6. Supplemental Materials

We provide additional qualitative and quantitative results.

6.1. Qualitative results

To qualitatively evaluate our self-supervised pre-training backbone, we used a K-means algorithm on the backbone's embedding right after the pre-training stage (without fine-tuning). In particular, we extracted the 2D-grid BEV embedding of each scene from 1% samples of Waymo [43] using the pre-trained backbone. We extracted each pixel embedding from the BEV of each scene to form a dataset. We used the K-means algorithm with 20 clusters to cluster these embeddings and project each BEV pixel's resulting cluster back onto their corresponding locations in the original point cloud. Fig. 6, Fig. 7, and Fig. 8 depict qualitative examples for Pedestrians, Cyclists, and Vehicles classes. It shows that even without any supervision the backbone's embedding encapsulates object awareness, where objects from the same category are clustered together. We provide more examples at the end of the supplemental material.

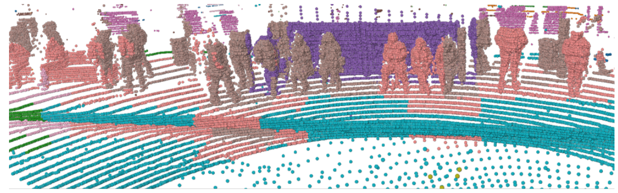

Figure 6. **Qualitative result.** A scene with pedestrians.

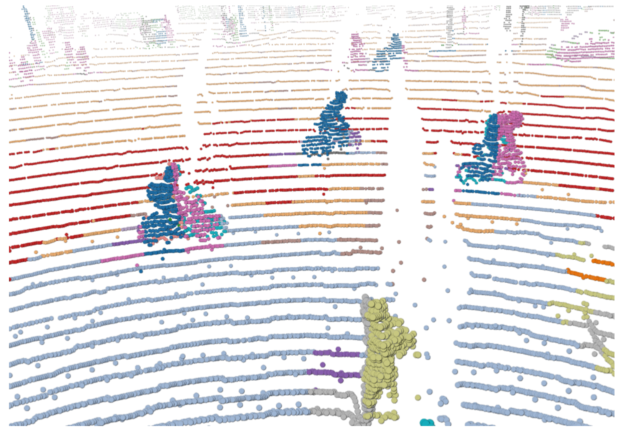

Figure 7. **Qualitative result.** A scene with cyclists.

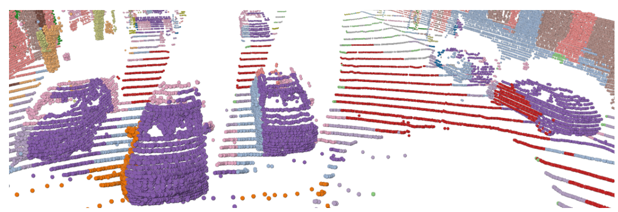

Figure 8. **Qualitative result.** A scene with vehicles.

6.2. Quantitative results

We present additional in-domain and out-of-domain quantitative results. Specifically, we provide 3D detection results on the KITTI dataset [14], including various difficulty levels. Additionally, we present 3D detection results for the Waymo dataset [43] on the Level-1 and Level-2 difficulty levels. Finally, we provide all the experiments related to data efficiency with and without frozen features.

6.2.1. Transfer learning on KITTI (out-of-domain).

We start by pre-training on the entire Waymo training split. Following this, we perform fine-tuning with varying amounts of labeled data from KITTI's train split and report the results on the entire validation split. Specifically, we partition the train set into 20%, 50%, and 100%, resulting in approximately $0.7K$, $1.9K$, and $3.7K$ scenes, respectively.

In Table 7, we present the results for the PV-RCNN [36] detector on the KITTI 3D detection benchmark. The results demonstrate that our approach consistently improves performance compared to training from scratch. Notably, the improvement is more significant when a smaller amount of labeled data is available. When using all training examples (100%), we achieve results on par with those presented in [55] and [3].

6.2.2. Object detection on Waymo dataset (in-domain)

As a complement to the results presented in Table 2, where we reported results for Waymo Level-2 difficulty, we now provide results for the Level-1 difficulty aswell. To ensure consistency with the common protocol [44], we conducted fine-tuning with 20% of labeled examples (approximately 31.6K scenes) from the training set, training for 30 epochs, and subsequently evaluated on the validation set.

The results are detailed in Table 8 and are compared to other state-of-the-art pre-training methods, namely GCC-3D [23] and ProposalContrast [55]. Notably, GCC-3D and ProposalContrast report results for Level-2 and are presented in this study for reference.

Our findings demonstrate that our approach not only improves performance over the baseline (training from scratch) but also outperforms other methods. It's worth noting that our approach achieves these improvements while exclusively utilizing the training split data for pre-training, in contrast to other methods that also leverage the validation split.

6.2.3. Data efficiency on Waymo dataset (in-domain)

In this experiment, we assess the performance of our pre-trained backbones in a data-efficient setting, where we employ different amounts of labeled data. Specifically, we di-

Labels	Method	mAP Mod.	*Car* Easy	*Car* Mod.	*Car* Hard	*Pedestrian* Easy	*Pedestrian* Mod.	*Pedestrian* Hard	*Cyclist* Easy	*Cyclist* Mod.	*Cyclist* Hard
20%	PV-RCNN	66.71	91.81	82.52	80.11	58.78	53.33	47.61	86.74	64.28	59.53
	ProposalContrast	68.13	91.96	82.65	80.15	62.58	55.05	50.06	88.58	66.68	62.32
	PatchContrast (Ours)	**70.75**	91.81	82.63	81.83	65.95	57.77	52.94	90.54	71.84	67.25
50%	PV-RCNN	69.63	91.77	82.68	81.9	63.70	57.10	52.77	89.77	69.12	64.61
	ProposalContrast	71.76	92.29	82.92	82.09	65.82	59.92	55.06	91.87	72.45	67.53
	PatchContrast (Ours)	**72.39**	91.78	84.47	82.23	68.21	60.76	54.84	90.59	71.94	67.37
100%	PV-RCNN	70.57	-	84.50	-	-	57.06	-	-	70.14	-
	GCC-3D	71.26	-	-	-	-	-	-	-	-	-
	STRL	71.46	-	84.70	-	-	57.80	-	-	71.88	-
	PointContrast	71.55	91.40	84.18	82.25	65.73	57.74	52.46	91.47	72.72	67.95
	ProposalContrast	72.92	92.45	84.72	82.47	68.43	60.36	55.01	92.77	73.69	69.51
	ALSO	72.96	-	84.68	-	-	60.16	-	-	74.04	-
	PatchContrast (Ours)	**72.97**	92.08	84.67	82.35	66.95	59.92	54.43	91.83	74.33	69.83

Table 7. **Transfer learning on KITTI.** Performance comparison on the KITTI validation set. The improvement is more significant when a smaller amount of labeled data is available.

