Just Add $100 More

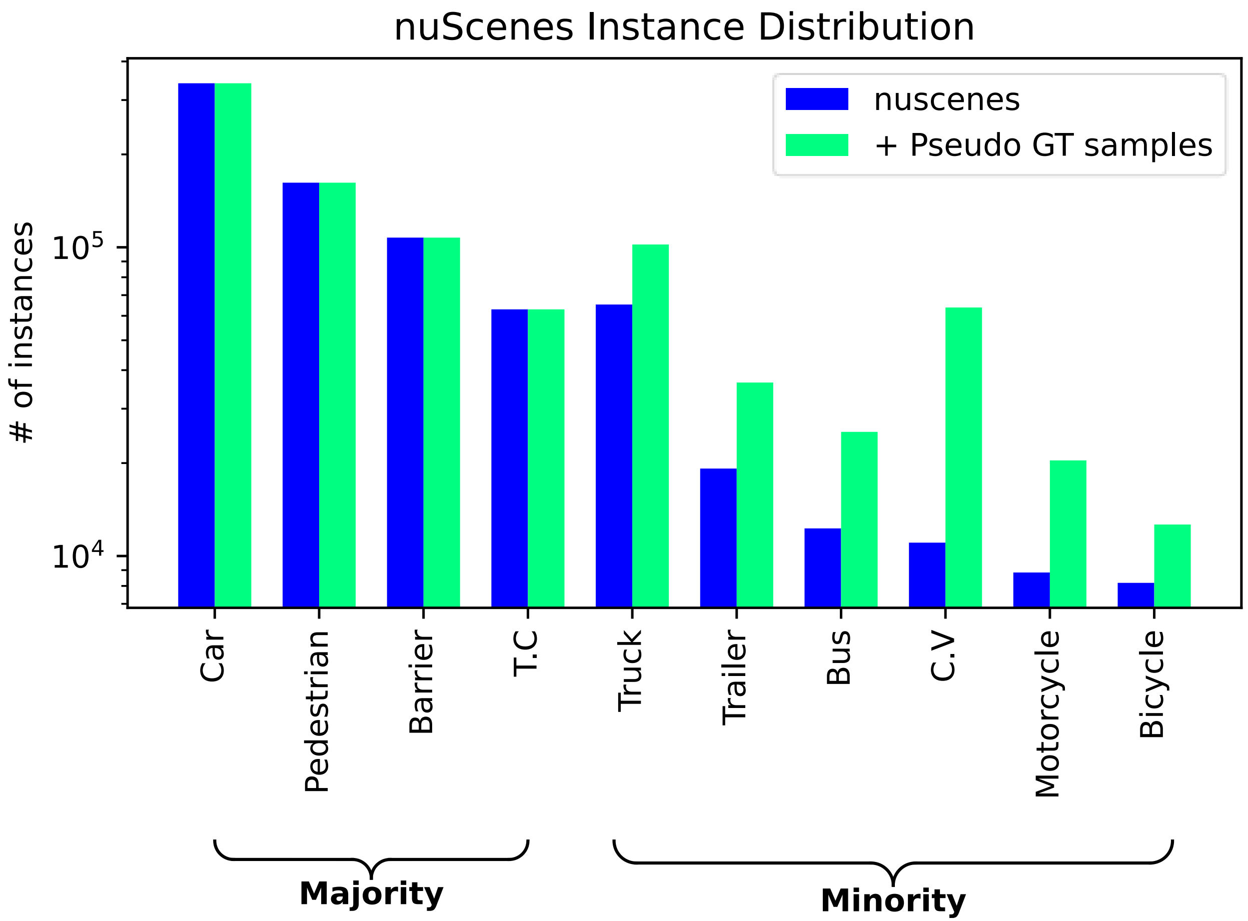
Augmenting Pseudo-LiDAR Point Cloud for Resolving Class-imbalance Problem

