Tilting at windmills

Data augmentation for deep pose estimation does not help with occlusion

Rafal Pytel
Osman S. Kayhan
Jan C. van Gemert

Delft University of Technology
Computer Vision Lab
Challenges: Occlusions

Dense, overlapping person instances
Challenges: Occlusions

Dense, overlapping person instances

Unusual poses with occlusions
How to check robustness to occlusion?
How to check robustness to occlusion?

A: Occlusion Attacks
Keypoint attacks
Occlusion attacks - keypoint level

Original Image

Occluded image (blackout nose)

Introduced occlusion

shifted face keypoint
Sensitivity to keypoint occlusion

Head keypoints are most vulnerable.
Part attacks
Occlusion attacks - part level

Affected area is a minimum box covering all keypoints of part.
Sensitivity to part occlusion

Parts with more keypoints occluded have largest loss in performance.
How can we mitigate occlusion problems?
Occlusion augmentations - Blurring

Original image  Blurring (K)  Blurring (P)
Occlusion augmentations - Cutout

Original image  Cutout (K)  Cutout (P)
Occlusion augmentations - PartMix

Original image  PartMix

Another left arm pasted instead of right leg
Do occlusion augmentation help?
Investigated approaches

Evaluation and experiments on two most popular datasets:

- **COCO**
- **MPII**

Investigated approaches:

- **HRNet**
- **SimpleBaseline**
- **HigherHRNet**
Do keypoint level augmentations help?

- Smaller step in improvement after every extra augmentation

<table>
<thead>
<tr>
<th>Augmentation</th>
<th>p</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>-</td>
<td>65.3</td>
</tr>
<tr>
<td>Baseline (flip, rot, scale)</td>
<td>-</td>
<td>73.9</td>
</tr>
<tr>
<td>Baseline (flip, rot, scale, half body)</td>
<td>-</td>
<td>74.3</td>
</tr>
<tr>
<td>Blurring (K)</td>
<td>0.5</td>
<td>74.5</td>
</tr>
<tr>
<td>Cutout (K)</td>
<td>0.5</td>
<td>74.5</td>
</tr>
</tbody>
</table>
Do keypoint level augmentations help?

- Smaller step in improvement after every extra augmentation
- Keypoint augmentations improves the performance slightly.

<table>
<thead>
<tr>
<th>Augmentation</th>
<th>p</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>-</td>
<td>65.3</td>
</tr>
<tr>
<td>Baseline (flip, rot, scale)</td>
<td>-</td>
<td>73.9</td>
</tr>
<tr>
<td>Baseline (flip, rot, scale, half body)</td>
<td>-</td>
<td>74.3</td>
</tr>
<tr>
<td><strong>Blurring (K)</strong></td>
<td>0.5</td>
<td>74.5</td>
</tr>
<tr>
<td><strong>Cutout (K)</strong></td>
<td>0.5</td>
<td>74.5</td>
</tr>
</tbody>
</table>
Do part level augmentations help?

Slight improvement for Cutout (+ 0.2%) and PartMix (+ 0.1%).

<table>
<thead>
<tr>
<th>Augmentation</th>
<th>p</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>-</td>
<td>65.3</td>
</tr>
<tr>
<td>Baseline (flip, rot, scale)</td>
<td>-</td>
<td>73.9</td>
</tr>
<tr>
<td>Baseline (flip, rot, scale, half body)</td>
<td>-</td>
<td>74.3</td>
</tr>
<tr>
<td>Blurring (P)</td>
<td>0.5</td>
<td>74.1</td>
</tr>
<tr>
<td><strong>Cutout (P)</strong></td>
<td>0.5</td>
<td>74.5</td>
</tr>
<tr>
<td><strong>PartMix</strong></td>
<td>0.5</td>
<td>74.4</td>
</tr>
</tbody>
</table>
How does detector influences results?

- Using ground truth bounding boxes shows largest improvement.
How does detector influences results?

- Using ground truth bounding boxes shows largest improvement.
- Proposed augmentations do not always improve results.
Do occlusion augmentations bring robustness?
Robustness against occlusion attacks

- Proposed augmentations make model slightly more robust.
Robustness against occlusion attacks

- Proposed augmentations make model slightly more robust.
- Data augmentation still **do not solve** the occlusion problem.
Qualitative examples - good

- **Ground truth**
- **Baseline**
- **T. Blurring (K)**

Wrong prediction of left wrist

Improved prediction after using proposed augmentation
Qualitative examples - no improvement

Wrong prediction for both baseline and proposed augmentation.
Qualitative examples - deprivation

Wrong annotation of left ankle after training with proposed augmentation
Conclusion

- For COCO dataset head keypoints are vulnerable to keypoint occlusion attacks.
Conclusion

- For COCO dataset head keypoints are vulnerable to keypoint occlusion attacks.
- Occlusions of parts with more keypoints affect overall performance the most.
Conclusion

- For COCO dataset head keypoints are vulnerable to keypoint occlusion attacks.
- Occlusions of parts with more keypoints affect overall performance the most.
- With all the bells and whistles current and proposed augmentation do not bring significant improvement.
Conclusion

- For COCO dataset head keypoints are vulnerable to keypoint occlusion attacks.
- Occlusions of parts with more keypoints affect overall performance the most.
- With all the bells and whistles current and proposed methods do not bring significant improvement.
- Person detectors influence a lot results of top down approaches, varying boost given by augmentation.
Tilting at windmills

Data augmentation for deep pose estimation does not help with occlusion

Thank you for your attention!

Official repository:
https://github.com/rpytel1/occlusion-vs-data-augmentations