Attribute Dominance: What Pops Out?

Naman Turakhia (Georgia Tech) Devi Parikh (Virginia Tech)

VirginiaTech Invent the Future®

Intuition

furry, white | smiling, teeth visible | white, male

sharp teeth, scary | wearing lipstick, heavy makeup | bearded, wearing glasses

- Certain attributes pop out more than others: strong presence, unusualness, absence of other more dominant attributes, etc.
- Humans tends to name the most dominant attributes first
- Order of naming attributes reveals information about the image
- Modeling attribute dominance can improve performance on human-centric applications: zero-shot learning, image search, description

Approach: Modeling Dominance

Model interplay between attributes

$$\hat{d}_t^m = \boldsymbol{w}_m^T \phi(\boldsymbol{x}_t)$$

Predicted dominance score of attribute a_m on image $\boldsymbol{x}_t$ | Parameters learnt via a linear regressor | Image features: attribute scores

$$pd_k^m(\boldsymbol{x}_t) = \frac{s_k^m(\boldsymbol{x}_t)}{\sum_{k=1}^{2M} s_k^m(\boldsymbol{x}_t)} \qquad s_k^m(\boldsymbol{x}_t) = \frac{1}{\log\left(|r^m(\boldsymbol{x}_t) - k| + 1\right) + 1}$$

Probability that attribute is k^{th} most dominant | Rank of attribute according to predicted dominance score

Approach: Zero-shot Learning

Appearance-based [Lampert 2009]

$$pa_{n'}(\boldsymbol{x}) \propto \prod_{m=1}^{M} pa^m(\boldsymbol{x})$$

Dominance-based

$$pd_{n'}(\boldsymbol{x}) \propto \prod_{k=1}^{K} pd_k^{m_k}(\boldsymbol{x})$$

Combined

$$p_{n'}(\boldsymbol{x}) = pa_{n'}(\boldsymbol{x}) pd_{n'}(\boldsymbol{x})$$

An image is more likely to belong to a category if
- ✧ Image satisfies the stated attribute presence
- ✧ Attributes named first by supervisor are most dominant in image

More natural interface for zero-shot learning: opportunity to leverage human tendencies

Approach: Image Search

$$p(\boldsymbol{x}) \propto \prod_{k=1}^{K} pa^{m_k}(\boldsymbol{x}) pd^{m}_{m_k}(\boldsymbol{x})$$

Probability that image is the target image | How well image satisfies attribute presence in query (~ Kumar 2010) | How well image's dominance pattern matches order of attributes in query

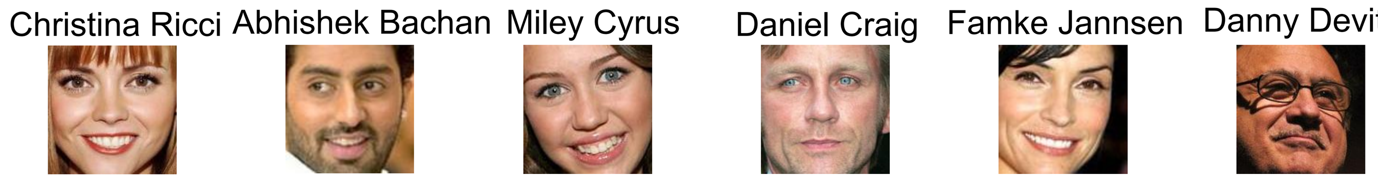
Approach: Textual Image Description

Task: Describe an input image using k attributes
Approach: Use k most dominant attributes

Data Collection

Public Figures Face Database (PubFig): 200 categories, 13 attributes

Christina Ricci, Abhishek Bachan, Miley Cyrus, Daniel Craig, Famke Jannsen, Danny Devito



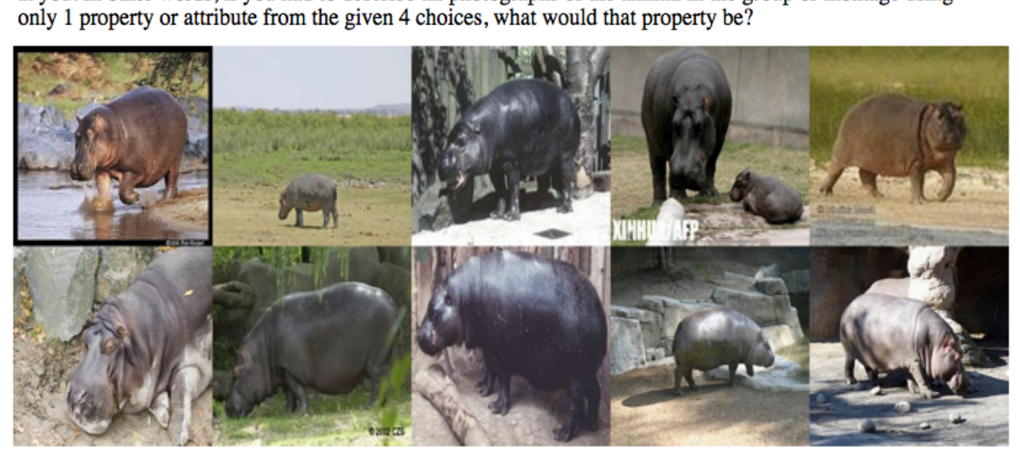
Animals with Attributes (AWA): 50 categories, 27 attributes

Bobcat, Dolphin, Giraffe, Lion, Polar bear, Zebra

What pops out? 6 subjects per question

Instructions:

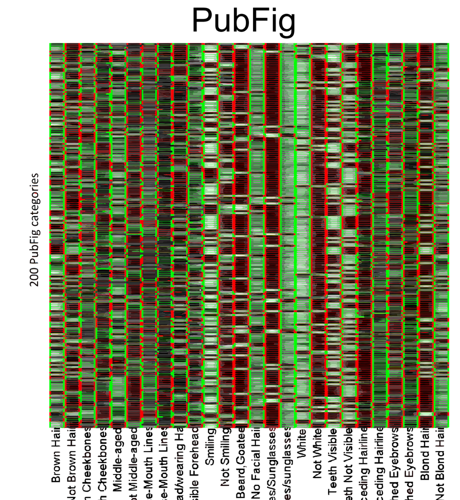
For each montage shown below, please tell us which 1 of the 4 properties/attributes of the animal pops out at you. In other words, if you had to describe all photographs of the animal in the group or montage using only 1 property or attribute from the given 4 choices, what would that property be?



- Has bulbous/bulging/round body
- Is a coastal animal
- Does not have bulbous body
- Is not a coastal animal

GT dominance: #subjects picked attribute

PubFig



AWA

Brighter intensities correspond to higher dominance. Green / red boundaries indicate whether the attribute is present / absent in that category.

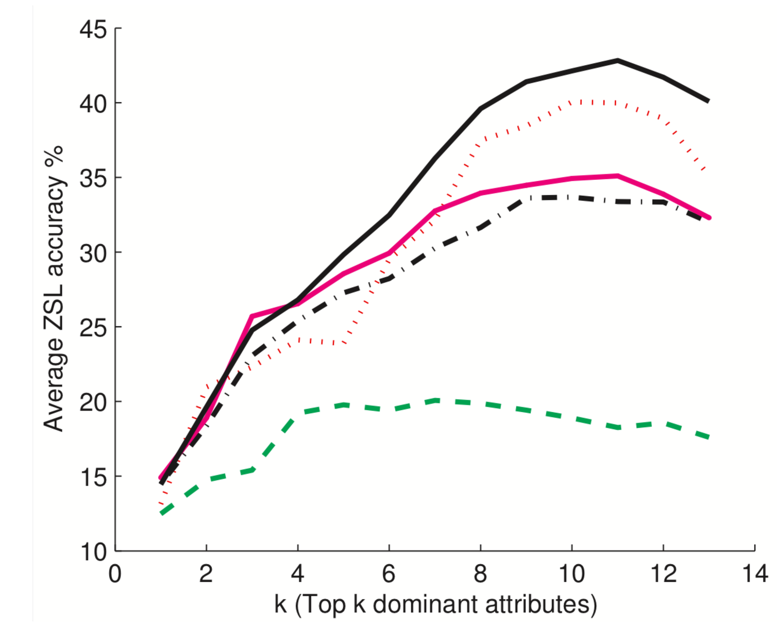
Observations:
- ✧ Same (present) attribute has different dominance in different categories
- ✧ Absence of attributes can be dominant

Correlation with: AWA PubFig (PubFig-subset)
- ✧ Ground truth relative attributes: (0.46)
- ✧ Ground truth global dominance: 0.54 0.50 (0.44)
- ✧ Ground truth TFIDF: 0.69 0.69
- ✧ Other subjects: 0.94 0.93 (0.93)
- ✧ *Predicted* dominance: 0.66 0.61 (0.68)

Results

Zero-shot Learning

PubFig: 69% with inter-subject



AWA: 68% with inter-subject
84% of [Lampert 2009] performance
with 32% attributes

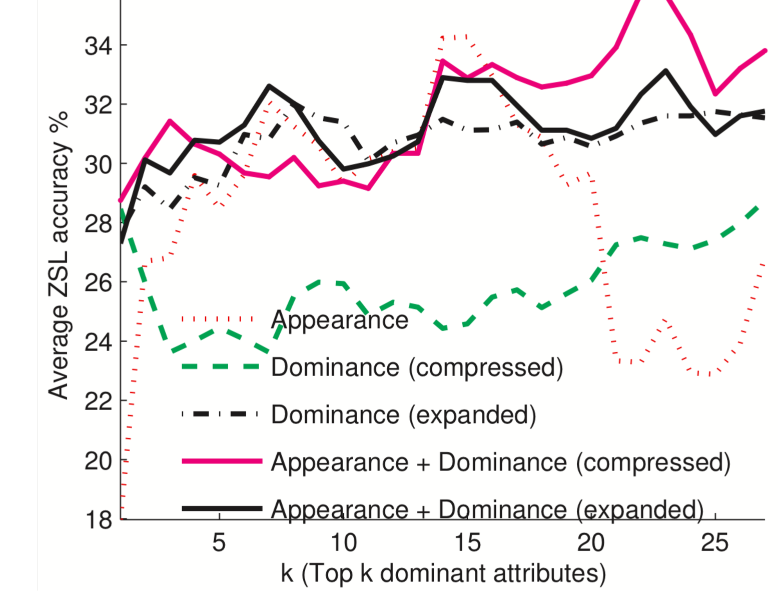
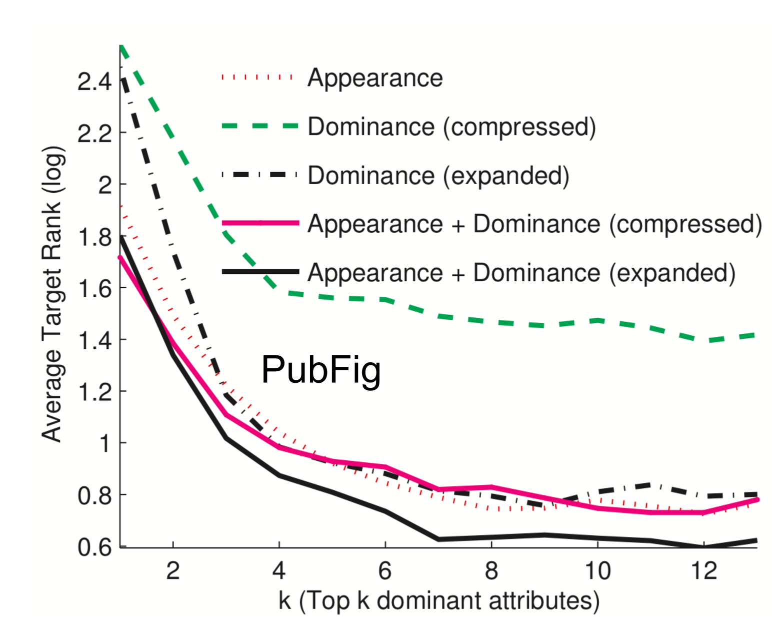


Image Search



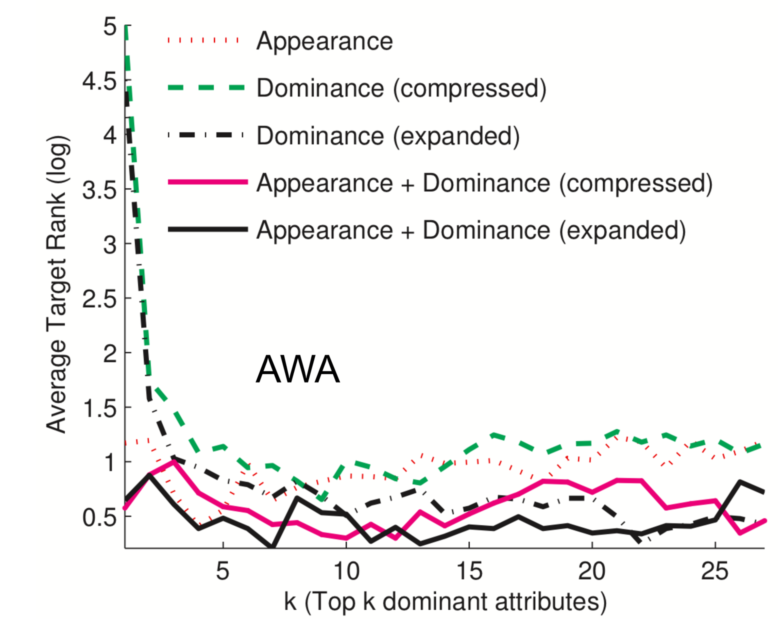
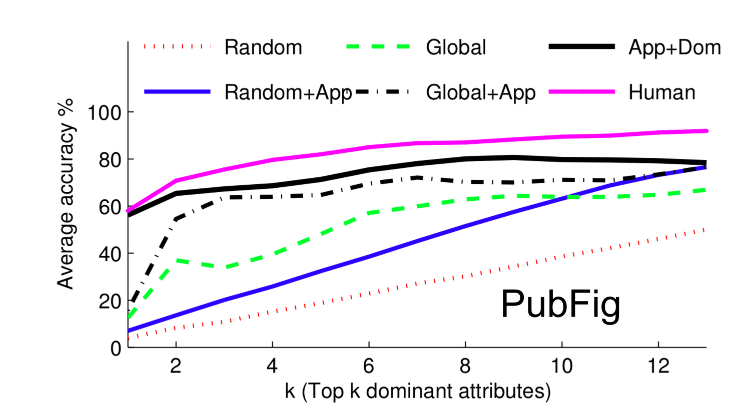


Image Description



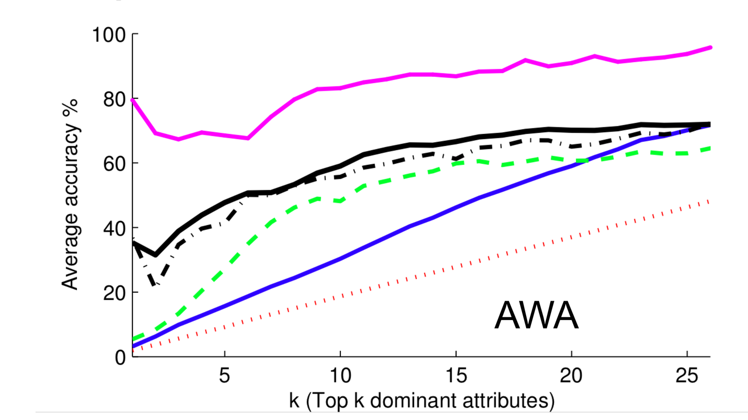


Image Description: User Studies

	PubFig	AWA
Random	5%	8%
Global	22%	28%
Ours	73%	64%

GT	PubFig	AWA
Random	2%	0%
Global	25%	16%
Ours	73%	84%

Validates our intuition!

Conclusion

- Attribute dominance: some attributes pop out more than others
- Contains information about image content
- Humans name dominant attributes first
- Leverage human tendencies to read between the lines
- Improved performance at human-centric applications: zero-shot learning, image search and image description