Method	Overall (L1) mAP	Overall (L1) mAPH	Overall (L2) mAP	Overall (L2) mAPH	*Vehicle* (L1) AP	*Vehicle* (L1) APH	*Vehicle* (L2) AP	*Vehicle* (L2) APH	*Ped.* (L1) AP	*Ped.* (L1) APH	*Ped.* (L2) AP	*Ped.* (L2) APH	*Cyc.* (L1) AP	*Cyc.* (L1) APH	*Cyc.* (L2) AP	*Cyc.* (L2) APH
PV-RCNN	71.09	66.74	64.84	60.86	75.41	74.74	67.44	66.80	71.98	61.24	63.70	53.95	65.88	64.25	63.39	61.82
GCC-3D	-	-	61.30	58.18	-	-	65.65	65.10	-	-	55.54	48.02	-	-	62.72	61.43
ProposalContrast	-	-	62.62	59.28	-	-	66.04	65.47	-	-	57.58	49.51	-	-	64.23	62.86
PatchContrast (Ours)	**74.59**	**70.46**	**67.91**	**64.14**	76.90	76.29	68.40	67.84	75.51	65.41	66.62	57.48	71.35	69.68	68.72	67.11
CenterPoint	72.66	69.99	66.48	64.01	72.76	72.23	64.91	64.42	74.19	67.96	66.03	60.34	71.04	69.79	68.49	67.28
GCC-3D	-	-	65.29	62.79	-	-	63.97	63.47	-	-	64.23	58.47	-	-	67.68	66.44
ProposalContrast	-	-	66.42	63.85	-	-	64.94	64.42	-	-	66.13	60.11	-	-	68.19	67.01
PatchContrast (Ours)	**73.22**	**70.58**	**67.02**	**64.57**	72.84	72.31	64.73	64.25	74.99	68.85	67.10	61.45	71.84	70.59	69.22	68.01

Table 8. **3D detection results on Waymo.** Performance comparison of Level-1 (L1) and Level-2 (L2) on the Waymo validation set, trained on 20% of the Waymo train set, demonstrates that we outperform previous state-of-the-art methods.

vided the Waymo training set into two groups comprising 399 sequences, equal to about $80K$ frames. The first 399 sequences were used for pre-training, while various amounts of labeled data from the remaining 399 sequences were utilized for fine-tuning.

We conducted fine-tuning for CenterPoints [56], PV-RCNN [36], and SECOND [49], using 1%, 5%, and 10% of Waymo's [43] train set for 12 epochs, followed by evaluation on the validation set. Each experiment was repeated 3 times for consistency.

In Table 9, we provide results for both Level-1 (L1) and Level-2 (L2), reporting both the averages and the standard deviations for each experiment. The results demonstrate that our pre-trained framework delivers substantial benefits, particularly when working with limited labeled data.

6.2.4. Data efficiency on Waymo with frozen features (in-domain)

Similar to the linear classification protocol utilized in 2D image domains [6, 18], we propose freezing the features of the backbone and training a detection head on top of them. This approach allows us to evaluate the feature embeddings before the fine-tuning process overwrites the backbone's weights.

As in the previous subsection, we partition the Waymo training set into two groups consisting of 399 sequences. The first 399 sequences are used for pre-training, and different amounts of labeled data from the remaining 399 sequences are employed for fine-tuning. We subsequently evaluate the model on the validation set.

Results for both Level-1 and Level-2 difficulties of Waymo, using different percentages of labeled data, are reported in Table 10. Each experiment was conducted 3 times for consistency, and we provided both the means and standard deviations for each experiment. For the detector, we adopt CenterPoint [56] and employ the 1x scheduler (12 epochs). The results demonstrate that our embedding captures meaningful information about 3D point cloud scenes without any fine-tuning while outperforming the previous state-of-the-art method.

6.3. More qualitative results

Below we provide more qualitative results of our pre-trained backbone's embedding without any fine-tuning.