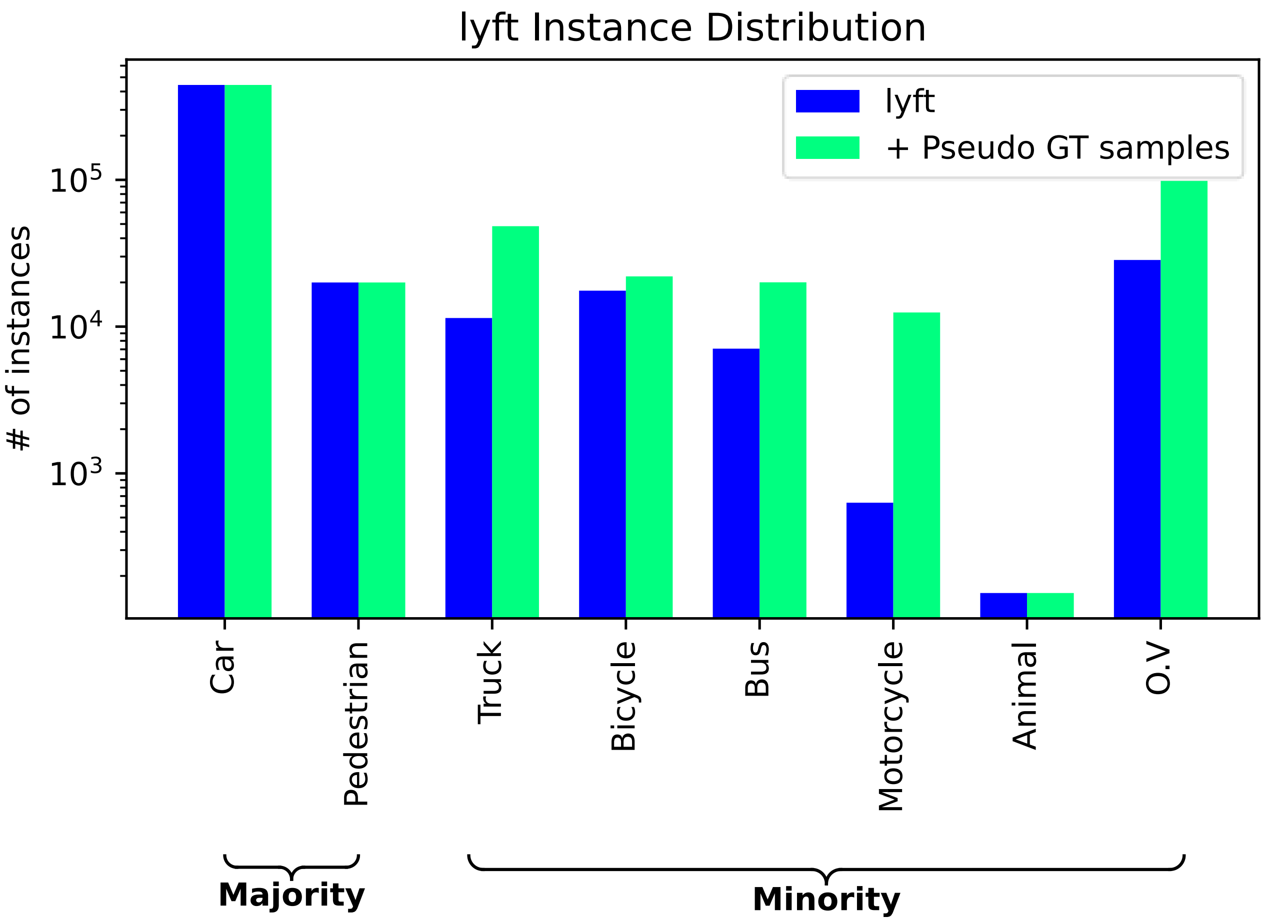
Mincheol Chang[1], Siyeong Lee[2], Jinkyu Kim[1], Namil Kim[2]

Korea University[1], NAVER LABS[2]

