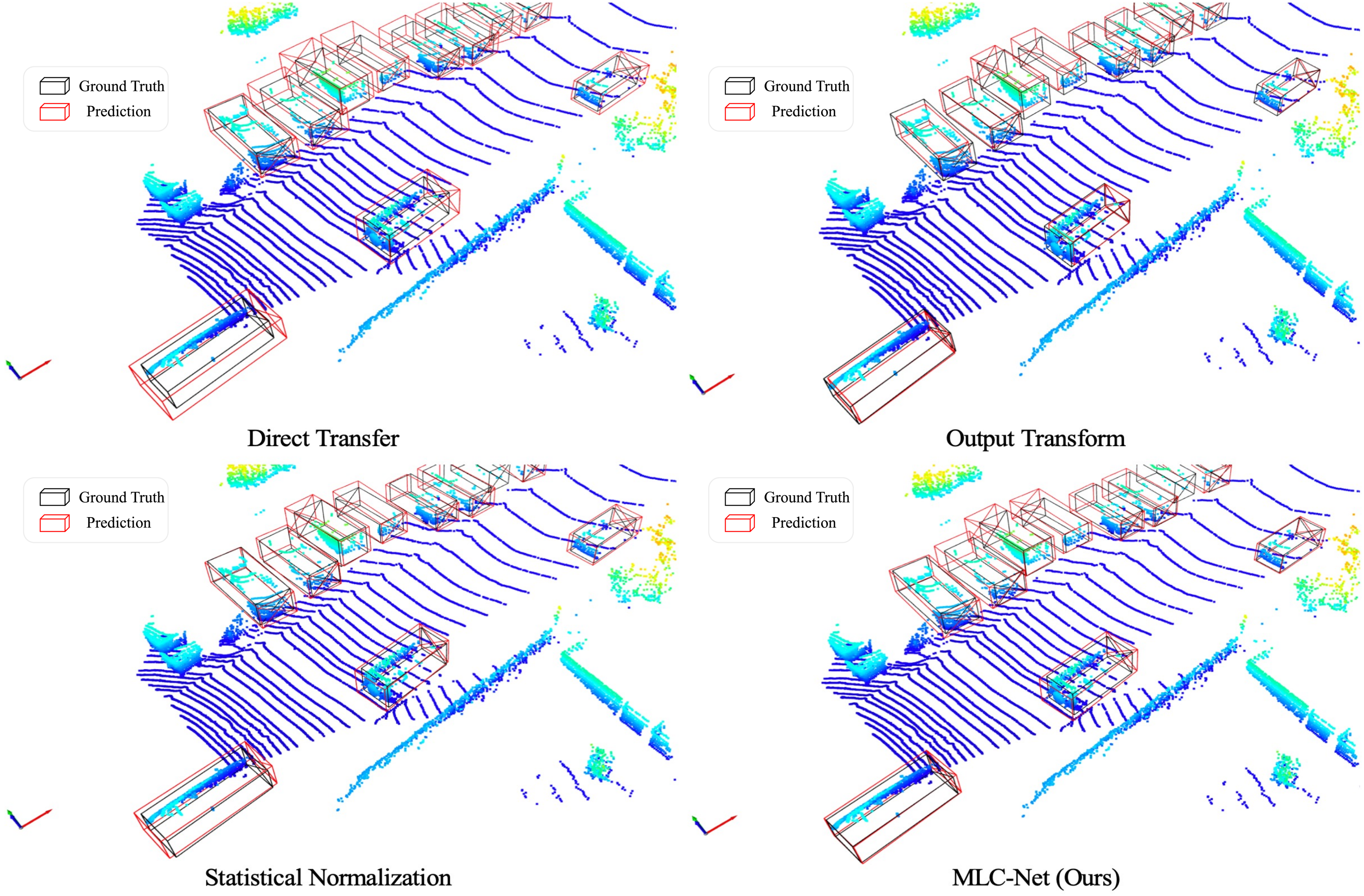
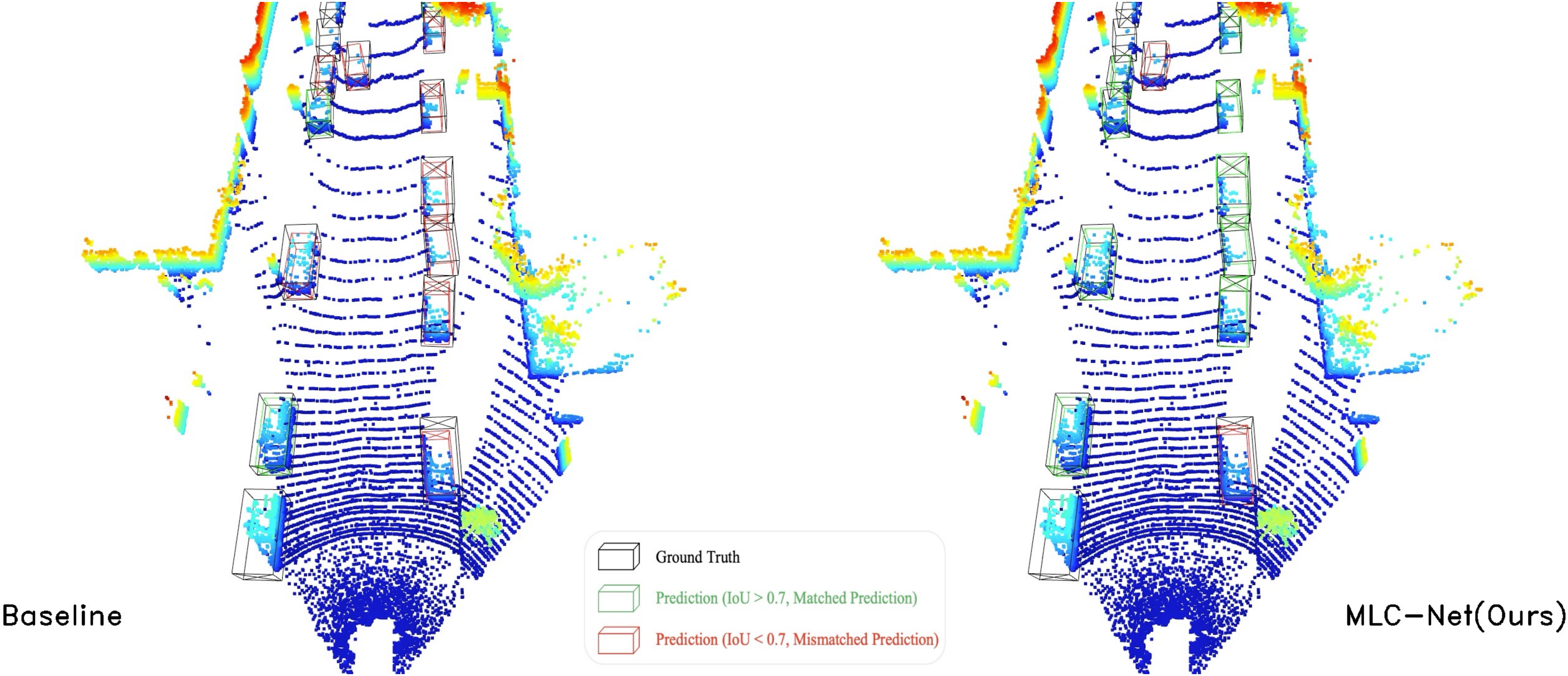
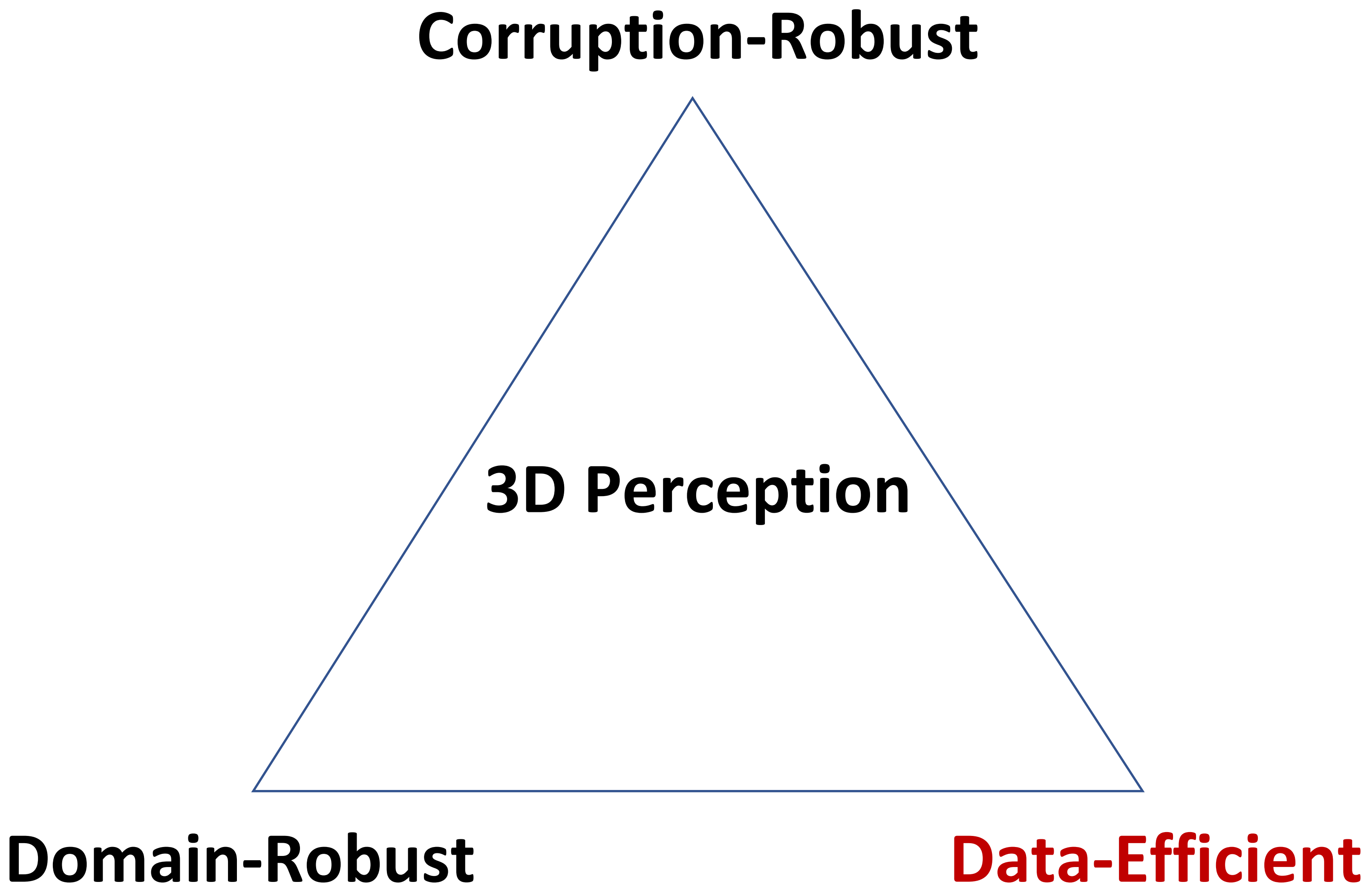