Labels	Method	*Vehicle* (L1) AP	*Vehicle* (L1) APH	*Vehicle* (L2) AP	*Vehicle* (L2) APH	*Ped.* (L1) AP	*Ped.* (L1) APH	*Ped.* (L2) AP	*Ped.* (L2) APH	*Cyc.* (L1) AP	*Cyc.* (L1) APH	*Cyc.* (L2) AP	*Cyc.* (L2) APH	Average and std (L1) AP	Average and std (L1) APH	Average and std (L2) AP	Average and std (L2) APH
1%	CenterPoint	33.50 31.89 30.92	32.82 31.23 30.28	28.81 27.42 26.57	28.22 26.85 26.01	38.16 37.49 38.12	28.80 28.38 26.67	32.52 31.95 32.40	24.52 24.17 22.66	22.45 24.16 20.88	21.32 22.95 19.46	21.59 23.23 20.08	20.50 22.07 18.72	30.84±0.76	26.88±1.22	27.17±0.72	23.75±1.11
	Ours	44.48 43.01 42.81	43.75 42.28 42.07	38.47 37.17 37.00	37.84 36.54 36.36	46.31 42.64 43.82	34.83 31.57 31.73	39.76 36.61 37.71	29.88 27.06 27.28	25.80 31.32 29.33	24.10 29.30 27.34	24.82 30.12 28.21	23.18 28.18 26.29	**38.84±0.17**	**34.11±0.35**	**34.43±0.18**	**30.29±0.31**
	PV-RCNN	48.47 47.72 49.47	38.75 40.54 41.60	41.89 41.20 42.78	33.50 35.01 36.05	29.39 24.85 31.03	14.87 12.58 15.50	24.53 20.70 25.93	12.41 10.48 12.96	17.73 11.65 17.17	9.60 6.43 10.02	17.05 11.21 16.51	9.23 6.19 9.63	30.83±2.41	21.1±1.26	26.87±2.18	18.38±1.16
	Ours	54.84 55.68 57.78	38.40 45.46 52.61	47.58 48.31 50.21	33.30 39.41 45.67	39.50 40.35 44.37	18.93 20.51 21.28	33.25 33.98 37.50	15.93 17.27 17.99	31.77 31.74 35.04	18.34 20.34 20.68	30.55 30.52 33.69	17.63 19.56 19.89	**43.45±1.99**	**28.50±3.16**	**38.40±1.81**	**25.18±2.79**
	SECOND	40.39 38.08 39.73	37.19 34.02 36.28	34.80 32.78 34.21	32.04 29.28 31.22	29.67 27.77 29.19	14.49 14.37 15.11	25.00 23.35 24.59	12.21 12.08 12.73	2.52 3.06 2.76	1.18 1.44 1.46	2.42 2.94 2.66	1.14 1.39 1.40	23.69 ±0.64	17.28 ±0.58	20.31 ±0.55	14.83 ±0.50
	Ours	46.42 44.57 46.59	42.82 40.44 43.52	40.19 38.51 40.31	37.07 34.93 37.65	36.23 33.27 37.58	18.13 17.30 20.03	30.70 28.15 31.85	15.36 14.63 16.97	22.38 13.67 27.13	12.91 7.01 17.68	21.52 13.14 26.09	12.41 6.74 17.01	**34.21 ±3.37**	**24.43 ±2.75**	**30.05 ±3.14**	**21.42 ±2.56**
5%	CenterPoint	56.65 57.59 57.34	55.96 56.85 56.62	49.36 50.19 49.95	48.75 49.55 49.32	55.43 55.86 54.32	47.81 48.16 47.06	48.2 48.59 47.23	41.51 41.83 40.86	53.57 54.26 52.48	52.18 52.89 51.18	51.52 52.18 50.48	50.18 50.87 49.22	55.28±0.60	52.08±0.52	49.74±0.55	46.90±0.48
	Ours	59.18 59.28 59.51	58.48 58.55 58.81	51.69 51.73 51.96	51.08 51.09 51.34	55.24 54.96 55.31	47.54 47.23 47.61	48.17 47.93 48.26	41.39 41.12 41.48	57.39 55.89 56.04	55.89 54.50 54.69	55.21 53.77 53.93	53.76 52.44 52.62	**56.98±0.28**	**53.7±0.27**	**51.40±0.27**	**48.48±0.26**
	PV-RCNN	67.28 67.31 67.56	65.97 66.04 66.31	58.88 58.92 59.16	57.71 57.79 58.05	56.89 56.8 56.98	28.89 28.26 28.24	48.62 48.55 48.70	24.69 24.16 24.14	50.89 49.92 49.80	29.55 25.94 27.61	48.94 48.01 47.90	28.42 24.95 26.55	58.16±0.18	40.76±0.70	51.97±0.16	36.27±0.65
	Ours	68.72 68.70 68.50	67.69 67.73 67.43	60.28 60.28 60.05	59.35 59.41 59.09	58.49 59.09 58.86	29.88 29.92 30.04	50.14 50.67 50.50	25.61 25.65 25.78	55.66 54.95 55.75	32.62 29.36 30.58	53.54 52.85 53.62	31.38 28.23 29.42	**60.97±0.06**	**42.81±0.54**	**54.66±0.06**	**38.21±0.52**
	SECOND	56.30 55.33 55.53	55.22 54.23 54.35	48.93 48.10 48.27	47.99 47.15 47.23	46.67 47.03 46.75	24.82 23.71 23.33	39.84 40.12 39.91	21.17 20.23 19.91	34.64 34.17 34.38	17.60 20.10 18.58	33.32 32.86 33.07	16.93 19.33 17.87	45.64±0.20	32.44±0.31	40.49±0.18	28.65±0.29
	Ours	59.44 58.93 58.23	58.45 57.95 57.13	51.84 51.39 50.79	50.97 50.53 49.82	50.22 49.67 49.72	24.72 25.07 23.91	43.10 42.62 42.65	21.22 21.51 20.52	42.83 43.40 41.31	20.28 23.45 22.35	41.20 41.74 39.73	19.50 22.56 21.50	**50.42±0.58**	**34.81±0.59**	**45.00±0.54**	**30.90±0.55**
10%	CenterPoint	62.16 62.41 62.31	61.49 61.75 61.63	54.44 54.62 54.55	53.85 54.04 53.96	62.08 61.96 61.07	55.14 54.97 54.06	54.45 54.42 53.48	48.27 48.18 47.25	61.36 61.04 60.21	59.84 59.60 58.71	59.03 58.74 57.92	57.57 57.35 56.49	61.62±0.37	58.58±0.38	55.74±0.37	53.00±0.37
	Ours	63.04 63.28 63.57	62.36 62.62 62.91	55.24 55.48 55.72	54.65 54.89 55.14	62.07 61.95 61.66	54.78 54.58 54.26	54.40 54.34 54.12	47.92 47.78 47.53	61.49 61.03 61.24	60.09 59.62 59.82	59.17 58.73 58.94	57.82 57.37 57.57	**62.15±0.06**	**59.01±0.07**	**56.24±0.05**	**53.41±0.06**
	PV-RCNN	70.36 70.05 70.55	69.45 69.11 69.65	61.79 61.52 61.99	60.98 60.66 61.18	61.92 60.74 61.46	30.42 30.36 30.68	53.26 52.14 52.82	26.17 26.06 26.37	56.95 55.77 57.74	31.44 30.62 28.91	54.77 53.63 55.54	30.24 29.45 27.81	62.84±0.57	43.4±0.35	56.38±0.55	38.77±0.34
	Ours	70.81 70.79 70.91	70.00 69.98 70.09	62.26 62.24 62.39	61.52 61.51 61.64	62.35 62.56 62.68	31.08 32.06 29.81	53.70 53.89 53.97	26.77 27.61 25.68	59.38 58.94 58.78	32.74 30.27 30.72	57.12 56.71 56.55	31.50 29.12 29.56	**64.13±0.04**	**44.08±0.53**	**57.65±0.04**	**39.43±0.48**
	SECOND	61.03 60.17 60.80	60.04 59.22 59.89	53.32 52.52 53.08	52.45 51.68 52.27	53.15 51.36 52.55	27.24 26.52 27.33	45.71 44.10 45.13	23.42 22.77 23.47	43.74 39.68 42.59	25.07 20.75 23.81	42.07 38.17 40.96	24.11 19.96 22.90	51.67±1.15	36.65±1.02	46.12±1.08	32.56±0.97
	Ours	62.67 62.90 62.12	61.78 62.08 61.24	54.86 55.03 54.37	54.06 54.31 53.59	54.74 54.33 53.95	27.67 27.57 26.52	47.23 46.85 46.50	23.88 23.76 22.87	47.56 46.70 46.27	26.57 26.48 25.01	45.76 44.93 44.51	25.56 25.47 24.06	**54.58±0.44**	**38.32±0.63**	**48.89±0.41**	**34.17±0.58**

Table 9. **Data efficiency on Waymo.** All of the experiments.

