In the supplementary material, we provide a discussion on the potential negative impact and limitations of PDV in Appendix A and Appendix B, respectively. We also provide additional detailed results on the Waymo Open Dataset [6] in Appendix C as well as on the KITTI dataset [2] in Appendix D. Finally, we provide visualizations of the point density variations across distance in Appendix E.

A. Potential Negative Impact

As autonomous driving progresses towards public use, PDV has the potential to be used for surveillance of civilians. Privacy is a large concern in the digital age, and it is important to ensure that this information is not used without consent or to maliciously track people's location.

Unlike other domains where safety concerns are minimal, reliable 3D object detection in autonomous driving is extremely important. Objects that are missed or even misclassified could have detrimental effects in subsequent decision-making tasks for the vehicle, running a risk for both the passengers and other individuals sharing the road. It is important to take the results of 3D object detection methods on improved benchmarks as only one aspect for reliable 3D object detection as increased performance on a metric is not indicative of the practicality of a method in a real-world scenario.

B. Limitations

Voxel Resolution. PDV shows the most improvement compared to other methods on large input spaces when the voxel resolution is limited, such as on the Waymo Open Dataset [6]. However, when the voxel resolution is sufficiently high, the improvements from PDV are less significant, as is the case on the KITTI dataset [2]. At higher voxel resolutions, the difference between the voxel center and the voxel point centroids decreases, resulting in less performance gains. Additionally, at high voxel resolutions, each occupied voxel represents a decent approximation of the point density, which can be captured through standard convolutions. Thus, PDV is most suited for when low voxel resolutions are necessary, such as for 360° detection, where the point density feature can be fully exploited for better performance.

LiDAR Dependency. PDV uses point density as an additional encoding feature, and therefore relies on the LiDAR specifications for accurate 3D object detection. Although PDV is tested on the KITTI dataset [2] and the Waymo Open Dataset [6], which have different LiDAR characteristics, it is necessary to expand other datasets to ensure that PDV is robust to different LiDAR sampling patterns. Some examples include Nuscenes [1], which uses a 32 beam LiDAR, and Cirrus [8], which uses a LiDAR designed for long range detection.

Method	Pedestrian LEVEL 1 mAP/mAPH		
	0-30m	30-50m	50m-Inf
PDV (Ours)	80.32/73.60	72.97/63.28	61.69/50.07
Method	Pedestrian LEVEL 2 mAP/mAPH		
	0-30m	30-50m	50m-Inf
PDV (Ours)	75.26/68.82	65.78/56.85	47.46/38.30

Table 9. Performance comparison on the Waymo Open Dataset with 202 validation sequences for 3D pedestrian (IoU = 0.5) detection across distance.

Method	Cyclist LEVEL 1 mAP/mAPH		
	0-30m	30-50m	50m-Inf
PDV (Ours)	80.86/79.83	62.61/61.45	46.23/44.12
Method	Cyclist LEVEL 2 mAP/mAPH		
	0-30m	30-50m	50m-Inf
PDV (Ours)	80.42/79.40	58.95/57.87	43.05/41.09

Table 10. Performance comparison on the Waymo Open Dataset with 202 validation sequences for 3D cyclist (IoU = 0.5) detection across distance.

Adverse Weather Conditions. PDV uses point density to provide additional encoding in the second stage. If there is a significant distribution shift in the LiDAR point density during inference time, there may be significant degradation in performance. For example, as highlighted in SPG [9], the Waymo Kirkland dataset [6] has a different distribution of points, where rainy weather results in a significant number of points missing on objects. Since PDV relies on point density, a sudden shift in point density variations due to weather may cause severe degradation in performance. A future improvement to PDV is to ensure robustness to potential shifts in point density distributions during inference.

C. More Waymo Dataset Results

We show the Waymo Open Dataset validation results for the pedestrian and cyclist classes across distance for PDV using the first LiDAR return only in Table 9 and Table 10, respectively. The distance evaluation is broken down into three categories: 0 to 30 meters, 30 to 50 meters, and beyond 50 meters. We also show the multi-class results for PDV using first and second LiDAR return in Table 11. Overall, PDV utilizes the additional points from the second LiDAR return better than PV-RCNN++, providing the largest performance increase on the cyclist class.

D. More KITTI Dataset Results

We provide KITTI dataset results on the *val* set on 3D and BEV for $AP|_{R_{11}}$ and $AP|_{R_{40}}$ in Table 12, Table 13, and Table 14. PDV's performance is compared to PV-RCNN [4] and CT3D [3].