Figure 3: Qualitative results on KITTI validation dataset for Waymo to KITTI transfer.

Video Demo

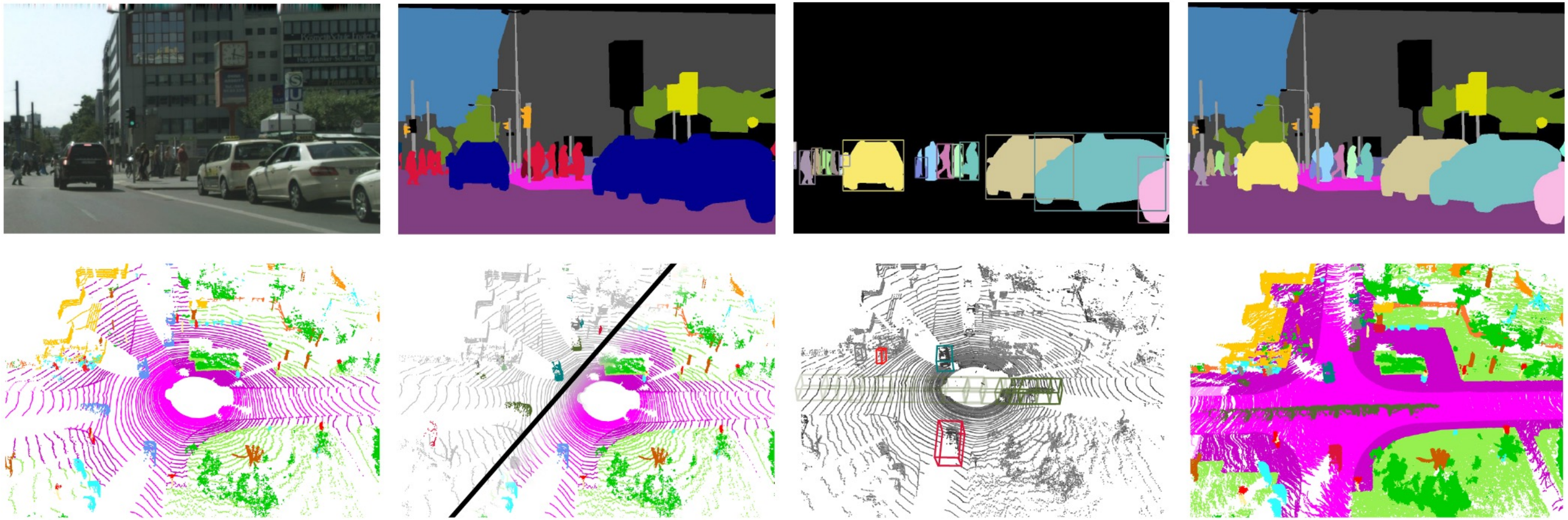

Baseline
Ground Truth
Prediction (IoU > 0.7, Matched Prediction)
Prediction (IoU < 0.7, Mismatched Prediction)
MLC−Net(Ours)


Corruption-Robust
3D Perception
Domain-Robust
Data-Efficient

LaserMix for Semi-Supervised LiDAR Semantic Segmentation

Lingdong Kong* Jiawei Ren* Liang Pan Ziwei Liu
S-Lab, Nanyang Technological University

AV Perception



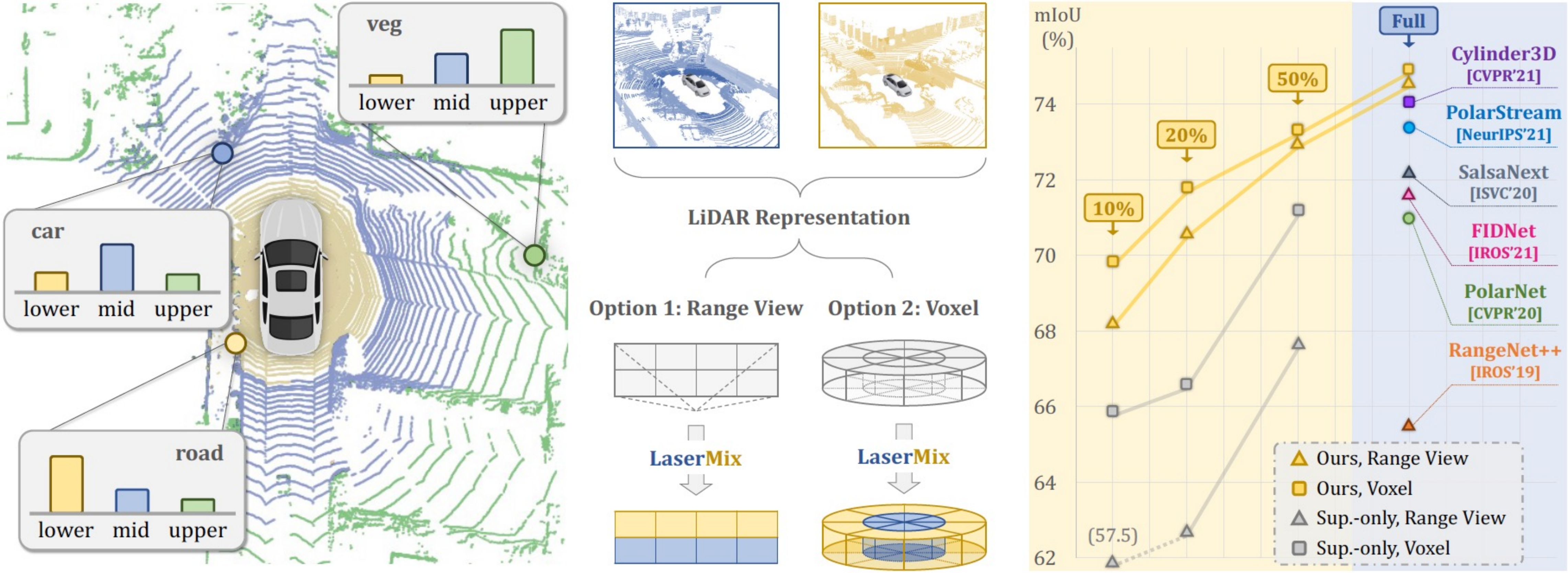
From left to right:

- LiDAR semantic segmentation
- LiDAR panoptic segmentation
- 3D object detection
- 4D LiDAR panoptic segmentation

Why LiDAR sensors?

