Supplementary Materials for SFPNet

Yanbo Wang[1,2], Wentao Zhao[1,2], Chuan Cao[1,2], Tianchen Deng[1,2], Jingchuan Wang[1,2], and Weidong Chen[1,2,†]

[1] Institute of Medical Robotics and Department of Automation, Shanghai Jiao Tong University, China
[2] Key Laboratory of System Control and Information Processing, Ministry of Education, Shanghai 200240, China
{yanbowang319,wentaozhao,alex008,dengtianchen,jchwang,wdchen}@sjtu.edu.cn

1 Introduction

In this supplementary materials, we provide our **dataset details** about sensors, scenes, annotation process and label distributions in Sec. 2. Additional method details are demonstrated in Sec. 3. More experiment results and network analysis are given in Sec. 4. Limitations and future works are discussed in Sec. 5.

2 Dataset: SeMantic InDustry

2.1 Scenes

Many applications rely on the crucial aspect of comprehending semantic scenes. However, most existing benchmarks [3,4,17,24] focus on driving scenes. To fill the gap in public dataset of industrial outdoor scenes for **robotic application**, we collect a total of 38904 frames of hybrid-solid LiDAR data in different substations and have annotated 25 categories as shown in Fig. 2. Overall comparison with previous benchmarks is shown in Tab. 1.

2.2 Sensors

Fig. 1 shows the sensors equipped on our industrial robot used to collect S.MID. To the best of our knowledge, S.MID is the first large-scale outdoor **hybrid-solid LiDAR semantic segmentation dataset**. In addition to the features shown in the figures, Livox Mid-360 is much more cost-effective compared to traditional mechanical spinning LiDAR.

In accordance with the illustration provided in Fig. 1 and Fig. 1 (b) in the main text, Livox Mid-360 is suitable for industrial robots involving scene understanding tasks since it covers a broader range of scenes with **non-repetitive scanning mode**. However, it is a double-edged sword. This mode will also make the point cloud relatively **sparse** and **randomly distributed**. Therefore, the single-frame hybrid-solid LiDAR segmentation task **brings more challenges to network design**.

[†]Corresponding Author.

Table 1: Semantic LiDAR dataset comparison. Frames† for train/val/test. Number of classes ‡ for single frame evaluation and annotated total number in brackets.

Datasets	Frames†	LiDAR	Types of LiDAR	Classes‡	Applications
nuScenes	28130/6019/6008	Velodyne-HDL-32E	Mechanical Spinning LiDAR	16 (32)	Autonomous Vehicle
SemanticKITTI	19130/4071/20351	Velodyne-HDL-64E	Mechanical Spinning LiDAR	19 (34)	Autonomous Vehicle
S.MID	13101/5000/20803	Livox Mid-360	**Hybrid-Solid LiDAR**	14 (25)	**Industrial Robot**

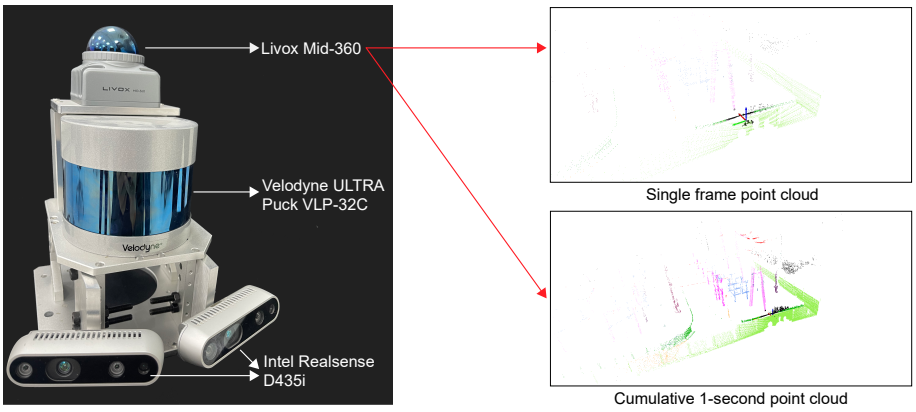


Fig. 1: Sensors and comparison between single frame and cumulative 1-second point clouds for Livox Mid-360. Although the single-frame point cloud is relatively sparse, the cumulative point cloud can better express the scene in the vertical direction. Please also note that only data collected by Livox Mid-360 and the corresponding labels are used in this research and have been released with S.MID.

