

PointGS: Semantic-Consistent Unsupervised 3D Point Cloud Segmentation with 3D Gaussian Splatting

Supplementary Material

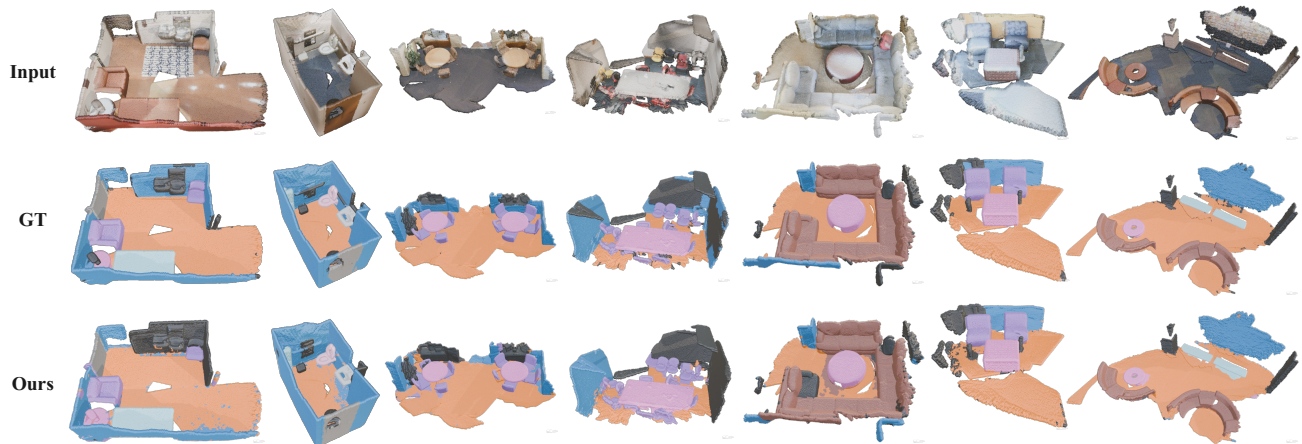


Figure 1. Qualitative comparison with Ground Truth of unsupervised segmentation on the ScanNet validation set. Each color represents one semantic class.

1. Additional Experiments

Due to space constraints in the main text, some of the experiments are placed in this supplementary material.

1.1. Qualitative Experiment on ScanNet-v2

To verify the performance on the ScanNet-v2 dataset, we conduct additional visual experiments on this dataset. As shown in Fig. 1, the performance of our method on the ScanNet dataset is very close to the Ground Truth. For more open rooms, we directly projected the scene; for more enclosed rooms, we first split the scene into two parts and then performed projection, following the strategy used for S3DIS.

1.2. Additional Contrast Experiment

We present qualitative results of the direct point projection method without 3D-GS as a comparison to our approach. The qualitative experimental results are presented in Fig. 2. As can be seen from Fig. 2, although the general semantic boundaries can be identified, the semantic ambiguity occurs due to the confusion between foreground and background. Since the semantics were directly grown from the initial sparse point cloud onto the original point cloud, the areas of semantic confusion present as large patches.

1.3. Additional Experiment on S3DIS

We also test the performance of S3DIS Area 5 in terms of mIoU for each category. The per-category mIoU from label 0 to 12 are respectively 68.4/ 74.5/ 57.5/ 40.2/ 51.6/ 44.9/ 43.6/ 39.0/ 50.1/ 65.5/ 38.3/ 45.3/ 31.5. The mIoU of S3DIS 6-fold validation is 45.8.

**PointGS
w/o 3DGS**



**PointGS
(Ours)**

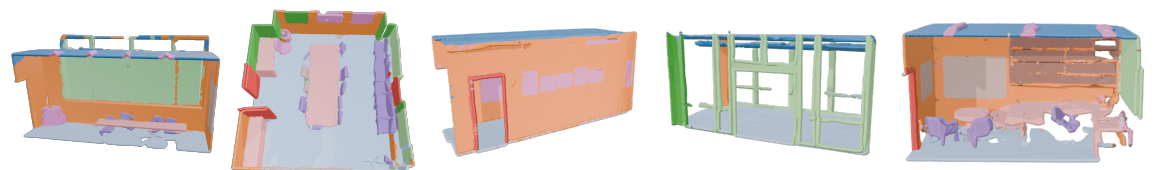


Figure 2. Qualitative comparison of different ablation settings on the ScanNet validation set.