VisDA Detection Challenge: Honorable Mention Talk

Siqi Yang, Arnold Wiliem and Brian C. Lovell
The University of Queensland, Australia
Task: Domain Adaptation for Object Detection

- Synthetic data (Train)
- Real data (Validation)
- Real data (Test)

Adaptation
Motivation

• Domain shift

**Image-level domain shift:**
Image scale, image style, context information for objects, etc.

**Instance-level domain shift:**
Object appearance, size, etc.

To minimize the domain discrepancy on both levels
Network Overview

Faster R-CNN

Image-level representations

RPN

Instance-level adaptation

Conv

Feature Level Domain Adaptation

GRL: Gradient Reverse Layer
Image Level Domain Adaptation

Faster R-CNN

Conv
RPN
Image-level representations

GRL
Conv
Image-level domain classifier
{Source, Target}

cls. reg.
Implementation Details

- ImageNet pre-trained VGG-16
- Our baseline model:
  - Caffe version of Faster R-CNN
- Training:
  - Training data: source training set + target test set
  - Single-scale training: [600]
  - Loss function: $L = L_{det} + \lambda(L_{img} + L_{ins})$, where $\lambda = 0.1$
- Testing:
  - Image pyramid inference: [150, 300]
  - Ensemble of source-only model and adapted model
## Evaluation on Test Set

<table>
<thead>
<tr>
<th>Model</th>
<th>aero-plane</th>
<th>bi-cycle</th>
<th>bus</th>
<th>car</th>
<th>horse</th>
<th>knife</th>
<th>motor-cycle</th>
<th>person</th>
<th>plant</th>
<th>skateboard</th>
<th>train</th>
<th>truck</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>3.1</td>
<td>17.2</td>
<td>15.5</td>
<td>29.6</td>
<td>17.5</td>
<td>0.7</td>
<td>22.2</td>
<td>3.3</td>
<td>14.0</td>
<td>5.6</td>
<td>2.1</td>
<td>2.4</td>
<td>11.1</td>
</tr>
<tr>
<td>Adapt</td>
<td>3.2</td>
<td>17.2</td>
<td>15.5</td>
<td>29.6</td>
<td>17.5</td>
<td>0.7</td>
<td>22.2</td>
<td>3.8</td>
<td>14.0</td>
<td>9.4</td>
<td>9.9</td>
<td>2.4</td>
<td>12.1</td>
</tr>
</tbody>
</table>
Thank you.