Generalized Few-Shot Point Cloud Segmentation Via Geometric Words (Supplementary Material)

Yating Xu[1] Conghui Hu[1] Na Zhao[2*] Gim Hee Lee[1]
[1]Department of Computer Science, National University of Singapore
[2]Singapore University of Technology and Design
xu.yating@u.nus.edu conghui@nus.edu.sg na_zhao@sutd.edu.sg gimhee.lee@nus.edu.sg

1. Experimental Results on 3-shot Setting

To further validate the effectiveness of our model, we compare our method with the baselines under the 3-shot setting on S3DIS and ScanNet in Tab. 1 and Tab. 2, respectively. The results consistently illustrate that our model outperforms all baselines by a large margin on novel class segmentation, and achieves the best overall performance.

Methods	mIoU-B	mIoU-N	mIoU-A	HM
Fully Supervised	76.51	58.69	68.29	66.42
attMPTI [3]	36.28	13.32	25.68	19.28
PIFS [1]	54.34	20.00	38.53	29.23
CAPL [2]	**73.66**	33.39	55.05	45.76
Ours	73.55	**41.55**	**58.78**	**53.04**

Table 1. Results on **S3DIS** under 3-shot setting.

Methods	mIoU-B	mIoU-N	mIoU-A	HM
Fully Supervised	43.12	37.04	41.34	39.85
attMPTI [3]	16.78	2.42	12.47	4.24
PIFS [1]	35.97	2.86	26.04	5.31
CAPL [2]	38.32	13.65	30.92	20.05
Ours	**40.22**	**17.90**	**33.52**	**24.72**

Table 2. Results on **ScanNet** under 3-shot setting.

2. t-SNE Visualization

Fig. 1 displays the t-SNE visualization for S3DIS under the 5-shot setting. The difference between the left and right figures is whether the geometric-aware semantic representation (GSR) is employed. By using GSR, the representation of novel classes are more discriminative.

*Na Zhao was concurrently a visiting professor at the National University of Singapore when this work was done.

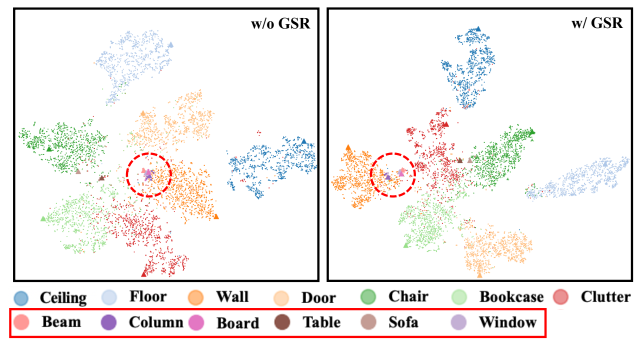

Figure 1. **t-SNE visualization on 5-shot setting of S3DIS.** Small dots represent point features and triangle represents weights for semantic prototypes P_{ori}. Novel classes are indicated with the red rectangle. Best viewed zoomed-in.

3. Further Analysis on ScanNet

To comprehensively evaluate the performance of our framework on GFS-3DSeg, we provide further analysis on ScanNet in this section.

3.1. Ablation Study

Tab. 3 shows the ablation study on ScanNet. Both geometric-aware semantic representation (GSR) and geometric-guided re-weighting (GRW) are beneficial to novel class generalization, and our full model with both GSR and GRW performs the best regarding overall segmentation accuracy.

3.2. Qualitative Results

The qualitative results in Fig. 2 demonstrate that our model can segment novel classes (Picture in the first row, Toilet and Sink in the second row) more precisely than CAPL [2]. Concurrently, we can still maintain good segmentation performance on base classes.

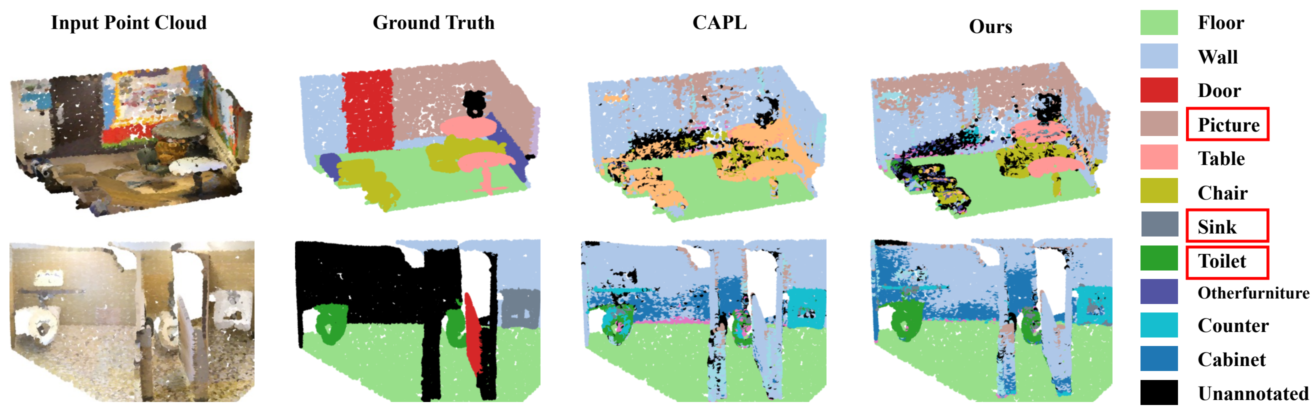

Figure 2. **Qualitative comparison on 5-shot setting of ScanNet**. Target novel classes are marked with red rectangles. The target novel class in the first row is **picture**. The target novel classes in the second row are **toilet** and **sink**.