Method	Veh. (LEVEL_1)		Veh. (LEVEL_2)		Ped. (LEVEL_1)		Ped. (LEVEL_2)		Cyc. (LEVEL_1)		Cyc. (LEVEL_2)	
	mAP	mAPH	mAP	mAPH	mAP	mAPH	mAP	mAPH	mAP	mAPH	mAP	mAPH
PV-RCNN† [4]	77.51	76.89	68.98	68.41	75.01	65.65	66.04	57.61	67.81	66.35	65.39	63.98
PV-RCNN++ [5]	78.79	78.21	70.26	69.71	76.67	67.15	**68.51**	59.72	68.98	67.63	66.48	65.17
PDV (Ours)	**79.43**	**78.89**	**70.47**	**69.98**	**76.94**	**68.09**	68.07	**60.00**	**71.45**	**70.18**	**69.11**	**67.88**
Improvement	*+0.64*	*+0.68*	*+0.21*	*+0.27*	*+0.27*	*+0.94*	*-0.44*	*+0.28*	*+2.47*	*+2.55*	*+2.63*	*+2.71*

Table 11. Performance comparison on the Waymo Open Dataset with 202 validation sequences for 3D vehicle (IoU = 0.7), pedestrian (IoU = 0.5) and cyclist (IoU = 0.5) detection using first and second LiDAR return. †: Results are from [5].

Method	Car 3D (R_{11})			Pedestrian 3D (R_{11})			Cyclist 3D (R_{11})		
	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
PV-RCNN⋆ [4]	89.35	83.69	78.70	64.60	57.90	53.23	85.22	70.47	65.75
CT3D⋆ [3]	89.11	**85.04**	78.76	61.74	56.28	52.51	85.04	71.71	68.05
PDV (Ours)	**89.52**	84.03	**79.09**	**65.83**	**61.18**	**55.87**	**90.48**	**73.23**	**69.55**
Improvement	*+0.41*	*-1.01*	*+0.33*	*+1.23*	*+3.28*	*+2.64*	*+5.26*	*+1.52*	*+1.50*

Table 12. 3D detection results on the KITTI *val* set for car, pedestrian, and cyclist classes using $AP|_{R_{11}}$. ⋆: Results are taken from publicly released models [3, 4, 7].

Method	Car BEV (R_{11})			Pedestrian BEV (R_{11})			Cyclist BEV (R_{11})		
	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
PV-RCNN⋆ [4]	90.09	87.90	87.41	67.01	61.38	56.10	86.79	73.55	69.69
CT3D⋆ [3]	90.25	88.18	87.78	64.23	59.84	55.76	**90.94**	73.68	**71.21**
PDV (Ours)	**90.33**	**88.33**	**87.91**	**69.01**	**63.54**	**59.46**	90.77	**73.75**	**71.21**
Improvement	*+0.08*	*+0.15*	*+0.13*	*+2.00*	*+2.16*	*+3.36*	*-0.17*	*+0.07*	*+0.00*

Table 13. BEV detection results on the KITTI *val* set for car, pedestrian, and cyclist classes using $AP|_{R_{11}}$. ⋆: Results are taken from publicly released models [3, 4, 7].

Method	Car BEV (R_{40})			Pedestrian BEV (R_{40})			Cyclist BEV (R_{40})		
	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
PV-RCNN⋆ [4]	93.02	90.33	88.53	67.97	60.52	55.80	91.02	74.54	69.92
CT3D⋆ [3]	**95.92**	**91.35**	89.29	64.41	59.18	54.86	92.60	75.40	71.31
PDV (Ours)	93.60	91.14	**90.74**	**69.40**	**63.42**	**58.70**	**93.09**	**76.08**	**71.46**
Improvement	*-2.32*	*-0.21*	*+1.45*	*+1.43*	*+2.90*	*+2.90*	*+0.49*	*+0.68*	*+0.15*

Table 14. BEV detection results on the KITTI *val* set for car, pedestrian, and cyclist classes using $AP|_{R_{40}}$. ⋆: Results are taken from publicly released models [3, 4, 7].

E. Point Density Distance Plots

Figure 7 shows point density across distance plots for PV-RCNN [4] and PDV. By using the relationship between distance and point density (number of points within each final bounding box prediction), PDV effectively reduces the number of false positives outside the distribution of training samples across distance.

