

Hierarchical Aggregation for 3D Instance Segmentation

Supplementary Material

Shaoyu Chen¹ Jiemin Fang^{2,1} Qian Zhang³ Wenyu Liu¹ Xinggong Wang^{1†}

¹ School of EIC, Huazhong University of Science & Technology

² Institute of AI, Huazhong University of Science & Technology

³ Horizon Robotics

{shaoyuchen, jaminfang, liuwuy, xgwang}@hust.edu.cn {qian01.zhang}@horizon.ai

In this supplementary material, we provide the detailed inference latency of main individual components of HAIS and compare it with other methods. Besides, we provide more qualitative results to demonstrate the effectiveness of HAIS.

A. Detailed Inference Time

Tab. 1 shows the inference time of main components of different methods. Compared with other methods which require time-consuming clustering and post processing procedures, our HAIS keeps a much more efficient pipeline. The point-wise prediction network, point aggregation, set aggregation and intra-instance prediction network takes 172, 125, 4 and 109 ms, respectively.

B. Additional Qualitative Results

We show more qualitative results on the validation split of the ScanNet v2 dataset in Fig. 1. The predicted center shift vectors of some points are not accurate and a large amount of instance fragments come into being. By introducing the hierarchical aggregation and intra-instance prediction, HAIS generates fine-grained instance predictions.

References

- [1] Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas A. Funkhouser, and Matthias Nießner. ScanNet: Richly-annotated 3d reconstructions of indoor scenes. In *CVPR*, 2017.
- [2] Lei Han, Tian Zheng, Lan Xu, and Lu Fang. Occuseg: Occupancy-aware 3d instance segmentation. In *CVPR*, 2020.
- [3] Li Jiang, Hengshuang Zhao, Shaoshuai Shi, Shu Liu, Chi-Wing Fu, and Jiaya Jia. Pointgroup: Dual-set point grouping for 3d instance segmentation. In *CVPR*, 2020.
- [4] Shih-Hung Liu, Shang-Yi Yu, Shao-Chi Wu, Hwann-Tzong Chen, and Tyng-Luh Liu. Learning gaussian instance segmentation in point clouds. *arXiv:2007.09860*, 2020.
- [5] Weiyue Wang, Ronald Yu, Qianguai Huang, and Ulrich Neumann. SGPN: similarity group proposal network for 3d point cloud instance segmentation. In *CVPR*, 2018.
- [6] Xinlong Wang, Shu Liu, Xiaoyong Shen, Chunhua Shen, and Jiaya Jia. Associatively segmenting instances and semantics in point clouds. In *CVPR*, 2019.
- [7] Bo Yang, Jianan Wang, Ronald Clark, Qingyong Hu, Sen Wang, Andrew Markham, and Niki Trigoni. Learning object bounding boxes for 3d instance segmentation on point clouds. In *NeurIPS*, 2019.
- [8] Li Yi, Wang Zhao, He Wang, Minhyuk Sung, and Leonidas J. Guibas. GSPN: generative shape proposal network for 3d instance segmentation in point cloud. In *CVPR*, 2019.

Table 1. Inference time of main components of different methods on the ScanNet v2 validation set. For fair comparison, data in this table is measured on the same type of GPU (Titan X).

Method	Component inference time (msec)	Per frame inference time (msec)
SGPN [5]	backbone (GPU): 2080 group merging (CPU): 149000 block merging (CPU): 7119	158439
ASIS [6]	backbone (GPU): 2083 mean shift (CPU): 172711 block merging (CPU): 7119	181913
GSPN [8]	backbone (GPU): 1612 point sampling (GPU): 9559 neighbour search (CPU): 1500	12702
3D-BoNet [7]	backbone (GPU): 2083 SCN (GPU): 667 block merging (CPU): 7119	9202
OccuSeg [2]	backbone GPU): 189 supervoxel (CPU): 1202 clustering (GPU+CPU): 513	1904
PointGroup [3]	backbone (GPU): 128 clustering (GPU+CPU):221 ScoreNet (GPU): 103	452
GICN [4]	backbone (GPU): 1497 SCN (GPU): 667 block merging(CPU): 7119	8615
H AIS	point-wise prediction(GPU): 172 point aggregation(GPU+CPU): 125 set aggregation(GPU): 4 intra-instance prediction(GPU): 109	410

