Amodal Instance Segmentation

Ke Li
{ke.li,malik}@eecs.berkeley.edu

Jitendra Malik

Task
• Goal: Identify both visible and occluded portions of each object instance.

Image  (Modal) Instance Segmentation  Amodal Instance Segmentation

• Solving this task would enable sophisticated occlusion reasoning. Given the modal and amodal masks, we can infer:
  – Amodal bounding box
  – Presence of occlusion
  – Extent, boundary and region of occlusion
  – Relative depth ordering
• Consequently, amodal instance segmentation is harder than these tasks.

Challenge
• Lack of supervised training data: no amodal segmentation annotations are available.
• We must train a model without supervised training data.
• Observation: Hard to undo occlusion, but easy to generate occlusion.
• We train the model on synthetic occlusions generated from standard (modal) instance segmentation data.

Method
• The neural net architecture is the same as that used by Iterative Instance Segmentation [1], which is based on the hypercolumn architecture.

• Initially, we use a detection system and a modal instance segmentation system to produce modal bounding boxes and heatmaps for each object.
• In each iteration, our system takes the patch inside the current amodal bounding box and the modal heatmap as input and predicts the amodal heatmap for a larger patch.
• Based on the heatmap prediction, it decides how to expand the amodal bounding box in the subsequent iteration.

Generating Training Data
• Step 1: Randomly crop a patch from an image and its corresponding modal mask.

• Step 2: Overlay random objects and retain the original mask (which is now amodal).

• Step 3: Rescale the composite patch and sample a modal bounding box.

• Generated samples:

Results

Occlusion Prediction:

On Small Set of Annotated Images:

<table>
<thead>
<tr>
<th>Method</th>
<th>mAP at 50% IoU</th>
<th>mAP at 70% IoU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faster R-CNN + IS [1]</td>
<td>34.1</td>
<td>14.0</td>
</tr>
<tr>
<td>Faster R-CNN + AIS</td>
<td>45.2</td>
<td>22.6</td>
</tr>
</tbody>
</table>

On PASCAL 3D+ (Rigid Objects):

<table>
<thead>
<tr>
<th>Method</th>
<th>mAP at 50% IoU</th>
<th>mAP at 70% IoU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faster R-CNN + IS [1]</td>
<td>37.4</td>
<td>15.9</td>
</tr>
<tr>
<td>Faster R-CNN + AIS</td>
<td>44.0</td>
<td>20.9</td>
</tr>
</tbody>
</table>

References