Contributions

- Propose an algorithm to detect fashion items a person in the image is wearing or carrying in the form of bounding boxes. This task is different from semantic-segmentation based approaches.
- Combine state-of-the-object detection pipeline (R-CNN) with bounding box geometry priors and body pose information to significantly improve the performance.
- Automatically learn pose-dependent bounding box location priors.

Detection scores

For each proposal bounding box and for each fashion item class, compute a score by combining appearance-based posterior from SVM, location prior, aspect ratio prior and perimeter prior.

\[ p(y_i | f, c, a, r, t) \propto p(y_i | f) p(c | y_i, t) p(a | y_i) p(r | y_i, t) \]

- \( y_i \in \{1, -1\} \): existence of fashion item
- \( f \): image features extracted from the bounding box
- \( c = (c_1, c_2) \): center position of the bounding box
- \( a \): log aspect ratio of the bounding box
- \( r \): log perimeter of the bounding box
- \( t \): body joints positions

Experiments

- Create detection dataset from Fashionista dataset by redefining fashion item classes (56 to 10 classes) and generating bounding box annotations.
- Use AlexNet pre-trained on ImageNet classification task to extract features from each bounding box.
- Automatically learn pose-dependent bounding box location priors.

<table>
<thead>
<tr>
<th>Methods</th>
<th>mAP</th>
<th>Bag</th>
<th>Belt</th>
<th>Glasses</th>
<th>Hat</th>
<th>Pants</th>
<th>Left Shoe</th>
<th>Right Shoe</th>
<th>Shorts</th>
<th>Skirt</th>
<th>Tights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>31.1</td>
<td>22.5</td>
<td>14.2</td>
<td>22.2</td>
<td>36.1</td>
<td>57.0</td>
<td>28.5</td>
<td>32.5</td>
<td>37.4</td>
<td>20.3</td>
<td>40.6</td>
</tr>
<tr>
<td>w/o geometric priors</td>
<td>22.4</td>
<td>19.4</td>
<td>6.0</td>
<td>13.0</td>
<td>28.9</td>
<td>72.2</td>
<td>24.2</td>
<td>24.1</td>
<td>34.5</td>
<td>15.2</td>
<td>31.7</td>
</tr>
<tr>
<td>w/o appearance</td>
<td>17.8</td>
<td>14.3</td>
<td>7.1</td>
<td>8.9</td>
<td>50.7</td>
<td>20.5</td>
<td>25.4</td>
<td>15.6</td>
<td>18.8</td>
<td>22.3</td>
<td></td>
</tr>
</tbody>
</table>