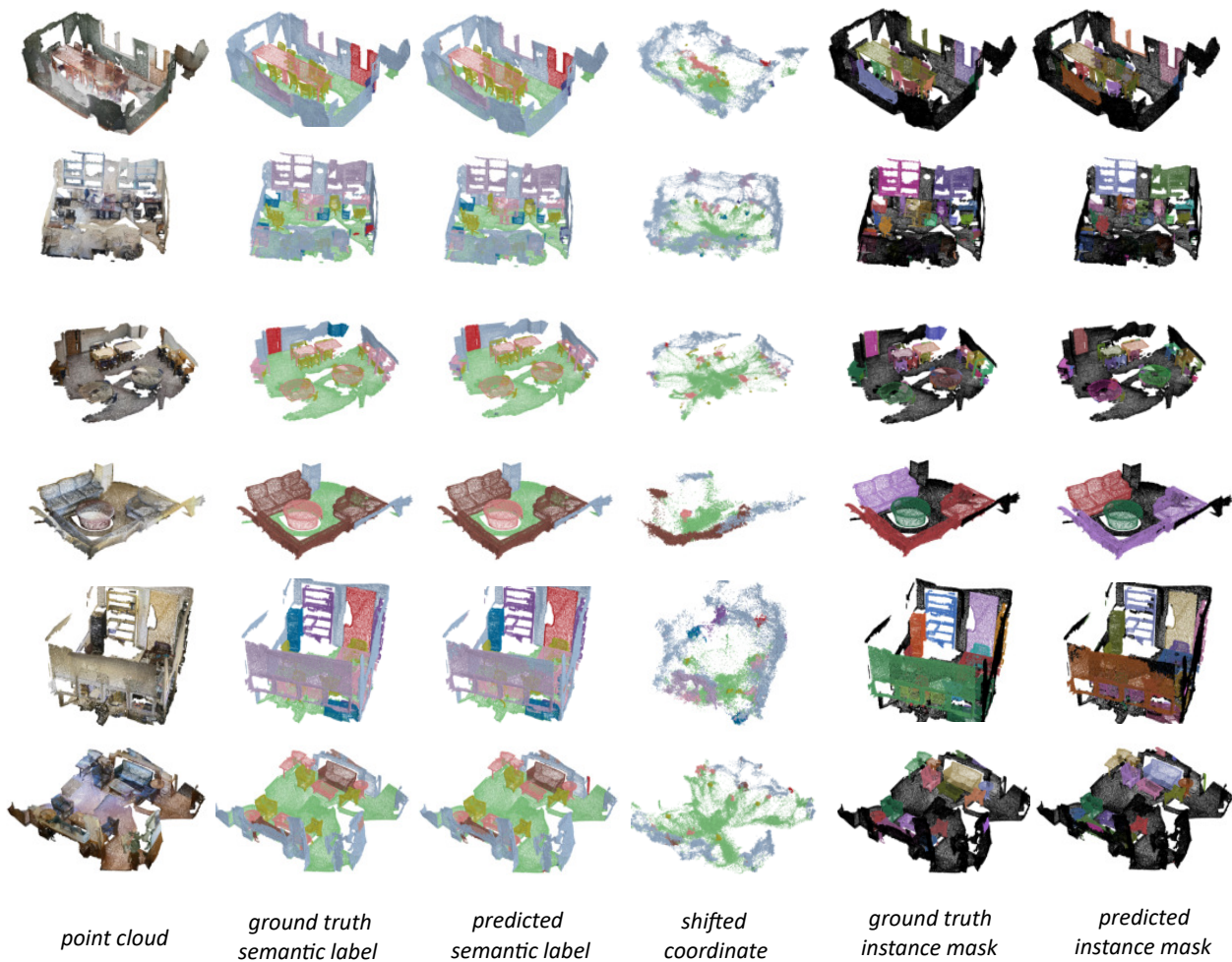


Figure 1. Qualitative results on ScanNet v2 [1]. From left to right: input point cloud, ground truth semantic label, predicted semantic label, shifted coordinate, ground truth instance mask and predicted instance mask. Best viewed in color.