Person Detection

• Dalal and Triggs ‘05
  – Learn to classify pedestrians vs. background
  – HOG + linear SVM
  – Doesn’t account for variations in body pose and viewpoint
Poselets

Poselets capture part of the pose from a given viewpoint

[Bourdev & Malik, ICCV09]
Poselets

But how are we going to create training examples of poselets? [Bourdev & Malik, ICCV09]
How do we train a poselet for a given pose configuration?

[Bourdev & Malik, ICCV09]
Finding correspondences at training time

Given part of a human pose

How do we find a similar pose configuration in the training set? [Bourdev & Malik, ICCV09]
Finding correspondences at training time

We use keypoints to annotate the joints, eyes, nose, etc. of people

[Bourdev & Malik, ICCV09]
Finding correspondences at training time

Residual Error

[Bourdev & Malik, ICCV09]
Training poselet classifiers

Residual Error: 0.15  0.20  0.10  0.85  0.15  0.35

1. Given a seed patch
2. Find the closest patch for every other person
3. Sort them by residual error
4. Threshold them

[Bourdev & Malik, ICCV09]
Training poselet classifiers

1. Given a seed patch
2. Find the closest patch for every other person
3. Sort them by residual error
4. Threshold them
5. Use them as positive training examples to train a linear SVM with HOG features

[Bourdev & Malik, ICCV09]
Which poselets should we train?

• Choose thousands of random windows, generate poselet candidates, train linear SVMs

• Select a small set of poselets that are:
  – Individually effective
  – Complementary

[Bourdev & Malik, ICCV09]
Person Detection Using Poselets

<table>
<thead>
<tr>
<th></th>
<th>100 poselets</th>
<th>200 poselets</th>
<th>[12]</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOC 2007</td>
<td>45.6%</td>
<td>46.9%</td>
<td>36.8%</td>
</tr>
<tr>
<td>VOC 2008</td>
<td>54.1%</td>
<td>52.6%</td>
<td>43.1%</td>
</tr>
</tbody>
</table>

Other Uses of Poselets

• **Object Segmentation (CVPR 2011)**
  – Predict area using “soft masks”
  – Deformation to match image edges
  – Extends poselets to other objects

• **Activity Recognition (CVPR 2011)**
  – Recognition from a single image
  – Use which poselets fired and to what extent to predict the activity
Describing People: A Poselet-Based Approach to Attribute Classification

L. Bourdev, S. Maji, and J. Malik

ICCV 2011
Random 1¼% of the test set

[Bourdev et al., ICCV11]
How poselets help in high-level vision

[Image: Various images of people in different poses and appearances, illustrating the concept of poselets for high-level vision tasks.]

[Bourdev et al., ICCV11]
Gender recognition with poselets
Attribute Classification Overview

Given a test image

Poselet Activations

[Bourdev et al., ICCV11]
Features

- Pyramid HOG
- LAB histogram
- Skin features
  - Hands-skin
  - Legs-skin

[Image: Features diagram with examples of Poselet patch, Skin mask, and Arms mask.]

[Bourdev et al., ICCV11]
Poselet Level Classification

- Linear SVM + Sigmoid
- Some attributes are associated with a particular body part

[Bourdev et al., ICCV11]
Person-level Classification

- Another Linear SVM + Sigmoid

[Bourdev et al., ICCV11]
Context-Level Classification

- **Context-level Classifiers**
  - IS MALE?
  - LONG PANTS?
  - HAS HAT?

- **Person-level Classifiers**
  - IS MALE?
  - LONG PANTS?
  - HAS HAT?

- **Poselet-level Classifiers**
  - IS MALE?
  - LONG PANTS?
  - HAS HAT?

- **Features**
  - Poselet Activations

[Bourdev et al., ICCV11]
Results
Random 1¼% of the test set

[Bourdev et al., ICCV11]
Precision/recall on our test set

![Graphs showing precision/recall for different labels](image)

<table>
<thead>
<tr>
<th>Label</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>is_male</td>
<td></td>
</tr>
<tr>
<td>has_long_hair</td>
<td></td>
</tr>
<tr>
<td>has_glasses</td>
<td></td>
</tr>
<tr>
<td>has_hat</td>
<td></td>
</tr>
<tr>
<td>has_long_sleeves</td>
<td></td>
</tr>
<tr>
<td>has_t-shirt</td>
<td></td>
</tr>
<tr>
<td>has_long_pants</td>
<td></td>
</tr>
<tr>
<td>has_jeans</td>
<td></td>
</tr>
<tr>
<td>has_shorts</td>
<td></td>
</tr>
</tbody>
</table>

SPM Poselets-No-context Poselets-Full

[Bourdev et al., ICCV11]
Gender Recognition vs. Cognitec

- Cognitec is one of the leading face recognition companies (according to latest NIST test)

<table>
<thead>
<tr>
<th></th>
<th>Cognitec</th>
<th>Poselets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>Frontal faces only</td>
<td>Any view</td>
</tr>
<tr>
<td>Features</td>
<td>Proprietary biometric</td>
<td>Standard HOG</td>
</tr>
<tr>
<td>Min resolution</td>
<td>60 pixel face width</td>
<td>Any resolution</td>
</tr>
<tr>
<td>Performance</td>
<td>75.0% AP</td>
<td>82.4% AP</td>
</tr>
</tbody>
</table>

- Upper bound for any gender recognizer based on frontal faces: Max AP=80.5% vs. 82.4%

[Bourdev et al., ICCV11]
Highest-scoring “is male”

[Image: Ten pictures of men and women, each with a caption indicating their gender]

Lowest-scoring “is male”  

[Bourdev et al., ICCV11]
Highest-scoring “has long hair”

Lowest-scoring “has long hair”

[Bourdev et al., ICCV11]
Highest-scoring “wears a hat”

[Image]

Lowest-scoring “wears a hat”  
[Bourdev et al., ICCV11]
Highest-scoring “wears glasses”

Lowest-scoring “wears glasses”

[Bourdev et al., ICCV11]
Highest-scoring “long sleeves”

Lowest-scoring “long sleeves”

[Bourdev et al., ICCV11]
most confused "is-male"

most confused "has-shorts"

most confused "has-t-shirt"

[Bourdev et al., ICCV11]
“100 Special Moments” by Jason Salavon
Highly-Weighted Poselets

is-male

has-long-hair

has-glasses
Which poselets are discriminative for gender?

Preferred by human subjects

Preferred by our system

[Bourdev et al., ICCV11]
Describing people

[Bourdev et al., ICCV11]
Poselets website

http://eecs.berkeley.edu/~lbourdev/poselets

- The set of published poselet papers
- H3D data set + Matlab tools
- Java3D annotation tool + video tutorial
- Matlab code to detect people using poselets
- Latest trained poselets

[Bourdev et al., ICCV11]