Motivation

Class Distribution on 3D Object benchmarks

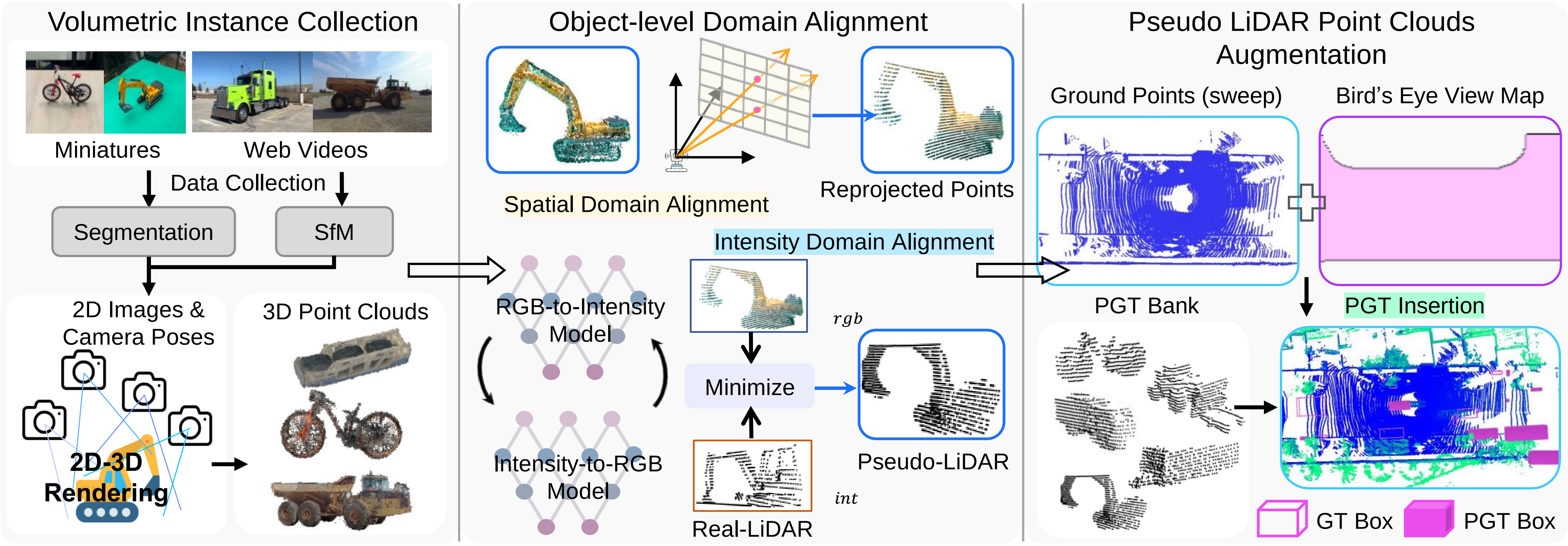




Pseudo Ground Truth Augmentation

Overview

For What? Balance the performance gap across classes.
How? Reconstuct 3D objects from two sources: miniature videos and public real-world videos.

