JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Uni3D: A Unified Baseline for Multi-dataset 3D Object Detection

Bo Zhang [1], Jiakang Yuan [2], Botian Shi [1], Tao Chen [2],
Yikang Li [1], Yu Qiao [1]

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

復旦大學
FUDAN UNIVERSITY

MDF: Multi-dataset Domain Fusion

➢Multi-dataset Domain Fusion：

- To learn robust representations that can generalize on multiple 3D perception datasets or tasks.
- To improve the reusability across different 3D datasets or domains or different manufacturers.
- To design an effective module which can easily be integrated into the existing 3D models such as PV-RCNN to enable them to train from multi-datasets.

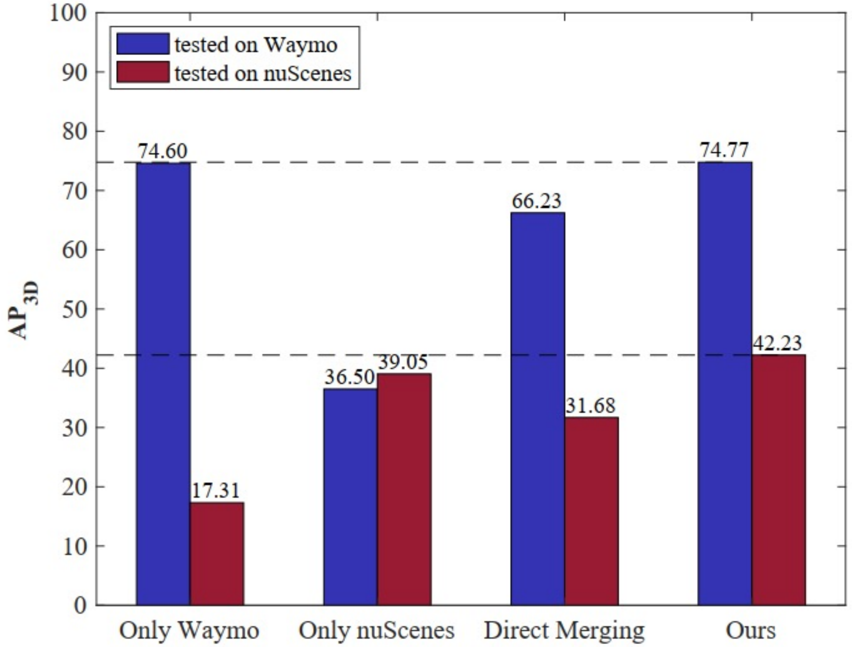
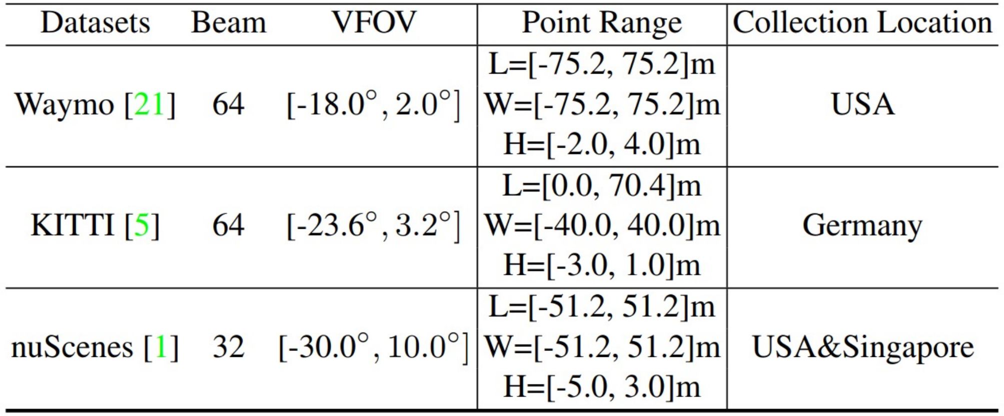


Figure 1. Challenges in training a detector from multiple datasets: 1) Only Waymo and Only nuScenes refer to the baseline detector trained on each individual dataset. 2) Direct Merging represents that we simply merge Waymo and nuScenes and train the detector on the merged dataset. 3) Ours denotes that the baseline detector is trained using the proposed method on the merged dataset.

MDF: Multi-dataset Domain Fusion

➢Major Challenges：

- Point Range is different for different dataset
- ROI-size is different for different dataset



Datasets	Beam	VFOV	Point Range	Collection Location

Waymo [21]	64	[-18.0°, 2.0°]	L=[-75.2, 75.2]m
			W=[-75.2, 75.2]m
			H=[-2.0, 4.0]m
				USA

KITTI [5]	64	[-23.6°, 3.2°]	L=[0.0, 70.4]m
			W=[-40.0, 40.0]m
			H=[-3.0, 1.0]m
				Germany

nuScenes [1]	32	[-30.0°, 10.0°]	L=[-51.2, 51.2]m
			W=[-51.2, 51.2]m
			H=[-5.0, 3.0]m
				USA&Singapore

Table 1. Overview of 3D autonomous driving dataset differences. VFOV denotes vertical field of view, and L, W, and H represent the length, width, and height of LiDAR range, respectively.

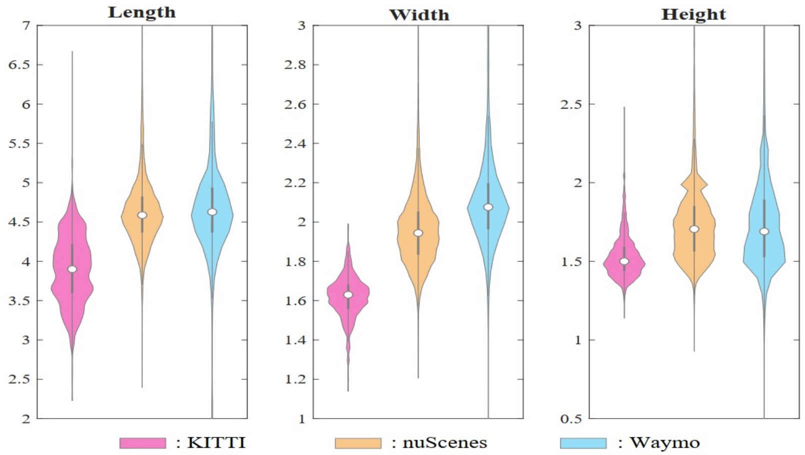
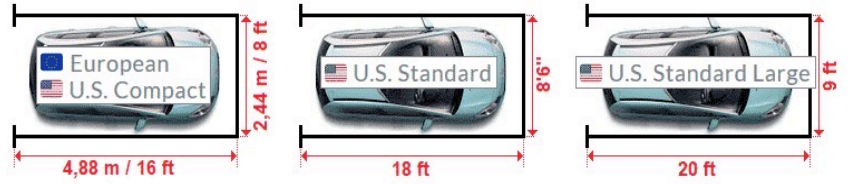


Figure 2. The statistical distribution differences of object size (Length, Width, and Height) across different datasets. For better illustrate the differences, we pick up the values within the range of [2.0, 7.0], [1.0, 3.0], and [0.5, 3.0] for Length, Width, and Height.