2.3 Annotation Process

Considering the safety inspection tasks of robots and the common objects found in substations, we have annotated a total of 25 categories under professional guidance. Acknowledging the tools and annotation strategies provided by previous researchers [3], we first develop a high-precision LiDAR-inertial SLAM system based on hybrid-state LiDAR for initial mapping. Subsequently, through manual correction, high-precision maps for annotation purposes are obtained as shown in Fig. 2.

Due to the presence of specialized equipment within the substations, there is a requirement for the annotators' expertise compared to that of annotators for autonomous driving datasets. Following training conducted by professionals, our dataset's labels have been carefully annotated.

2.4 Label Distributions

For single-frame segmentation task, we merge the annotated labels into 14 classes (*knife switch, main transformer, arrester, voltage transformer, busbar, switch, current transformer, scaffold, support column, road, other-ground, fence, fire*

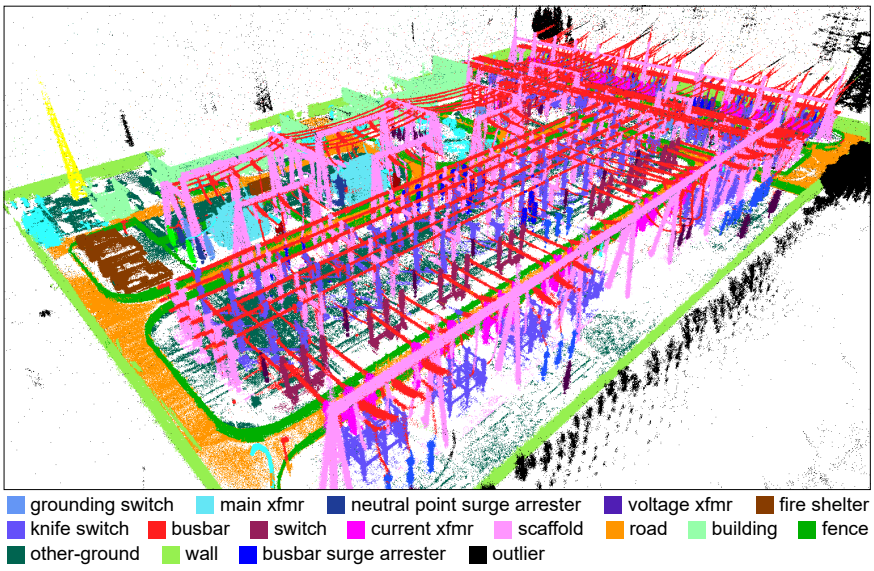


Fig. 2: Example of maps built in the annotation process.

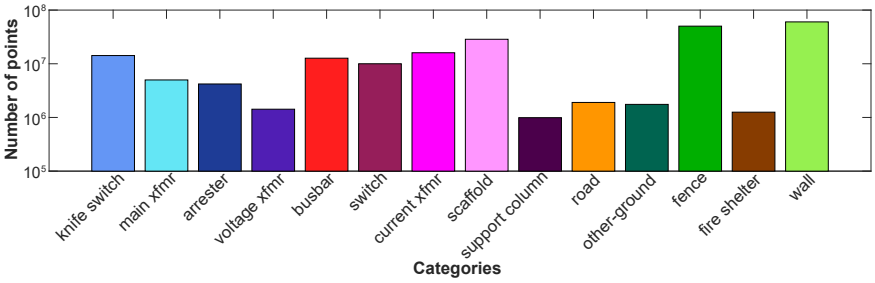


Fig. 3: Label distributions.

shelter, wall). The label distributions are shown in Fig. 3. The imbalanced count of classes is common in substation scenes. Hence, similar to imbalanced class distributions observed in autonomous driving datasets, addressing the issue of imbalanced class distribution in S.MID is an integral aspect that methods must contend with.