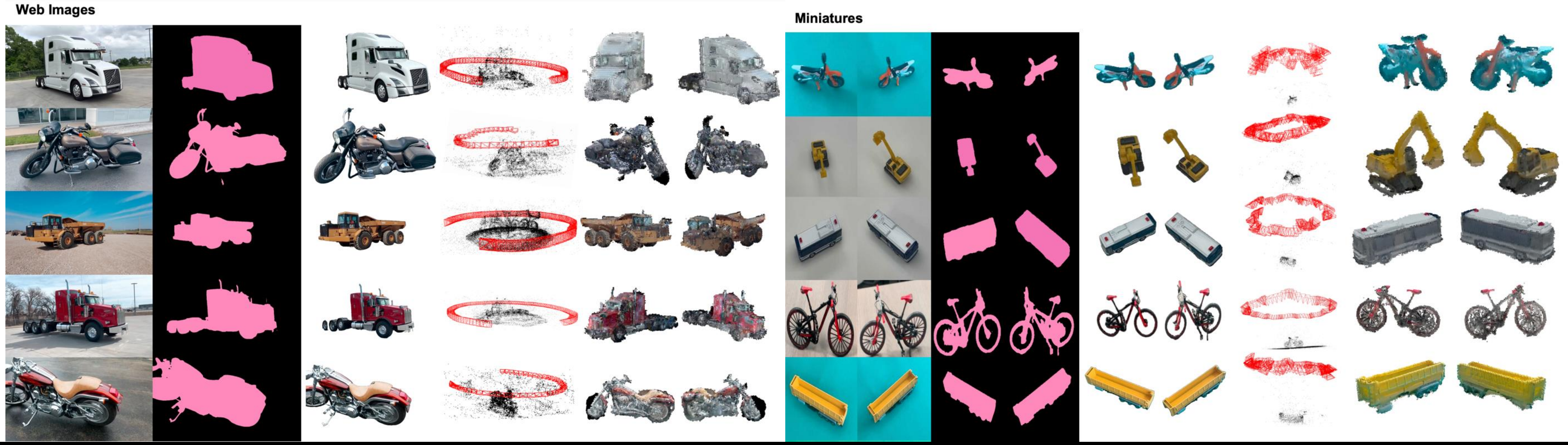


Volumetric 3D Instance Collection

From input images to a rendered point cloud

1. Foreground segmentation (e.g., SAM)
2. Camera parameters

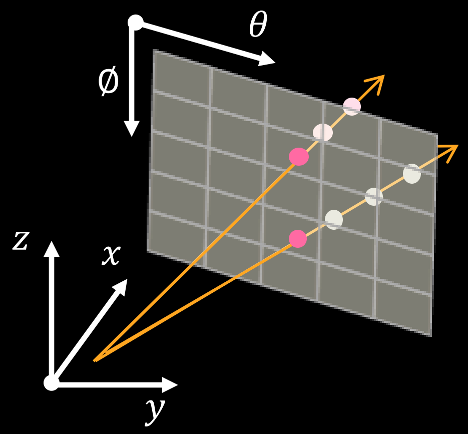
» RGB-colored point cloud

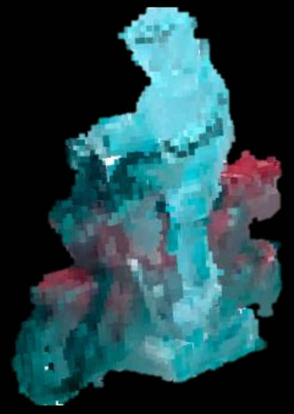


Object level alignment

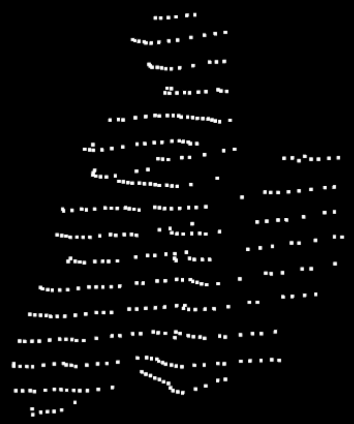
Domain alignment between rendered point cloud and real-world point cloud