Standard dimensions of car parking spaces

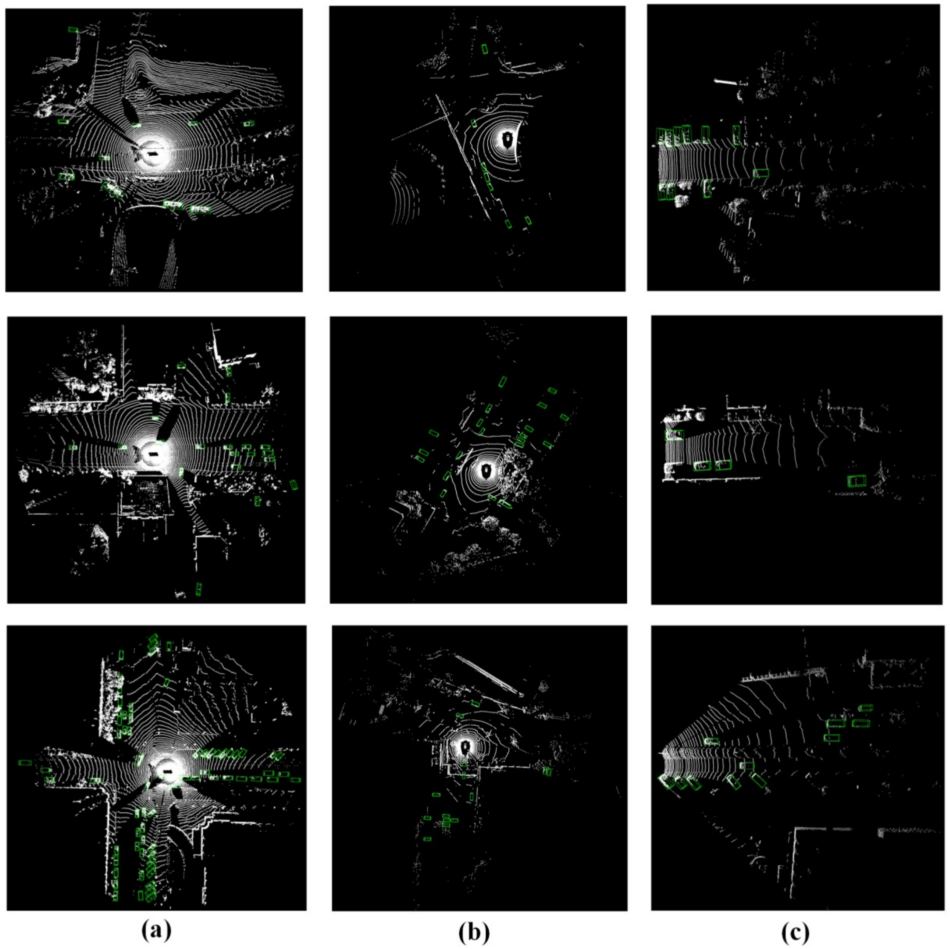


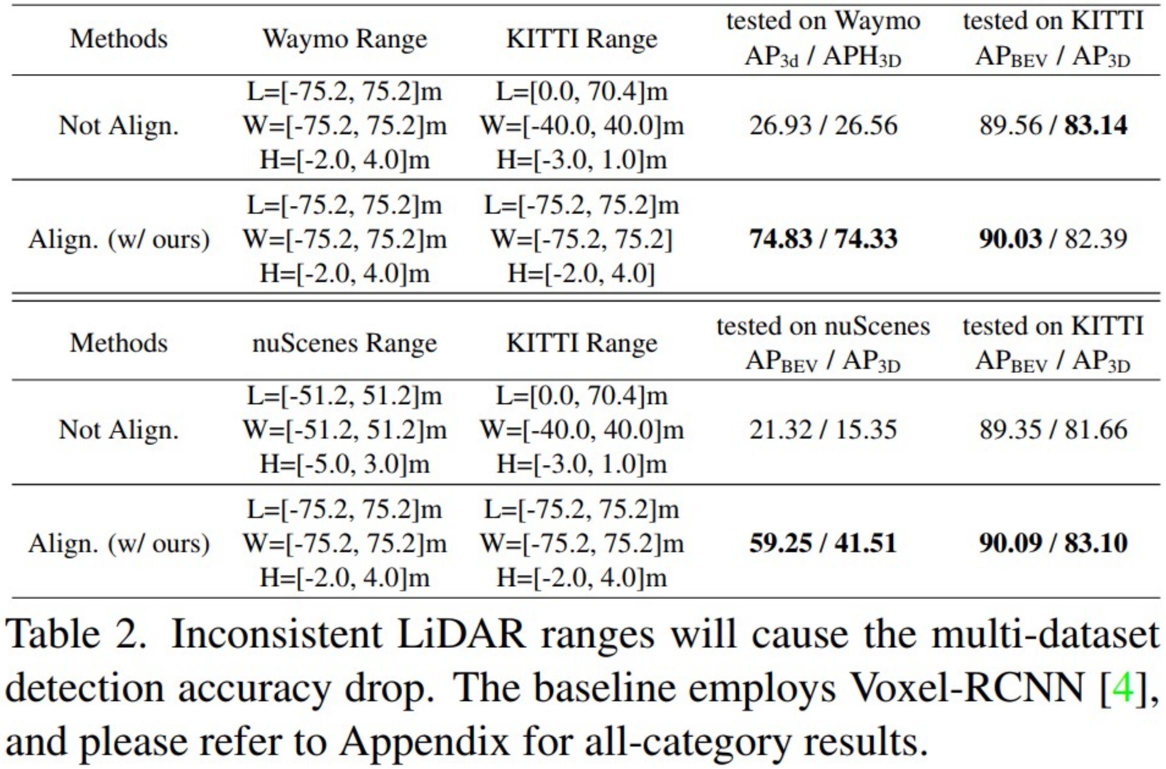
Parallel parking space :

Width = European / U.K. : 2,44 meters ; U.S. Compact : 8 ft ; U.S. Standard : 8'6" ; U.S. Standard Large : 9 ft
Length = European / U.K. : 4,88 meters ; U.S. Compact : 16 ft ; U.S. Standard : 18 ft ; U.S. Standard Large : 20 ft
Minimum height = European : 2,13 meters ; U.S. Compact : 7 ft.

