Ensemble of Exemplar SVMs for Object Detection and Beyond

Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros

http://www.cs.cmu.edu/~tmalisie/projects/iccv11/
Exemplar SVM for Object Detection

Overview

- Identifies object rather than object class
- One positive training example
- One classifier for each exemplar
- More “what it is not” than “what it is similar to”
- Can adapt features to exemplar
Exemplar SVM for Object Detection

Training

● Represent exemplar $E$ using HOG template $x_E$
● Each exemplar defines HOG dimensions based on bounding box aspect ratio
  ○ 8x8 cells
  ○ ~100 cells per exemplar
● Negative samples $N_E$ taken from images with no category instances
● Attempt to separate $x_E$ from all $N_E$ by largest possible margin
Exemplar SVM for Object Detection

Training

- Training linear SVM equivalent to learning weight vector $w_E$ on $x_E$
- Solve convex optimization problem

$$\Omega_E(w, b) = ||w||^2 + C_1 h(w^T x_E + b) + C_2 \sum_{x \in \mathcal{N}_E} h(-w^T x - b)$$
Exemplar SVM for Object Detection

Hard Negative Mining

● Accuracy of Exemplar SVM relies on large (millions) number of negative examples
● Computing an SVM on so much data is costly
Exemplar SVM for Object Detection

Hard Negative Mining

- SVM is driven by small number of examples near decision boundary
Exemplar SVM for Object Detection

Hard Negative Mining

- Finding set of “hard” negatives is important for efficiency
Exemplar SVM for Object Detection

Hard Negative Mining

- Train SVM on subset of negatives
- Find objects where current model performs poorly
  - Near boundary
  - Misclassified
- Train new SVM on these “hard” examples
- Iterate
Exemplar SVM for Object Detection

Calibration

- Exemplar SVMs are independently trained
- Outputs from multiple SVMs will not be commensurate
- SVMs must be calibrated
Exemplar SVM for Object Detection

Calibration

- Calibration done by fitting sigmoid function to outputs from validation set (Platt, 1999)
- Calibrated SVM output is a probability
Exemplar SVM for Object Detection

Calibration

- Calibration usually done on validation set with positive and negative examples
- Don’t know which members of validation set should be positive instances a priori
  - Can’t just use same-class positives
  - Have to be visually similar
- Each exemplar must find its own positives
Exemplar SVM for Object Detection

Calibration

- Use Exemplar-SVM detector on members of validation set
- Compute overlap score between detections and ground truths
  - High ( > 0.5) overlap positive
  - Low ( < 0.2) overlap negative
Exemplar SVM for Object Detection

Calibration

- Calibration acts as rescaling and shifting of decision boundary
Exemplar SVM for Object Detection

Object Detection

- Each SVM applied to image via sliding windows
- Use exemplar co-occurrence instead of non-maxima suppression
- Overlapping SVM scores pooled to form context feature
- Final score is a weighted sum of SVM score and context score
# Exemplar SVM for Object Detection

## Object Detection - Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional NN + Calibration</td>
<td>0.110</td>
</tr>
<tr>
<td>Local Distance Function + Calibration</td>
<td>0.157</td>
</tr>
<tr>
<td><strong>Exemplar-SVMs + Calibration</strong></td>
<td><strong>0.198</strong></td>
</tr>
<tr>
<td><strong>Exemplar-SVMs + Co-occurrence</strong></td>
<td><strong>0.227</strong></td>
</tr>
<tr>
<td>One SVM per category (Dalal and Triggs 2005)</td>
<td>0.097</td>
</tr>
<tr>
<td>Deformable Part Model (Felzenszwalb et al 2010)</td>
<td>0.266</td>
</tr>
</tbody>
</table>
Exemplar SVM for Object Detection

