Attributes

Phuong Pham

January 29, 15
Describing Objects by Their Attributes

Based on Derek Hoiem’s slide
Describe objects at a more abstract level

“Gradient changes from bottom to top. Especially, many corners at the top left region.”
vs.
“Large, angry animal with pointy teeth”
Why Infer Properties (if we can)

1. We want detailed information about objects
2. We want to be able to infer something about unfamiliar objects (name is *unknown*)
3. We want to make comparisons between objects or categories

Familiar Objects

<table>
<thead>
<tr>
<th>Has Stripes</th>
<th>Has Ears</th>
<th>Has Eyes</th>
<th>Has Four Legs</th>
<th>Has Mane</th>
<th>Has Tail</th>
<th>Has Snout</th>
<th>Brown</th>
<th>Muscular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

New Object

<table>
<thead>
<tr>
<th>Has Stripes (like cat)</th>
<th>Has Mane and Tail (like horse)</th>
<th>Has Snout (like horse and dog)</th>
</tr>
</thead>
</table>
Attributes (64)

- **Visible parts**: “has wheels”, “has snout”, “has eyes”
- **Visible materials or material properties**: “made of metal”, “shiny”, “clear”, “made of plastic”
- **Shape**: “3D boxy”, “round”

**Shape**: Horizontal Cylinder
**Part**: Wing, Propeller, Window, Wheel
**Material**: Metal, Glass
Attribute dataset

• **a-Pascal**: 20 categories from PASCAL 2008 trainval dataset (10K object images)

• **a-Yahoo**: 12 new categories from Yahoo image search
Attribute labels are somewhat ambiguous

- Agreement among “experts” (authors) 84.3
- Between experts and Turk labelers 81.4
- Among Turk labelers 84.1
Base features

• Spatial pyramid histograms of quantized
  • Color and texture for materials
  • Histograms of gradients (HOG) for parts
  • Canny edges for shape
Discriminative Attributes

a random split
Learning (semantic) attributes

Simplest approach: Train classifier using all features for each attribute independently.

Base features

“Has Wheels”

“No Wheels Visible”
Correlated attributes

Big Problem: Many attributes are strongly correlated through the object category

Most things that “have wheels” are “made of metal”

When we try to learn “has wheels”, we may accidentally learn “made of metal”

Has Wheels, Made of Metal?
Feature selection

• Select features that can distinguish between two classes
  • Things that have the attribute (e.g., wheels)
  • Things that do not, but have similar attributes to those that do

“Has Wheels” vs. “No Wheels”

L1 logistic regression

- Car Wheel Features
- Boat Wheel Features
- Plane Wheel Features
- All Wheel Features
Assigning attributes

Train and Test on Same Classes from PASCAL (*)
Assigning attributes
Train on 20 PASCAL classes + Test on 12 different Yahoo classes
Naming familiar objects

• Use attribute predictions as features

• Train linear SVM to categorize objects

<table>
<thead>
<tr>
<th>PASCAL 2008</th>
<th>Base Features</th>
<th>Semantic Attributes</th>
<th>All Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification Accuracy</td>
<td>58.5%</td>
<td>54.6%</td>
<td>59.4%</td>
</tr>
<tr>
<td>Class-normalized Accuracy</td>
<td>35.5%</td>
<td>28.4%</td>
<td>37.7%</td>
</tr>
</tbody>
</table>

Attributes not big help when sufficient data
Unusual attributes

752 reports
68% are correct
Learning to identify new objects

Textual only = 100 examples in semantic attribute space
= 8 examples in base features space
= 3 examples in semantic & discriminative attribute space
Take away points

• Semantic attributes are helpful for certain tasks:
  • Learning from textual description only
  • Describe unknown objects
  • Identify unusual attributes

• And it is *feasible* to learn semantic attributes from base features and sufficient datasets.

• Its shortcomings?
Relative attributes

(> absolute attributes?)

Based on Devi Parikh’s slides
Using attributes describe a MULE
Comparison is helpful

Has tail

Is furry

Has four-legs

Legs shorter than horses’

Tail longer than donkeys’
Absolute comparison (smiling)
Relative comparison (smiling)
Relative attributes

- Enhanced human-machine communication
- More informative
- Natural for humans

Contributions

- Relative attributes
  - Allow relating images and categories to each other
  - Learn ranking function for each attribute
- Novel applications
  - Zero-shot learning from attribute comparisons
  - Automatically generating relative image descriptions
Learning Relative Attributes

For each attribute $a_m$, open supervision is

$$O_m: \{ ( \sim, \ldots ) \},$$

$$S_m: \{ ( \sim, \ldots ) \},$$
Learning Relative Attributes

Learn a scoring function $r_m(x_i) = w_m^T x_i$

that best satisfies constraints:

$\forall (i, j) \in O_m : w_m^T x_i > w_m^T x_j$

$\forall (i, j) \in S_m : w_m^T x_i = w_m^T x_j$
Max-margin learning to rank formulation

\[
\begin{align*}
\min & \quad \left( \frac{1}{2} \| \mathbf{w}_m^T \|_2^2 + C \left( \sum \xi_{ij}^2 + \sum \gamma_{ij}^2 \right) \right) \\
\text{s.t} & \quad \mathbf{w}_m^T (\mathbf{x}_i - \mathbf{x}_j) \geq 1 - \xi_{ij}, \forall (i, j) \in O_m \\
& \quad |\mathbf{w}_m^T (\mathbf{x}_i - \mathbf{x}_j)| \leq \gamma_{ij}, \forall (i, j) \in S_m \\
& \quad \xi_{ij} \geq 0; \gamma_{ij} \geq 0
\end{align*}
\]

Based on [Joachims 2002]

Image → Relative Attribute Score
Features

• Gist descriptor
• Lab color histogram
Datasets

Outdoor Scene Recognition (OSR)  
[Oliva 2001]

- 8 classes, ~2700 images, Gist
- 6 attributes: open, natural, etc.

Public Figures Face (PubFig)  
[Kumar 2009]

- 8 classes, ~800 images, Gist+color
- 11 attributes: white, chubby, etc.
Relative zero-shot learning ($\mathbb{R}^n \rightarrow \mathbb{R}^m$)

Can predict new classes based on their relationships to existing classes – without training images

$$c_i^{(s)} \sim \mathcal{N}(\mu_i^{(s)}, \Sigma_i^{(s)}) \, (S, J, H)$$

$$\mu_{jm}^{(u)} = \frac{1}{2}(\mu_{im}^{(s)} + \mu_{km}^{(s)})$$

$$\mu_{jm}^{(u)} = \mu_{im}^{(s)} + d_m$$

Age: Hugh $\succ$ Clive $\succ$ Scarlett $\succ$ Jared $\succ$ Miley

Smiling: Miley $\succ$ Jared
An attribute is more discriminative when used relatively

**DAP:** binary attr

**SRA:** rel attr (classifier score)

**Proposed:** rel attr (ranked score)

An attribute is more discriminative when used relatively
Describe image results

Conventional binary description: *not dense*

**Dense:**

**Not dense:**

* more dense than

* less dense than
Human Studies: Which Image is Being Described?

Binary (existing):
Not natural
Not open
Has perspective

Relative (ours):
More natural than insidecity
Less natural than highway
More open than street
Less open than coast
Has more perspective than highway
Has less perspective than insidecity

18 subjects
Test cases:
10 OSR, 20 PubFig
Take away points

Relative attributes learnt as ranking functions
  • Natural and accurate zero-shot learning of novel concepts by relating them to existing concepts
  • Precise image descriptions for human interpretation
  • The approach has been explored by many research projects after that.
Questions