MDF: Multi-dataset Domain Fusion

➢ Performance Drop with Inconsistent Point Range：


(a) (b) (c)



Methods	Waymo Range	KITTI Range	tested on Waymo AP_{3d} / APH_{3D}	tested on KITTI AP_{BEV} / AP_{3D}

Not Align.	L=[-75.2, 75.2]m
	W=[-75.2, 75.2]m
	H=[-2.0, 4.0]m
		L=[0.0, 70.4]m
		W=[-40.0, 40.0]m
		H=[-3.0, 1.0]m
			26.93 / 26.56	89.56 / **83.14**

Align. (w/ ours)	L=[-75.2, 75.2]m
	W=[-75.2, 75.2]m
	H=[-2.0, 4.0]m
		L=[-75.2, 75.2]m
		W=[-75.2, 75.2]
		H=[-2.0, 4.0]
			74.83 / 74.33	**90.03** / 82.39

Methods	nuScenes Range	KITTI Range	tested on nuScenes AP_{BEV} / AP_{3D}	tested on KITTI AP_{BEV} / AP_{3D}

Not Align.	L=[-51.2, 51.2]m
	W=[-51.2, 51.2]m
	H=[-5.0, 3.0]m
		L=[0.0, 70.4]m
		W=[-40.0, 40.0]m
		H=[-3.0, 1.0]m
			21.32 / 15.35	89.35 / 81.66

Align. (w/ ours)	L=[-75.2, 75.2]m
	W=[-75.2, 75.2]m
	H=[-2.0, 4.0]m
		L=[-75.2, 75.2]m
		W=[-75.2, 75.2]m
		H=[-2.0, 4.0]m
			59.25 / 41.51	**90.09 / 83.10**

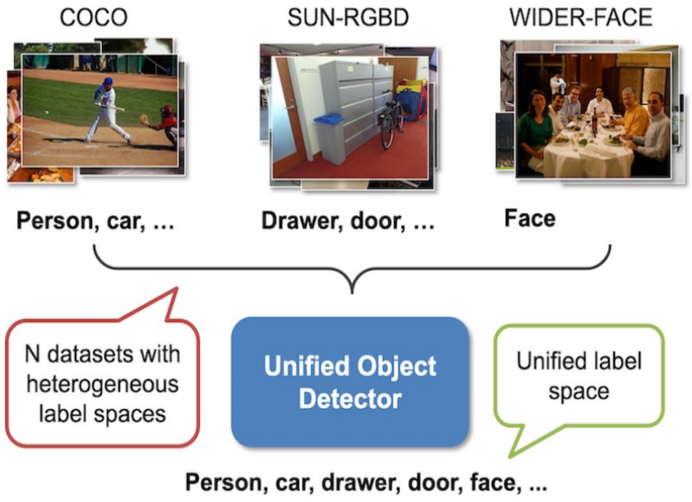
Table 2. Inconsistent LiDAR ranges will cause the multi-dataset detection accuracy drop. The baseline employs Voxel-RCNN [4], and please refer to Appendix for all-category results.

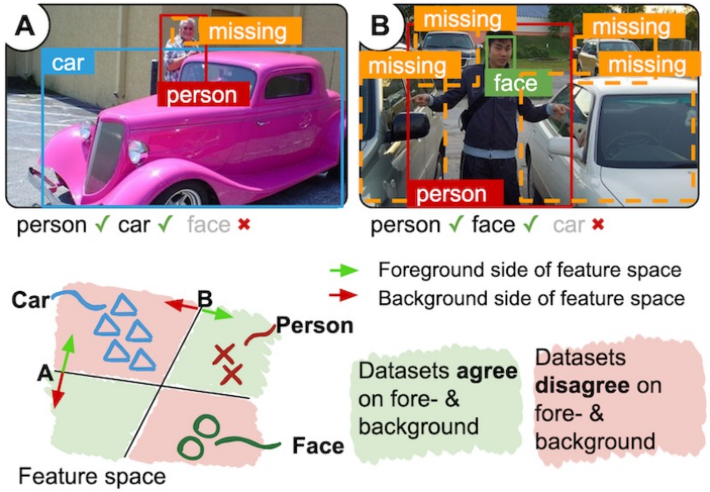
MDF: Multi-dataset Domain Fusion

➢Major Challenges：

- Cross-dataset Point Data Distribution Differences

- Cross-dataset Semantic Taxonomy Differences




A
car
missing
person
person ✓ car ✓ face ✖
B
missing
missing
face
missing
missing
person
person ✓ face ✓ car ✖
Foreground side of feature space
Background side of feature space
Car
B
Person
A
Face
Feature space
Datasets agree on fore- & background
Datasets disagree on fore- & background

➢Our Solution：

- Statistics-level Alignment: We design a dataset-specific BN layer that can replace BN module in 3D or 2D Backbone, to achieve an effective distribution of point cloud representations.

- Taxonomy-level Alignment: We design a Semantic-level Feature Coupling-and-Recoupling (C.R.) module that insert the off-the-shelf 3D detector to achieve the taxonomy-level alignment.

Uni3D: A Unified Baseline for Multi-dataset 3D Object Detection

➢Uni3D Framework

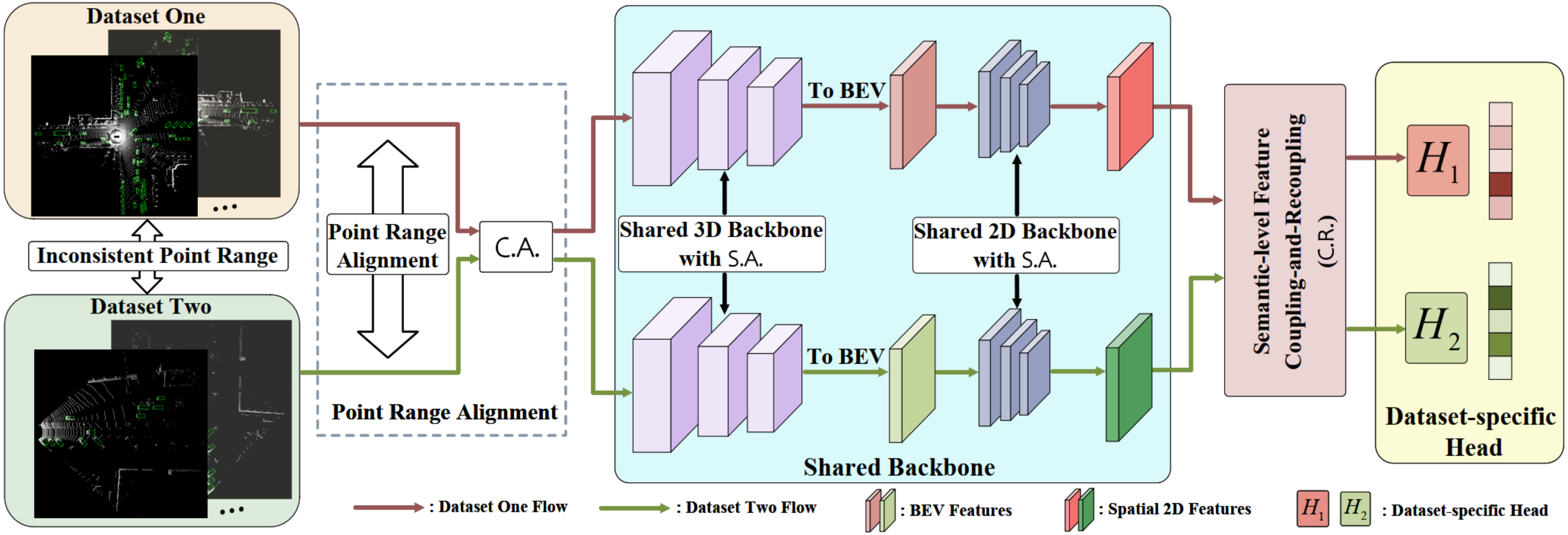


Figure 3. The overview of Uni3D including: 1) point range alignment, 2) parameter-shared 3D and 2D backbones with data-level correction operation, 3) semantic-level feature coupling-and-recoupling module, and 4) dataset-specific detection heads. `C.A.` denotes Coordinate-origin Alignment to reduce the adverse effects caused by point range alignment, and `S.A.` is the designed Statistics-level Alignment.