1 Sensor-distribution alignment

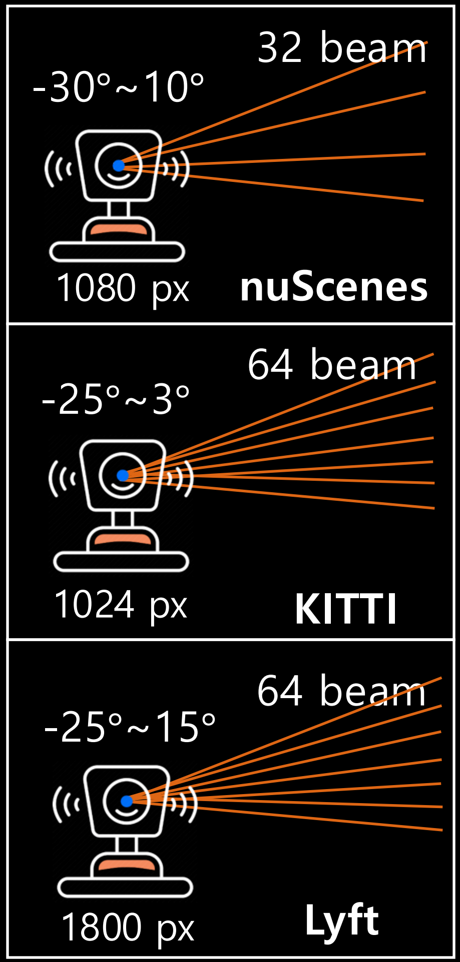




Dense RGB points

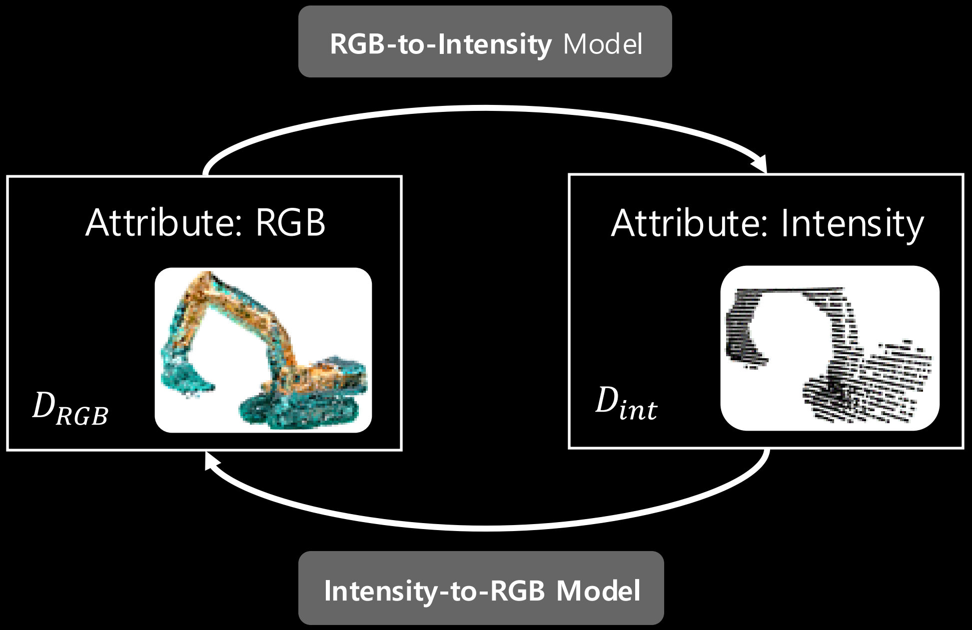


Simulated points



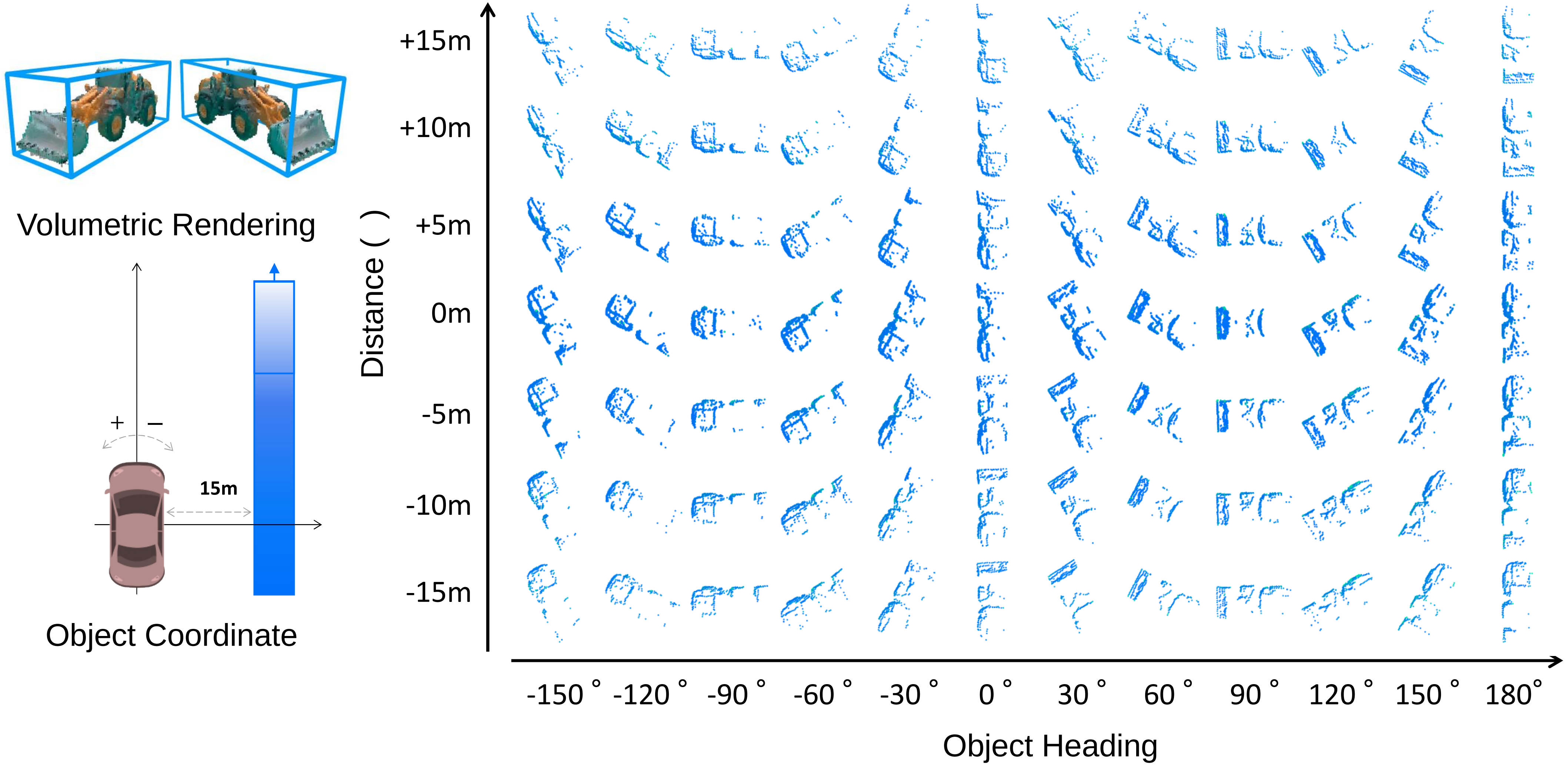
2 Sensor-property alignment

Unpaired Multi-domain Attribute Translation between colored point cloud and point intensity.



