Polynomial Building segmentation

- A fundamental task for disaster management, urban planning, and geographical information updating, etc.
- The pixel-wise segmentation methods in most studies produce building extraction results in raster format.
- Polynomial building segmentation approaches produce more realistic building polygons in the desirable vector format for practical applications.

Limitations of the existing methods

- Relying on a perfect segmentation map to guarantee the quality of vectorization;
- Requiring a complex post-processing procedure;
- Generating inaccurate vertices with a fixed quantity, a wrong sequential order, self-intersections, etc.

Our proposed approach

- A multi-task segmentation network for joint semantic and geometric learning via three relevant tasks.
- A rule-based vertex generation module to bridge the gap between the image and the graph-based network.
- A polynomial refinement network to automatically move the polygon vertices into more accurate locations.
- Generating building polygons with a flexible quantity of vertices that are in a proper sequential order.

Multi-task segmentation network

- Building footprint segmentation, multi-class corner prediction, and edge orientation prediction.
- Formulated as pixel-wise classification problems and trained jointly with the cross entropy loss.

Vertex generation module (VGM)

- The initial vertex set is obtained by densely extracting each pixel from the segmentation contour.
- The corner and edge orientation criterions are designed for selecting a set of valid vertices.

Polygon refinement network (PRN)

- GGNN-based model utilizes extra information e.g., the feature of each vertex and their relations.
- Predicting a displacement for each vertex to produce the final result with more accurate vertices.

Experimental result evaluation on building segmentation

- Evaluated on two popular datasets: (1) The CrowdAI mapping challenge dataset; (2) The Vegas dataset of the SpaceNet building dataset.
- For building segmentation results, our method improves the F1-score of current state-of-the-art by 1.5%, 0.4%, and 2.1% under different IoU thresholds.

<table>
<thead>
<tr>
<th>Method</th>
<th>AP</th>
<th>AP50</th>
<th>AP75</th>
<th>AR</th>
<th>AR50</th>
<th>AR75</th>
<th>F1</th>
<th>F150</th>
<th>F175</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mask-RCNN (He et al. 2017)</td>
<td>41.9</td>
<td>57.5</td>
<td>48.8</td>
<td>47.6</td>
<td>70.8</td>
<td>55.5</td>
<td>44.6</td>
<td>69.1</td>
<td>51.9</td>
</tr>
<tr>
<td>PANet (Lin et al. 2018)</td>
<td>50.7</td>
<td>67.9</td>
<td>32.6</td>
<td>54.4</td>
<td>74.5</td>
<td>65.2</td>
<td>52.5</td>
<td>74.2</td>
<td>63.9</td>
</tr>
<tr>
<td>PolyMapper (Li et al. 2019)</td>
<td>55.7</td>
<td>86.0</td>
<td>61.1</td>
<td>62.1</td>
<td>88.6</td>
<td>71.4</td>
<td>58.7</td>
<td>87.3</td>
<td>68.1</td>
</tr>
<tr>
<td>FrameField (Girard et al. 2020)</td>
<td>50.5</td>
<td>76.2</td>
<td>59.3</td>
<td>55.3</td>
<td>78.1</td>
<td>64.0</td>
<td>52.8</td>
<td>77.3</td>
<td>61.6</td>
</tr>
<tr>
<td>ASIP (Li et al. 2020)</td>
<td>65.8</td>
<td>87.6</td>
<td>73.4</td>
<td>78.7</td>
<td>94.3</td>
<td>86.1</td>
<td>71.7</td>
<td>90.8</td>
<td>79.2</td>
</tr>
<tr>
<td>Ours</td>
<td>73.8</td>
<td>92.0</td>
<td>81.9</td>
<td>72.6</td>
<td>90.5</td>
<td>80.7</td>
<td>73.2</td>
<td>91.2</td>
<td>81.3</td>
</tr>
</tbody>
</table>

Results of vertex prediction and ablation study

- For vertex prediction results, our method achieves the F1-score gain of 6.64% and 6.82% compared with ASIP.
- Our method produces more accurate polygon vertices in terms of locations, quantities, angles, etc.
- The VGM produces better F1-score compared with the Baseline via filtering out the invalid vertices.
- The PRN further improves the vertex F1-scores by adjusting the vertices to more accurate locations.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Baseline</th>
<th>+ VGM</th>
<th>+ PRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASIP</td>
<td>51.13</td>
<td>73.55</td>
<td>60.32</td>
</tr>
<tr>
<td>Ours</td>
<td>64.25</td>
<td>69.90</td>
<td>66.96</td>
</tr>
</tbody>
</table>

Qualitative comparison of ASIP (top) and our method (bottom)

An example of our prediction results at different stages

Baseline + VGM + PRN