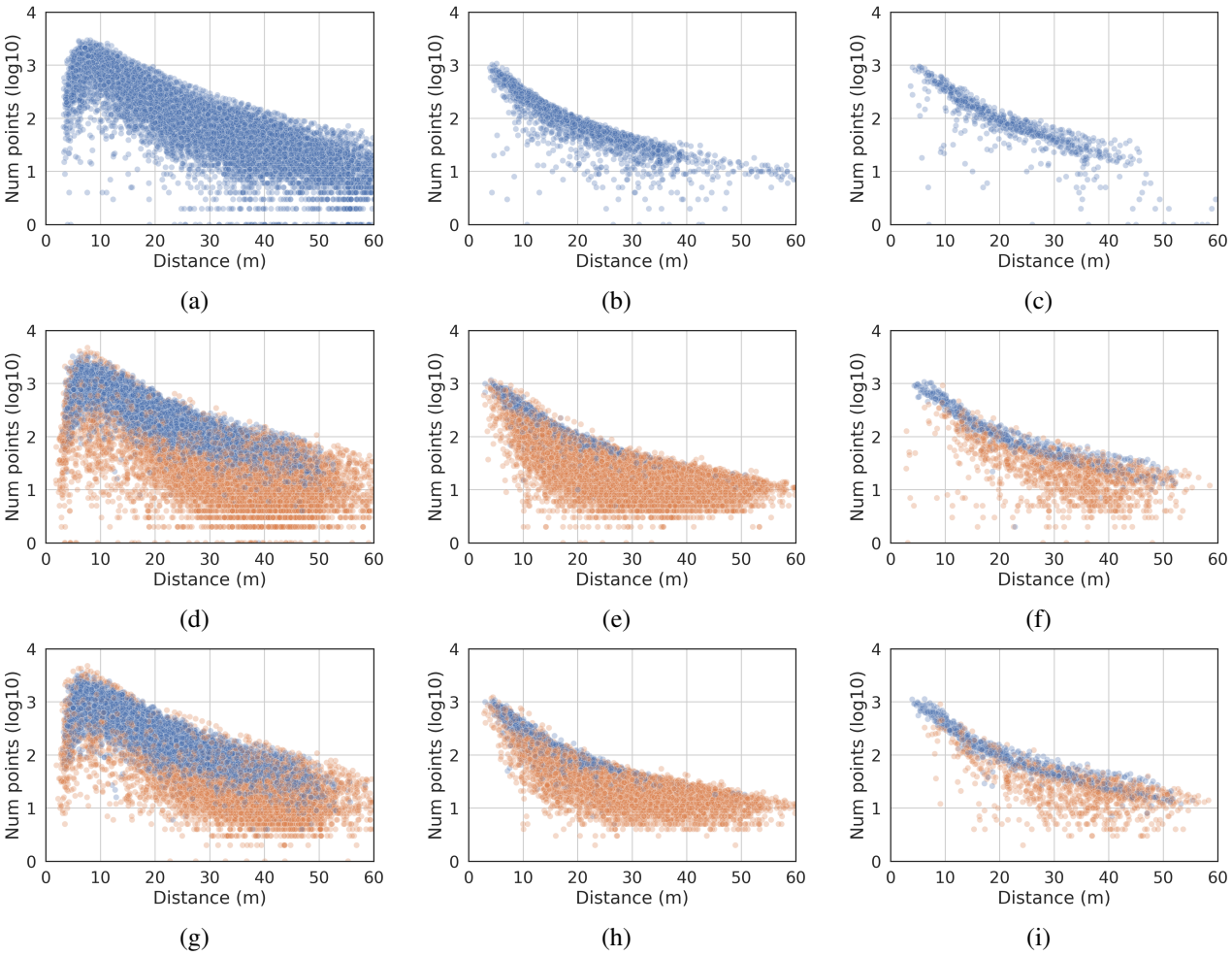

Figure 7. From left to right column: Number of points in ground truth boxes across distance for cars, pedestrians and cyclists on the KITTI dataset. The first row shows the distribution of training samples on the *train* split. The second and third row show the predictions of PV-RCNN [4] and PDV on the KITTI *val* split, respectively. Blue predictions are true positives while orange predictions are false positives for cars (IoU < 0.7), pedestrians (IoU < 0.5), and cyclists (IoU < 0.5).

References

[1] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. *CVPR*, 2020. 1

[2] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. *CVPR*, 2012. 1

[3] Hualian Sheng, Sijia Cai, Yuan Liu, Bing Deng, Jianqiang Huang, Xian-Sheng Hua, and Min-Jian Zhao. Improving 3d object detection with channel-wise transformer. *ICCV*, 2021. 1, 2

[4] 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. *CVPR*, 2020. 1, 2, 3

[5] Shaoshuai Shi, Li Jiang, Jiajun Deng, Zhe Wang, Chaoxu Guo, Jianping Shi, Xiaogang Wang, and Hongsheng Li. Pv-rcnn++: Point-voxel feature set abstraction with local vector representation for 3d object detection. *arXiv preprint*, 2021. 2

[6] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. *CVPR*, 2020. 1

[7] OpenPCDet Development Team. Openpcdet: An open-source toolbox for 3d object detection from point clouds. https://github.com/open-mmlab/OpenPCDet, 2020. 2

[8] Ze Wang, Sihao Ding, Ying Li, Jonas Fenn, Sohini Roychowdhury, Andreas Wallin, Lane Martin, Scott Ryvola, Guillermo Sapiro, and Qiang Qiu. Cirrus: A long-range bi-pattern lidar dataset. *ICRA*, 2021. 1

[9] Qiangeng Xu, Yin Zhou, Weiyue Wang, Charles R Qi, and Dragomir Anguelov. Spg: Unsupervised domain adaptation for 3d object detection via semantic point generation. *ICCV*, 2021. 1