Object level alignment

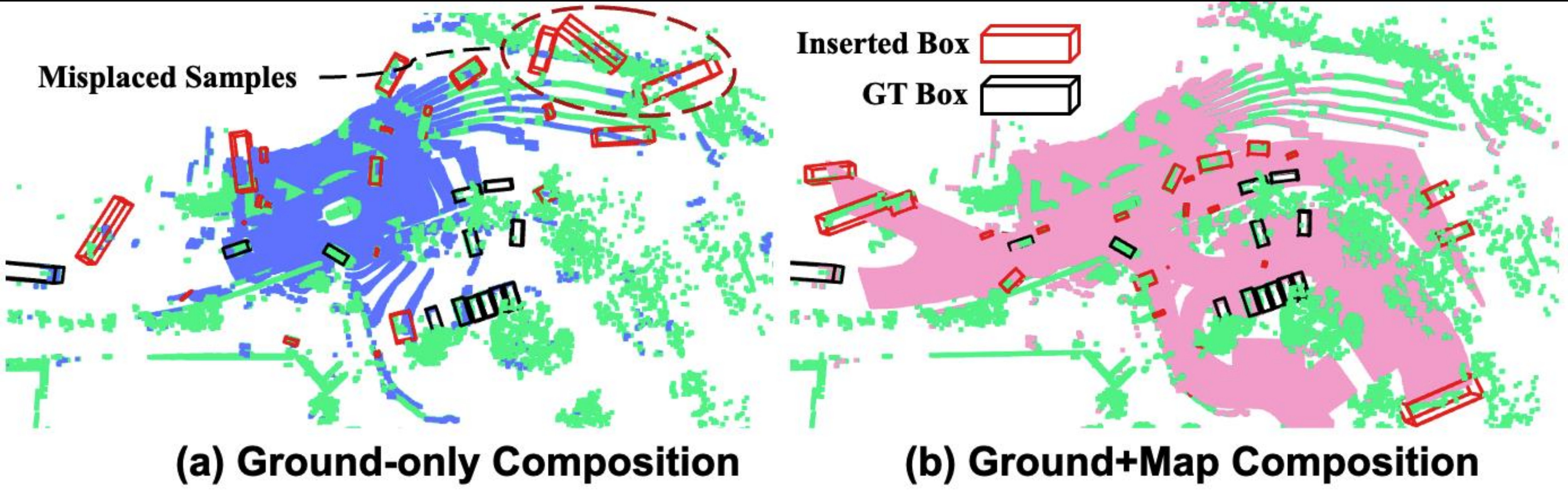
Samples of our pseudo LiDAR bank



Object insertion

Ground and Map priors based placement

For What? Compose the scene more realistically.
How? Create a rasterized map with a 0.128m per pixel resolution around the ego vehicle by using map and estimated ground areas.



Experiments

Detection performance comparison on nuScenes test set

Abbr. C.V: Construction Vehicle, Ped: Pedestrian, T.C: Traffic Cone, M.C: Motorcycle, B.C: Bicycle.
†: our reproduction. ‡: test time augmentation

Model	Aug.	Minority classes						mAP	NDS
		Bus	C.V	Trailer	Truck	M.C	B.C		
CP-Voxel‡	GT-Aug	64.4	**31.0**	60.0	47.2	65.7	41.0	63.8	68.7
	Real-Aug†	64.5	29.0	60.1	57.3	72.2	47.1	65.8	71.3
	PGT-Aug	**68.1**	29.0	**61.7**	**57.7**	**74.0**	**48.6**	**67.1** (+1.3%)	**72.3** (+1.0%)
Transfusion-L	GT-Aug	63.7	29.0	58.7	46.3	67.1	**44.2**	63.9	68.6
	Real-Aug†	64.3	**31.0**	60.0	47.3	65.7	41.0	63.8	68.7
	PGT-Aug	**67.3**	30.1	**60.2**	**56.9**	**68.2**	40.6	**65.1** (+1.3%)	**69.9** (+1.2%)

[GT-Aug] Yan Yan, et. al. SECOND: sparsely embedded convolutional detection. SENSORS 2018

[Real-Aug] Jinglin Zhan, et. al. Real-aug: Realistic scene synthesis for lidar augmentation in 3d object detection. ArXiV 2023

