1. Pseudo-code of Voting LPs

We provide the pseudo-code of voting LPs in Alg. 1. No hyperparameter need be tuned in the algorithm. Ablation experiment Tab. 5 in the main paper verifies the effectiveness of our algorithm.

Algorithm 1: Voting LPs

Data: Distance vector: $Dist\_p$
The number of core point: N_{cp}
The number of noise point: N_{np}
Count points number vector for instance: Ins_c
Instances: $I \in \{I^1, I^2, \ldots, I^{N_{\text{I}}}\}$
Core points coordinates:
$P_c \in \left\{P_c^1, P_c^2, \ldots, P_c^{N_{\text{cp}}}\right\}$
Noise points coordinates:
$P_{np} \in \left\{P_{np}^1, P_{np}^2, \ldots, P_{np}^{N_{\text{np}}}\right\}$
Predicted semantic labels of core pints:
$S_c \in \left\{S_c^1, S_c^2, \ldots, S_c^{N_{\text{cp}}}\right\}$
Predicted semantic labels of noise pints:
$S_{np} \in \left\{S_{np}^1, S_{np}^2, \ldots, S_{np}^{N_{\text{np}}}\right\}$
Predicted instance labels of core pints:
$I_c \in \left\{I_c^1, I_c^2, \ldots, I_c^{N_{\text{cp}}}\right\}$
Merging radius for each semantic class:
$\alpha m_c \; c \in \{chair, cabinet, sofa, etc.\}$

for i, p_{np} **in** enumerate(P_{np})
    $s_{np} = S_{np}^i$
    for j, p_c **in** enumerate(P_c)
        $s_c = S_c^j$,$i_c = I_c^j$
        $Dist\_p$.pushback($\left\|P_{np}^i - P_c^j\right\|$)
        if $s_c = s_{np}$ & & $\left\|P_{np}^i - P_c^j\right\| < \alpha m_{s_{np}}$
            $Ins_c[i_c]$ ++
    if max(Ins_c) == 0
        min_idx = argmin($Dist\_p$)
        Assign p_{np} to $I_c^{min\_idx}$
        continue
    min_idx = argmax(Ins_c)
    Assign p_{np} to I^{min_idx}
    Empty $Dist\_p$, Ins_c

2. ScanNetV2 Benchmark Challenge

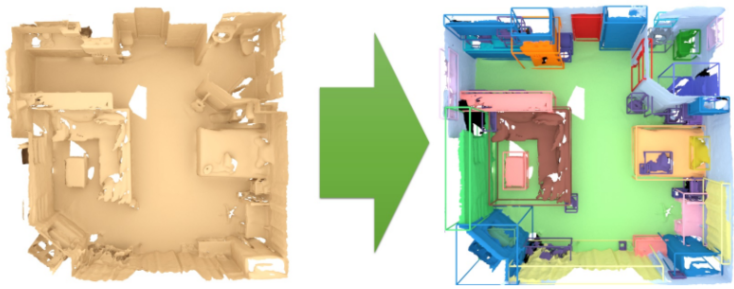
PBNet ranks 1st on mAP metric of ScanNetV2 3D instance segmentation challenge, by January 2023. In this challenge, the ground-truth of the hidden test set is unreleased. Participants are limited to submitting predictions to the official server once for two weeks to get metric values. Fig. 1 is the screenshot proof.

3. Visualization Results

In Fig. 2 and Fig. 3, we provide the visualization results of PBNet on ScanNetV2 and S3DIS datasets. Each column from left to right is described as follows: Input, semantic ground-truth (Sem. GT), semantic prediction (Sem. Pred.), instance ground-truth (Ins. GT), and instance prediction (Ins. Pred.).

3D Semantic Instance Benchmark

The 3D semantic instance prediction task involves detecting and segmenting the object in an 3D scan mesh.



