

LiDAR-CS Dataset: LiDAR Point Cloud Dataset with Cross-Sensors for 3D Object Detection

Supplementary Material

October 9, 2022

1 Overview

In this supplementary material, we provide more evaluation results and analysis on the LiDAR-CS Dataset. First of all, we give more experimental analysis on the cross-sensors settings. Then, we extend the LiDAR-CS dataset by providing 70,000 frames point cloud with multiple settings such as different installation heights. In addition, comprehensive experiments are set to verify the influences of different settings. Finally, more experimental results and analyses on both cross-sensors and cross-settings are given at the end of the document.

2 Analysis on Cross-Sensors Evaluation

For better understanding, we give the qualitative comparison results in Fig. 1 and more detailed quantitative results are provided in Tab. 2 of the main paper. In Fig. 1, the horizontal and vertical axis represent different sensors and the vertical axis denotes the sensor used for training and horizontal axis denotes the sensor used for testing. In each block, the gray value represents the normalized mean mAP over the four baseline detectors e.g., PointPillars [1], SECOND [2], PV-RCNN [3], and CenterPoint [4]. Here, the PointRCNN is removed due to the abnormal performances on this type of dataset. To be clear, the darker the block, the higher the detection performance. From this figure, we can find that the diag-

onal blocks give relatively better performances compared with other blocks along the same row or same column. This means that the training and testing on the same source data give the best performance. In addition, we interestingly find that the ONCE-40 and VLD-64 sensors share similar performances. We argue that the two sensors may share the same scan line distribution rather than the total scan lines being different.

The detailed evaluation results including all the categories are provided in Tab. 2. However, we find that VLD-128 performs badly in the cross-evaluation, though it can provide 128 channels' point cloud. We guess this comes from the special vertical angles distribution. VLD-128 is designed to capture the context with a 200m range, however, the scan lines within 70m are far more sparse than VLD-64. While only objects within 70m are evaluated in our metric, therefore VLD-128 loses the superiority in cross-evaluation.

3 Cross-Settings Data

Not only the LiDAR sensor affects the distribution of scanned point cloud, the setting of the LiDAR sensor, such as the height and pitch angle, also affects the final points distribution. The endpoints of the laser beam varied while the sensor is set at different heights and angles. In this section, we explore the influence of different sensors settings for 3D object detection performance.

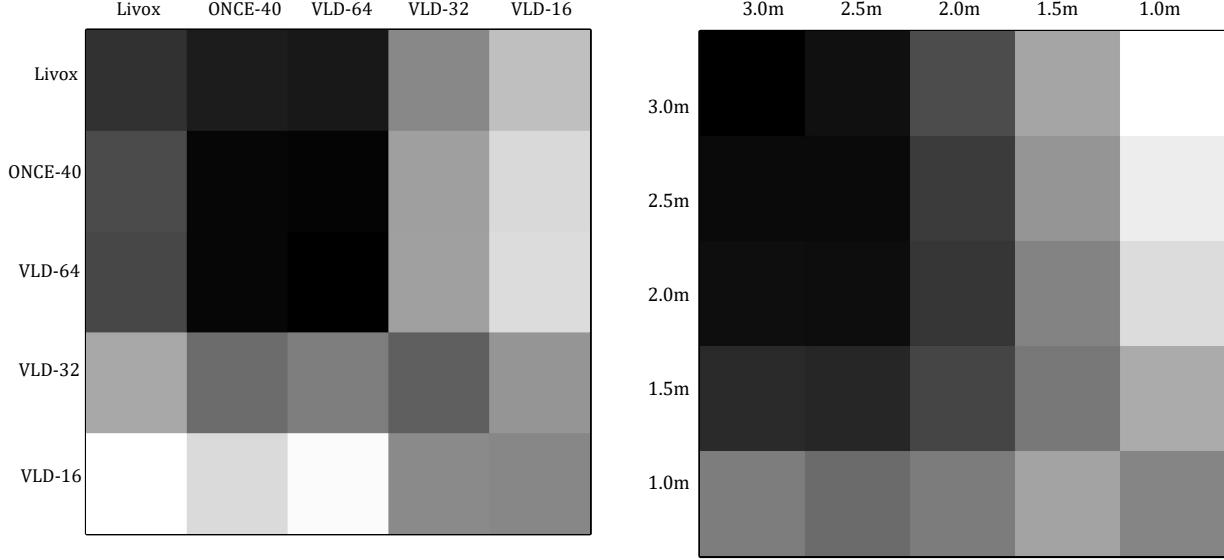


Figure 1: Qualitative cross-validation comparison results on different sensors, each block indicates the mean mAP over the four baseline detectors except for the PointRCNN. The darker block means a higher mean mAP. Here, the vertical axis denotes the sensor used for training and horizontal axis denotes the sensor used for testing.