[CP-Voxel] Tianwei Yin, et. al. Center-based 3d object detection and tracking. CVPR 2021

[Transfusion-L] Xuyang Bai, et.al. Transfusion: Robust lidar-camera fusion for 3d object detection with transformers. CVPR 2022

Experiments

Detection performance comparison on Kitti and Lyft val datasets

Kitti ***val*** datasets

Model	Class	Target classes			Other classes						mAP
		Cyclist			Car			Pedestrian			
	# of objects in *val*	290	262	56	2980	5082	3116	1139	605	434	
	Difficulty	Easy	Mod.	Hard	Easy	Mod.	Hard	Easy	Mod.	Hard	
SECOND	GT-Aug	62.3	55.3	49.8	**91.4**	**82.4**	**79.6**	87.1	67.9	63.8	68.5
	PGT-Aug	**63.3**	**56.2**	**50.2**	90.7	82.1	79.3	**90.3**	**72.1**	**67.7**	**70.1** (+1.6%)

Lyft ***val*** datasets

Abbr. O.V: Other Vehicle, Ped: Pedestrian, M.C: Motorcycle, B.C: Bicycle.

Model	Class	Target classes					Other classes		mAP
		Truck	Bus	O.V	M.C	B.C	Car	Ped.	
	# of objects in *val*	2,721	1,653	4,920	187	3,347	91529	4952	
CP-Voxel	GT-Aug	19.15	20.48	31.91	**4.54**	5.31	**37.14**	6.00	13.84
	PGT-Aug	**19.85**	**21.11**	**31.99**	4.39	**5.48**	37.11	**6.12**	**14.01** (+0.17%)

[SECOND, GT-Aug] Yan Yan, et. al. SECOND: sparsely embedded convolutional detection. SENSORS 2018

[CP-Voxel] Tianwei Yin, et. al. Center-based 3d object detection and tracking. CVPR 2021

Experiments

Comparison between other methods that aim for class imbalance problems

Abbr. C.V: Construction Vehicle, Ped: Pedestrian, T.C: Traffic Cone, M.C: Motorcycle, B.C: Bicycle.
†: our reproduction.

Model	Aug.	Majority classes				Minority classes						mAP
		Car	Ped	Barrier	T.C	Bus	C.V	Trailer	Truck	M.C	B.C	
PointPillars	CBLoss†	82.7	**74.7**	54.0	52.1	51.2	61.7	17.9	30.5	48.7	20.5	49.4
	DWA	81.0	72.3	50.2	50.1	49.0	63.4	10.7	34.3	32.9	6.9	44.6
	PGT-Aug	**83.0**	71.8	54.8	51.1	**54.9**	**69.7**	20.2	**39.5**	49.6	14.5	50.9
	PGT Aug + CBLoss	82.7	**74.7**	**56.5**	**55.9**	54.4	68.7	**20.9**	34.1	**53.5**	**20.5**	**52.2** (+1.3%)

[PointPillars] Alex H. Lang, et. al. Point- pillars: Fast encoders for object detection from point clouds. CVPR 2019

[CBLoss] Yin Cui, et. al. Class-balanced loss based on effective number of samples. CVPR 2019

[DWA] Daeun Lee and Jinkyu Kim. Resolving class imbalance for lidar-based object detector by dynamic weight average and contextual ground truth sampling. WACV 2023

Experiments

Quality of Pseudo LiDAR Point Clouds. FID scores between given samples and nuScenes samples

