AttentionMask: Attentive, Efficient Object Proposal Generation Focusing on Small Objects

Motivation

Task: Object Proposal Generation
Goal: Generating class-agnostic object candidates
Problem 1: State-of-the-art systems often miss small objects
Problem 2: Simply adding a module to detect small objects is impossible due to inefficient use of resources
Our idea: Starting from [1], we focus processing on relevant parts of the image to save resources — use those resources to better detect small objects

Results (for more results scan the QR code)


Evaluation on MS COCO

- Evaluation against several state-of-the-art systems
- Average recall (AR@# of proposals) is used as evaluation measure

<table>
<thead>
<tr>
<th>Method</th>
<th>All@100</th>
<th>All@10</th>
<th>All@100</th>
<th>All@10</th>
<th>All@100</th>
<th>All@10</th>
<th>All@100</th>
<th>All@10</th>
<th>All@100</th>
<th>All@10</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCG [4]</td>
<td>0.077</td>
<td>0.186</td>
<td>0.299</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.484</td>
<td>1.35s</td>
<td>1.35s</td>
<td>1.35s</td>
<td>48m</td>
</tr>
<tr>
<td>SharpMaskZoom [2]</td>
<td>0.151</td>
<td>0.246</td>
<td>0.372</td>
<td>0.093</td>
<td>0.389</td>
<td>0.469</td>
<td>1.38s</td>
<td>1.35s</td>
<td>1.35s</td>
<td>1.35s</td>
<td>48m</td>
</tr>
<tr>
<td>SharpMask [2]</td>
<td>0.154</td>
<td>0.278</td>
<td>0.365</td>
<td>0.065</td>
<td>0.399</td>
<td>0.518</td>
<td>2.09s</td>
<td>2.09s</td>
<td>2.09s</td>
<td>2.09s</td>
<td>48m</td>
</tr>
<tr>
<td>Instance-CNN [3]</td>
<td>0.156</td>
<td>0.217</td>
<td>0.383</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.437</td>
<td>2.09s</td>
<td>2.09s</td>
<td>2.09s</td>
<td>48m</td>
</tr>
<tr>
<td>FastMask [1]</td>
<td>0.156</td>
<td>0.313</td>
<td>0.696</td>
<td>0.106</td>
<td>0.406</td>
<td>0.517</td>
<td>0.33s</td>
<td>0.33s</td>
<td>0.33s</td>
<td>0.33s</td>
<td>48m</td>
</tr>
<tr>
<td>AttentionMask</td>
<td>0.140</td>
<td>0.349</td>
<td>0.444</td>
<td>0.162</td>
<td>0.621</td>
<td>0.560</td>
<td>0.23s</td>
<td>0.23s</td>
<td>0.23s</td>
<td>0.23s</td>
<td>48m</td>
</tr>
</tbody>
</table>

AttentionMask beats all state-of-the-art methods across all categories including runtime!

References


Paper + Code

Christian Wilms and Simone Frintrop
{wilms,frintrop}@informatik.uni-hamburg.de

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG
DEPARTMENT OF INFORMATICS
COMPUTER VISION GROUP