- Accurate depth sensing
- Robust at low-light conditions
- Dense perceptions
- …

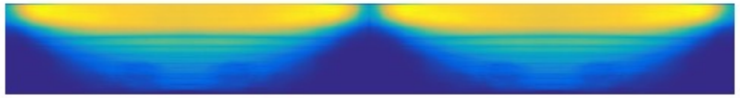
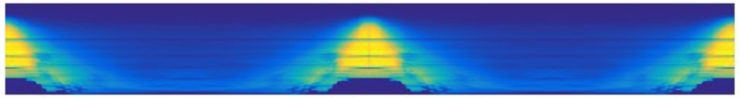
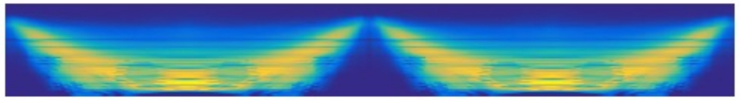
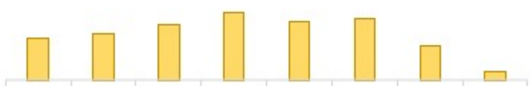
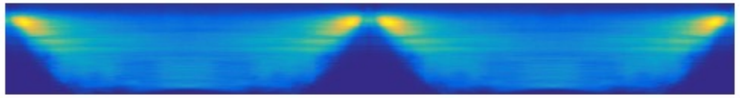
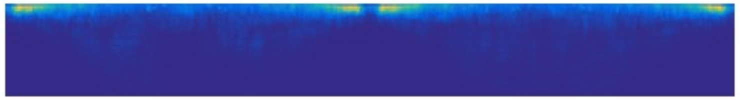
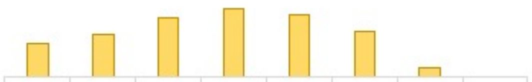
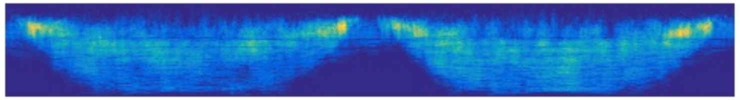
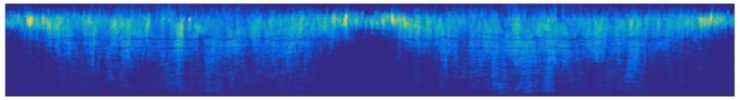
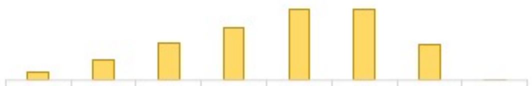
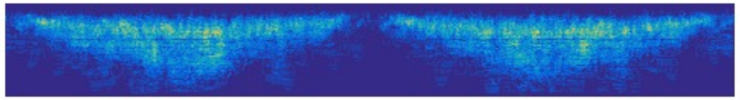
Overview



(a) (b) (c)

(a) Motivation. Semantic scene priors are overt for each category in LiDAR point clouds.
(b) Generalizability. LaserMix can be added into various popular LiDAR representations.
(c) Effectiveness. LaserMix helps to improve both semi- and fully-supervised settings.

Spatial Prior

Class	Type	Proportion	Distribution	Heatmap
vegetation	static	24.825%		
road	static	22.545%		
sidewalk	static	16.353%		
car	dynamic	4.657%		
traffic-sign	static	0.061%		
motorcycle	dynamic	0.045%		
person	dynamic	0.036%		
bicycle	dynamic	0.018%		

Certain class tends to appear at certain areas around the ego-vehicle!

Motivation

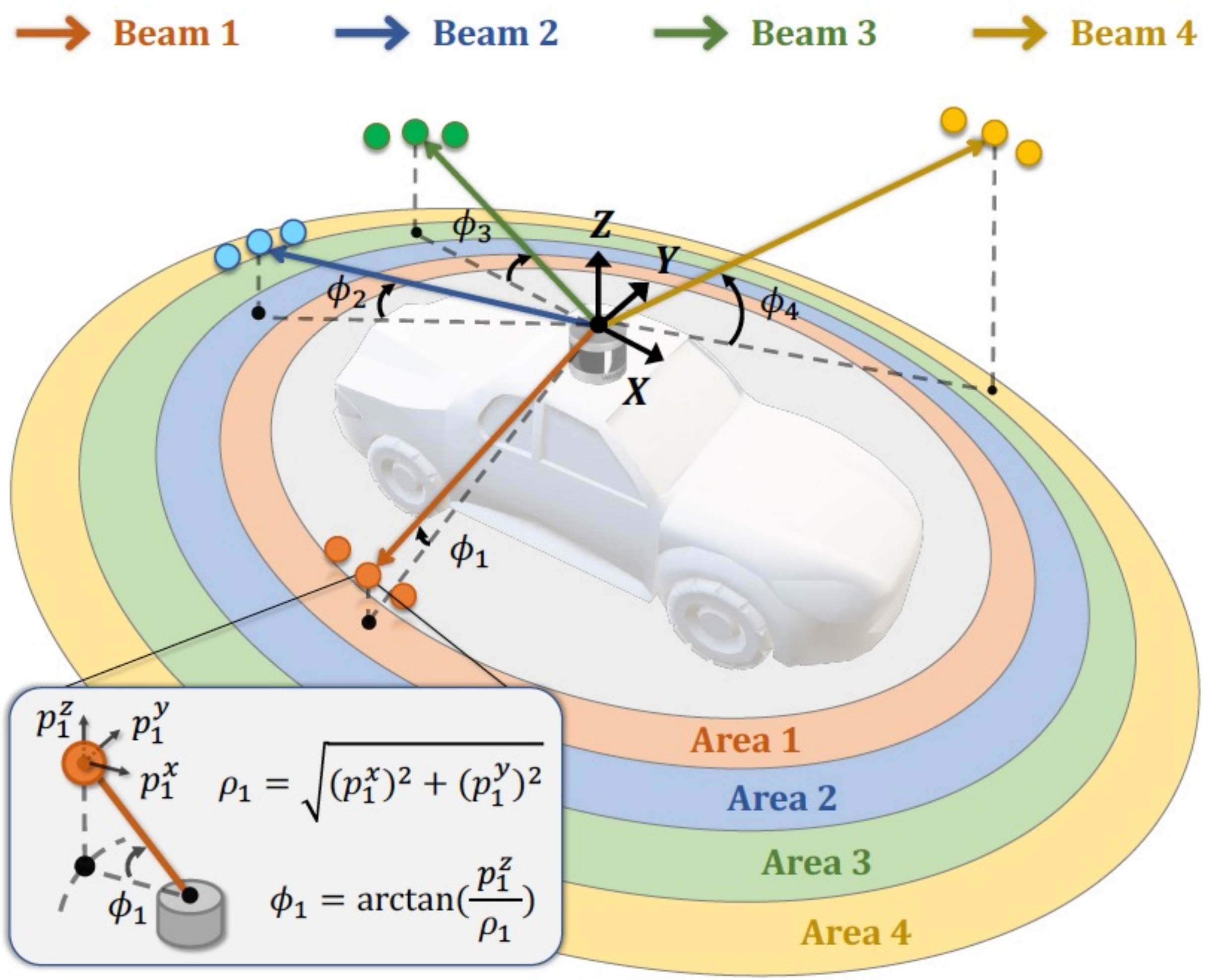
- We target on the less-explored semi-supervised LiDAR segmentation.

- Our goal is to leverage the abundant raw LiDAR scans for training accurate segmentation models.

- We make advantages of the spatial prior in LiDAR scenes for effective learning with semi supervisions.

- **TL;DR** - LaserMix leverages the strong spatial prior of driving scenes to construct low-variation areas via laser beam mixing, and encourages models to make confident and consistent predictions before and after mixing.