3 Additional Method Details

3.1 Overall Framework

Following the previous work [11, 30], we adopt a U-Net [16] structure as shown in Fig. 4. We firstly apply regular voxelization to form a sparse tensor $X \in \mathbb{R}^{N \times C_{in}}$. Our sparse focal point module is introduced in down stages and central

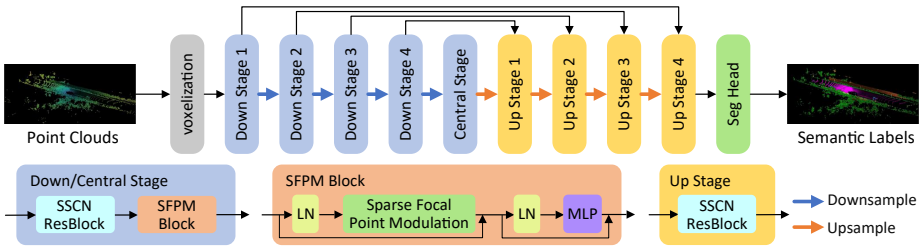


Fig. 4: Overall Framework. Our network employs an encoder-decoder structure with four down/up stages and one central stage. Similar to the transformer [21], our sparse focal point block consists of core modulator SFPM, layer normalization, and MLP as feed-forward network.

stage. After traversing through the backbone with skip connections, we employ a simple projection head to get the segmentation result. Due to the long-tailed data distribution in the prevalent LiDAR semantic segmentation datasets, we adopt focal loss [12] to address the issue of class imbalance.

3.2 Properties Discussion

Proof of *translation invariance* can be found in Sec. 3.1 in the main text. Here, we provide an extension analysis of *explicit locality with contextual learning*. The realization of our aggregation step $\kappa_{focal}(\cdot)$ is achieved through linear projection and Eqs. (4) – (6). The set of increasing kernels of SubMconv layers in Eq. (4) provides explicit locality and the operations before and after it will preserve this property (element-wise multiplication or channel-wise calculation). By using the gate mechanism described in Eq. (5), the input-dependent multi-level context from Eq. (4) can be adaptively aggregated. Additionally, Eq. (5) provides a "soft shape" in the sparse space through corresponding gate weight for each position i. Heuristic thinking: When dealing with diverse point cloud distributions, varying densities in each scan, and distinct classes, a qualified feature encoder exhibits varying dependencies across different contextual levels and positions within sparse space.

4 Additional experiments

More segmentation results on SemanticKITTI val and test sets are displayed in Tabs. 3 and 4 and nuScenes test set in Tab. 5 and S.MID test set in Tab. 2 Additional ablation study on S.MID in Tab. 6. More visual comparisons between SphereFormer [11] and ours on S.MID val set are shown in Fig. 5. More network analysis results are shown in Fig. 6.

Since most of the previous training techniques and augmentation methods such as Cutmix [11,26], Lasermix [10], Polarmix [23] and post-processing [9] are designed for mechanical spinning LiDAR, in order to ensure the consistency of

Table 2: Results of our proposed method and SOTA LiDAR Segmentation methods on S.MID test set. Note that all results are obtained from open source code with carefully chosen parameters.

Methods	mIoU	knife switch	main xfmr	arrester	voltage xfmr	busbar	switch	current xfmr	scaffold	sup column	road	other-ground	fence	fire shelter	wall
Cylinder3D [30]	68.1	82.9	69.8	74.8	44.1	79.1	92.9	93.5	79.9	54.7	57.0	37.9	77.6	28.4	81.0
SphereFormer [11]	68.3	84.2	71.5	75.5	49.8	80.1	96.6	96.7	86.6	47.5	60.8	40.1	74.7	8.9	83.4
Ours	**70.9**	88.8	90.4	85.2	50.4	76.1	97.1	96.9	89.2	60.2	57.6	29.7	83.1	1.2	87.3

Table 3: Results of our proposed method and state-of-the-art LiDAR Segmentation methods on SemanticKITTI val set. Note that all results are obtained from the literature.