Based on 3DTrans Codebase

Uni3D: A Unified Baseline for Multi-dataset 3D Object Detection

➢ Semantic-level Feature Coupling-and-Recoupling （C.R.）：

$$f_{cat}^{bev} = [f_i^{bev}, \ldots, f_j^{bev}],$$
$$\hat{f}_{shared}^{bev} = [M_{shared} \odot \phi_d(Conv(f_{cat}^{bev}))]\, f_{cat}^{bev}, \tag{4}$$

$$\hat{f}_i^{bev} = SE_i(\hat{f}_{shared}^{bev}) + f_i^{bev},$$
$$\hat{f}_j^{bev} = SE_j(\hat{f}_{shared}^{bev}) + f_j^{bev}, \tag{5}$$

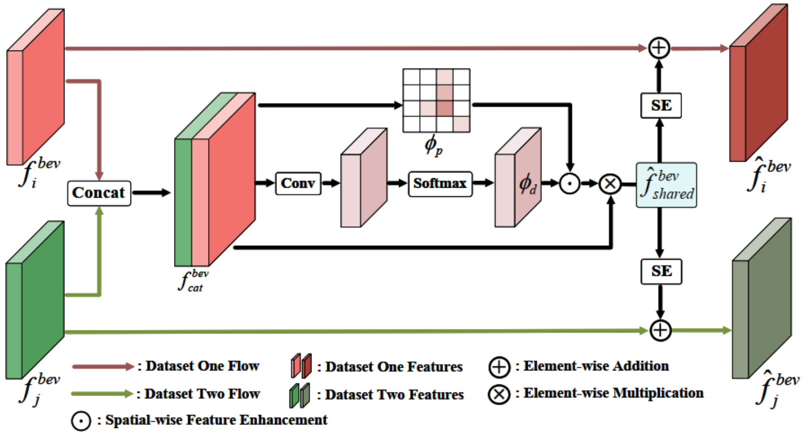


Figure 4. Semantic-level feature coupling-and-recoupling during the multi-dataset training stage.

Model	Option	Tested on Waymo			Tested on nuScenes		
		Vehicle	Pedestrian	Cyclist	Car	Pedestrian	Cyclist
Voxel-RCNN (Direct Merging)	-	66.67 / 66.23	60.36 / 54.08	52.03 / 51.25	51.40 / 31.68	15.04 / 9.99	5.40 / 3.87
Voxel-RCNN (w/ C.A.+S.A.+C.R.)	**BEV feature copy**	**75.26 / 74.77**	**75.46 / 68.75**	**65.02 / 63.12**	60.18 / **42.23**	**30.08 / 24.37**	**14.60** / 12.32
oxel-RCNN (w/ C.A.+S.A.+C.R.)	**BEV feature mask**	73.78 / 73.29	72.67 / 66.32	64.20 / 62.81	**60.37** / 40.66	29.57 / 23.51	14.13 / **12.42**

Table 11. Options for inference usage of the C.R. module: The model is jointly trained on Waymo and nuScenes, and evaluated on the validation of Waymo and nuScenes.

Model	Ensemble?	Vehicle	Pedestrian	Cyclist
Voxel-RCNN (w/ C.A.+S.A.)	No	75.16 / 74.67	74.83 / 68.07	64.68 / **63.73**
Voxel-RCNN (w/ C.A.+S.A.+E.N.)	Yes	70.59 / 70.07	73.56 / 66.86	62.75 / 61.82
Voxel-RCNN (w/ C.A.+S.A.+C.R.)	No	**75.26 / 74.77**	**75.46 / 68.75**	**65.02** / 63.12

Table 12. Comparisons against model ensemble: The model is trained on Waymo and nuScenes, and evaluated on the validation of Waymo. E.N. is the dataset-level model ensemble.

Uni3D: A Unified Baseline for Multi-dataset 3D Object Detection

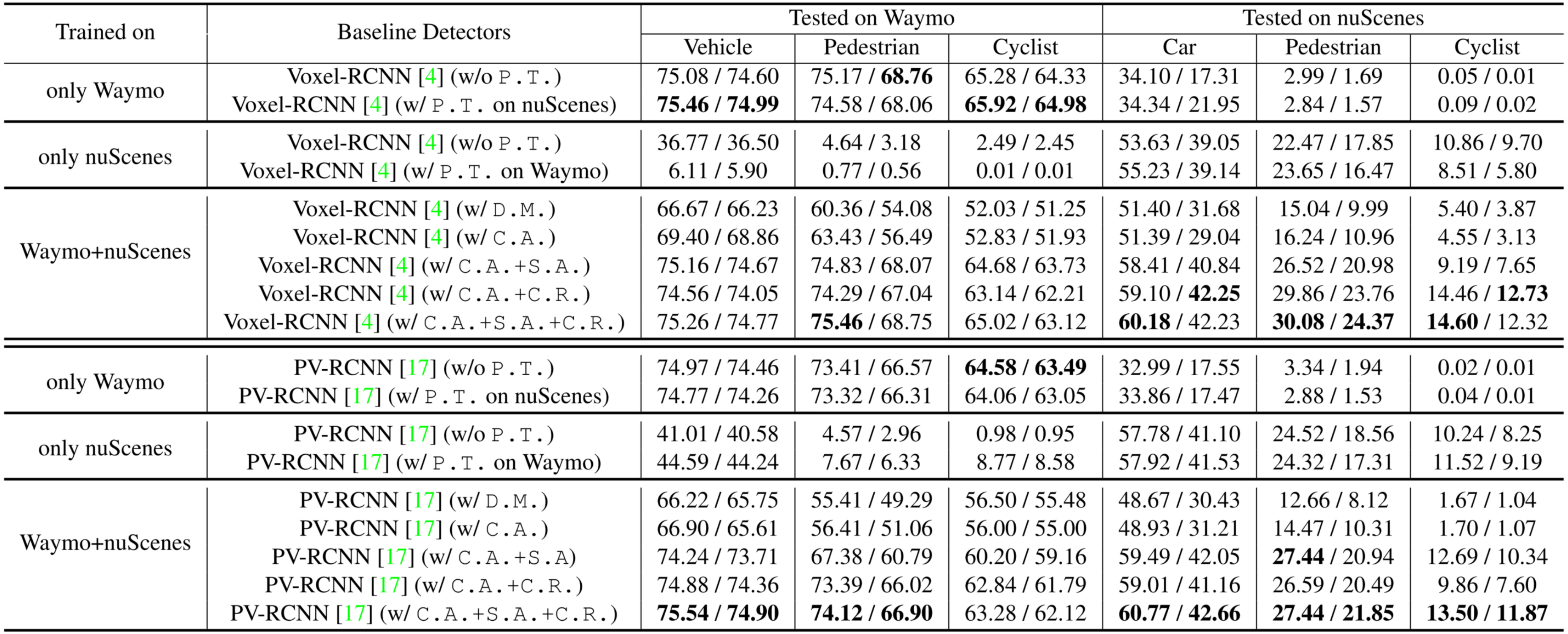
➢Our Results:

Baseline Model Design

- w/o P.T.(Single-dataset): Traditional Single-dataset 3D Object Detection Pipeline
- P.T.(Pre-training): Traditional Single-dataset 3D Object Detection Pipeline, Pretrained on Another Dataset
- D.M.(Direct Merging): Directly Merging Multi-datasets and Using the Above Singe-dataset Pipeline
- C.A.: Coordinate-origin Alignment,: To align point-cloud-range from different LiDAR sensor
- S.A.: Statistics-level Alignment: To reduce the data-level domain shift
- C.R.: Coupling-and-Recoupling baseline: To reduce the semantic-level domain shift

Uni3D: Our Results

➢Jointly train Waymo and nuScenes:

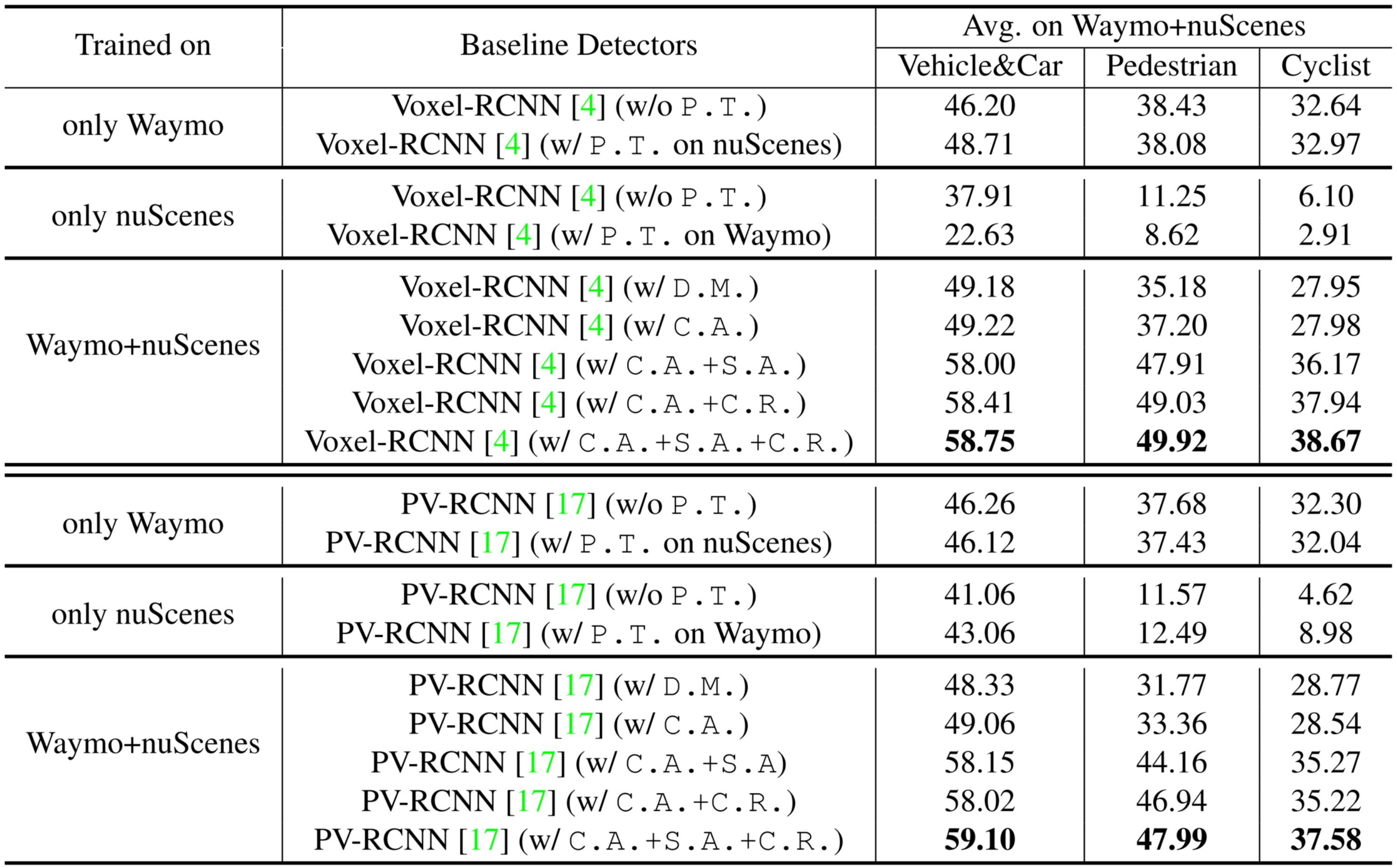


Trained on	Baseline Detectors	Tested on Waymo			Tested on nuScenes		
		Vehicle	Pedestrian	Cyclist	Car	Pedestrian	Cyclist
only Waymo	Voxel-RCNN [4] (w/o P.T.)	75.08 / 74.60	75.17 / **68.76**	65.28 / 64.33	34.10 / 17.31	2.99 / 1.69	0.05 / 0.01
	Voxel-RCNN [4] (w/ P.T. on nuScenes)	**75.46 / 74.99**	74.58 / 68.06	**65.92 / 64.98**	34.34 / 21.95	2.84 / 1.57	0.09 / 0.02
only nuScenes	Voxel-RCNN [4] (w/o P.T.)	36.77 / 36.50	4.64 / 3.18	2.49 / 2.45	53.63 / 39.05	22.47 / 17.85	10.86 / 9.70
	Voxel-RCNN [4] (w/ P.T. on Waymo)	6.11 / 5.90	0.77 / 0.56	0.01 / 0.01	55.23 / 39.14	23.65 / 16.47	8.51 / 5.80
Waymo+nuScenes	Voxel-RCNN [4] (w/ D.M.)	66.67 / 66.23	60.36 / 54.08	52.03 / 51.25	51.40 / 31.68	15.04 / 9.99	5.40 / 3.87
	Voxel-RCNN [4] (w/ C.A.)	69.40 / 68.86	63.43 / 56.49	52.83 / 51.93	51.39 / 29.04	16.24 / 10.96	4.55 / 3.13
	Voxel-RCNN [4] (w/ C.A.+S.A.)	75.16 / 74.67	74.83 / 68.07	64.68 / 63.73	58.41 / 40.84	26.52 / 20.98	9.19 / 7.65
	Voxel-RCNN [4] (w/ C.A.+C.R.)	74.56 / 74.05	74.29 / 67.04	63.14 / 62.21	59.10 / **42.25**	29.86 / 23.76	14.46 / **12.73**
	Voxel-RCNN [4] (w/ C.A.+S.A.+C.R.)	75.26 / 74.77	**75.46** / 68.75	65.02 / 63.12	**60.18** / 42.23	**30.08 / 24.37**	**14.60** / 12.32
only Waymo	PV-RCNN [17] (w/o P.T.)	74.97 / 74.46	73.41 / 66.57	**64.58 / 63.49**	32.99 / 17.55	3.34 / 1.94	0.02 / 0.01
	PV-RCNN [17] (w/ P.T. on nuScenes)	74.77 / 74.26	73.32 / 66.31	64.06 / 63.05	33.86 / 17.47	2.88 / 1.53	0.04 / 0.01
only nuScenes	PV-RCNN [17] (w/o P.T.)	41.01 / 40.58	4.57 / 2.96	0.98 / 0.95	57.78 / 41.10	24.52 / 18.56	10.24 / 8.25
	PV-RCNN [17] (w/ P.T. on Waymo)	44.59 / 44.24	7.67 / 6.33	8.77 / 8.58	57.92 / 41.53	24.32 / 17.31	11.52 / 9.19
Waymo+nuScenes	PV-RCNN [17] (w/ D.M.)	66.22 / 65.75	55.41 / 49.29	56.50 / 55.48	48.67 / 30.43	12.66 / 8.12	1.67 / 1.04
	PV-RCNN [17] (w/ C.A.)	66.90 / 65.61	56.41 / 51.06	56.00 / 55.00	48.93 / 31.21	14.47 / 10.31	1.70 / 1.07
	PV-RCNN [17] (w/ C.A.+S.A)	74.24 / 73.71	67.38 / 60.79	60.20 / 59.16	59.49 / 42.05	**27.44** / 20.94	12.69 / 10.34
	PV-RCNN [17] (w/ C.A.+C.R.)	74.88 / 74.36	73.39 / 66.02	62.84 / 61.79	59.01 / 41.16	26.59 / 20.49	9.86 / 7.60
	PV-RCNN [17] (w/ C.A.+S.A.+C.R.)	**75.54 / 74.90**	**74.12 / 66.90**	63.28 / 62.12	**60.77 / 42.66**	**27.44 / 21.85**	**13.50 / 11.87**