3.1 Dataset Details

Considering that the LiDAR devices are usually placed vertically on the roof of the vehicle, it will have more practical significance to simulate different heights of LiDAR sensors, since different types of vehicles have different heights such as trucks and minibusses, etc. Five different heights are evaluated here as **1.0m**, **1.5m**, **2.0m**, **2.5m**, **3.0m** respectively. For a fair comparison, the VLD-64 is selected as the default LiDAR device here. Specifically, each of the groups has 14,000 frames and it shares the same scenarios and data splitting with cross-sensors data as introduced in the main paper. An example of the cross-settings data is shown in Fig. 3.

3.2 Cross-Settings Evaluation

We follow the same metric in main paper to evaluate the domain gaps across different LiDAR set-

tings. Five mainstream detectors are employed here, including **PointPillars** [5], **SECOND** [2], **PointRCNN** [6], **PV-RCNN** [3] and **CenterPoint** [4]. All the detectors are trained independently on the training split of each point cloud group and evaluated on the testing split of all the groups, respectively. The simplified table and complete table can be found in Tab. 1 and Tab. 3, respectively.

Furthermore, qualitative evaluation results with image representation are shown in Fig. 3 for better analysis. Some interesting conclusions can be conducted here: **1)** for each testing data (or for each column), the best performance comes from the model trained on the same sensor, which shares the same conclusion as in the cross-sensors dataset. **2)** for each training data (or for each row), testing data captured from sensors set in higher positions gives better performance, due to the better perspective for objects in the far distance.