Laser Partition & Mixing



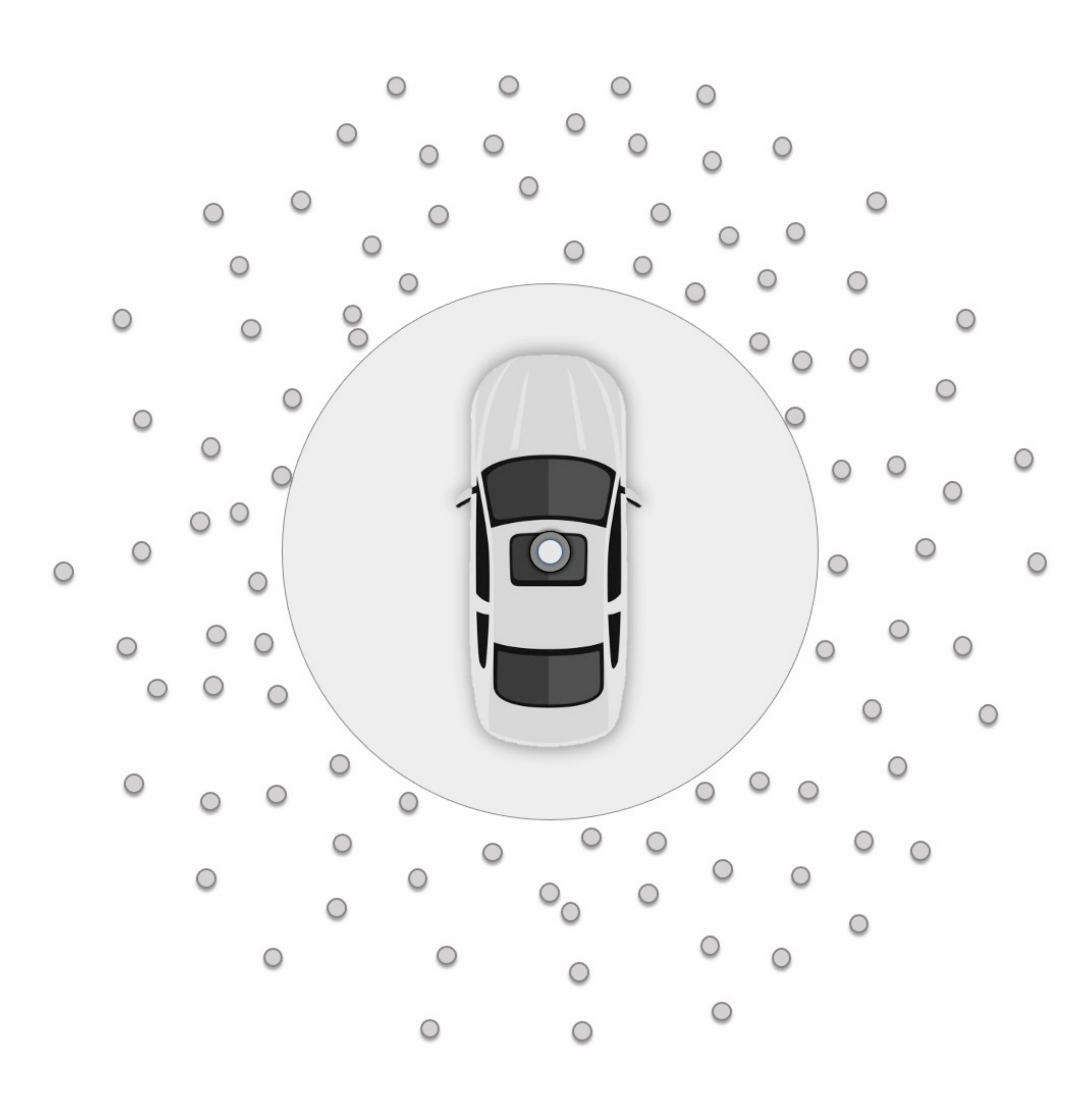
- Inclination:

$$\phi_i = \arctan(\frac{p_i^z}{\sqrt{(p_i^x)^2+(p_i^y)^2}})$$

- Depth: $\rho_i = \sqrt{(p_i^x)^2+(p_i^y)^2}$

- Azimuth: $\alpha_i = \arctan(\frac{p_i^y}{p_i^x})$

Laser Partition & Mixing



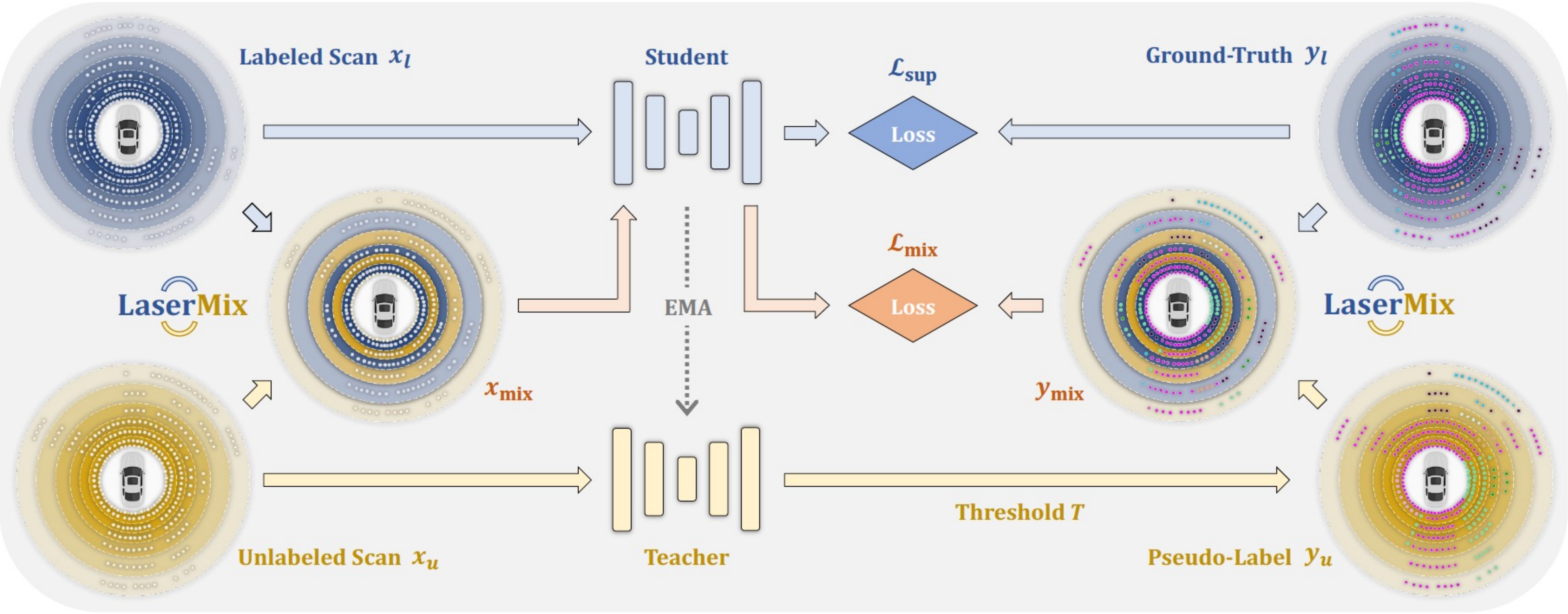
- Inclination:

$$\phi_i = \arctan\left(\frac{p_i^z}{\sqrt{(p_i^x)^2+(p_i^y)^2}}\right)$$

- Depth: $\rho_i = \sqrt{(p_i^x)^2+(p_i^y)^2}$

- Azimuth: $\alpha_i = \arctan\left(\frac{p_i^y}{p_i^x}\right)$

Consistency Regularization



Consistency Regularization

Algorithm 1 Pseudo-code for one training iteration in our SSL framework.