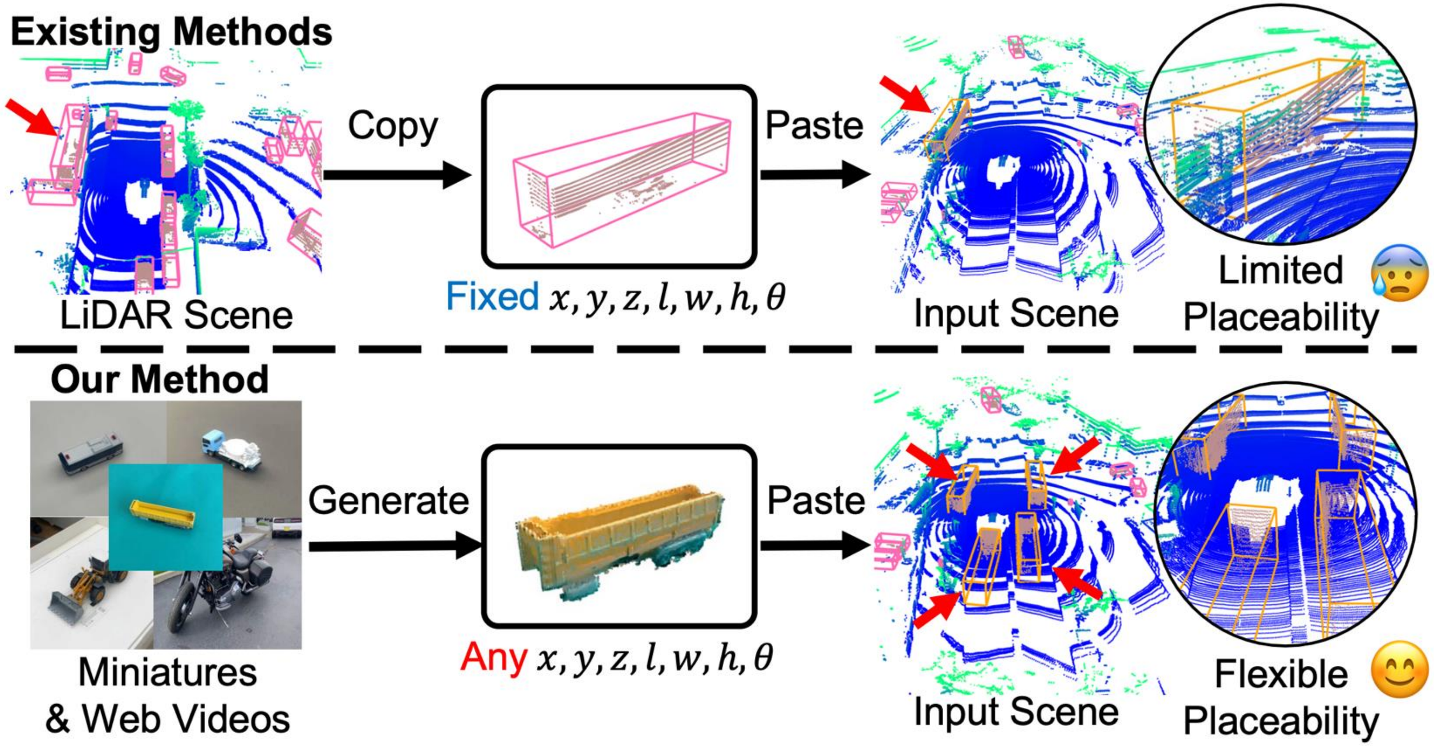
Abbr. G.S: Gaussian Splatting

	Pseudo-LiDAR Point Clouds						Lyft	A2D2
Volumetric 3D type	Plenoxels					G.S	-	-
Azimuth Resolution (px)	3600	1080	1080	1080	1080	1080	-	-
RGB features	✓	✗	✗	✓	✓	✓	-	-
Group intensity Loss	✓	✗	✓	✗	✓	✓	-	-
Bus	17.7	14.6	13.1	**13.0**	13.2	**11.2**	8.7	19.8
Construction Vehicle	**7.0**	7.5	7.6	7.5	7.6	**7.6**	-	6.0
Trailer	20.6	12.7	11.9	**11.9**	12.2	**13.7**	-	36.5
Truck	8.9	8.4	7.6	7.6	**7.3**	**6.9**	6.6	13.4
Motorcycle	20.7	2.5	7.0	7.2	**3.7**	**1.3**	3.0	10.1
Bicycle	9.0	3.3	2.2	2.4	**2.1**	**1.8**	1.8	0.7
Avg. FID Score	14.2	8.2	8.3	8.3	**7.7**	**7.1**	4.8	14.4
mAP (*for all 10 classes*)	63.40	63.44	63.48	63.41	**63.52**	**63.77**	63.45	63.17
NDS (*for all 10 classes*)	68.83	68.99	69.02	68.87	**69.11**	**69.35**	68.88	68.73

[Plenoxels] SaraFridovich-Keil, et. al. Plenoxels: Radiance fields without neural networks , CVPR 2022

[G.S] Bernhard Kerbl, et. Al, 3d gaussian splatting for real-time radiance field rendering. SIGGRAPH 2023

Recap: Comparison with existing methods (GT-Aug and Real-Aug)



Cost-effective pipeline to effectively generate and augment pseudo-LiDAR samples

For more extra explanation and experimental results
Please check the paper and the supplemental webpage.

https://just-add-100-more.github.io/