Train	Val	VLD-64 (3.0m)				VLD-64 (2.5m)				VLD-64 (2.0m)				VLD-64 (1.5m)				VLD-64 (1.0m)			
		mAP	Car	Truck	Ped.																
VLD-64 (3.0m)	PointPillar	67.37	88.52	90.38	31.46	66.48	86.40	90.08	31.05	63.47	81.92	85.95	30.64	58.60	75.58	78.94	26.70	52.70	72.97	63.80	13.68
	SECOND	70.12	86.35	90.51	32.50	69.15	84.38	90.17	31.76	65.88	80.13	85.96	30.92	61.63	75.44	78.90	28.48	57.08	72.87	61.20	22.36
	PointRCNN	46.20	65.91	75.85	12.02	45.37	63.90	71.87	13.76	44.14	59.86	69.79	13.90	41.11	55.84	59.61	14.86	46.43	63.53	54.50	19.34
	PV-RCNN	73.93	91.24	92.95	30.62	72.88	89.38	90.98	31.00	70.60	87.09	88.85	30.00	66.36	80.86	82.09	28.03	62.39	78.40	66.86	23.25
	CenterPoint	81.08	88.56	90.73	60.98	80.81	88.31	90.51	61.16	77.10	84.08	86.33	58.85	72.26	77.70	78.92	55.81	68.31	75.06	63.09	52.41
VLD-64 (2.5m)	PointPillar	66.51	86.59	90.35	31.50	66.32	86.46	90.33	31.31	64.06	82.16	87.89	31.48	59.19	75.89	80.80	26.85	53.45	73.46	65.66	14.49
	SECOND	69.59	84.44	90.05	31.18	69.58	84.45	90.19	31.77	67.24	81.94	87.84	31.81	62.44	75.70	78.87	29.37	57.94	73.15	62.60	23.20
	PointRCNN	46.63	65.91	75.84	12.21	46.16	63.90	73.83	13.85	43.72	57.90	69.79	13.86	40.82	55.82	61.59	13.54	46.43	63.53	54.61	19.39
	PV-RCNN	73.55	91.20	92.75	30.39	73.65	91.27	92.75	30.86	71.64	87.19	90.71	31.13	67.40	82.81	82.19	29.19	63.47	80.37	66.93	24.32
	CenterPoint	80.91	88.51	90.39	60.63	80.84	88.45	90.62	61.07	77.59	84.21	88.34	58.93	73.00	79.60	79.32	56.61	69.29	76.86	63.32	53.76
VLD-64 (2.0m)	PointPillar	65.78	86.26	89.63	30.43	65.93	86.22	90.05	30.80	64.16	82.19	88.04	31.76	60.39	77.78	81.61	28.23	54.32	75.36	65.90	16.69
	SECOND	69.45	84.24	89.73	30.68	69.31	84.27	89.75	31.06	67.30	82.09	87.85	31.28	63.03	75.84	81.04	28.74	58.15	73.41	63.33	22.37
	PointRCNN	44.94	65.87	71.78	11.76	45.00	63.89	71.76	12.24	41.69	57.89	63.80	13.69	38.72	53.81	55.65	13.55	43.88	63.42	52.80	16.48
	PV-RCNN	73.61	91.20	92.82	30.51	73.88	91.28	92.97	31.36	71.97	89.15	90.95	30.74	68.74	84.83	86.36	29.90	64.91	82.52	71.28	24.95
	CenterPoint	80.77	88.41	90.57	60.58	80.70	88.39	90.28	60.63	78.00	86.08	88.18	59.23	73.53	79.86	81.42	56.98	70.14	77.51	65.68	54.48
VLD-64 (1.5m)	PointPillar	62.76	83.92	86.87	28.13	63.53	84.00	87.40	28.94	62.80	81.78	87.68	29.10	60.41	77.82	83.31	29.04	56.93	77.25	71.85	21.12
	SECOND	68.94	83.94	89.61	31.08	69.07	84.00	89.85	31.42	67.16	81.85	88.02	32.05	64.57	77.73	83.66	31.23	61.73	75.51	71.87	27.95
	PointRCNN	42.59	63.83	71.49	10.32	42.96	63.84	71.61	12.07	41.32	59.82	67.65	12.12	36.41	51.86	55.70	12.18	42.81	61.65	57.00	15.20
	PV-RCNN	72.49	89.15	91.81	29.91	72.76	89.22	92.32	29.98	71.19	87.15	90.47	30.44	68.83	84.87	86.29	30.22	66.58	82.82	75.23	27.72
	CenterPoint	79.85	86.22	89.86	59.94	79.49	86.20	89.73	60.09	77.32	84.04	87.86	59.04	74.18	79.88	83.41	57.63	72.33	79.13	72.03	57.65
VLD-64 (1.0m)	PointPillar	54.58	74.65	71.20	23.69	57.18	76.57	75.57	25.05	58.15	75.14	79.14	26.80	57.83	73.37	80.26	26.58	60.10	77.33	78.40	26.79
	SECOND	65.37	79.26	84.24	28.19	65.56	79.43	86.20	28.36	63.94	77.40	84.66	28.53	61.34	73.48	80.65	28.73	62.32	75.77	76.27	28.60
	PointRCNN	36.77	59.67	58.77	6.29	36.93	59.72	59.10	6.64	37.25	57.74	59.18	8.29	32.79	51.80	49.40	7.90	34.19	53.77	45.36	9.44
	PV-RCNN	69.46	84.88	88.22	28.68	70.42	86.86	88.83	30.07	69.51	84.88	89.12	30.09	67.46	82.69	84.98	30.25	69.11	84.70	81.19	30.95
	CenterPoint	77.51	83.68	86.93	57.52	77.61	83.69	87.25	58.15	75.52	81.57	86.72	56.46	72.59	77.53	82.49	54.85	73.78	79.73	77.78	58.12

Table 1: Cross evaluation on LiDAR-CS benchmark under five different LiDAR heights with five baseline detectors. “Ped.” is short for “Pedestrian”. The categories of “Bicyclist” and “Motorcyclist” have not been given in this table, but the **mAP** is still computed over all the five categories.

Analysis. The first conclusion is easy for understanding. Keep the testing data the same, the model trained with the same source data gives the best performance. For the second conclusion, we keep the trained model the same and change the testing data. In this case, the LiDAR sensor with a higher height gives better performance. This is because higher height represents a better viewpoint and less occlusion. All of these factors result in better perception results.