Input: Shuffled labeled batch $(X_l, Y_l) = \{(x_l^{(b)}, y_l^{(b)}); b \in (1, \ldots, B)\}$, shuffled unlabeled batch $X_u = \{x_u^{(b)}; b \in (1, \ldots, B)\}$, threshold T, loss weights λ_{mix} and λ_{mt}, Student and Teacher nets.

for $b = 1$ **to** B **do**

  $x_{\text{mix}}^{(2b-1)}, x_{\text{mix}}^{(2b)} = \text{LaserMix}(x_l^{(b)}, x_u^{(b)})$ *// LaserMix data*

end for

$X_{\text{mix}} = \{x_{\text{mix}}^{(i)}; i \in (1, \ldots, 2B)\}$

$S_l, S_u, S_{\text{mix}} = \text{Student}\big(\text{Concat}(X_l, X_u, X_{\text{mix}})\big)$ *// Student net prediction scores*

$\hat{S}_l, \hat{S}_u = \text{Teacher}\big(\text{Concat}(X_l, X_u)\big)$ *// Teacher net prediction scores*

$Y_u = \text{PseudoLabel}(\hat{S}_u, T)$ *// Produce pseudo-label from scores larger than threshold*

for $b = 1$ **to** B **do**

  $y_{\text{mix}}^{(2b-1)}, y_{\text{mix}}^{(2b)} = \text{LaserMix}(y_l^{(b)}, y_u^{(b)})$ *// LaserMix label*

end for

$Y_{\text{mix}} = \{y_{\text{mix}}^{(i)}; i \in (1, \ldots, 2B)\}$

$L_{\text{sup}} = \text{CrossEntropy}(S_l, Y_l)$ *// Supervised loss*

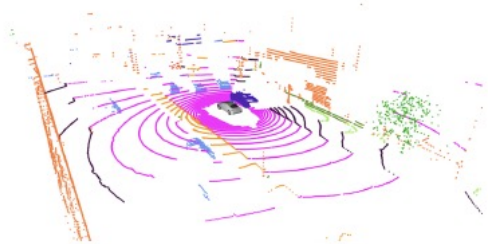
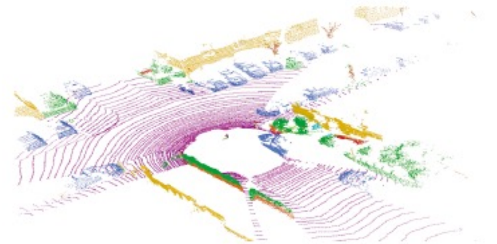
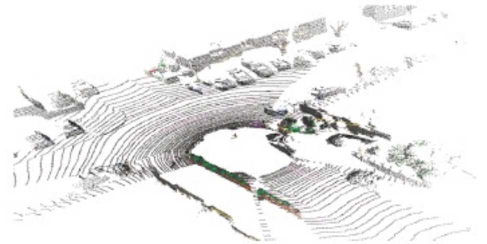
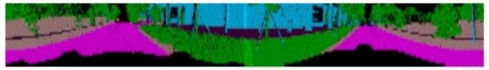
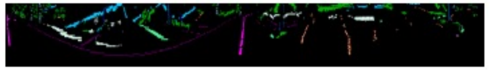
$L_{\text{mix}} = \text{CrossEntropy}(S_{\text{mix}}, Y_{\text{mix}})$ *// Mix loss*

$L_{\text{mt}} = \text{L2}\big(\text{Concat}(S_l, S_u), \text{Concat}(\hat{S}_l, \hat{S}_u)\big)$ *// Mean Teacher loss*

$L = L_{\text{sup}} + \lambda_{\text{mix}} L_{\text{mix}} + \lambda_{\text{mt}} L_{\text{mt}}$

Backward(L), Update(Student), UpdateEMA(Teacher)

Settings

	nuScenes [15]	SemanticKITTI [16]	ScribbleKITTI [4]
Vis.			
#Class	16	19	19
#Train	29130	19130	19130
#Val	6019	4071	4071
Res. (RV)	32×1920	64×2048	64×2048
Res. (voxel)	$[240, 180, 20]$	$[240, 180, 20]$	$[240, 180, 20]$
#Beam	32	64	64
$[\phi_{\text{up}}, \phi_{\text{low}}]$	$[10°, -30°]$	$[3°, -25°]$	$[3°, -25°]$
$[p^x_{\text{max}}, p^x_{\text{min}}]$	$[50m, -50m]$	$[50m, -50m]$	$[50m, -50m]$
$[p^y_{\text{max}}, p^y_{\text{min}}]$	$[50m, -50m]$	$[50m, -50m]$	$[50m, -50m]$
$[p^z_{\text{max}}, p^z_{\text{min}}]$	$[3m, -5m]$	$[2m, -4m]$	$[2m, -4m]$
#Label	100%	100%	8.06%
Intensity			
Range			
Semantics			

High-res LiDAR:
- SemanticKITTI
- Denser scenes

Low-res LiDAR:
- nuScenes
- Sparser scenes

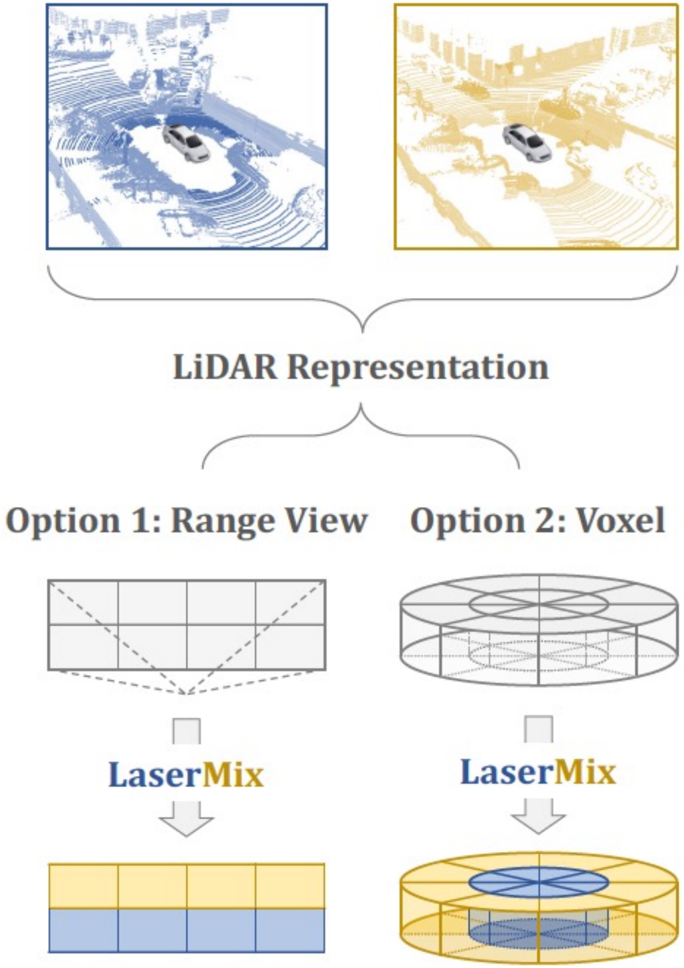
Weak supervision:
- ScribbleKITTI
- Sparse labels

Settings

- **Range View**
  - Backbone: FIDNet [IROS'21]
  - # Param: 6.05M
  - 6 x 32 x 1920 (nuScenes)
  - 6 x 64 x 2048 (SemanticKITTI/ScribbleKITTI)

- **Voxel**
  - Backbone: Cylinder3D [CVPR'21]
  - # Param: 28.13M
  - [240, 180, 20]

- **Data Split**
  - 1%, 10%, 20%, 50% (labeled)
  - Random sampling
  - Assume the remaining ones are unlabeled



Y. Zhao, et al. "FIDNet: LiDAR point cloud semantic segmentation with fully interpolation decoding," *IROS*, 2021.
X. Zhu, et al. "Cylindrical and asymmetrical 3D convolution networks for LiDAR segmentation," *CVPR*, 2021.

