Symbiotic Segmentation and Part Localization for Fine-Grained Categorization
Yuning Chai, Victor Lempitsky and Andrew Zisserman

Overview

Deformable model learning
GrabCut initialization

Defining all terms:

$E(p, f, c) = \alpha E^{DPM}(p|W, I) + \beta E^{GC}(f, c|I) + E^{S}(p, f, S)$

where $E^{DPM}(p|W, I)$ measures the DPM energy, $E^{GC}(f, c|I)$ the GrabCut energy, and $E^{S}(p, f, S)$ the symbiotic term.

Optimization

Updating part localizations $p$:

$\min_p \alpha E^{DPM}(p|W, I) + E^{S}(p, f, S)$

Updating foreground segmentation $f$ and color models $c$:

$\min_f \beta E^{GC}(f, c|I) + E^{S}(p, f, S)$

The modified energy can still be minimized exactly via graph cut.

Training

1. Train the root DPM model $w_0$ using standard DPM training
2. Run GrabCut on all images, using the center as foreground initialization, to compute a set of foreground masks for each training image.
3. Compute the root saliency $s_0$ using results from 1 and 2
4. Multiply the interestingness measure of the DPM by $s_0$ to initialize the parts
5. Train the parts of DPM $w_i$
6. Compute the parts saliency masks $s_i$

Quantitative Results

Compared to state-of-the-art

<table>
<thead>
<tr>
<th>Method</th>
<th>Bird11</th>
<th>Birds 10</th>
<th>Dogs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yang et al.</td>
<td>70.0</td>
<td>45.9</td>
<td>41.3</td>
</tr>
<tr>
<td>GrabCut</td>
<td>68.0</td>
<td>45.3</td>
<td>40.1</td>
</tr>
<tr>
<td>Symbiotic</td>
<td>59.4</td>
<td>47.3</td>
<td>45.6</td>
</tr>
</tbody>
</table>

- **Mean accuracy (mA)** is shown in all tables.
- "Symb*" is the model with classifiers trained on image sets not augmented by left-right mirroring.

Compared to baselines

- (0.2, 0.4, 0.6) are the direct baselines to (0.3, 0.5, 0.7). The symbiotic optimization is the only difference between each pair.
- All the improvements are due to the fact that part localization and segmentation processes assist each other within the proposed symbiotic model.