Object Detection - Results

-
Exemplar SVM for Object Detection

Object Detection - Results

<table>
<thead>
<tr>
<th>Exemplar</th>
<th>w</th>
<th>Averaged Detections</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Exemplar" /></td>
<td><img src="image2" alt="w" /></td>
<td><img src="image3" alt="Averaged Detections" /></td>
</tr>
<tr>
<td><img src="image4" alt="Exemplar" /></td>
<td><img src="image5" alt="w" /></td>
<td><img src="image6" alt="Averaged Detections" /></td>
</tr>
<tr>
<td><img src="image7" alt="Exemplar" /></td>
<td><img src="image8" alt="w" /></td>
<td><img src="image9" alt="Averaged Detections" /></td>
</tr>
<tr>
<td><img src="image10" alt="Exemplar" /></td>
<td><img src="image11" alt="w" /></td>
<td><img src="image12" alt="Averaged Detections" /></td>
</tr>
</tbody>
</table>

Average of first 20 detections

Average of first 10 detections
Exemplar SVM for Object Detection

Object Detection - Results

●
Exemplar SVM for Object Detection

Object Detection - Results
Exemplar SVM for Object Detection

Object Detection - Results
Exemplar SVM for Object Detection

Object Detection - Results
Exemplar SVM for Object Detection

Segmentation & Geometry Estimation - Results
Exemplar SVM for Object Detection

3D Model Transfer - Results
Exemplar SVM for Object Detection

Related Object Priming - Results
Exemplar SVM for Object Detection

Summary

- **Pros**
  - Good detection rate
  - Highly parallelizable
  - Facilitates metadata transfer
  - No retraining when new exemplars are added

- **Cons**
  - Must store classifier for every exemplar
  - Expensive to apply all classifiers to query image
  - Can’t handle objects/viewpoints not seen before
Summary

- We’ve seen how categories can be used to label objects
- We’ve seen how objects can be labeled without using category-level detection
- Does output have to be a category?
Data-driven Visual Similarity for Cross-domain Image Retrieval

Abhinav Shrivastava, Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros

http://graphics.cs.cmu.edu/projects/crossDomainMatching/
Exemplar SVM for Object Detection

Again?

- A key point in Exemplar-SVM training was no in-class negatives
- What if they are allowed?
Exemplar SVM for Object Detection

Additional Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional NN + Calibration</td>
<td>0.110</td>
</tr>
<tr>
<td>Local Distance Function + Calibration</td>
<td>0.157</td>
</tr>
<tr>
<td>Exemplar-SVMs + Calibration</td>
<td>0.198</td>
</tr>
<tr>
<td>Exemplar-SVMs + Co-occurrence</td>
<td>0.227</td>
</tr>
<tr>
<td><em><em>Exemplar-SVMs</em> + Calibration</em>*</td>
<td><strong>0.142</strong></td>
</tr>
<tr>
<td><em><em>Exemplar-SVMs</em> + Co-occurrence</em>*</td>
<td><strong>0.197</strong></td>
</tr>
</tbody>
</table>
Exemplar SVM

In-class negatives

- In-class negatives do not completely destroy the effectiveness
- Allowing in-class negatives allows for more freedom in training and application
Exemplar SVM for Image Retrieval

- Many image retrieval systems are heavily influenced by
  - Time of day image was taken
  - Season image was taken
- These algorithms don’t always focus on what is unique in an image
Exemplar SVM for Image Retrieval
Exemplar SVM for Image Retrieval
Exemplar SVM for Image Retrieval
Exemplar SVM for Image Retrieval

- Learn “important parts” using an Exemplar-SVM
  - Query image is positive sample
  - Random images are negatives
- Helps to identify what in the image is unique vs the rest of the world
Exemplar SVM for Image Retrieval
Exemplar SVM for Image Retrieval

Uniqueness
Exemplar SVM for Image Retrieval

Results
Exemplar SVM for Image Retrieval
Cross-domain Matching
Exemplar SVM for Image Retrieval

Cross-domain Matching
Exemplar SVM for Image Retrieval

Cross-domain Matching
Exemplar SVM for Image Retrieval
Cross-domain Matching
Exemplar SVM for Image Retrieval

Cross-domain Matching
Summary

- Categories are useful because they
  - help us organize the world
  - encode a lot of useful information
- But they aren’t always necessary or appropriate for the task at hand