Comparative Studies

Repr.	Method	nuScenes [15] 1%	10%	20%	50%	SemanticKITTI [16] 1%	10%	20%	50%	ScribbleKITTI [4] 1%	10%	20%	50%
Range View	*Sup.-only*	38.3	57.5	62.7	67.6	36.2	52.2	55.9	57.2	33.1	47.7	49.9	52.5
	MeanTeacher [26]	42.1	60.4	65.4	69.4	37.5	53.1	56.1	57.4	34.2	49.8	51.6	53.3
	CBST [30]	40.9	60.5	64.3	69.3	39.9	53.4	56.1	56.9	35.7	50.7	52.7	54.6
	CutMix-Seg [29]	43.8	63.9	64.8	69.8	37.4	54.3	56.6	57.6	36.7	50.7	52.9	54.3
	CPS [13]	40.7	60.8	64.9	68.0	36.5	52.3	56.3	57.4	33.7	50.0	52.8	54.6
	LaserMix (Ours)	**49.5**	**68.2**	**70.6**	**73.0**	**43.4**	**58.8**	**59.4**	**61.4**	**38.3**	**54.4**	**55.6**	**58.7**
	$\Delta\uparrow$	+11.2	+10.7	+7.9	+5.4	+7.2	+6.6	+3.5	+4.2	+5.2	+6.7	+5.7	+6.2
Voxel	*Sup.-only*	50.9	65.9	66.6	71.2	45.4	56.1	57.8	58.7	39.2	48.0	52.1	53.8
	MeanTeacher [26]	51.6	66.0	67.1	71.7	45.4	57.1	59.2	60.0	41.0	50.1	52.8	53.9
	CBST [30]	53.0	66.5	69.6	71.6	48.8	58.3	59.4	59.7	41.5	50.6	53.3	54.5
	CPS [13]	52.9	66.3	70.0	72.5	46.7	58.7	59.6	60.5	41.4	51.8	53.9	54.8
	LaserMix (Ours)	**55.3**	**69.9**	**71.8**	**73.2**	**50.6**	**60.0**	**61.9**	**62.3**	**44.2**	**53.7**	**55.1**	**56.8**
	$\Delta\uparrow$	+4.4	+4.0	+5.2	+2.0	+5.2	+3.9	+4.1	+3.6	+5.0	+5.7	+3.0	+3.0

A. Tarvainen and H. Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," NeurIPS, 2017.

G. French, et al. "Semi-supervised semantic segmentation needs strong, high-dimensional perturbations," BMVC, 2020.

Y. Zou, et al. "Domain adaptation for semantic segmentation via class-balanced self-training," ECCV, 2018.

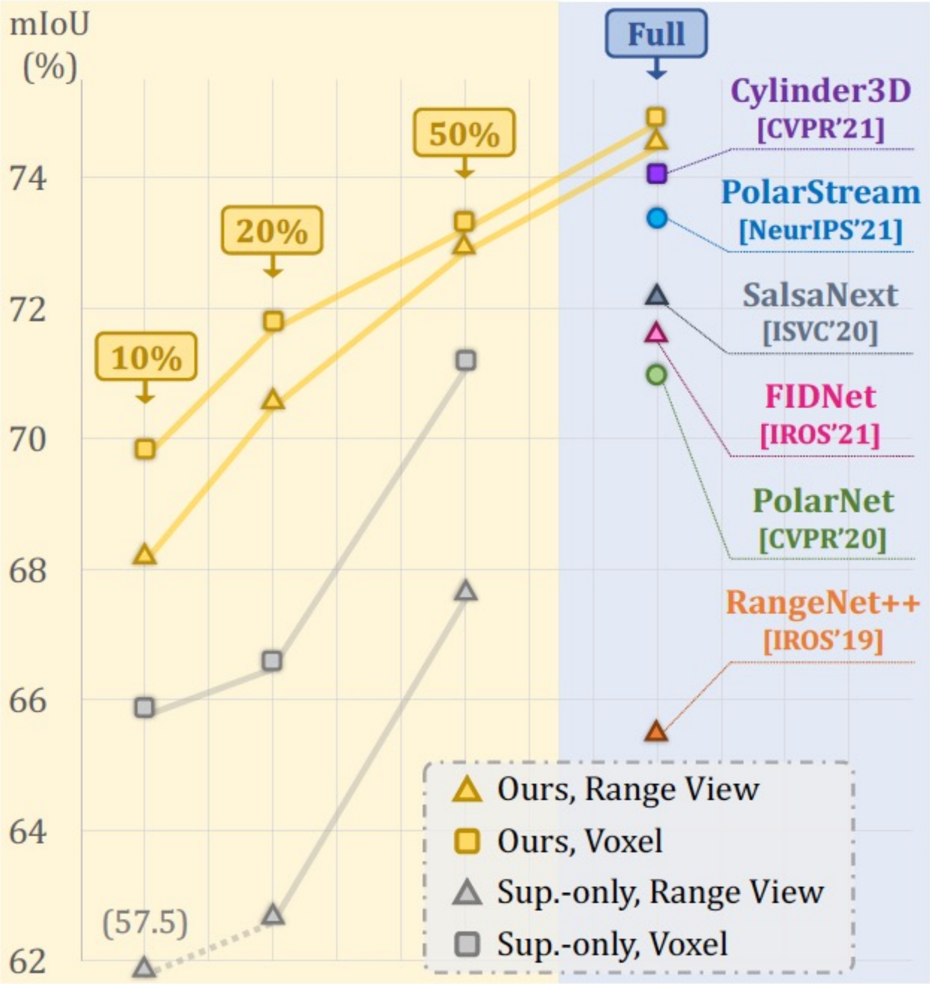
X. Chen, et al. "Semi-supervised semantic segmentation with cross pseudo supervision," CVPR, 2021.

Comparative Studies

Method	5%	10%	20%	30%	40%
GPC [14]	41.8	49.9	58.8	59.4	59.9
Ours (RV)	54.6	58.8	59.4	60.1	60.8
$\Delta\uparrow$	+12.8	+8.9	+0.6	+0.7	+0.9
Ours (Voxel)	56.7	60.0	61.9	62.1	62.3
$\Delta\uparrow$	+14.9	+10.1	+3.1	+1.7	+1.4

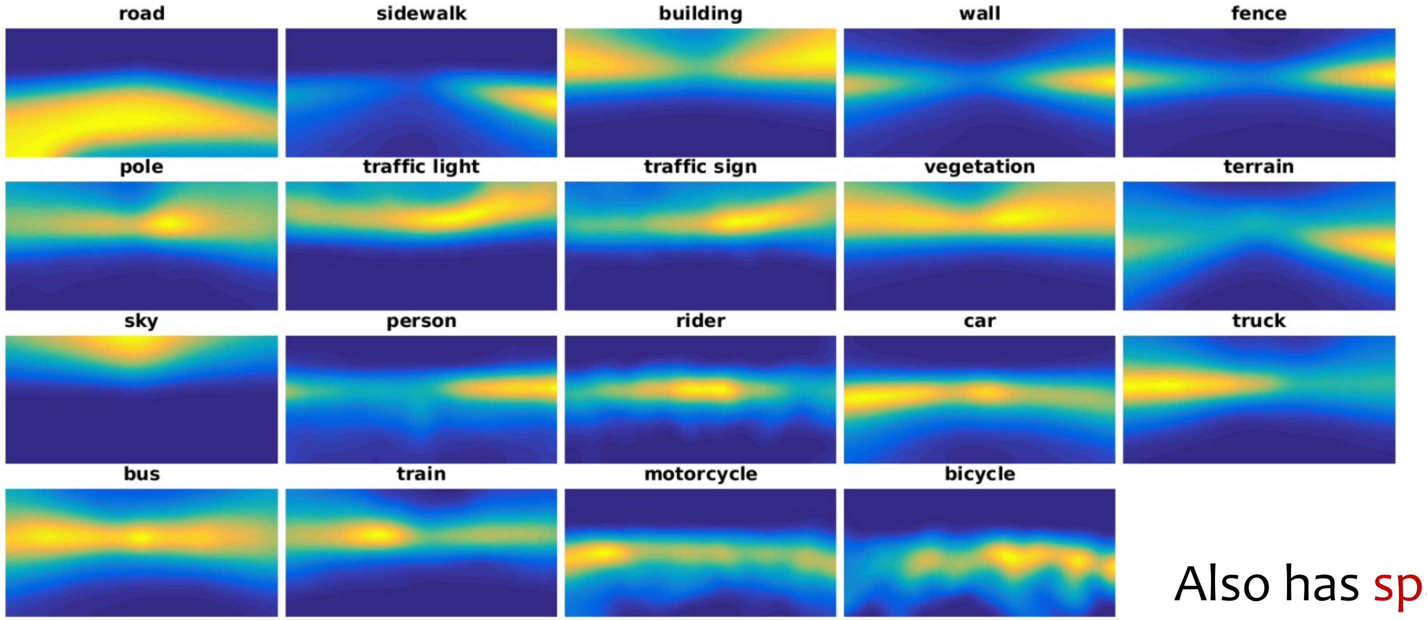
SemanticKITTI

nuScenes



L. Jiang, et al. "Guided point contrastive learning for semi-supervised point cloud semantic segmentation," *ICCV*, 2021.

