Relative Attributes
Devi Parikh and Kristen Grauman
ICCV 2011

Experiment Presentation
by Aishwarya Padmakumar
Motivation

Sometimes it is hard to make a binary decision on whether an image satisfies an attribute. Comparisons are easier.

Image Source: https://filebox.ece.vt.edu/~parikh/relative.html
Sources

Code and data for following experiments obtained from: https://filebox.ece.vt.edu/~parikh/relative.html#code
→ Public Figures Face Database – Images of famous public figures
→ Images of the same person are grouped into categories
→ Attributes – Male, White, Young, Smiling, Chubby, VisibleForehead, BushyEyebrows, NarrowEyes, PointyNose, BigLips, RoundFace
## Ground Truth - Ordering of categories for each attribute

<table>
<thead>
<tr>
<th>Class1</th>
<th>Attr1</th>
<th>Attr2</th>
<th>Attr3</th>
<th>Attr4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Class2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Class3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Evaluating Ranking

→ Obtain training pairs using category ranking
→ Train SVM to get a score for each image
→ Average scores of images in category to get score of category
→ Rank categories according to this score
→ Compare to ground truth ranking
Ranking Prediction

Attributes

Classes

Rank difference

0
1
2
3
4
5
6
7
Varying the amount of training data
<table>
<thead>
<tr>
<th>Classes</th>
<th>Attributes</th>
<th>Rank difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>
With 80% of original training data
With 60% of original training data
With 40% of original training data
What sort of mistakes does it make?
Failure cases - Who has bigger lips?
Failure cases - Who has bigger lips?

Supposed to be the person on the right but certain expressions don’t indicate this.
Failure cases - Who looks younger?
Failure cases - Who looks younger?

Supposed to be the girl on the right. Other images are more obvious but expression, make-up and lighting affect this.
Instance Level Supervision vs Category Level Supervision
WhittleSearch Dataset

- Collected by: Adriana Kovashka, Devi Parikh, and Kristen Grauman
- Dataset of shoes: 14658 images
- Provide extracted features (GIST and colour - LAB histogram), accounting for image orientation
→ Category level ordering and instance level comparisons for each attribute
→ Attributes: pointy-at-the-front, open, bright-in-color, covered-with-ornaments, shiny, high-at-the-heel, long-on-the-leg, formal, sporty, feminine
→ Instance level ordering obtained via MTurk
Instance level supervision vs category level supervision

![Bar chart showing accuracy comparison between instance and category level supervision for various attributes.]
Why does training on category level information do well sometimes?

- List of categories - athletic_shoes, boots, clogs, flats, high_heels, pumps, rain_boots, sneakers, stiletto, wedding_shoes
- Pointy, ornaments, long on the leg seem to be attributes that we can associate with some of these classes
- Possibly category level labels result in more well-separated data, making it easier to train an SVM (SVMs are sensitive to noise near the support vectors)
Some cases where the prediction differs

More Sporty: Ground truth - left; Instance based - left; Category based - right
Some cases where the prediction differs

More Sporty: Ground truth - left; Instance based - right; Category based - left
Some cases where the prediction differs

More Sporty: Ground truth - equal; Instance based - right; Category based - left
Usefulness of different types of features
Comparing GIST and colour features - instance comparisons
Comparing GIST and colour features - category comparisons
Where GIST and the combination differ

More Bright: Ground truth - right; GIST - equal; Both - right
Where GIST and the combination differ

More Bright: Ground truth - left; GIST - equal; Both - right
Where colour and the combination differ

More Bright: Ground truth - right; Colour - equal; Both - right
Where colour and the combination differ

More Bright: Ground truth - equal; Colour - equal; Both - left
Relative Descriptions
Relative Descriptions
Relative Descriptions - Pointy at the front
Relative Descriptions - open
Relative Descriptions - bright in colour
Relative Descriptions - covered in ornaments
Relative Descriptions - shiny
Relative Descriptions - high at the heel
Relative Descriptions - long on the leg
Relative Descriptions - formal
Relative Descriptions - sporty
Relative Descriptions - feminine
Varying the number of images in between
Leave 1/16 of the images in between (formal)
Leave 1/8 of the images in between (formal)
Leave 1/4 of the images in between (formal)
Leave 1/2 of the images in between (formal)
Thank You!