Object Detection using R-CNN Experiments

CS381V: Visual Recognition, Spring 2016
William Xie
Feb. 24, 2016
Fast R-CNN

- R-CNN: Girshick et al., CVPR 2013
- **Fast R-CNN**: Girshick, ICCV 2015
- Faster R-CNN: Ren et al., NIPS 2015
Fast R-CNN

- Implemented in modified Caffe, requires Matlab

- With VGG16
  Train: 9x faster than traditional R-CNN
  Test: 200x faster than R-CNN *

*https://github.com/rbgirshick/fast-rcnn
Fast R-CNN

• Available models: CaffeNet, VGG16, VGG_M_1024

• Trained with ImageNet (ILSVRC 2012), fine-tuned on PASCAL VOC 2007
PASCAL VOC

- 20 classes + background

CLASSES = ('__background__',
            'aeroplane', 'bicycle', 'bird', 'boat',
            'bottle', 'bus', 'car', 'cat', 'chair',
            'cow', 'diningtable', 'dog', 'horse',
            'motorbike', 'person', 'pottedplant',
            'sheep', 'sofa', 'train', 'tvmonitor')
Positive examples
Negative examples
• Each region of interest -> 21 scores, 21 boxes

• Non-maximum suppression and probability threshold
Region proposal

~2000 per image
(Selective search)
Detection and classification
Conv1

11 x 11
Conv2

5 x 5
Conv3

$3 \times 3$
Conv4

3 x 3
Conv5

3 x 3
Conv5
Running time

- CPU mode
- Intel Core i7-3770 @ 3.40 GHz (4 cores)
- CaffeNet
  - Pre-computed bounding boxes: ~8s / image
  - Single image level bounding box: ~1s / image
- VGG16 pre-computed: ~35s / image
Image level detection and classification

- No region proposals
- Input: 1 bounding box of the entire image
PASCAL
Imagenet
# Image classification accuracy

- Imagenet data, 100 images per class

<table>
<thead>
<tr>
<th></th>
<th>car</th>
<th>bottle</th>
<th>chair</th>
<th>tv</th>
<th>plant</th>
<th>person</th>
<th>cat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample data accuracy</td>
<td>87</td>
<td>45</td>
<td>19</td>
<td>87</td>
<td>76</td>
<td>72</td>
<td>69</td>
</tr>
<tr>
<td>VOC 07 with detection AP</td>
<td>74.2</td>
<td>36.5</td>
<td>34.4</td>
<td>64.8</td>
<td>33.4</td>
<td>58.7</td>
<td>67.6</td>
</tr>
</tbody>
</table>
Takeaway

• Works for image level classification
• Detection works without region proposal
• Class independent detection
• Detection is only as good as the classification
Questions?