Table 3. Results of joint training on Waymo and nuScenes datasets. Following the existing 3D object detection works [17, 29, 30], we report the car (Vehicle on Waymo), pedestrian, and cyclist results under IoU threshold of 0.7, 0.5, and 0.5, respectively, and utilize AP and APH of LEVEL 1 metric on Waymo, and AP_{BEV} and AP_{3D} over 40 recall positions on nuScenes. The best detection results are marked using **bold**. Due to the page limitation, the average accuracy of multiple datasets is reported in Appendix.

Uni3D: Our Results

➢(AVG) Jointly train Waymo and nuScenes:

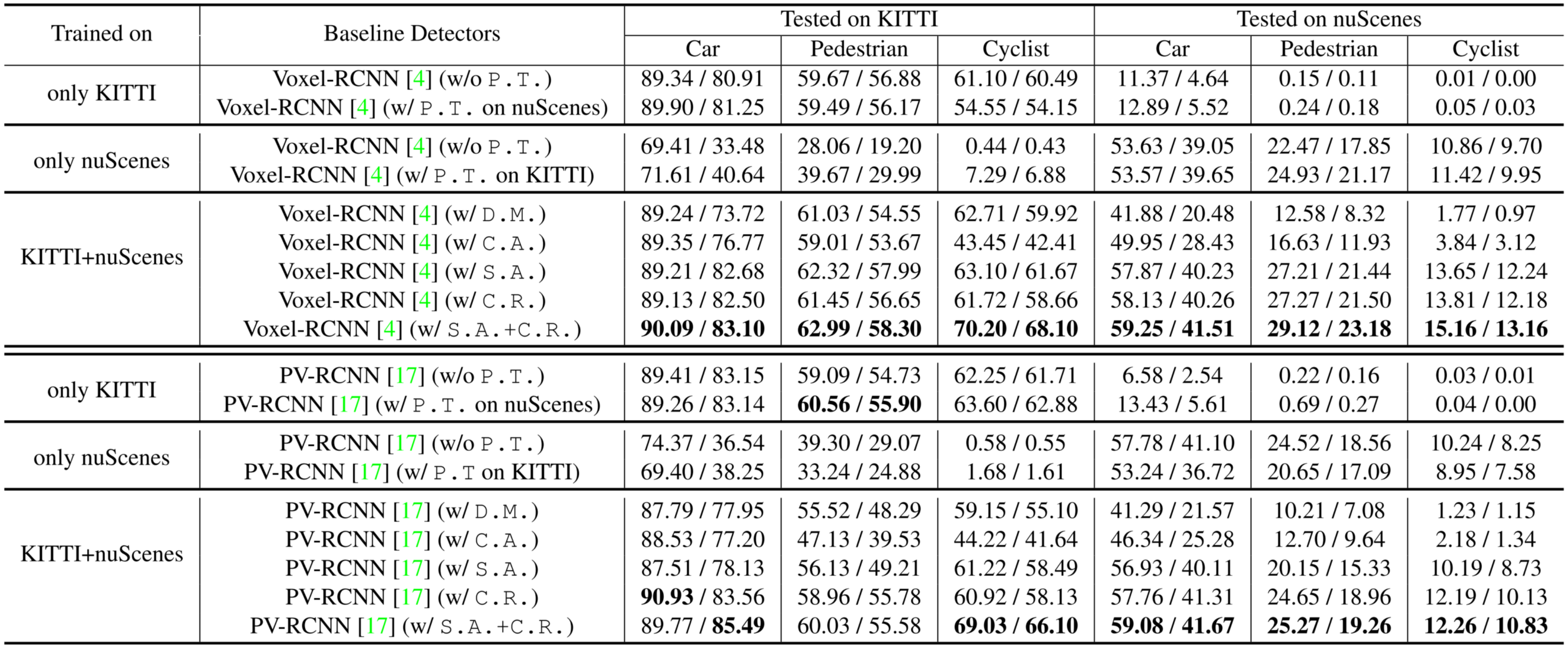


Trained on	Baseline Detectors	Avg. on Waymo+nuScenes		
		Vehicle&Car	Pedestrian	Cyclist
only Waymo	Voxel-RCNN [4] (w/o P.T.)	46.20	38.43	32.64
	Voxel-RCNN [4] (w/ P.T. on nuScenes)	48.71	38.08	32.97
only nuScenes	Voxel-RCNN [4] (w/o P.T.)	37.91	11.25	6.10
	Voxel-RCNN [4] (w/ P.T. on Waymo)	22.63	8.62	2.91
Waymo+nuScenes	Voxel-RCNN [4] (w/ D.M.)	49.18	35.18	27.95
	Voxel-RCNN [4] (w/ C.A.)	49.22	37.20	27.98
	Voxel-RCNN [4] (w/ C.A.+S.A.)	58.00	47.91	36.17
	Voxel-RCNN [4] (w/ C.A.+C.R.)	58.41	49.03	37.94
	Voxel-RCNN [4] (w/ C.A.+S.A.+C.R.)	**58.75**	**49.92**	**38.67**
only Waymo	PV-RCNN [17] (w/o P.T.)	46.26	37.68	32.30
	PV-RCNN [17] (w/ P.T. on nuScenes)	46.12	37.43	32.04
only nuScenes	PV-RCNN [17] (w/o P.T.)	41.06	11.57	4.62
	PV-RCNN [17] (w/ P.T. on Waymo)	43.06	12.49	8.98
Waymo+nuScenes	PV-RCNN [17] (w/ D.M.)	48.33	31.77	28.77
	PV-RCNN [17] (w/ C.A.)	49.06	33.36	28.54
	PV-RCNN [17] (w/ C.A.+S.A)	58.15	44.16	35.27
	PV-RCNN [17] (w/ C.A.+C.R.)	58.02	46.94	35.22
	PV-RCNN [17] (w/ C.A.+S.A.+C.R.)	**59.10**	**47.99**	**37.58**

Table 14. Average (Avg.) detection results of joint training on Waymo and nuScenes datasets. Here, we report the car (Vehicle on Waymo), pedestrian, and cyclist results under IoU threshold of 0.7, 0.5, and 0.5, respectively, and utilize AP of LEVEL 1 metric on Waymo, and AP_{3D} over 40 recall positions on nuScenes. The best detection results are marked using **bold**.