Labels	Method	*Vehicle* (L1) AP	*Vehicle* (L1) APH	*Vehicle* (L2) AP	*Vehicle* (L2) APH	*Ped.* (L1) AP	*Ped.* (L1) APH	*Ped.* (L2) AP	*Ped.* (L2) APH	*Cyc.* (L1) AP	*Cyc.* (L1) APH	*Cyc.* (L2) AP	*Cyc.* (L2) APH	Average and std (L1) AP	Average and std (L1) APH	Average and std (L2) AP	Average and std (L2) APH
1%	CenterPoint	9.81	9.24	8.38	7.89	9.59	4.88	7.97	4.06	0.68	0.41	0.66	0.39				
		8.60	8.08	7.34	6.90	9.75	4.93	8.10	4.10	0.67	0.35	0.64	0.34	6.27±0.47	4.47±0.36	5.30±0.39	3.80±0.30
		8.11	7.68	6.93	6.56	8.43	4.28	7.00	3.55	0.76	0.42	0.73	0.40				
	PropsalContrast	9.43	8.98	8.03	7.66	7.78	3.99	6.48	3.32	0.63	0.34	0.61	0.32				
		7.68	7.31	6.58	6.26	6.28	3.13	5.22	2.60	0.70	0.32	0.68	0.30	5.76±0.79	4.24±0.58	4.88±0.66	3.60±0.49
		9.64	9.19	8.22	7.84	9.11	4.62	7.54	3.83	0.57	0.29	0.55	0.28				
	Ours	35.66	34.61	30.69	29.78	27.48	14.40	23.24	12.18	6.21	3.81	5.97	3.66				
		32.94	31.97	28.30	27.47	25.89	13.50	21.87	11.40	7.99	3.89	7.68	3.74	**22.84±0.49**	**17.16±0.62**	**19.76±0.41**	**14.83±0.54**
		34.31	33.25	29.51	28.59	27.32	14.42	23.08	12.18	7.76	4.61	7.47	4.43				
5%	CenterPoint	29.84	28.75	25.63	24.69	17.25	8.94	14.46	7.49	4.86	3.43	4.68	3.30				
		28.27	27.05	24.27	23.23	17.46	9.15	14.63	7.66	3.91	2.75	3.76	2.65	17.14±0.53	13.46±0.41	14.75±0.46	11.60±0.36
		29.84	28.52	25.62	24.49	18.25	9.52	15.28	7.97	4.56	3.00	4.39	2.88				
	PropsalContrast	34.07	32.92	29.35	28.36	20.78	11.05	17.58	9.34	8.03	6.35	7.73	6.11				
		31.34	30.21	26.95	25.98	19.08	10.05	16.05	8.45	5.78	4.42	5.56	4.25	19.87±1.11	15.83±0.94	17.22±1.02	13.75±0.86
		32.13	30.98	27.64	26.65	19.88	10.51	16.71	8.83	7.70	5.96	7.41	5.74				
	Ours	51.00	49.98	44.23	43.33	38.82	22.62	33.18	19.33	30.75	27.20	29.57	26.15				
		51.16	50.11	44.35	43.43	38.94	22.34	33.27	19.09	28.21	24.47	27.13	23.54	**39.75±0.39**	**32.80±0.48**	**35.22±0.39**	**29.15±0.46**
		51.34	50.39	44.48	43.65	38.25	21.95	32.65	18.73	29.23	26.13	28.11	25.13				
10%	CenterPoint	35.82	34.78	30.84	29.94	21.05	11.39	17.73	9.59	7.24	5.29	6.96	5.08				
		35.12	33.99	30.24	29.27	21.28	11.38	17.92	9.59	8.26	6.44	7.95	6.19	21.55±0.18	17.19±0.07	18.69±0.17	14.92±0.08
		35.76	34.61	30.80	29.81	21.86	11.57	18.43	9.75	7.60	5.30	7.31	5.10				
	PropsalContrast	42.70	41.70	36.96	36.08	28.85	16.14	24.49	13.70	16.22	13.45	15.60	12.94				
		42.47	41.35	36.74	35.77	29.42	15.99	24.97	13.58	15.43	12.38	14.83	11.90	29.22±0.11	23.53±0.27	25.64±0.11	20.69±0.25
		42.13	41.02	36.45	35.48	29.08	16.02	24.65	13.58	16.74	13.70	16.09	13.17				
	Ours	55.07	54.17	47.88	47.10	44.28	26.92	38.04	23.12	36.85	33.34	35.44	32.06				
		55.32	54.45	48.07	47.31	43.88	26.99	37.70	23.17	35.95	32.69	34.57	31.44	**45.38±0.32**	**38.12±0.06**	**40.41±0.28**	**34.05±0.07**
		55.43	54.52	48.20	47.40	45.24	27.19	38.83	23.33	36.36	32.77	34.97	31.51				
20%	CenterPoint	39.74	38.70	34.31	33.40	23.93	13.16	20.22	11.11	10.97	8.63	10.55	8.30				
		38.17	37.09	32.93	32.00	24.98	13.79	21.10	11.64	9.91	7.90	9.53	7.59	25.00±0.72	20.13±0.52	21.80±0.67	17.58±0.49
		39.94	38.90	34.50	33.60	24.81	13.53	20.97	11.43	12.56	9.49	12.08	9.13				
	PropsalContrast	49.36	48.40	42.78	41.93	35.01	20.64	29.97	17.66	21.52	17.62	20.69	16.94				
		48.96	47.92	42.40	41.50	34.87	20.14	29.86	17.24	22.06	18.24	21.22	17.54	35.40±0.19	28.93±0.19	31.26±0.19	25.57±0.18
		49.77	48.78	43.13	42.27	34.27	19.89	29.35	17.03	22.81	18.75	21.94	18.03				
	Ours	57.27	56.38	49.89	49.11	47.75	30.33	41.08	26.08	37.65	34.26	36.20	32.94				
		57.84	57.09	50.42	49.76	47.74	30.35	41.17	26.15	37.07	33.94	35.64	32.64	**47.61±0.10**	**40.54±0.26**	**42.46±0.10**	**36.25±0.25**
		57.46	56.71	50.09	49.43	47.76	30.43	41.12	26.19	37.97	35.34	36.52	33.99				
50%	CenterPoint	47.06	45.93	40.83	39.84	29.66	16.74	25.19	14.21	16.16	12.77	15.54	12.28				
		45.84	44.69	39.75	38.74	30.35	16.92	25.75	14.36	17.63	14.76	16.95	14.19	31.27±0.30	25.44±0.29	27.47±0.28	22.40±0.27
		47.06	45.89	40.84	39.81	30.47	17.18	25.88	14.58	17.17	14.10	16.51	13.56				
	PropsalContrast	55.71	54.75	48.52	47.68	42.60	26.08	36.73	22.46	30.31	27.39	29.16	26.34				
		55.64	54.71	48.45	47.63	42.55	26.22	36.68	22.58	29.89	26.78	28.75	25.75	42.77±0.09	36.01±0.09	38.04±0.09	32.10±0.09
		55.40	54.46	48.24	47.41	42.06	26.15	36.28	22.54	30.76	27.52	29.59	26.46				
	Ours	62.30	61.61	54.59	53.97	53.01	35.48	46.03	30.77	42.71	39.85	41.08	38.33				
		62.94	62.23	55.19	54.56	53.98	36.28	46.90	31.48	43.44	40.56	41.77	39.00	**53.09±0.39**	**45.96±0.36**	**47.62±0.36**	**41.31±0.33**
		62.83	62.14	55.07	54.45	53.54	35.78	46.53	31.05	43.03	39.72	41.38	38.20				
100%	CenterPoint	49.85	48.73	43.34	42.35	32.18	18.23	27.38	15.50	20.33	17.04	19.55	16.38				
		49.07	47.92	42.65	41.64	32.85	18.79	27.94	15.98	20.15	17.32	19.38	16.65	34.16±0.16	28.00±0.01	30.12±0.14	24.75±0.01
		49.39	48.15	42.94	41.85	33.02	18.57	28.09	15.79	20.57	17.23	19.78	16.57				
	PropsalContrast	57.71	56.79	50.33	49.51	45.03	28.16	38.79	24.23	33.25	30.07	31.97	28.91				
		57.04	56.09	49.74	48.90	44.70	27.17	38.55	23.41	32.25	29.37	31.01	28.23	44.96±0.34	37.86±0.42	40.03±0.31	33.79±0.38
		57.35	56.42	50.01	49.19	44.98	27.88	38.78	24.01	32.30	28.78	31.06	27.67				
	Ours	65.36	64.67	57.50	56.88	57.13	38.31	49.82	33.36	44.98	41.84	43.26	40.24				
		64.97	64.25	57.16	56.52	57.44	39.28	50.11	34.21	45.91	42.75	44.15	41.11	**55.91±0.17**	**48.47±0.25**	**50.29±0.16**	**43.68±0.23**
		65.08	64.38	57.24	56.62	56.62	37.98	49.36	33.06	45.74	42.79	43.99	41.15				