Methods	mIoU	car	bicycle	motor.	truck	other-veh.	person	bicyclist	m.cyclist	road	parking	sidewalk	other-gro.	building	fence	vegetation	trunk	terrain	pole	traffic s.
SSCN [7]	66.6	96.3	44.6	76.3	89.6	58.6	77.3	91.3	0.0	94.3	51.7	81.8	1.2	91.0	62.5	88.3	70.2	75.3	64.6	51.4
SphereFormer [11]	69.0	97.0	53.4	77.2	95.1	67.0	78.2	93.7	0.0	95.2	55.5	83.1	2.8	91.0	60.4	89.2	72.5	76.9	66.3	55.9
Ours	**69.2**	97.2	53.2	80.2	93.1	70.6	75.4	91.5	0.0	95.2	56.3	83.4	3.3	92.2	66.8	89.3	72.6	76.7	65.0	51.9

the three different types of LiDAR experiments, we did not use any training techniques. In this situation, SFPNet still shows competitive results on mechanical spinning LiDAR test sets.

In both S.MID val (in the main text) and test set (Tab. 2), we can see that when the distribution pattern of point clouds changes, the performance of cubic and radial window attention will deteriorate or even become worse than that of the improved SSCN. This shows that SFPM can better cope with different types of LiDAR with various point distributions due to its adaptive mechanism.

5 Limitations and Future works

Our work focuses on the representational capabilities of the network on general LiDAR point clouds. However, data augmentation, training techniques and post-processing are also important topics for segmentation tasks. For instance, 3.7% ˜ 4.9% mIoU improvement for SSCN-based networks can be achieved on mechanical spinning LiDAR through Polarmix [23] .

Future works can be done to explore augmentation methods for general LiDAR point clouds. We will also extend our methods to more LiDAR point cloud tasks such as object detection and panoptic segmentation, and on fused various types of LiDAR datasets. Efficiency improvement will also be considered in the future.

Table 4: Results of our proposed method and state-of-the-art LiDAR Segmentation methods on SemanticKITTI test set. Note that all results are obtained from the literature. LiDAR-based methods in the table are listed by year of publication.