Comparative Studies



Method	1/16	1/8	1/4	1/2
MeanTeacher [26]	66.1	71.2	74.4	76.3
***w/* Ours**	68.7	72.3	75.7	76.8
$\Delta\uparrow$	+2.6	+1.1	+1.3	+0.5
CCT [11]	66.4	72.5	75.7	76.8
GCT [12]	65.8	71.3	75.3	77.1
CPS [13]	69.8	74.4	76.9	78.6
CPS-CutMix [13]	74.5	76.6	77.8	78.8
***w/* Ours**	75.5	77.1	78.3	79.1
$\Delta\uparrow$	+1.0	+0.5	+0.5	+0.3

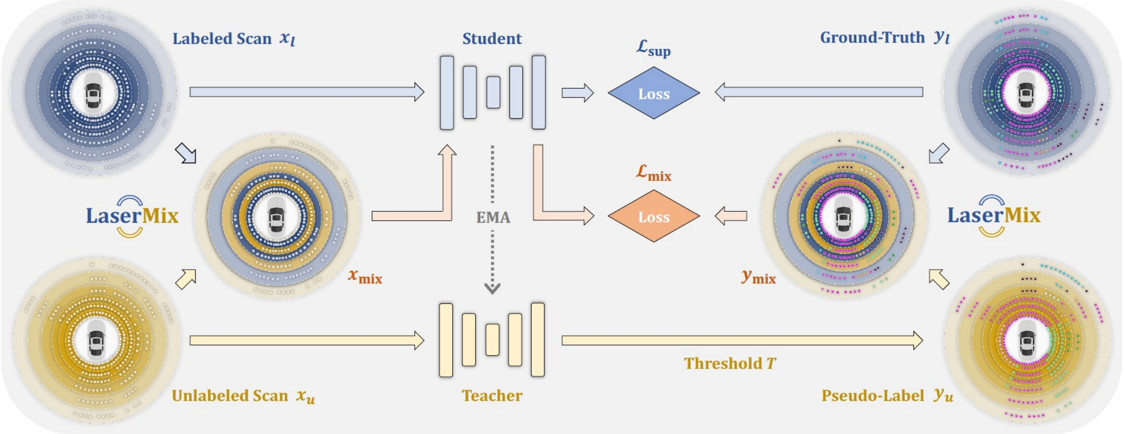
Cityscapes (RGB)

Also has spatial priors in scenes!

Y. Ouali, et al. "Semi-supervised semantic segmentation with cross-consistency training," CVPR, 2020.
Z. Ke, et al. "Guided collaborative training for pixel-wise semi-supervised learning," ECCV, 2020.

Ablation Studies

#	$\mathcal{L}_{\text{mt}}$	$\mathcal{L}_{\text{mix}}$	SS	TS	1%	10%	20%	50%
(1)	✓				42.1	60.4	65.4	69.4
(2)		✓	✓		45.6	64.3	67.8	71.6
	✓	✓	✓		47.0	65.5	69.5	72.0
(3)		✓		✓	46.0	64.1	69.5	72.3
	✓	✓		✓	49.5	68.2	70.6	73.0

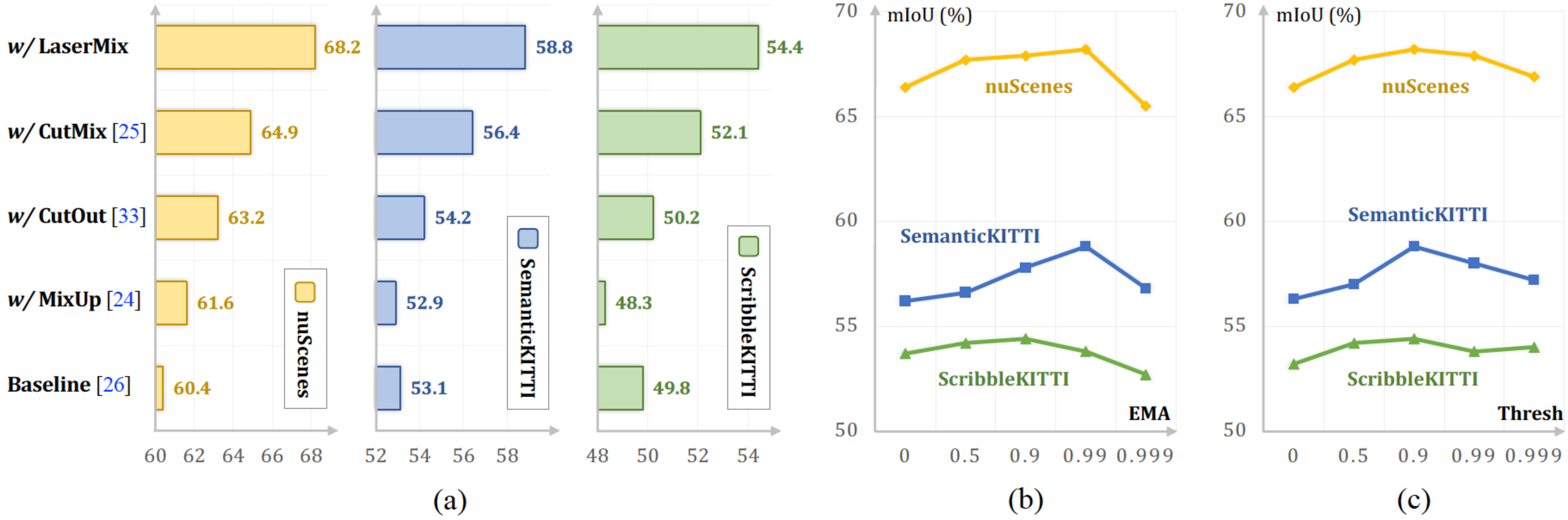


(1) Results of MeanTeacher.

(2) Results of LaserMix w/ student supervisions; much better than the counterpart.

(3) Results of LaserMix w/ teacher supervisions; much better than the counterpart.

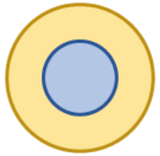
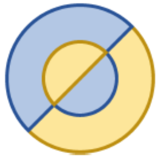
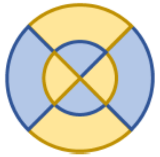
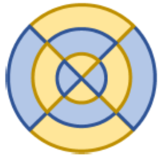
Ablation Studies



(a) Comparisons among different mixing techniques. (b) EMA. (c) Confidence threshold.

A. Nekrasov, et al. "Mix3D: Out-of-context data augmentation for 3D scenes," *3DV*, 2021.
S. Yun, et al. "Cutmix: Regularization strategy to train strong classifiers with localizable features," *ICCV*, 2019
T. DeVries and G. W. Taylor. "Improved regularization of convolutional neural networks with cutout," *arXiv*, 2017
H. Zhang, et al. "Mixup: Beyond empirical risk minimization," *ICLR*, 2018.

Ablation Studies

Baseline	$(1\alpha, 2\phi)$	$(1\alpha, 3\phi)$	$(1\alpha, 4\phi)$	$(1\alpha, 5\phi)$	$(1\alpha, 6\phi)$
					
60.4	63.5 (+3.1)	65.2 (+4.8)	66.5 (+6.1)	66.2 (+5.8)	65.4 (+5.0)
$(2\alpha, 1\phi)$	$(2\alpha, 2\phi)$	$(2\alpha, 3\phi)$	$(2\alpha, 4\phi)$	$(2\alpha, 5\phi)$	$(2\alpha, 6\phi)$
					
61.5 (+1.1)	63.3 (+2.9)	65.9 (+5.5)	66.1 (+5.7)	66.7 (+6.3)	65.3 (+4.9)
$(3\alpha, 1\phi)$	$(3\alpha, 2\phi)$	$(3\alpha, 3\phi)$	$(3\alpha, 4\phi)$	$(3\alpha, 5\phi)$	$(3\alpha, 6\phi)$
					
60.9 (+0.6)	64.2 (+3.8)	65.9 (+5.5)	66.3 (+5.9)	66.0 (+5.6)	65.2 (+4.8)
$(4\alpha, 1\phi)$	$(4\alpha, 2\phi)$	$(4\alpha, 3\phi)$	$(4\alpha, 4\phi)$	$(4\alpha, 5\phi)$	$(4\alpha, 6\phi)$
					
60.9 (+0.6)	64.7 (+4.3)	65.3 (+4.9)	65.6 (+5.2)	65.7 (+5.3)	65.2 (+4.8)