Table 10. **Data efficiency on Waymo with frozen features.** All of the experiments.

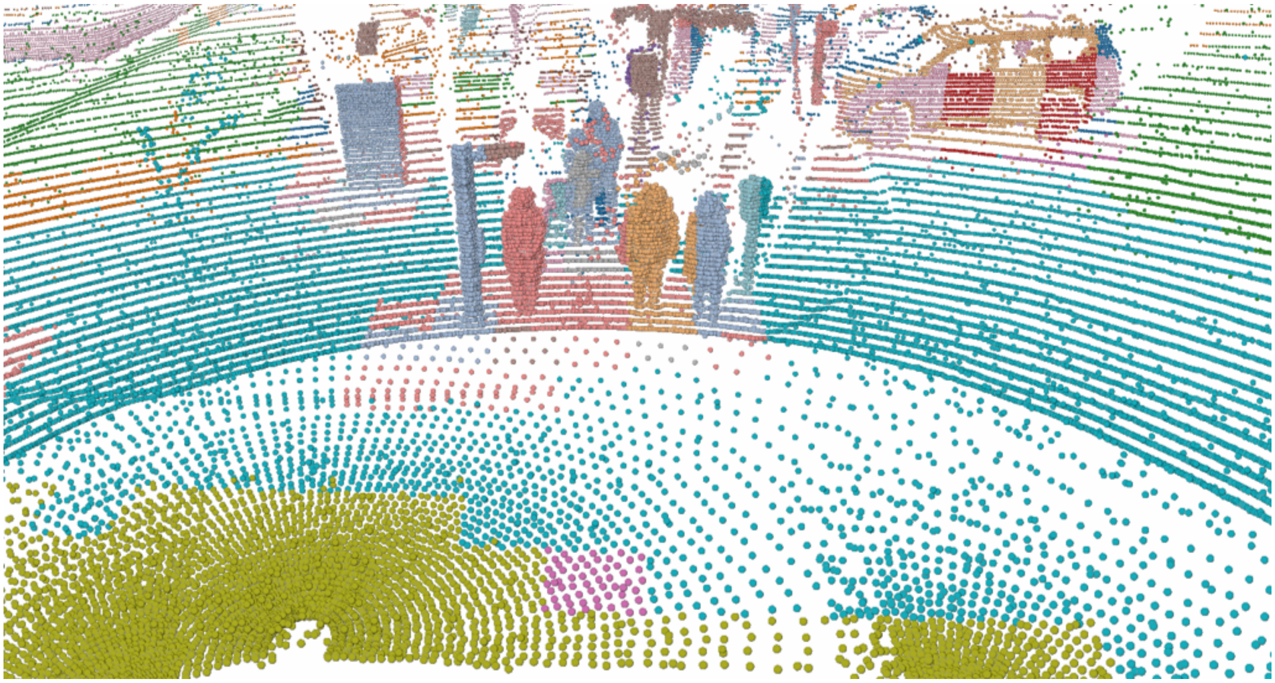

Figure 9. **Qualitative result.** Example 1.