Methods	mIoU	car	bicycle	motor.	truck	other-veh.	person	bicyclist	m.cyclist	road	parking	sidewalk	other-gro.	building	fence	vegetation	trunk	terrain	pole	traffic s.
PointNet++ [15]	20.1	53.7	1.9	0.2	0.9	0.2	0.9	1.0	0.0	72.0	18.7	41.8	5.6	62.3	16.9	46.5	13.8	30.0	6.0	8.9
TangentConv [19]	40.9	90.8	2.7	16.5	15.2	12.1	23.0	28.4	8.1	83.9	33.4	63.9	15.4	83.4	49.0	79.5	49.3	58.1	35.8	28.5
SqueezeSegV2 [22]	39.7	81.8	18.5	17.9	13.4	14.0	20.1	25.1	3.9	88.6	45.8	67.6	17.7	73.7	41.1	71.8	35.8	60.2	20.2	26.3
DarkNet53Seg [3]	49.9	86.4	24.5	32.7	25.5	22.6	36.2	33.6	4.7	91.8	64.8	74.6	27.9	84.1	55.0	78.3	50.1	64.0	38.9	52.2
RangeNet53++ [14]	52.2	91.4	25.7	34.4	25.7	23.0	38.3	38.8	4.8	91.8	65.0	75.2	27.8	87.4	58.6	80.5	55.1	64.6	47.9	55.9
KPConv [20]	58.8	95.0	30.2	42.5	33.4	44.3	61.5	61.6	11.8	90.3	61.3	72.7	31.5	90.5	64.2	84.8	69.2	69.1	56.4	47.4
3D-MiniNet [1]	55.8	90.5	42.3	42.1	28.5	29.4	47.8	44.1	14.5	91.6	64.2	74.5	25.4	89.4	60.8	82.8	60.8	66.7	48.0	56.6
SqueezeSegV3 [25]	55.9	92.5	38.7	36.5	29.6	33.0	45.6	46.2	20.1	91.7	63.4	74.8	26.4	89.0	59.4	82.0	58.7	65.4	49.6	58.9
PointASNL [28]	46.8	87.9	0.0	25.1	39.0	29.2	34.2	57.6	0.0	87.4	24.3	74.3	1.8	83.1	43.9	84.1	52.2	70.6	57.8	36.9
RandLA-Net [8]	55.9	94.2	29.8	32.2	43.9	39.1	48.4	47.4	9.4	90.5	61.8	74.0	24.5	89.7	60.4	83.8	63.6	68.6	51.0	50.7
PolarNet [29]	54.3	93.8	40.3	30.1	22.9	28.5	43.2	40.2	5.6	90.8	61.7	74.4	21.7	90.0	61.3	84.0	65.5	67.8	51.8	57.5
SPVNAS [18]	67.0	97.2	50.6	50.4	56.6	58.0	67.4	67.1	50.3	90.2	67.6	75.4	21.8	91.6	66.9	86.1	73.4	71.0	64.3	67.3
JS3C-Net [27]	66.0	95.8	59.3	52.9	54.3	46.0	69.5	65.4	39.9	88.9	61.9	72.1	31.9	92.5	70.8	84.5	69.8	67.9	60.7	68.7
Cylinder3D [30]	68.9	97.1	67.6	63.8	50.8	58.5	73.7	69.2	48.0	92.2	65.0	77.0	32.3	90.7	66.5	85.6	72.5	69.8	62.4	66.2
$(AF)^2$-S3Net [5]	70.8	94.3	63.0	81.4	40.2	40.0	76.4	81.7	77.7	92.0	66.8	76.2	45.8	92.5	69.6	78.6	68.0	63.1	64.0	73.3
RPVNet [26]	70.3	97.6	68.4	68.7	44.2	61.1	75.9	74.4	43.4	93.4	70.3	80.7	33.3	93.5	72.1	86.5	75.1	71.7	64.8	61.4
RangeViT-CS [2]	64.0	95.4	55.8	43.5	29.8	42.1	63.9	58.2	38.1	93.1	70.2	80.0	32.5	92.0	69.0	85.3	70.6	71.2	60.8	64.7
RangeFormer [9]	73.3	96.7	69.4	73.7	59.9	66.2	78.1	75.9	58.1	92.4	73.0	78.8	42.4	92.3	70.1	86.6	73.3	72.8	66.4	66.6
SphereFormer [11]	74.8	97.5	70.1	70.5	59.6	67.7	79.0	80.4	75.3	91.8	69.7	78.2	41.3	93.8	72.8	86.7	75.1	72.4	66.8	72.9
Ours	70.3	97.2	64.9	63.8	44.8	54.7	70.4	74.6	52.9	91.9	70.6	78.0	39.7	93.3	71.5	85.4	73.7	70.1	66.1	72.1

Table 5: Results of our proposed method and state-of-the-art LiDAR Segmentation methods on nuScenes test set. Note that all results are obtained from the literature. Methods in the table are listed by year of publication.