Uni3D: Our Results

➢Jointly train KITTI and nuScenes:

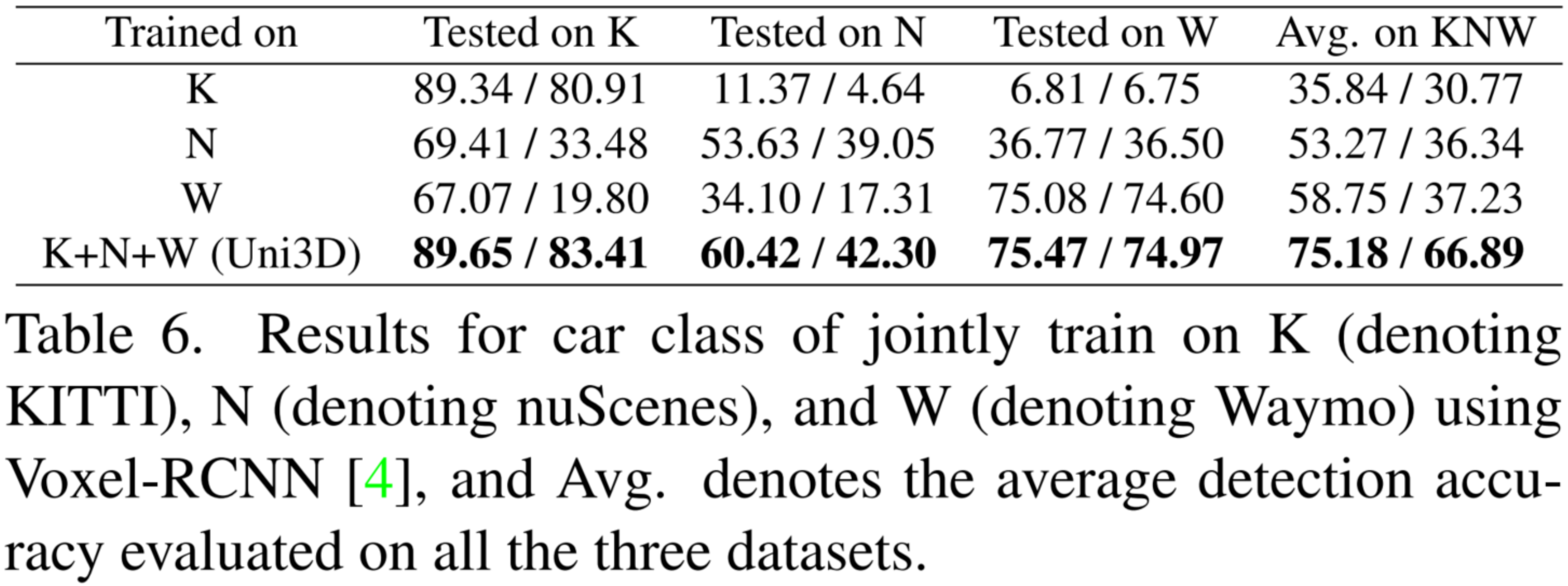


Trained on	Baseline Detectors	Tested on KITTI			Tested on nuScenes		
		Car	Pedestrian	Cyclist	Car	Pedestrian	Cyclist
only KITTI	Voxel-RCNN [4] (w/o P.T.)	89.34 / 80.91	59.67 / 56.88	61.10 / 60.49	11.37 / 4.64	0.15 / 0.11	0.01 / 0.00
	Voxel-RCNN [4] (w/ P.T. on nuScenes)	89.90 / 81.25	59.49 / 56.17	54.55 / 54.15	12.89 / 5.52	0.24 / 0.18	0.05 / 0.03
only nuScenes	Voxel-RCNN [4] (w/o P.T.)	69.41 / 33.48	28.06 / 19.20	0.44 / 0.43	53.63 / 39.05	22.47 / 17.85	10.86 / 9.70
	Voxel-RCNN [4] (w/ P.T. on KITTI)	71.61 / 40.64	39.67 / 29.99	7.29 / 6.88	53.57 / 39.65	24.93 / 21.17	11.42 / 9.95
KITTI+nuScenes	Voxel-RCNN [4] (w/ D.M.)	89.24 / 73.72	61.03 / 54.55	62.71 / 59.92	41.88 / 20.48	12.58 / 8.32	1.77 / 0.97
	Voxel-RCNN [4] (w/ C.A.)	89.35 / 76.77	59.01 / 53.67	43.45 / 42.41	49.95 / 28.43	16.63 / 11.93	3.84 / 3.12
	Voxel-RCNN [4] (w/ S.A.)	89.21 / 82.68	62.32 / 57.99	63.10 / 61.67	57.87 / 40.23	27.21 / 21.44	13.65 / 12.24
	Voxel-RCNN [4] (w/ C.R.)	89.13 / 82.50	61.45 / 56.65	61.72 / 58.66	58.13 / 40.26	27.27 / 21.50	13.81 / 12.18
	Voxel-RCNN [4] (w/ S.A.+C.R.)	**90.09 / 83.10**	**62.99 / 58.30**	**70.20 / 68.10**	**59.25 / 41.51**	**29.12 / 23.18**	**15.16 / 13.16**
only KITTI	PV-RCNN [17] (w/o P.T.)	89.41 / 83.15	59.09 / 54.73	62.25 / 61.71	6.58 / 2.54	0.22 / 0.16	0.03 / 0.01
	PV-RCNN [17] (w/ P.T. on nuScenes)	89.26 / 83.14	**60.56 / 55.90**	63.60 / 62.88	13.43 / 5.61	0.69 / 0.27	0.04 / 0.00
only nuScenes	PV-RCNN [17] (w/o P.T.)	74.37 / 36.54	39.30 / 29.07	0.58 / 0.55	57.78 / 41.10	24.52 / 18.56	10.24 / 8.25
	PV-RCNN [17] (w/ P.T on KITTI)	69.40 / 38.25	33.24 / 24.88	1.68 / 1.61	53.24 / 36.72	20.65 / 17.09	8.95 / 7.58
KITTI+nuScenes	PV-RCNN [17] (w/ D.M.)	87.79 / 77.95	55.52 / 48.29	59.15 / 55.10	41.29 / 21.57	10.21 / 7.08	1.23 / 1.15
	PV-RCNN [17] (w/ C.A.)	88.53 / 77.20	47.13 / 39.53	44.22 / 41.64	46.34 / 25.28	12.70 / 9.64	2.18 / 1.34
	PV-RCNN [17] (w/ S.A.)	87.51 / 78.13	56.13 / 49.21	61.22 / 58.49	56.93 / 40.11	20.15 / 15.33	10.19 / 8.73
	PV-RCNN [17] (w/ C.R.)	**90.93** / 83.56	58.96 / 55.78	60.92 / 58.13	57.76 / 41.31	24.65 / 18.96	12.19 / 10.13
	PV-RCNN [17] (w/ S.A.+C.R.)	89.77 / **85.49**	60.03 / 55.58	**69.03 / 66.10**	**59.08 / 41.67**	**25.27 / 19.26**	**12.26 / 10.83**

Table 4. Results of joint training on KITTI and nuScenes datasets. The experiment and evaluation settings follow Table 3.