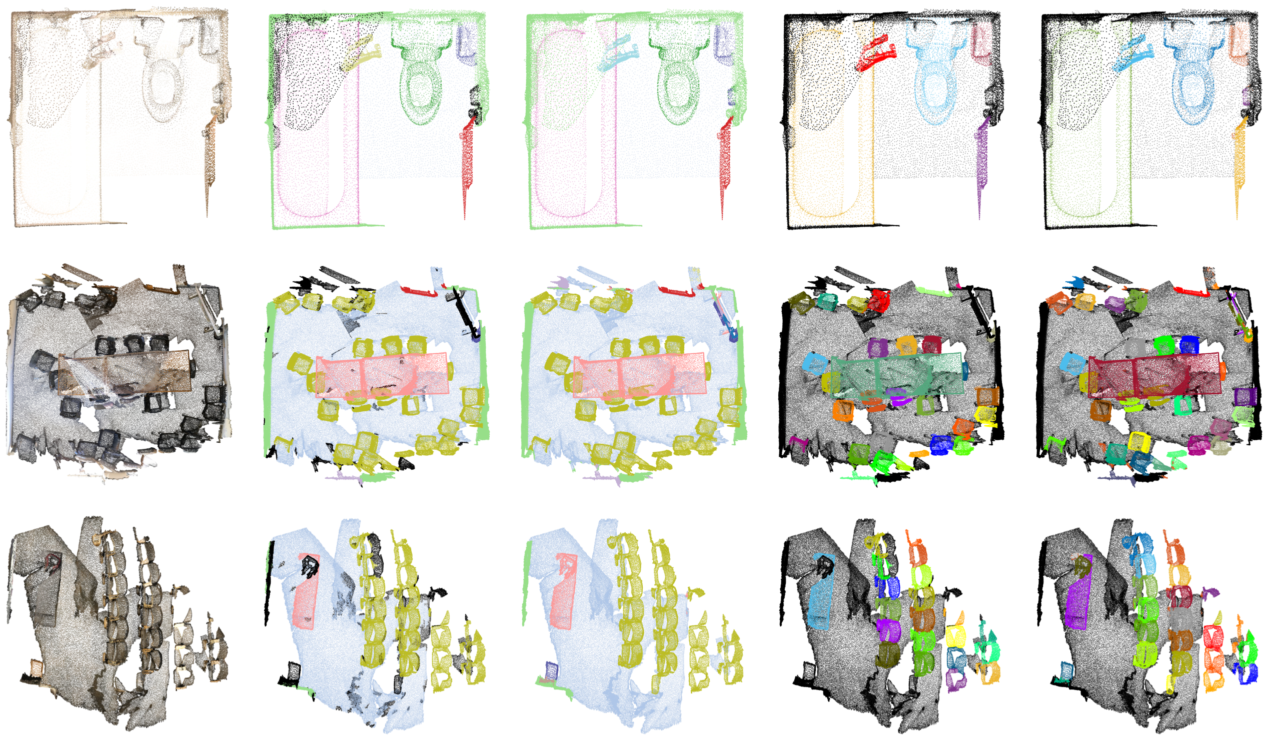
Evaluation and metrics
Our evaluation ranks all methods according to the average precision for each class. We report the mean average precision AP at overlap 0.25 (AP 25%), overlap 0.5 (AP 50%), and over overlaps in the range [0.5:0.95:0.05] (AP). Note that multiple predictions of the same ground truth instance are penalized as false positives.

This table lists the benchmark results for the 3D semantic instance scenario.