- [4] Tianwei Yin, Xingyi Zhou, and Philipp Krahenbuhl. Center-based 3d object detection and tracking. In *CVPR*, pages 11784–11793, 2021. [1](#), [2](#)
- [5] Alex H Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, and Oscar Beijbom. Pointpillars: Fast encoders for object detection from point clouds. In *CVPR*, pages 12697–12705, 2019. [2](#)
- [6] Shaoshuai Shi, Xiaogang Wang, and Hongsheng Li. Pointrcnn: 3d object proposal generation and detection from point cloud. In *CVPR*, pages 770–779, 2019. [2](#)

References

- [1] Alex H Lang, Sourabh Vora, Holger Caesar, Lubing Zhou, Jiong Yang, and Oscar Beijbom. Pointpillars: Fast encoders for object detection from point clouds. In *CVPR*, pages 12697–12705, 2019. [1](#)
- [2] Yan Yan, Yuxing Mao, and Bo Li. Second: Sparsely embedded convolutional detection. *Sensors*, 18(10):3337, 2018. [1](#), [2](#)
- [3] Shaoshuai Shi, Chaoxu Guo, Li Jiang, Zhe Wang, Jianping Shi, Xiaogang Wang, and Hongsheng Li. Pv-rcnn: Point-voxel feature set abstraction for 3d object detection. In *CVPR*, pages 10529–10538, 2020. [1](#), [2](#)

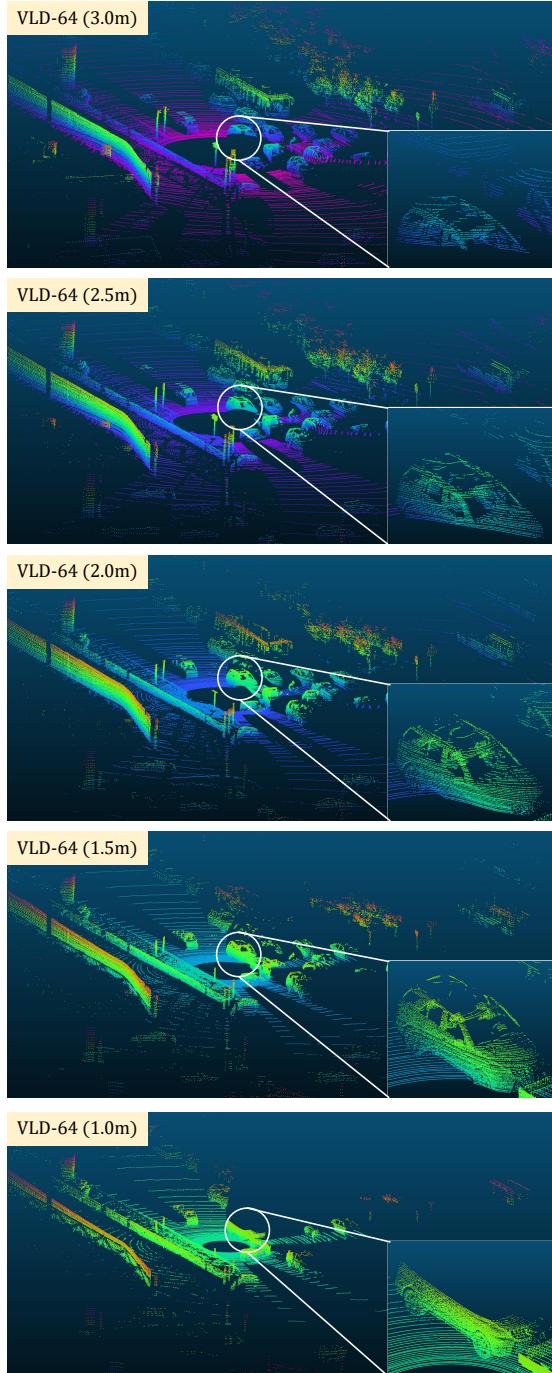


Figure 3: An example of the Cross-Settings dataset. All the point clouds are generated from the same scenario under different sensor heights. The points in the cycle are zoomed in and shown in the white boxes for a better view. The point clouds are colorized by the height of points. Better viewed in color.