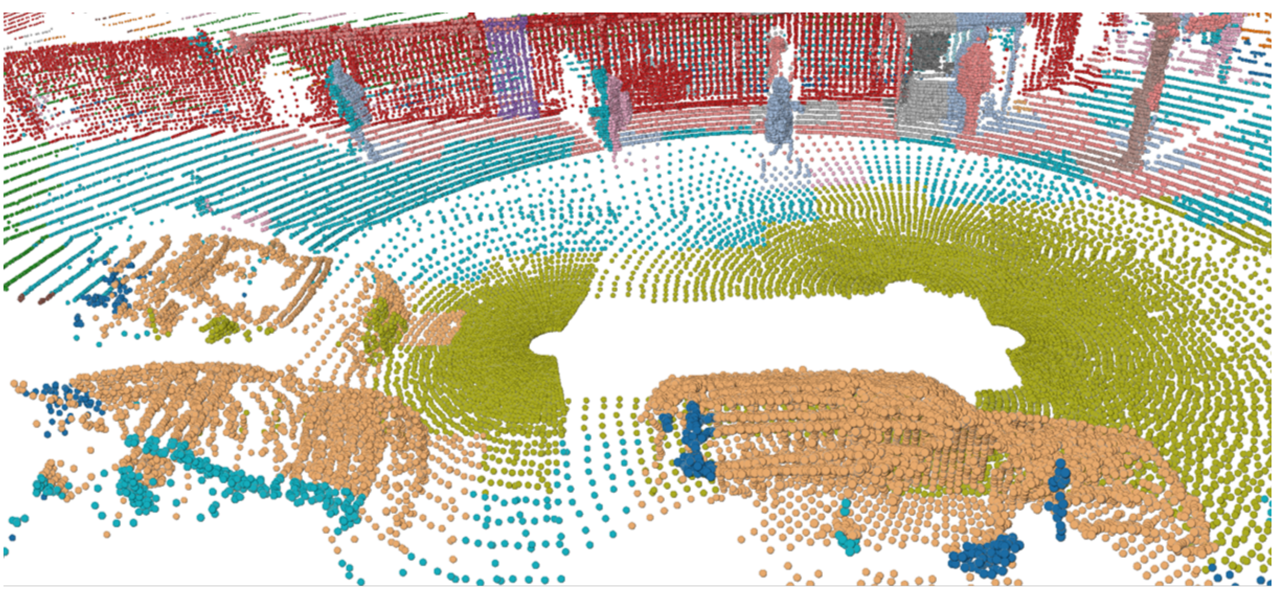

Figure 10. **Qualitative result.** Example 2.

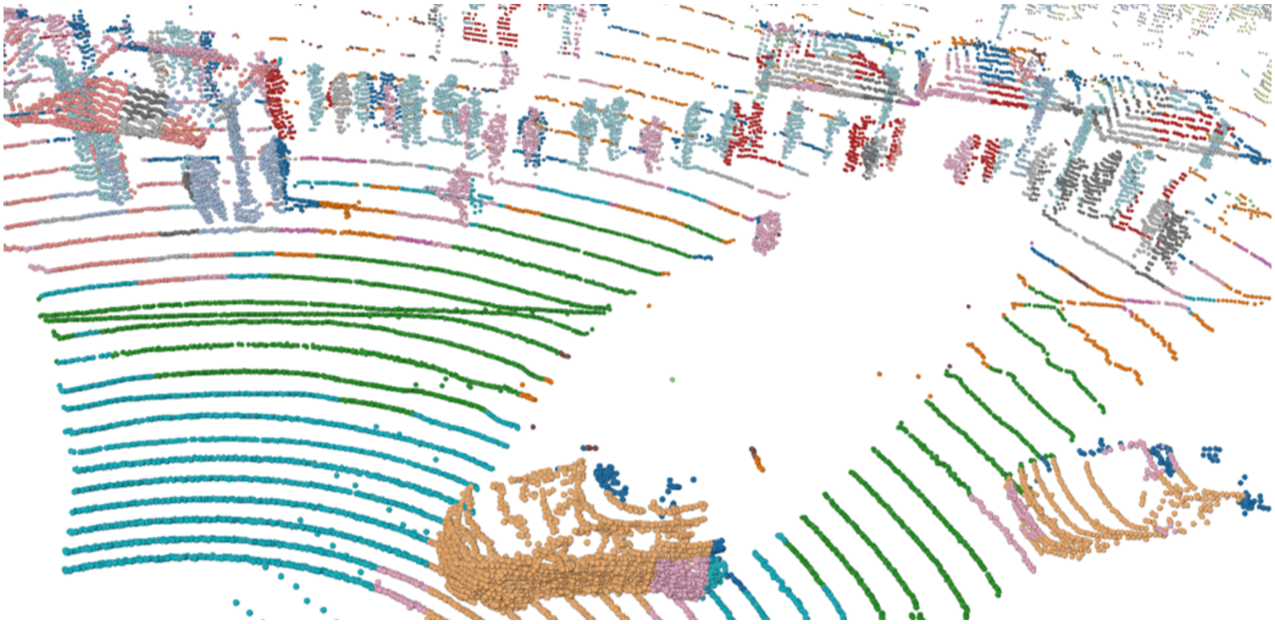

Figure 11. **Qualitative result.** Example 3.