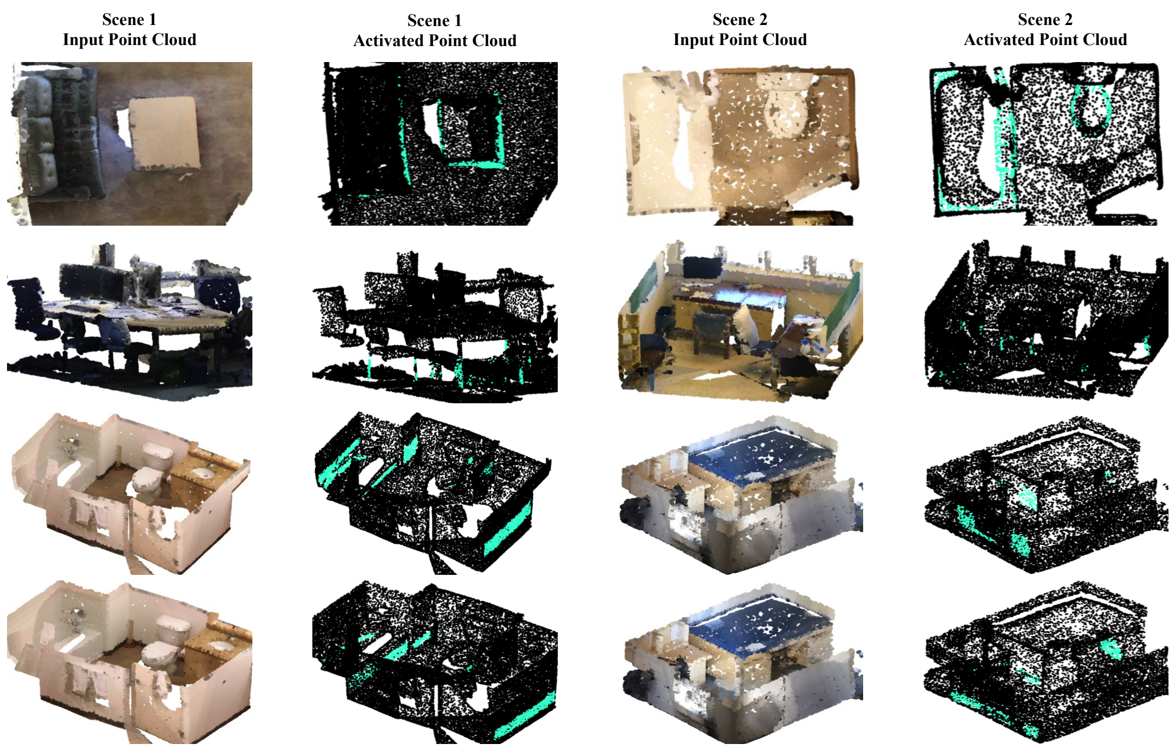

Figure 3. **Visualization of geometric words on ScanNet**. Each row shows the activated point cloud regarding the same geometric word in two different scenes. The activated points are colored green.

GSR	GRW	mIoU-B	mIoU-N	mIoU-A	HM
✗	✗	38.22	14.39	31.07	20.88
✓	✗	**40.21**	17.54	33.40	24.39
✓	✓	40.18	**18.58**	**33.70**	**25.39**

Table 3. Effectiveness of geometric-aware semantic representation (GSR) and geometric-guided re-weighting (GRW) on ScanNet.

3.3. Visualization of Geometric Words

Fig. 3 visualizes the geometric words (GWs) on ScanNet. Each row shows two activated point clouds regarding to the same geometric word in different scenes. In the first row, the edge of the sofa, table, bathtub and toilet are all activated when provided with the same GW. In the second row, the stick of chair and table are activated. It suggests that the GWs are able to represent shared geometric components **across different scenes and different classes**. Interestingly, we also find that GWs are height-aware. The activated parts in the third and fourth rows regarding two GWs represent vertical planes of different heights.

References

[1] Fabio Cermelli, Massimiliano Mancini, Yongqin Xian, Zeynep Akata, and Barbara Caputo. Prototype-based incremental few-shot semantic segmentation. *arXiv preprint arXiv:2012.01415*, 2020. 1

[2] Zhuotao Tian, Xin Lai, Li Jiang, Shu Liu, Michelle Shu, Hengshuang Zhao, and Jiaya Jia. Generalized few-shot semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11563–11572, 2022. 1

[3] Na Zhao, Tat-Seng Chua, and Gim Hee Lee. Few-shot 3d point cloud semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8873–8882, 2021. 1