Methods	Input	mIoU	FW mIoU	barrier	bicycle	bus	car	construction	motor	pedestrian	traffic cone	trailer	truck	driveable	other flat	sidewalk	terrain	manmade	vegetation
PolarNet [29]	L	69.4	87.4	72.2	16.8	77.0	86.5	51.1	69.7	64.8	54.1	69.7	63.5	96.6	67.1	77.7	72.1	78.1	84.5
AMVNet [13]	L	77.3	90.1	80.6	32.0	81.7	88.9	67.1	84.3	76.1	73.5	84.9	67.3	97.5	67.4	79.4	75.5	91.5	88.7
SPVCNN [18]	L	77.4	89.7	80.0	30.0	91.9	90.8	64.7	79.0	75.6	70.9	81.0	74.6	97.4	69.2	80.0	76.1	89.3	87.1
JS3C-Net [27]	L	73.6	88.1	80.1	26.2	87.8	84.5	55.2	72.6	71.3	66.3	76.8	71.2	96.8	64.5	76.9	74.1	87.5	86.1
Cylinder3D [30]	L	77.2	89.9	82.8	29.8	84.3	89.4	63.0	79.3	77.2	73.4	84.6	69.1	97.7	70.2	80.3	75.5	90.4	87.6
$(AF)^2$-S3Net [5]	L	78.3	88.5	78.9	52.2	89.9	84.2	77.4	74.3	77.3	72.0	83.9	73.8	97.1	66.5	77.5	74.0	87.7	86.8
PMF [31]	L+C	77.0	89.0	82.0	40.0	81.0	88.0	64.0	79.0	80.0	76.0	81.0	67.0	97.0	68.0	78.0	74.0	90.0	88.0
2D3DNet [6]	L+C	80.0	90.1	83.0	59.4	88.0	85.1	63.7	84.4	82.0	76.0	84.8	71.9	96.9	67.4	79.8	76.0	92.1	89.2
RangeFormer [9]	L	80.1	90.0	85.6	47.4	91.2	90.9	70.7	84.7	77.1	74.1	83.2	72.6	97.5	70.7	79.2	75.4	91.3	88.9
SphereFormer [11]	L	81.9	91.7	83.3	39.2	94.7	92.5	77.5	84.2	84.4	79.1	88.4	78.3	97.9	69.0	81.5	77.2	93.4	90.2
Ours	L	80.2	90.8	83.7	42.5	89.1	91.5	74.1	83.5	79.1	74.7	87.3	73.3	97.7	78.1	80.3	76.2	92.3	89.3

Table 6: Additional ablation study on S.MID val set.

	Basic blocks	Focal level = 2	Focal level = 3	Global Avg Pooling	mIoU
Optimal design	✓		✓	✓	71.9
Ablation 1	✓	✓		✓	69.9 (-2.0)
Ablation 2	✓		✓		69.8 (-2.1)
Ablation 3	✓				67.6 (-4.3)

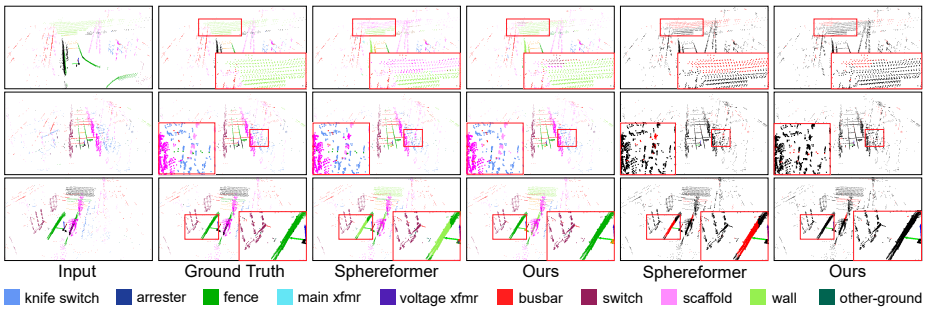
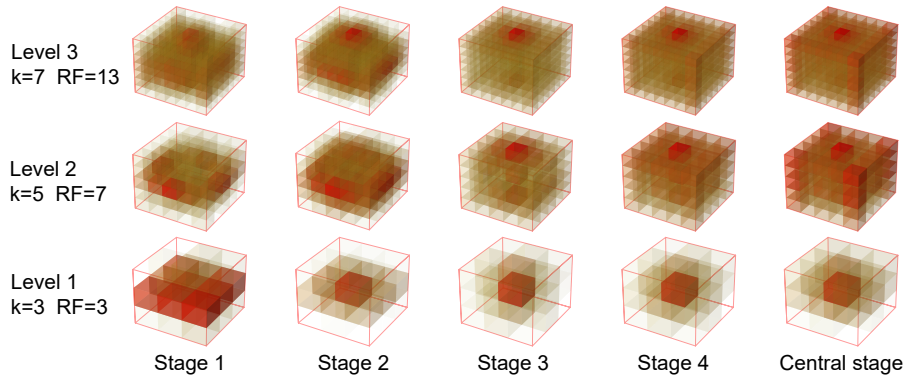
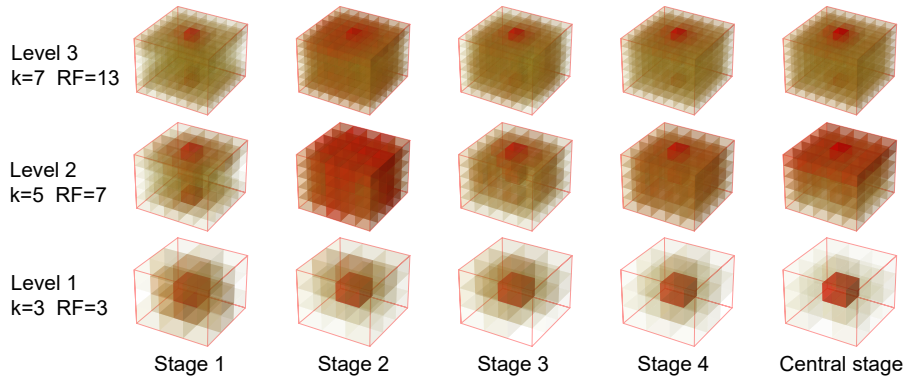


Fig. 5: Visual comparison between SphereFormer [11] and ours on S.MID val set. Details have been zoomed with red box. Difference maps are shown in the last two columns.



(a) SemanticKITTI.



(b) S.MID.

Fig. 6: Visualization of parameters of $SubMconv_{3d}^{l}$ at three focal levels in four down stages and central stage in SFPNet. SemanticKITTI shows similar patterns to nuScenes as demonstrate in the main text. S.MID shows a special pattern in the vertical direction due to the particularity of its point cloud.

References

1. Alonso, I., Riazuelo, L., Montesano, L., Murillo, A., Valada, A., Asfour, T.: 3d-mininet: Learning a 2d representation from point clouds for fast and efficient 3d lidar semantic segmentation. IEEE Robotics and Automation Letters **PP**, 1–1 (07 2020). https://doi.org/10.1109/LRA.2020.3007440
2. Ando, A., Gidaris, S., Bursuc, A., Puy, G., Boulch, A., Marlet, R.: Rangevit: Towards vision transformers for 3d semantic segmentation in autonomous driving. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5240–5250 (2023)
3. Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., Gall, J.: Semantickitti: A dataset for semantic scene understanding of lidar sequences. In: ICCV. pp. 9297–9307 (2019)
4. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: nuscenes: A multimodal dataset for autonomous driving. In: CVPR. pp. 11621–11631 (2020)
5. Cheng, R., Razani, R., Taghavi, E., Li, E., Liu, B.: 2-s3net: Attentive feature fusion with adaptive feature selection for sparse semantic segmentation network. In: CVPR. pp. 12547–12556 (2021)
6. Genova, K., Yin, X., Kundu, A., Pantofaru, C., Cole, F., Sud, A., Brewington, B., Shucker, B., Funkhouser, T.: Learning 3d semantic segmentation with only 2d image supervision. pp. 361–372 (12 2021). https://doi.org/10.1109/3DV53792.2021.00046
7. Graham, B., Engelcke, M., Van Der Maaten, L.: 3d semantic segmentation with submanifold sparse convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 9224–9232 (2018)
8. Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A.: Randla-net: Efficient semantic segmentation of large-scale point clouds. In: CVPR. pp. 11108–11117 (2020)
9. Kong, L., Liu, Y., Chen, R., Ma, Y., Zhu, X., Li, Y., Hou, Y., Qiao, Y., Liu, Z.: Rethinking range view representation for lidar segmentation. In: ICCV. pp. 228–240 (2023)
10. Kong, L., Ren, J., Pan, L., Liu, Z.: Lasermix for semi-supervised lidar semantic segmentation. In: CVPR. pp. 21705–21715 (2023)
11. Lai, X., Chen, Y., Lu, F., Liu, J., Jia, J.: Spherical transformer for lidar-based 3d recognition. In: CVPR. pp. 17545–17555 (2023)
12. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: ICCV. pp. 2980–2988 (2017)
13. Liong, V.E., Nguyen, T.N.T., Widjaja, S.A., Sharma, D., Chong, Z.J.: Amvnet: Assertion-based multi-view fusion network for lidar semantic segmentation. ArXiv **abs/2012.04934** (2020), https://api.semanticscholar.org/CorpusID:228063957
14. Milioto, A., Vizzo, I., Behley, J., Stachniss, C.: Rangenet++: Fast and accurate lidar semantic segmentation. In: 2019 IEEE/RSJ international conference on intelligent robots and systems (IROS). pp. 4213–4220. IEEE (2019)
15. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In: NeurIPS. vol. 30 (2017)
16. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)

17. Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., et al.: Scalability in perception for autonomous driving: Waymo open dataset. In: CVPR. pp. 2446–2454 (2020)
18. Tang, H., Liu, Z., Zhao, S., Lin, Y., Lin, J., Wang, H., Han, S.: Searching efficient 3d architectures with sparse point-voxel convolution. In: ECCV. pp. 685–702. Springer (2020)
19. Tatarchenko, M., Park, J., Koltun, V., Zhou, Q.Y.: Tangent convolutions for dense prediction in 3d. In: CVPR. pp. 3887–3896 (2018)
20. Thomas, H., Qi, C.R., Deschaud, J.E., Marcotegui, B., Goulette, F., Guibas, L.J.: Kpconv: Flexible and deformable convolution for point clouds. In: ICCV. pp. 6411–6420 (2019)
21. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. NeurIPS **30** (2017)
22. Wu, B., Zhou, X., Zhao, S., Yue, X., Keutzer, K.: Squeezesegv2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a lidar point cloud. In: 2019 international conference on robotics and automation (ICRA). pp. 4376–4382. IEEE (2019)
23. Xiao, A., Huang, J., Guan, D., Cui, K., Lu, S., Shao, L.: Polarmix: A general data augmentation technique for lidar point clouds. NeurIPS **35**, 11035–11048 (2022)
24. Xiao, P., Shao, Z., Hao, S., Zhang, Z., Chai, X., Jiao, J., Li, Z., Wu, J., Sun, K., Jiang, K., et al.: Pandaset: Advanced sensor suite dataset for autonomous driving. In: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC). pp. 3095–3101. IEEE (2021)
25. Xu, C., Wu, B., Wang, Z., Zhan, W., Vajda, P., Keutzer, K., Tomizuka, M.: Squeezesegv3: Spatially-adaptive convolution for efficient point-cloud segmentation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVIII 16. pp. 1–19. Springer (2020)
26. Xu, J., Zhang, R., Dou, J., Zhu, Y., Sun, J., Pu, S.: Rpvnet: A deep and efficient range-point-voxel fusion network for lidar point cloud segmentation. In: ICCV. pp. 16024–16033 (2021)
27. Yan, X., Gao, J., Li, J., Zhang, R., Li, Z., Huang, R., Cui, S.: Sparse single sweep lidar point cloud segmentation via learning contextual shape priors from scene completion. In: AAAI. vol. 35, pp. 3101–3109 (2021)
28. Yan, X., Zheng, C., Li, Z., Wang, S., Cui, S.: Pointasnl: Robust point clouds processing using nonlocal neural networks with adaptive sampling. In: CVPR. pp. 5589–5598 (2020)
29. Zhang, Y., Zhou, Z., David, P., Yue, X., Xi, Z., Gong, B., Foroosh, H.: Polarnet: An improved grid representation for online lidar point clouds semantic segmentation. In: CVPR. pp. 9601–9610 (2020)
30. Zhu, X., Zhou, H., Wang, T., Hong, F., Ma, Y., Li, W., Li, H., Lin, D.: Cylindrical and asymmetrical 3d convolution networks for lidar segmentation. In: CVPR. pp. 9939–9948 (2021)
31. Zhuang, Z., Li, R., Jia, K., Wang, Q., Li, Y., Tan, M.: Perception-aware multi-sensor fusion for 3d lidar semantic segmentation. pp. 16260–16270 (10 2021). https://doi.org/10.1109/ICCV48922.2021.01597