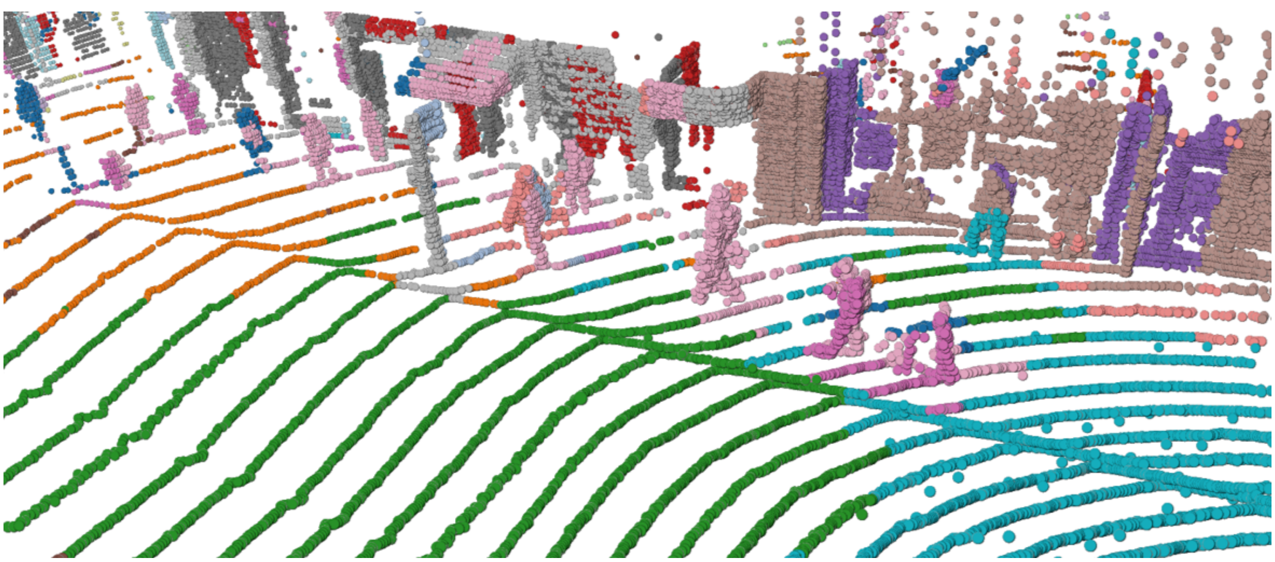

Figure 12. **Qualitative result.** Example 4.

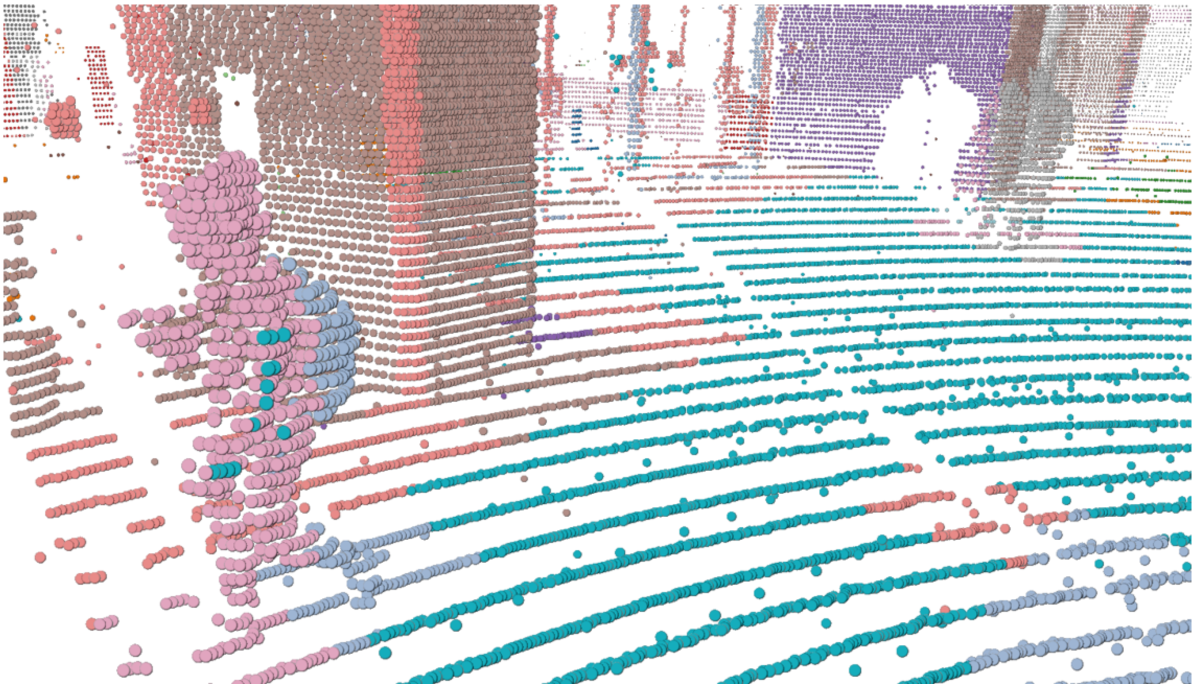

Figure 13. **Qualitative result.** Example 5.

