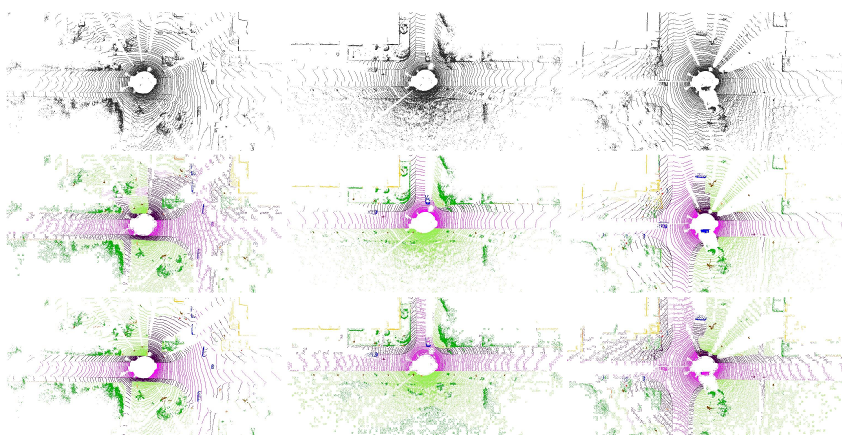
A Cylindrical Convolution Network for Dense Top-View Semantic Segmentation with LiDAR Point Clouds Supplementary Material

Jiacheng Lu[1], Shuo Gu[1] (✉), Cheng-Zhong Xu[2], and Hui Kong[2] (✉)

[1] PCA Lab, Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, and Jiangsu Key Lab of Image and Video Understanding for Social Security, School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
{lujiacheng,shuogu}@njust.edu.cn

[2] University of Macau, Macau, China {czxu,huikong}@um.edu.mo

This supplementary material provides additional visual results that cannot be included in the paper submission due to space limitations. In the first section, we show visual results for sparse predictions of our method on SemanticKITTI dataset [1]. In the second section, we display more visual results on SemanticKITTI dataset, including comparisons with the results of previous methods. Moreover, a demonstration video is included in the same zip file as the supplementary material.



(a) (b) (c)

Fig. 1. Qualitative results generated on the SemanticKITTI validation set. From top to bottom in each column, we display the input point cloud, the ground truth, the prediction from our method, respectively.

1 Visual Results for Sparse Predictions

In the main text, we have shown dense semantic prediction results. To fully demonstrate the effectiveness of our model, we present sparse semantic prediction results here.

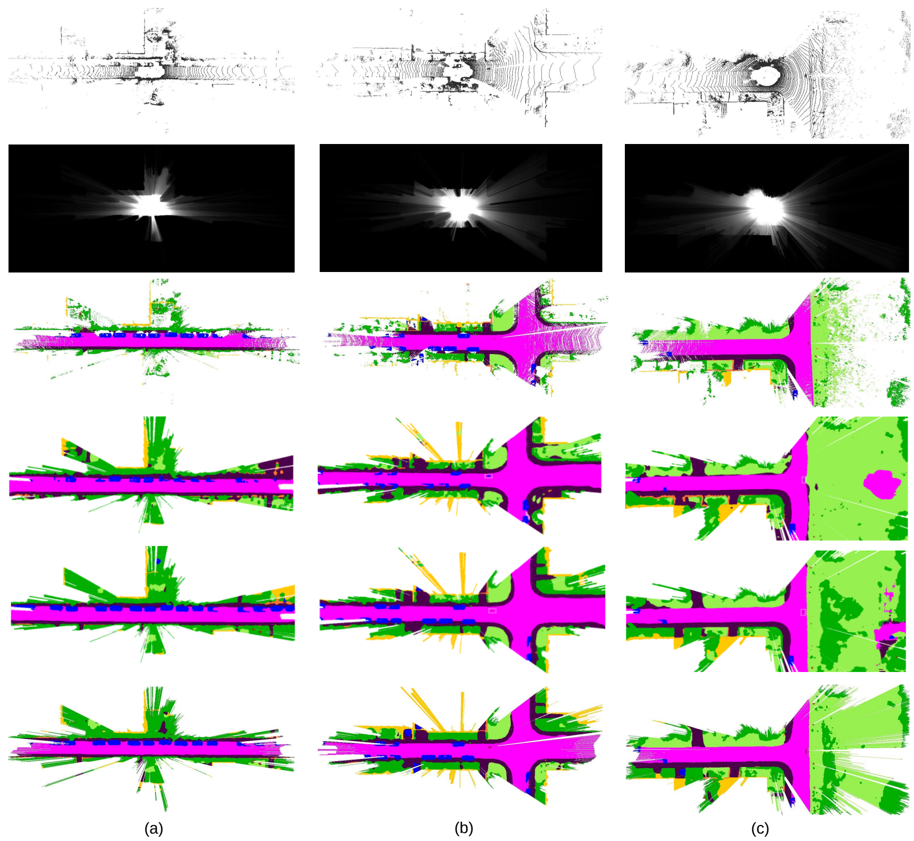


Fig. 2. Qualitative results generated on the SemanticKITTI validation set. From top to bottom in each column, we display the input point cloud, the 2D occupancy map, the ground truth, the prediction from Bieder *et al.* [2], PillarSeg [3] and our method, respectively. The unobserved areas were erased using the observability map as in [2]