Metric: AP

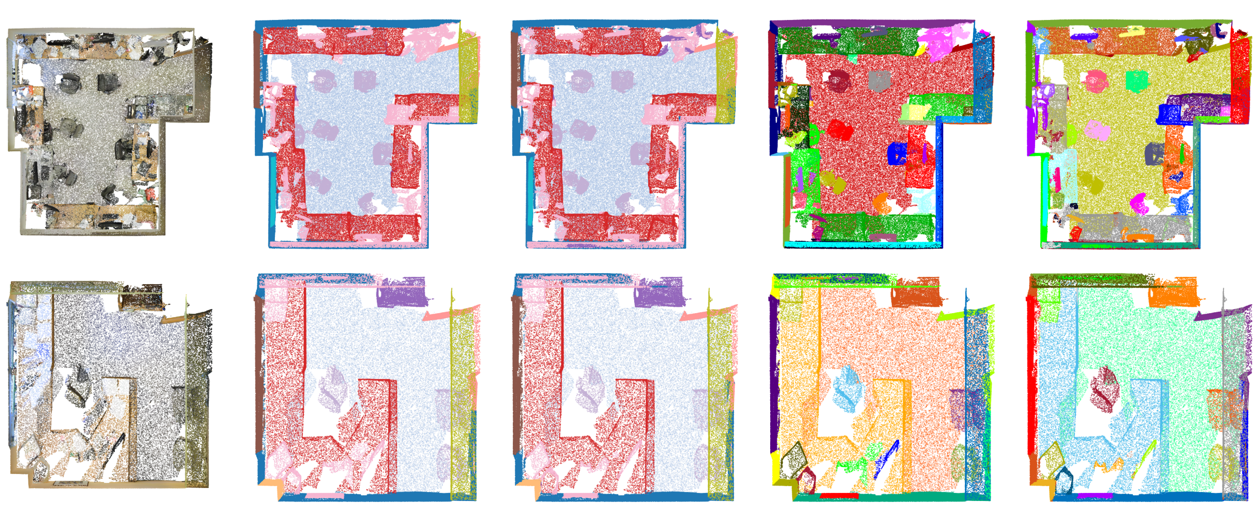
Method	Info	avg ap	bathtub	bed	bookshelf	cabinet	chair	counter	curtain	desk	door	otherfurniture	picture	refrigerator	shower curtain	
PBNet	P	**0.573** 1	0.926 2	0.575 8	**0.619** 1	**0.472** 1	0.736 4	0.239 3	0.487 21	0.383 2	0.459 2	0.506 5	0.533 6	0.585 4	0.767 7	0
Mask3D		0.566 2	0.926 2	0.597 4	0.408 16	0.420 2	0.737 3	0.239 2	0.598 7	**0.386** 1	0.458 3	**0.549** 1	0.568 4	**0.716** 1	0.601 21	0
Jonas Schult, Francis Engelmann, Alexander Hermans, Or Litany, Siyu Tang, Bastian Leibe: Mask3D for 3D Semantic Instance Segmentation.																
GraphCut		0.552 3	**1.000** 1	0.611 3	0.438 12	0.392 5	0.714 5	0.139 6	0.598 8	0.327 4	0.389 5	0.510 4	**0.598** 1	0.427 19	0.754 10	0
SPFormer	P	0.549 4	0.745 10	**0.640** 1	0.484 5	0.395 4	0.739 2	**0.311** 1	0.566 12	0.335 3	**0.468** 1	0.492 6	0.555 5	0.478 11	0.747 12	0
Sun Jiahao, Qing Chunmei, Tan Junpeng, Xu Xiangmin: Superpoint Transformer for 3D Scene Instance Segmentation. AAAI 2023																
DKNet		0.532 5	0.815 6	0.624 2	0.517 3	0.377 7	**0.749** 1	0.107 8	0.509 18	0.304 6	0.437 4	0.475 7	0.581 2	0.539 7	0.775 6	
Yizheng Wu, Min Shi, Shuaiyuan Du, Hao Lu, Zhiguo Cao, Weicai Zhong: 3D Instances as 1D Kernels. ECCV 2022																
IPCA-Inst		0.520 6	0.889 4	0.551 11	0.548 2	0.418 3	0.665 14	0.064 15	0.585 9	0.260 12	0.277 15	0.471 9	0.500 7	0.644 2	0.785 4	0
SoftGroup++		0.513 7	0.704 16	0.578 7	0.398 17	0.363 11	0.704 6	0.061 16	0.647 4	0.297 10	0.378 8	0.537 2	0.343 9	0.614 3	0.828 3	
SSTNet	P	0.506 8	0.738 13	0.549 12	0.497 4	0.316 15	0.693 9	0.178 5	0.377 28	0.198 17	0.330 9	0.463 10	0.576 3	0.515 9	0.857 2	**0**
Zhihao Liang, Zhihao Li, Songcen Xu, Mingkui Tan, Kui Jia: Instance Segmentation in 3D Scenes using Semantic Superpoint Tree Networks. ICCV2021																
SoftGroup	P	0.504 9	0.667 22	0.579 6	0.372 19	0.381 6	0.694 8	0.072 12	0.677 2	0.303 7	0.387 6	0.531 3	0.319 13	0.582 5	0.754 9	
Thang Vu, Kookhoi Kim, Tung M. Luu, Xuan Thanh Nguyen, Chang D. Yoo: SoftGroup for 3D Instance Segmentaiton on Point Clouds. CVPR 2022 [Oral]																
OccuSeg+instance		0.486 10	0.802 7	0.536 14	0.428 14	0.369 8	0.702 7	0.205 4	0.331 32	0.301 8	0.379 7	0.474 8	0.327 10	0.437 15	**0.862** 1	0
Lei Han, Tian Zheng, Lan Xu, Lu Fang: OccuSeg: Occupancy-aware 3D Instance Segmentation. CVPR2020																

Figure 1. ScanNetV2 Benchmark Challenge



Input Sem. GT Sem. Pred. Ins. GT Ins. Pred.

Figure 2. Visualization Results of PBNet on ScanNetV2.



Input Sem. GT Sem. Pred. Ins. GT Ins. Pred.

Figure 3. Visualization Results of PBNet on S3DIS.