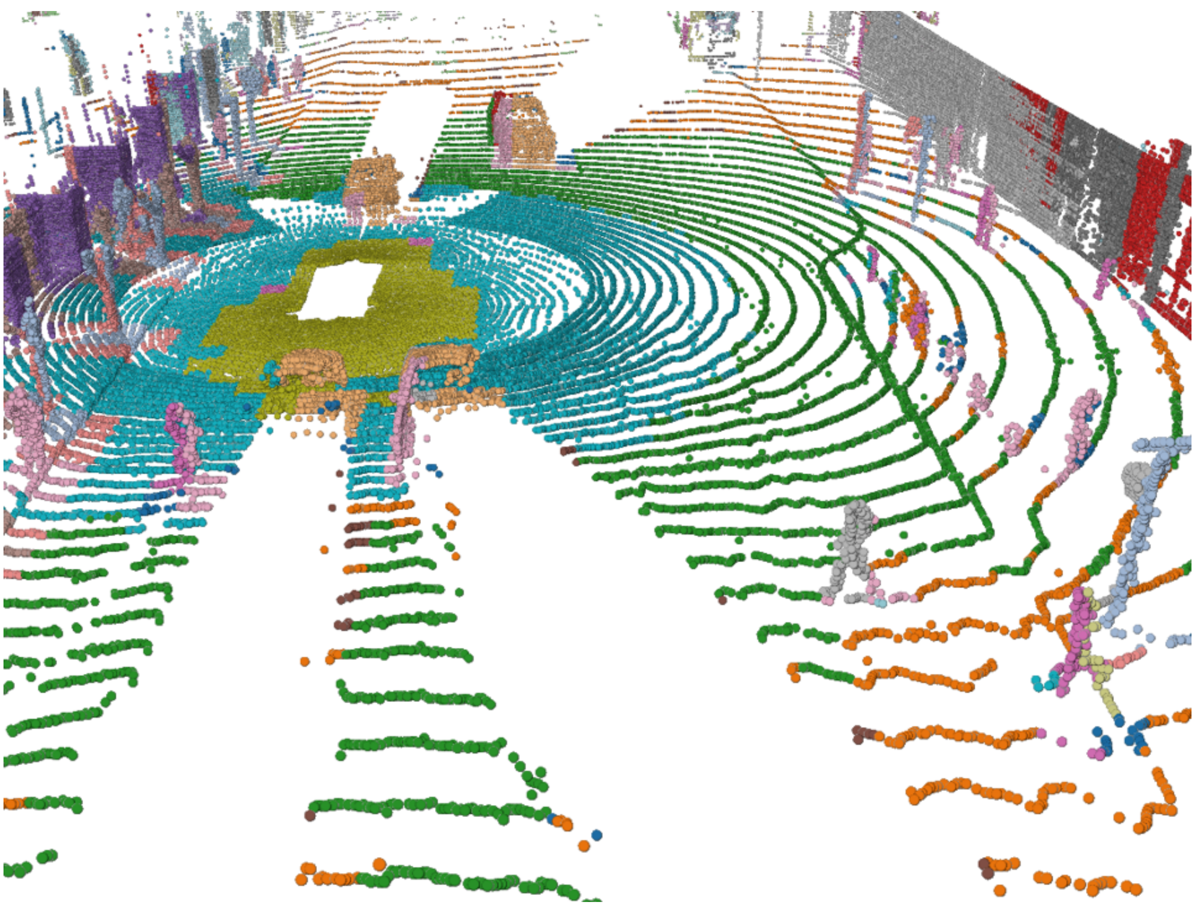

Figure 14. **Qualitative result.** Example 6.

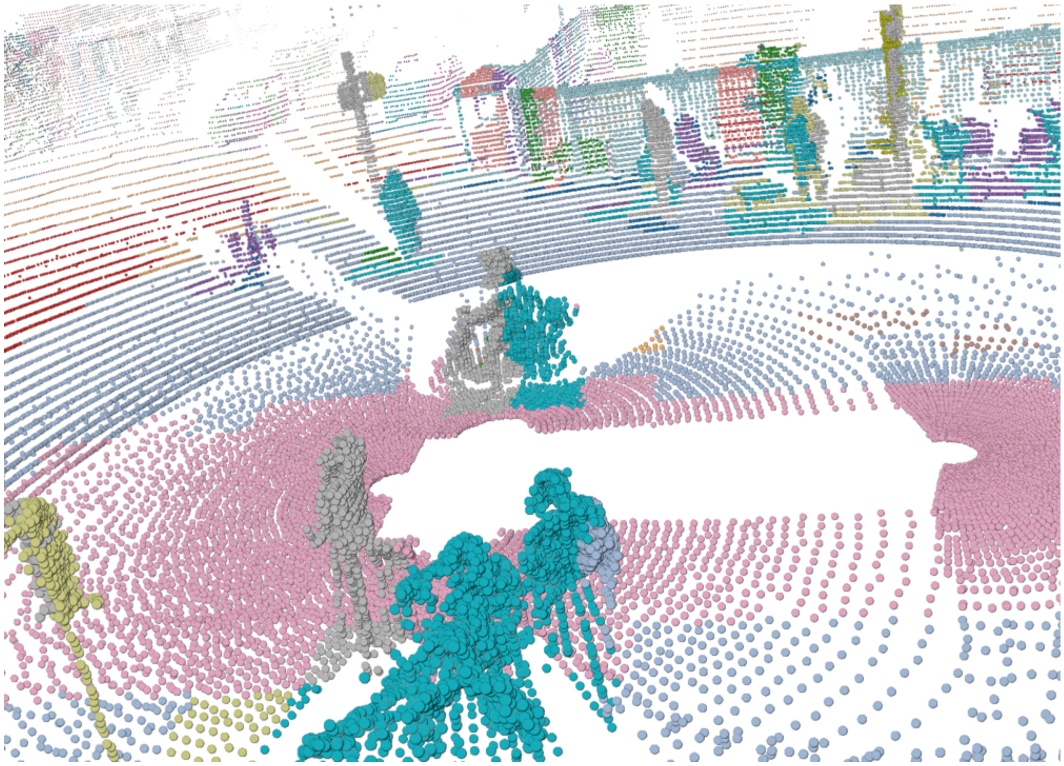

Figure 15. **Qualitative result.** Example 9.

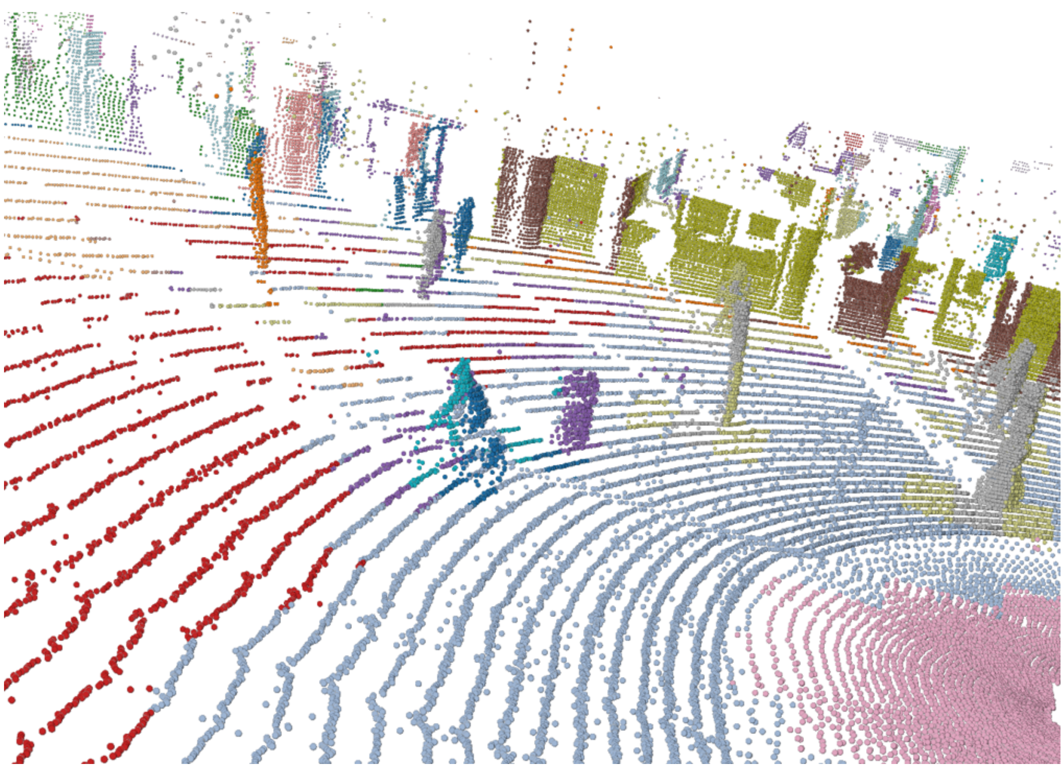

Figure 16. **Qualitative result.** Example 10.