Uni3D: Discussions

- ➢ Uni3D: Jointly train KITTI, nuScenes, and Waymo

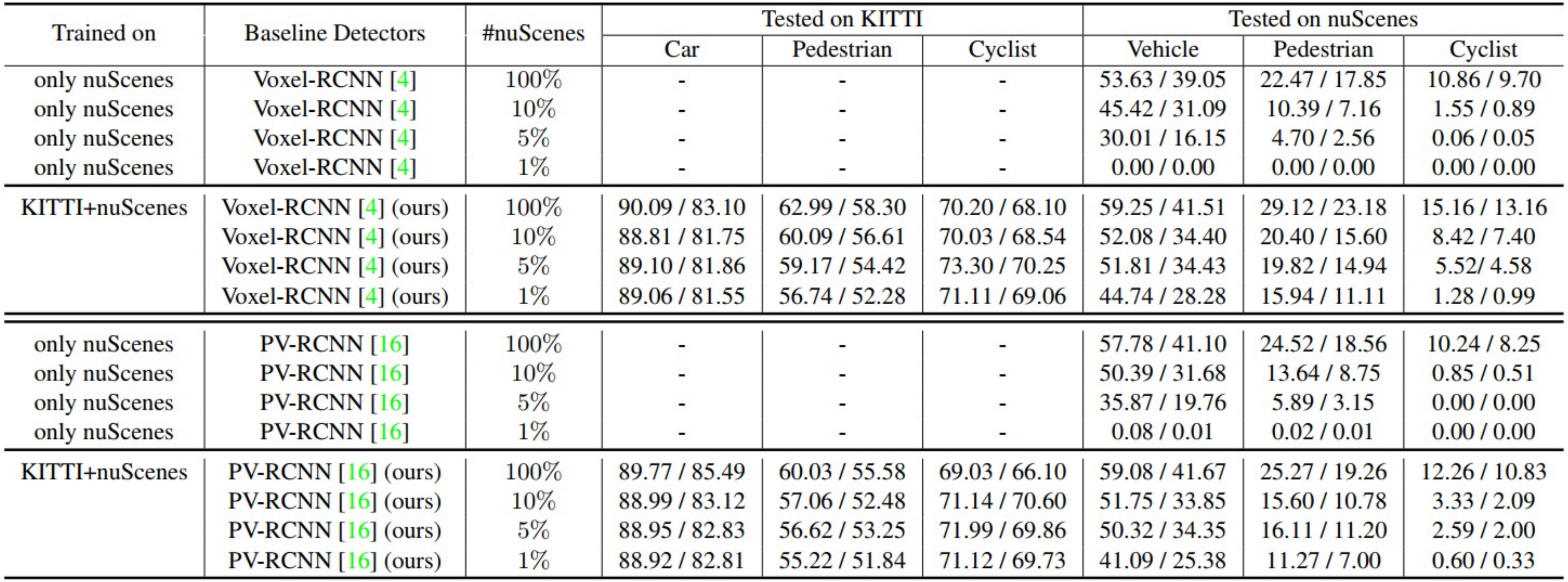


Trained on	Tested on K	Tested on N	Tested on W	Avg. on KNW
K	89.34 / 80.91	11.37 / 4.64	6.81 / 6.75	35.84 / 30.77
N	69.41 / 33.48	53.63 / 39.05	36.77 / 36.50	53.27 / 36.34
W	67.07 / 19.80	34.10 / 17.31	75.08 / 74.60	58.75 / 37.23
K+N+W (Uni3D)	**89.65 / 83.41**	**60.42 / 42.30**	**75.47 / 74.97**	**75.18 / 66.89**

Table 6. Results for car class of jointly train on K (denoting KITTI), N (denoting nuScenes), and W (denoting Waymo) using Voxel-RCNN [4], and Avg. denotes the average detection accuracy evaluated on all the three datasets.

Uni3D: Discussions

➢ Uni3D: Reduce the Data Acquisition Cost



Trained on	Baseline Detectors	#nuScenes	Tested on KITTI			Tested on nuScenes		
			Car	Pedestrian	Cyclist	Vehicle	Pedestrian	Cyclist
only nuScenes	Voxel-RCNN [4]	100%	-	-	-	53.63 / 39.05	22.47 / 17.85	10.86 / 9.70
only nuScenes	Voxel-RCNN [4]	10%	-	-	-	45.42 / 31.09	10.39 / 7.16	1.55 / 0.89
only nuScenes	Voxel-RCNN [4]	5%	-	-	-	30.01 / 16.15	4.70 / 2.56	0.06 / 0.05
only nuScenes	Voxel-RCNN [4]	1%	-	-	-	0.00 / 0.00	0.00 / 0.00	0.00 / 0.00
KITTI+nuScenes	Voxel-RCNN [4] (ours)	100%	90.09 / 83.10	62.99 / 58.30	70.20 / 68.10	59.25 / 41.51	29.12 / 23.18	15.16 / 13.16
	Voxel-RCNN [4] (ours)	10%	88.81 / 81.75	60.09 / 56.61	70.03 / 68.54	52.08 / 34.40	20.40 / 15.60	8.42 / 7.40
	Voxel-RCNN [4] (ours)	5%	89.10 / 81.86	59.17 / 54.42	73.30 / 70.25	51.81 / 34.43	19.82 / 14.94	5.52/ 4.58
	Voxel-RCNN [4] (ours)	1%	89.06 / 81.55	56.74 / 52.28	71.11 / 69.06	44.74 / 28.28	15.94 / 11.11	1.28 / 0.99
only nuScenes	PV-RCNN [16]	100%	-	-	-	57.78 / 41.10	24.52 / 18.56	10.24 / 8.25
only nuScenes	PV-RCNN [16]	10%	-	-	-	50.39 / 31.68	13.64 / 8.75	0.85 / 0.51
only nuScenes	PV-RCNN [16]	5%	-	-	-	35.87 / 19.76	5.89 / 3.15	0.00 / 0.00
only nuScenes	PV-RCNN [16]	1%	-	-	-	0.08 / 0.01	0.02 / 0.01	0.00 / 0.00
KITTI+nuScenes	PV-RCNN [16] (ours)	100%	89.77 / 85.49	60.03 / 55.58	69.03 / 66.10	59.08 / 41.67	25.27 / 19.26	12.26 / 10.83
	PV-RCNN [16] (ours)	10%	88.99 / 83.12	57.06 / 52.48	71.14 / 70.60	51.75 / 33.85	15.60 / 10.78	3.33 / 2.09
	PV-RCNN [16] (ours)	5%	88.95 / 82.83	56.62 / 53.25	71.99 / 69.86	50.32 / 34.35	16.11 / 11.20	2.59 / 2.00
	PV-RCNN [16] (ours)	1%	88.92 / 82.81	55.22 / 51.84	71.12 / 69.73	41.09 / 25.38	11.27 / 7.00	0.60 / 0.33

Table 7. Results of reducing the number of samples in nuScenes dataset under the nuScenes-KITTI consolidation setting.

Only 5% nuScenes data is available

Thanks for Listening

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory