Relative Attributes

Devi Parikh (TTIC) and Kristen Grauman (UT Austin)

1. Main Idea

Motivation:

Categorical (binary) attributes are restrictive and can be unnatural

Natural ? Not Natural

Smiling ? Not Smiling

Proposed idea: Relative Attributes

- ✧ Richer communication between humans and machines
- ✧ Describe images or categories relatively e.g. "dogs are furrier than giraffes" "find less congested downtown Chicago scene than  "
- ✧ Learn a ranking function for each attribute

Enables new applications

- ✧ Novel zero-shot learning from attribute comparisons
- ✧ Precise automatically generated textual descriptions of images

2. Learning Relative Attributes

For each attribute a_m, Supervision is O_m: $\{(\text{image} \succ \text{image}), \ldots\}$, S_m: $\{(\text{image} \sim \text{image}), \ldots\}$

Learn a scoring function $r_m(\boldsymbol{x_i}) = \boldsymbol{w_m^T x_i}$ that best satisfies constraints:

$$\forall (i,j) \in O_m : \boldsymbol{w_m^T x_i} > \boldsymbol{w_m^T x_j} \qquad \forall (i,j) \in S_m : \boldsymbol{w_m^T x_i} = \boldsymbol{w_m^T x_j}$$

Max-margin learning to rank formulation

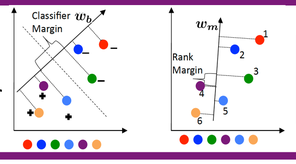
$$\min \quad \left(\frac{1}{2}||\boldsymbol{w_m^T}||_2^2 + C\left(\sum \xi_{ij}^2 + \sum \gamma_{ij}^2\right)\right)$$

Based on [Joachims, 2002]

$$\text{s.t} \quad \boldsymbol{w_m^T}(\boldsymbol{x_i} - \boldsymbol{x_j}) \geq 1 - \xi_{ij}, \forall (i,j) \in O_m \qquad \xi_{ij} \geq 0; \gamma_{ij} \geq 0$$
$$|\boldsymbol{w_m^T}(\boldsymbol{x_i} - \boldsymbol{x_j})| \leq \gamma_{ij}, \forall (i,j) \in S_m$$

3. Ranking Function vs. Binary Classifier Score

How do learned ranking functions differ from classifier outputs?



% correctly ordered pairs	Classifier	Ranker
Outdoor scenes	80%	89%
Celebrity faces	67%	82%

4. Relative Zero-shot Learning

Learnt relative attributes Age:  Smiling: 

- ✧ Training: Images from S seen categories and descriptions of U unseen categories
- ✧ Testing: Categorize image into one of N (=S+U) categories (max-likelihood)

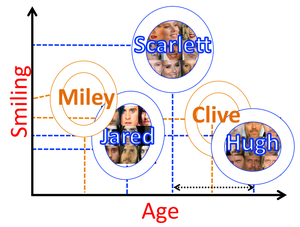
Unseen categories

Age: Hugh ≻ Clive ≻ Scarlett

Jared ≻ Miley

Smiling: Miley ≻ Jared

Relative attributes space



Teaching novel concepts simply by relating them to known concepts!

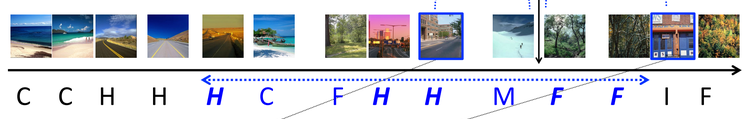
- ✧ Need not use all attributes
- ✧ Need not relate to all S seen categories

5. Describing Images Relatively

Learnt relative attributes Density:  ...

Auto - generate textual description of: 

Relative attributes space



Relative description:

"more dense than  , less dense than  "

"more dense than ***Highways***, less dense than ***Forests***"

Conventional binary description:

"not dense" Not dense:  Dense: 

6. Datasets

Outdoor Scene Recognition (OSR): 2688 images, 8 categories: coast (C), forest (F), highway (H), inside-city (I), mountain (M), open-country (O), street (S) and tall-building (T), gist features

Public Figure Face (PubFig): 772 images, 8 categories: Alex Rodriguez (A), Clive Owen (C), Hugh Laurie (H), Jared Leto (J), Miley Cyrus (M), Scarlett Johansson (S), Viggo Mortensen (V) and Zac Efron (Z), gist and color features

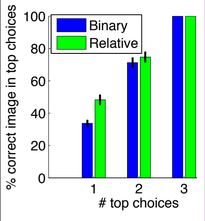
	Binary	Relative
OSR	T I S H C O M F	
natural	0 0 0 0 1 1 1 1	T≺I∼S≺H≺C∼O∼M≺F
open	0 0 0 1 1 1 1 0	T∼F≺I∼S≺M≺H∼C∼O
perspective	1 1 1 1 0 0 0 0	O≺C≺M∼F≺H≺I≺S≺T
large-objects	1 1 1 0 0 0 0 0	F≺O∼M≺I∼S≺H∼C≺T
diagonal-plane	1 1 1 1 0 0 0 0	F≺O∼M≺C≺I∼S≺H≺T
close-depth	1 1 1 1 0 0 0 1	C≺M≺O≺T∼I∼S∼H∼F
PubFig	A C H J M S V Z	
Masculine-looking	1 1 1 1 0 1 1 1	S≺M≺Z≺V≺J≺A≺H≺C
White	0 1 1 1 1 1 1 1	A≺C≺H≺Z≺J≺S≺M≺V
Young	0 0 0 0 1 1 0 1	V≺H≺C≺J≺A≺S≺Z≺M
Smiling	1 1 1 0 1 1 0 1	J≺V≺H≺A∼C≺S∼Z≺M
Chubby	1 0 0 0 0 0 0 0	V≺J≺H≺C≺Z≺M≺S≺A
Visible-forehead	1 1 1 0 1 1 1 0	J≺Z≺M≺S≺A∼C∼H∼V
Bushy-eyebrows	0 1 0 1 0 0 0 0	M≺S≺Z≺V≺H≺A≺C≺J
Narrow-eyes	0 1 1 0 0 0 1 1	M≺J≺S≺A≺H≺C≺V≺Z
Pointy-nose	0 0 1 0 0 0 0 1	A≺C≺J∼M∼V≺S≺Z≺H
Big-lips	1 0 0 0 1 1 0 0	H≺J≺V≺Z≺C≺M≺A≺S
Round-face	1 0 0 0 1 1 0 0	H≺V≺J≺C≺Z≺A≺S≺M

7. Image Description Results

Example descriptions

	Binary descriptions	Relative descriptions
	not natural, not open, perspective	more natural than tallbuilding; less natural than forest; more open than tallbuilding; less open than coast; more perspective than tallbuilding;
	not natural, not open, perspective	more natural than insidecity; less natural than highway; more open than street; less open than coast; more perspective than highway; less perspective than insidecity
	natural, open, perspective	more natural than tallbuilding; less natural than mountain; more open than mountain; less perspective than opencountry;
	White, not Smiling, VisibleForehead	more White than AlexRodriguez; more Smiling than JaredLeto; less Smiling than ZacEfron; more VisibleForehead than JaredLeto; less VisibleForehead than MileyCyrus
	White, not Smiling, not VisibleForehead	more White than AlexRodriguez; less White than MileyCyrus; less Smiling than HughLaurie; more VisibleForehead than ZacEfron; less VisibleForehead than MileyCyrus
	not Young, BushyEyebrows, RoundFace	more Young than CliveOwen; less Young than ScarlettJohansson; more BushyEyebrows than ZacEfron; less BushyEyebrows than AlexRodriguez; more RoundFace than CliveOwen; less RoundFace than ZacEfron

Human study



18 subjects
30 test cases

Which image is...?

Binary: Smiling, not VisibleForehead

Smiling VisibleForehead

Not Smiling Not VisibleForehead

Relative:

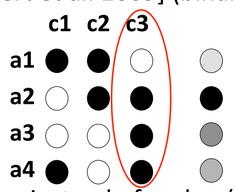
More Smiling than More VisibleForehead than

Less Smiling than Less VisibleForehead than

8. Zero-shot Learning Results

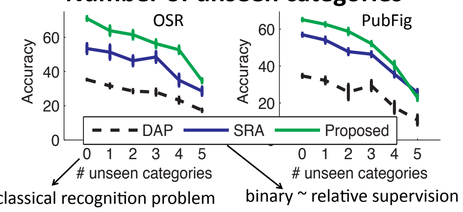
Baselines:

- ✧ Direct Attribute Prediction (DAP) [Lampert et al. 2009] (binary)

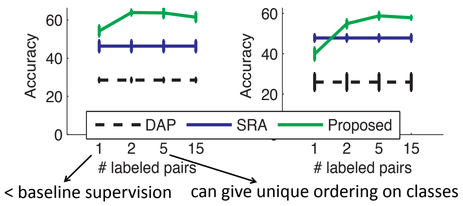


- ✧ Classifier instead of ranker (SRA)

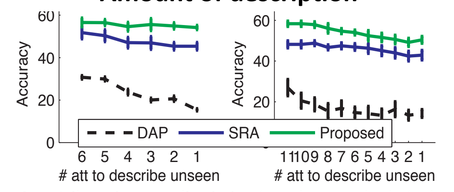
Number of unseen categories



Amt. of labeled data to learn attributes

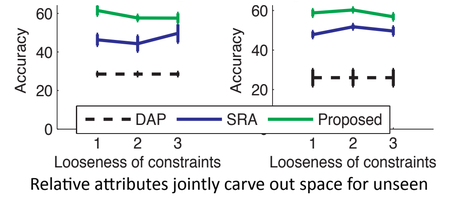


Amount of description



An attribute is more discriminative when used relatively

Quality of description



Relative attributes jointly carve out space for unseen category

9. Contributions

- ✧ Relative attributes
- ✧ Richer semantic communication between humans and machines
- ✧ Novel applications: zero-shot learning via comparisons and automatic relative image description

Data available online!

http://ttic.uchicago.edu/~dparikh/relative.html