Diagnosing Error in Object Detectors

Derek Hoiem
Yodsawalai Chodpathumwan
Qieyun Dai
(presented by Yuduo Wu)

Most of the slides are from Derek Hoiem's ECCV 2012 presentation
Object detection is a collection of problems

**Intra-class Variation** for “Airplane”

<table>
<thead>
<tr>
<th>Occlusion</th>
<th>Shape</th>
<th>Viewpoint</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Occlusion Image" /></td>
<td><img src="image2" alt="Shape Image" /></td>
<td><img src="image3" alt="Viewpoint Image" /></td>
<td><img src="image4" alt="Distance Image" /></td>
</tr>
</tbody>
</table>
Object detection is a collection of problems

Confusing Distractors for “Airplane”

Background

Similar Categories

Dissimilar Categories

Localization Error
How to **evaluate** object detectors?

- Detector analysis tool is important
- **Average Precision (AP)**
  - Good summary statistic for quick comparison
  - Not a good driver of research

<table>
<thead>
<tr>
<th></th>
<th>aero</th>
<th>bike</th>
<th>bird</th>
<th>boat</th>
<th>bottle</th>
<th>bus</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) base</td>
<td>.290</td>
<td>.546</td>
<td>.006</td>
<td>.134</td>
<td>.262</td>
<td>.394</td>
</tr>
<tr>
<td>b) BB</td>
<td>.287</td>
<td>.551</td>
<td>.006</td>
<td>.145</td>
<td>.265</td>
<td>.397</td>
</tr>
<tr>
<td>c) context</td>
<td>.328</td>
<td>.568</td>
<td>.025</td>
<td>.168</td>
<td>.285</td>
<td>.397</td>
</tr>
</tbody>
</table>

Typical evaluation through comparison of AP numbers

- **Tools to evaluate detectors:**
  - where detectors fail and succeed
  - potential impact of particular improvements

figs from Felzenszwalb et al. 2010
Detectors Analyzed as Examples on VOC 2007

**Deformable Parts Model (DPM)**
- Sliding window
- Mixture of HOG templates with latent HOG parts

**Multiple Kernel Learning (MKL)**
- Jumping window
- Various spatial pyramid bag of words features combined with MKL

Felzenszwalb et al. 2010 (v4)
Vedaldi et al. 2009
Top false positives: Airplane (DPM)

- Background: 27%
- Localization: 29%
- Similar Objects: 33% (Bird, Boat, Car)
- Other Objects: 11%

AP = 0.36

Impact of Removing/Fixing FPs

- L
- S
- B

0 0.05 0.1 0.15
Top false positives: Dog (DPM)

AP = 0.03

- Background: 23%
- Similar Objects: 50% (Person, Cat, Horse)
- Localization: 17%
- Other Objects: 10%

Impact of Removing/Fixing FPs

1. False positives
2. Localization
3. Similar Objects
4. Background
5. Other Objects
6. Impact of Removing/Fixing FPs
7. Person
8. Cat
9. Horse
10. Dog
Top false positives: Dog (MKL)

- Similar Objects: 74% (Cow, Person, Sheep, Horse)
- Background: 4%
- Localization: 17%
- Other Objects: 5%

Impact of Removing/Fixing FPs

Top 5 False Positives

AP = 0.17
Summary of False Positive Analysis

**DPM v4** (FGMR 2010)

**MKL** (Vedaldi et al. 2009)

**Biggest Improvement:** Localization Error

**Small Improvement:** Background Dissimilar Objects

MKL more reasonable than the DPM detectors
Analysis of **object characteristics**

Additional annotations for seven categories: occlusion level, parts visible, sides visible

**Example of occlusion for aeroplane class**

- **Level of occlusion:** 2 (moderate)
- **Parts visible:** bike body, handlebars, wheel
- **Parts not visible:** seat
- **View:** side visible (front, top, etc., not visible)
- **Area:** 3233 pixels
- **Aspect Ratio (w/h):** 1.24
Normalized Average Precision

- Average precision is **sensitive** to number of positive examples

\[
\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}
\]

\[
\text{TruePositive} = \text{Recall} \times N_j
\]

- **Normalized** average precision:
  - replace variable \( N_j \) with **fixed** \( N \)
Object characteristics: Aeroplane
Object characteristics: Aeroplane

Occlusion: poor robustness to occlusion, but little impact on overall performance
Object characteristics: Aeroplane

Size: strong preference for average to above average sized airplanes
Aspect Ratio: 2-3x better at detecting wide (side) views than tall views

Object characteristics:
- Aeroplane
- Tall
- X-Tall
- Medium
- Wide
- X-Wide

Easier (Wide) → Harder (Tall)
**Object characteristics: Aeroplane**

**Sides/Parts**: best performance = direct side view with all parts visible

---

![Graph showing performance metrics for different sides and parts of an aeroplane](image-url)
Summarizing Detector Performance

DPM (v4): Sensitivity and Impact

Avg. Performance of Best Case

Avg. Overall Performance

Avg. Performance of Worst Case

- occ: 0.103 ± 0.03
- trn: 0.323 ± 0.02
- size: 0.445 ± 0.05
- asp: 0.435 ± 0.04
- view: 0.375 ± 0.03
- part: 0.362 ± 0.02
Summarizing Detector Performance

Best, Average, Worst Case

DPM (FGMR 2010)
MKL (Vedaldi et al. 2009)

Impact
Sensitivity

occlusion  trunc  size  aspect  view  part_vis
Summarizing Detector Performance

Best, Average, Worst Case

Occlusion: high sensitivity, low potential impact

DPM (FGMR 2010)
MKL (Vedaldi et al. 2009)
Summarizing Detector Performance

Best, Average, Worst Case

DPM (FGMR 2010)
MKL (Vedaldi et al. 2009)

MKL more sensitive to size
Conclusions

• **Most errors are reasonable**
  – Localization error and confusion with similar objects
  – Misdetection of occluded or small objects

• **Large improvements in specific areas** (e.g., remove all background FPs or robustness to occlusion) has small impact in overall AP
  – More specific analysis should be standard

• **Our code and annotations** are available online
  – Automatic generation of analysis summary from standard annotations
More Information:
www.cs.illinois.edu/homes/dhoiem/publications/detectionAnalysis_eccv12.tar.gz
or:
http://videolectures.net/eccv2012_hoiem_detectors/?q=eccv%202012

Thank you!