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: $\text{Dist}_p$
- The number of core point: $N_{cp}$
- The number of noise point: $N_{np}$
- Count points number vector for instance: $\text{Ins}_{ic}$
- Instances: $I \in \{I_1, I_2, \ldots, I_{N_I}\}$
- Core points coordinates: $P_c \in \{P_c^1, P_c^2, \ldots, P_c^{N_{cp}}\}$
- Noise points coordinates: $P_{np} \in \{P_{np}^1, P_{np}^2, \ldots, P_{np}^{N_{np}}\}$
- Predicted semantic labels of core pints: $S_c \in \{S_c^1, S_c^2, \ldots, S_c^{N_{cp}}\}$
- Predicted semantic labels of noise pints: $S_{np} \in \{S_{np}^1, S_{np}^2, \ldots, S_{np}^{N_{np}}\}$
- Predicted instance labels of core pints: $I_c \in \{I_c^1, I_c^2, \ldots, I_c^{N_{cp}}\}$
- Merging radius for each semantic class: $\alpha_m \in \{\text{chair, cabinet, sofa, etc.}\}$

1. for $i, p_{np}$ in enumerate($P_{np}$)
   2. $s_{np} = S_{np}^i$
   3. for $j, p_c$ in enumerate($P_c$)
      4. $s_c = S_c^j, i_c = I_c^j$
      5. $\text{Dist}_p$.pushback($\|P_{np}^i - P_c^j\|$)
      6. if $s_c = s_{np}$ & $\|P_{np}^i - P_c^j\| < \alpha_m_{snp}$
         $\text{Ins}_{ic}[i_c]++$
      7. if max($\text{Ins}_{ic}$) == 0
         $\text{min_idx} = \text{argmin}(\text{Dist}_p)$
         Assign $p_{np}$ to $I_{ic}^{\text{min_idx}}$
         continue
      8. $\text{min_idx} = \text{argmax}(\text{Ins}_{ic})$
      9. Assign $p_{np}$ to $I_{ic}^{\text{min_idx}}$

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 0.5 (AP 50%), overlap 0.5 (AP 75%), and overlap 0.5 (AP 90%). 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.

<table>
<thead>
<tr>
<th>Method</th>
<th>Info</th>
<th>avg ap</th>
<th>bath tub</th>
<th>bed</th>
<th>bookshelf</th>
<th>cabinet</th>
<th>chair</th>
<th>curtain</th>
<th>desk</th>
<th>door</th>
<th>other furniture</th>
<th>picture</th>
<th>refrigerator</th>
<th>shower</th>
<th>curtain</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRNet</td>
<td>[P]</td>
<td>0.573</td>
<td>0.935 2</td>
<td>0.575 8</td>
<td>0.619 1</td>
<td>0.472 1</td>
<td>0.736 4</td>
<td>0.279 3</td>
<td>0.487</td>
<td>3.583 2</td>
<td>0.459 2</td>
<td>0.506 5</td>
<td>0.533 6</td>
<td>0.585 4</td>
<td>0.767 7</td>
</tr>
<tr>
<td>Mask3D</td>
<td>[P]</td>
<td>0.566 2</td>
<td>0.506 2</td>
<td>0.557 4</td>
<td>0.494 16</td>
<td>0.426 2</td>
<td>0.737 3</td>
<td>0.239 2</td>
<td>0.687</td>
<td>3.308 1</td>
<td>0.458 3</td>
<td>0.540 1</td>
<td>0.568 4</td>
<td>0.716 1</td>
<td>0.601 0</td>
</tr>
<tr>
<td>Jia et al.</td>
<td>[P]</td>
<td>0.549 4</td>
<td>0.745</td>
<td>0.640 1</td>
<td>0.484 5</td>
<td>0.595 4</td>
<td>0.793 2</td>
<td>0.511 1</td>
<td>0.568</td>
<td>3.355 3</td>
<td>0.468 1</td>
<td>0.492 5</td>
<td>0.555 5</td>
<td>0.476 11</td>
<td>0.747 0</td>
</tr>
<tr>
<td>Sun et al.</td>
<td>[P]</td>
<td>0.530 3</td>
<td>1.100 1</td>
<td>0.611 3</td>
<td>0.438 12</td>
<td>0.392 5</td>
<td>0.714 4</td>
<td>0.139 6</td>
<td>0.998 8</td>
<td>0.327 4</td>
<td>0.389 5</td>
<td>0.510 4</td>
<td>0.991 8</td>
<td>0.427 19</td>
<td>0.754 0</td>
</tr>
<tr>
<td>GraphCut</td>
<td>[P]</td>
<td>0.500 3</td>
<td>0.815 4</td>
<td>0.624 2</td>
<td>0.517 3</td>
<td>0.377 1</td>
<td>0.789 4</td>
<td>0.107 6</td>
<td>0.509</td>
<td>0.304 4</td>
<td>0.437 4</td>
<td>0.475 7</td>
<td>0.581 2</td>
<td>0.509 7</td>
<td>0.775 6</td>
</tr>
<tr>
<td>Yeh et al.</td>
<td>[P]</td>
<td>0.500 3</td>
<td>0.815 4</td>
<td>0.624 2</td>
<td>0.517 3</td>
<td>0.377 1</td>
<td>0.789 4</td>
<td>0.107 6</td>
<td>0.509</td>
<td>0.304 4</td>
<td>0.437 4</td>
<td>0.475 7</td>
<td>0.581 2</td>
<td>0.509 7</td>
<td>0.775 6</td>
</tr>
<tr>
<td>SoftGroup++</td>
<td>[P]</td>
<td>0.513 7</td>
<td>0.704</td>
<td>0.578 7</td>
<td>0.391 17</td>
<td>0.363 3</td>
<td>0.704 6</td>
<td>0.661</td>
<td>0.648 4</td>
<td>0.297</td>
<td>0.376 8</td>
<td>0.532 7</td>
<td>0.343 9</td>
<td>0.514 3</td>
<td>0.828 3</td>
</tr>
<tr>
<td>SSTNet</td>
<td>[P]</td>
<td>0.506 8</td>
<td>0.738</td>
<td>0.549</td>
<td>0.497 4</td>
<td>0.316</td>
<td>0.693 8</td>
<td>0.178 5</td>
<td>0.377</td>
<td>0.197 6</td>
<td>0.350 8</td>
<td>0.453 19</td>
<td>0.576 3</td>
<td>0.515 9</td>
<td>0.657 2</td>
</tr>
<tr>
<td>SoftGroup</td>
<td>[P]</td>
<td>0.500 3</td>
<td>0.815 4</td>
<td>0.624 2</td>
<td>0.517 3</td>
<td>0.377 1</td>
<td>0.789 4</td>
<td>0.107 6</td>
<td>0.509</td>
<td>0.304 4</td>
<td>0.437 4</td>
<td>0.475 7</td>
<td>0.581 2</td>
<td>0.509 7</td>
<td>0.775 6</td>
</tr>
<tr>
<td>Xi et al.</td>
<td>[P]</td>
<td>0.506 8</td>
<td>0.738</td>
<td>0.549</td>
<td>0.497 4</td>
<td>0.316</td>
<td>0.693 8</td>
<td>0.178 5</td>
<td>0.377</td>
<td>0.197 6</td>
<td>0.350 8</td>
<td>0.453 19</td>
<td>0.576 3</td>
<td>0.515 9</td>
<td>0.657 2</td>
</tr>
</tbody>
</table>

Figure 1. ScanNetV2 Benchmark Challenge
Figure 2. Visualization Results of PBNet on ScanNetV2.

Figure 3. Visualization Results of PBNet on S3DIS.