2 Addition Visual Results on SemanticKITTI

Here we show two groups of comparisons with the results for Bieder *et al.* [2], PillarSeg [3], MASS [4] and our method on SemanticKITTI. For a fair compar-

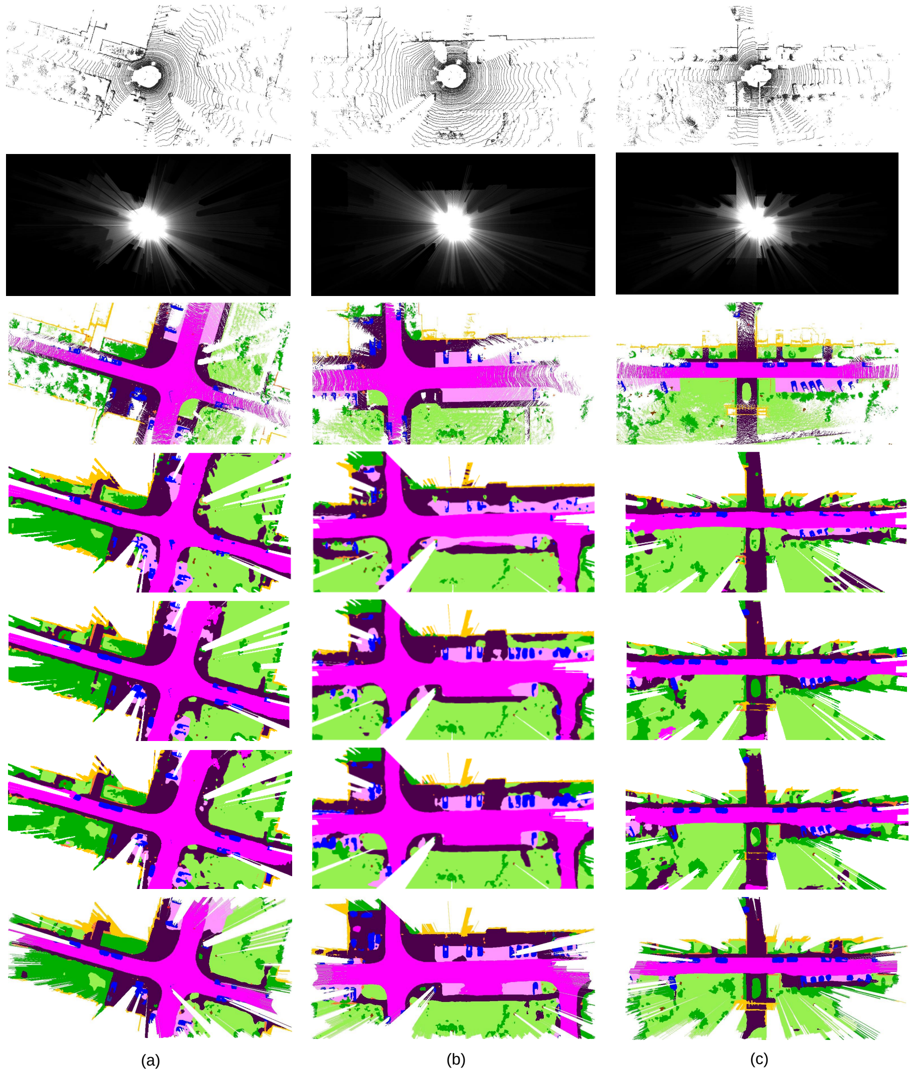


Fig. 3. Qualitative results generated on the SemanticKITTI [1] validation set. From top to bottom in each column, we display the input point cloud, the 2D occupancy map, the ground truth, the prediction from Bieder *et al.* [2], PillarSeg [3], MASS [4] and our method, respectively. The unobserved areas were erased using the observability map as in [2]

ison, the unobservable regions in our predictions are also filtered out using the observability map as in [2].

As shown in Fig 2 and Fig 3, our method is able to produce very similar results to the ground truth for challenging urban scenes. Compared with other methods, our method achieves a higher level of accuracy, especially for the prediction of small volume objects.

References

1. 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: Proceedings of the IEEE/CVF International Conference on Computer Vision. (2019) 9297–9307
2. Bieder, F., Wirges, S., Janosovits, J., Richter, S., Wang, Z., Stiller, C.: Exploiting multi-layer grid maps for surround-view semantic segmentation of sparse lidar data. In: 2020 IEEE Intelligent Vehicles Symposium (IV), IEEE (2020) 1892–1898
3. Fei, J., Peng, K., Heidenreich, P., Bieder, F., Stiller, C.: Pillarsegnet: Pillar-based semantic grid map estimation using sparse lidar data. In: 2021 IEEE Intelligent Vehicles Symposium (IV), IEEE (2021) 838–844
4. Peng, K., Fei, J., Yang, K., Roitberg, A., Zhang, J., Bieder, F., Heidenreich, P., Stiller, C., Stiefelhagen, R.: Mass: Multi-attentional semantic segmentation of lidar data for dense top-view understanding. IEEE Transactions on Intelligent Transportation Systems (2022)