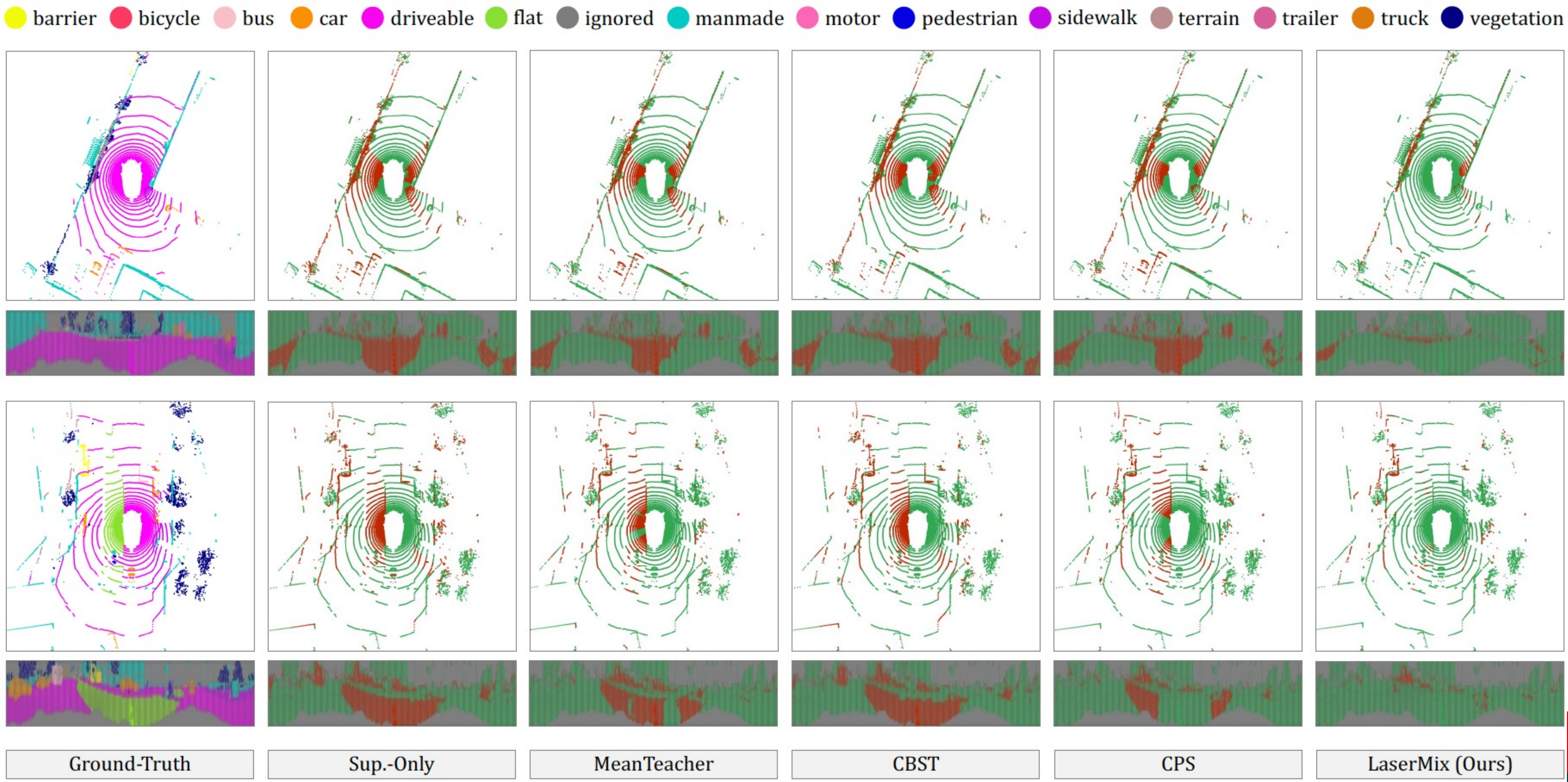
- Inclination:

$$\phi_i = \arctan\left(\frac{p_i^z}{\sqrt{(p_i^x)^2+(p_i^y)^2}}\right)$$

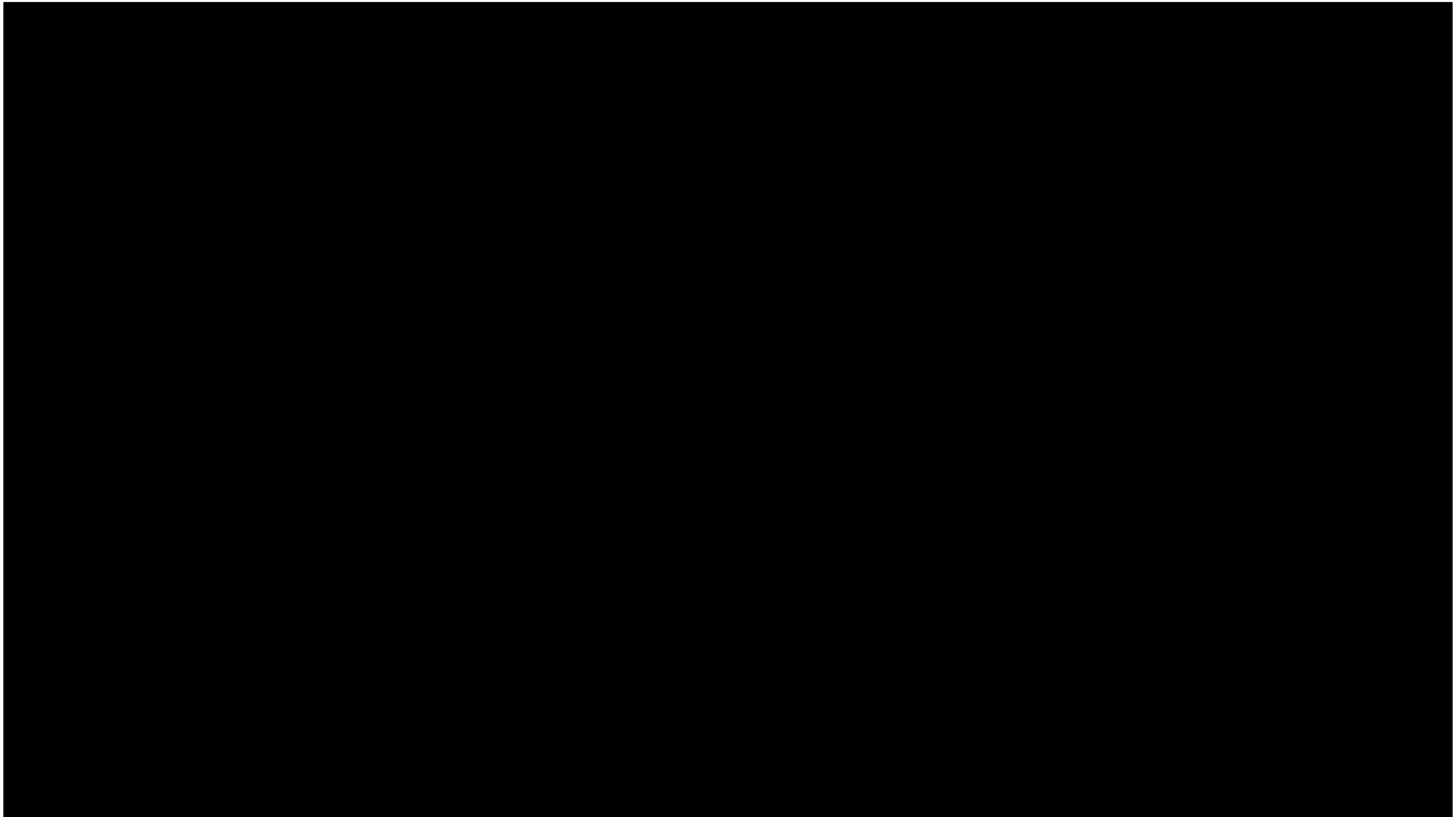
- Depth: $\rho_i = \sqrt{(p_i^x)^2+(p_i^y)^2}$

- Azimuth: $\alpha_i = \arctan\left(\frac{p_i^y}{p_i^x}\right)$

Visual Comparison



Know More About LaserMix



Know More About LaserMix

LaserMix

- **Paper:** https://arxiv.org/abs/2207.00026
- **Code:** https://github.com/ldkong1205/LaserMix
- **Tutorial:** https://zhuanlan.zhihu.com/p/528689803
- **Project Page:** https://ldkong.com/LaserMix

Corruption-Robust

Thank you!

Domain-Robust

Data-Efficient