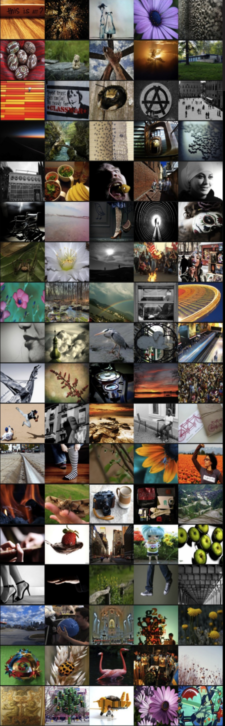


Learning using attributes

Thomas Mensink

Computer Vision by Learning, March 28th 11:30-12:15

Image Classification: Visual examples

Which image shows an axolotl?




Image Classification: Visual examples

Which image shows an axolotl?

Traindata:

Image Classification: Visual examples

Which image shows an axolotl?

Traindata:

We can classify based on visual examples

Image Classification: Textual descriptions

Which image shows an aye-aye?

Image Classification: Textual descriptions

Which image shows an aye-aye?




Description, Aye-aye . . .

- is nocturnal
- lives in trees
- has large eyes
- has long middle fingers

Image Classification: Textual descriptions

Which image shows an aye-aye?

Description, Aye-aye . . .

- is nocturnal
- lives in trees
- has large eyes
- has long middle fingers

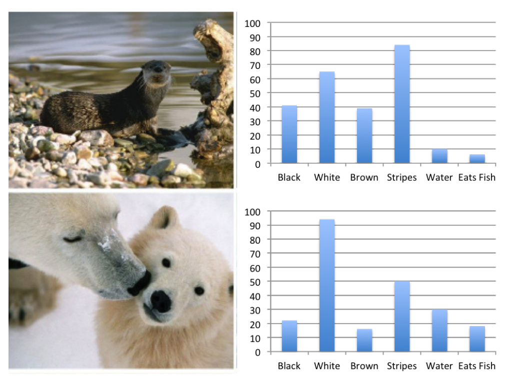
We can classify based on textual descriptions

Attribute-Based Classification

Definition

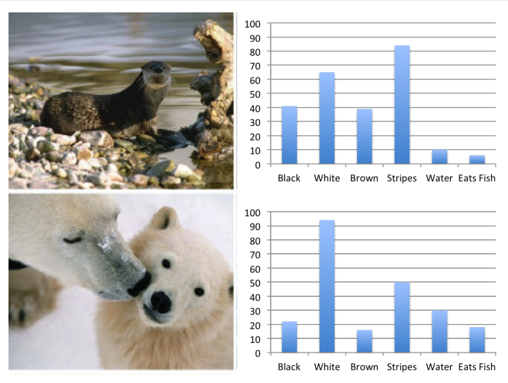
Classification using a *class description* in terms of semantic properties or *attributes*

Attribute-Based Classification: Properties



- Semantic interpretable representation
- Dimension reduction:
  1. high-dimensional low-level features
  2. low-dimensional semantic representation

Attribute-Based Classification: Requirements



- Vocabulary of Attributes and Attribute-to-class Mapping
- Attribute predictors
- Learning model to make decision

Zero-shot recognition

- **Goal:** Classify images into classes which we have never seen
- **Assumption 1:** Text descriptions of unseen+related classes
- **Assumption 2:** Visual examples from related classes.

Zero-shot recognition (2)

1. Vocabulary of attributes and class descriptions:
   Aye-ayes have properties X, and Y, but not Z
2. Train classifiers for each attibute **X, Y, Z.**
   From visual examples of related classes
3. Make image attributes predictions:

$$\Rightarrow \quad P(X|\mathrm{img}) = 0.8 \quad P(Y|\mathrm{img}) = 0.3 \quad P(Z|\mathrm{img}) = 0.6$$

4. Combine into decision: *this image is not an Aye-aye*

Zero-shot recognition (2)

1. Vocabulary of attributes and class descriptions:
   Aye-ayes have properties X, and Y, but not Z
2. Train classifiers for each attibute **X, Y, Z.**
   From visual examples of related classes
3. Make image attributes predictions:

$$\Rightarrow \quad \begin{aligned} P(X|\mathrm{img}) &= 0.8 \\ P(Y|\mathrm{img}) &= 0.3 \\ P(Z|\mathrm{img}) &= 0.6 \end{aligned}$$

4. Combine into decision: *this image is not an Aye-aye*

Zero-shot recognition (2)

1. Vocabulary of attributes and class descriptions:
   Aye-ayes have properties X, and Y, but not Z

2. Train classifiers for each attibute **X, Y, Z.**
   From visual examples of related classes

3. Make image attributes predictions:

$$\Rightarrow \quad \begin{aligned} P(X|\mathrm{img}) &= 0.8 \\ P(Y|\mathrm{img}) &= 0.3 \\ P(Z|\mathrm{img}) &= 0.6 \end{aligned}$$

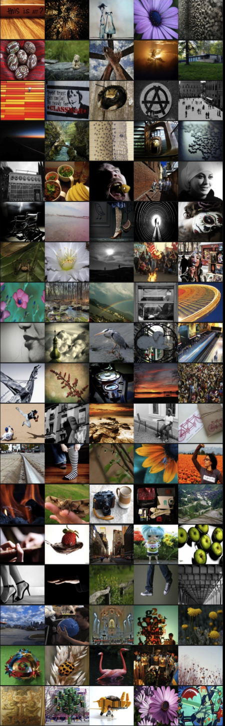
4. Combine into decision: *this image is not an Aye-aye*

Zero-shot recognition (3)

- **Goal:** Classify images into classes which we have never seen
- **Assumption 1:** Text descriptions of unseen+related classes
- **Assumption 2:** Visual examples from related classes.
- **Solution:** Attribute-based zero-shot classification **[Lampert CVPR'09]**

1. Construct and train attribute classifiers
2. Convert image to attribute representation
3. Use attribute-to-class mapping for final decision

Outline



2. Attribute Vocabulary

What are good attributes?

Good attributes. . .

- . . . are task and category dependent;
- . . . class discriminative, but not class specific;
- . . . interpretable by humans; and
- . . . detectable by computers

Quiz: What are good attributes?

Possible attributes

- is grey?
- is made of atoms?
- lives in Amsterdam?
- eat fish?
- has a SIFT descriptor with empty bin 3?
- number of wheels?

Attributes for Animal Classification

AwA dataset: 30K images, 50 classes, 85 attributes **[Lampert CVPR'09]**

black	pads	strong	arctic
white	paws	weak	coastal
cyan	longleg	muscle	desert
brown	longneck	active	bush
gray	tail	inactive	plains
orange	chewteeth	nocturnal	forest
red	meatteeth	hibernate	fields
yellow	buckteeth	agility	jungle
patches	strainteeth	fish	mountains
spots	horns	meat	ocean
stripes	claws	plankton	ground
furry	tusks	vegetation	water
hairless	bipedal	insects	tree
toughskin	quadrapedal	forager	cave
big	flys	grazer	fierce
small	hops	hunter	timid
bulbous	swims	scavenger	smart
lean	tunnels	skimmer	group
flippers	walks	stalker	solitary
hands	fast	newworld	nestspot
hooves	slow	oldworld	domestic

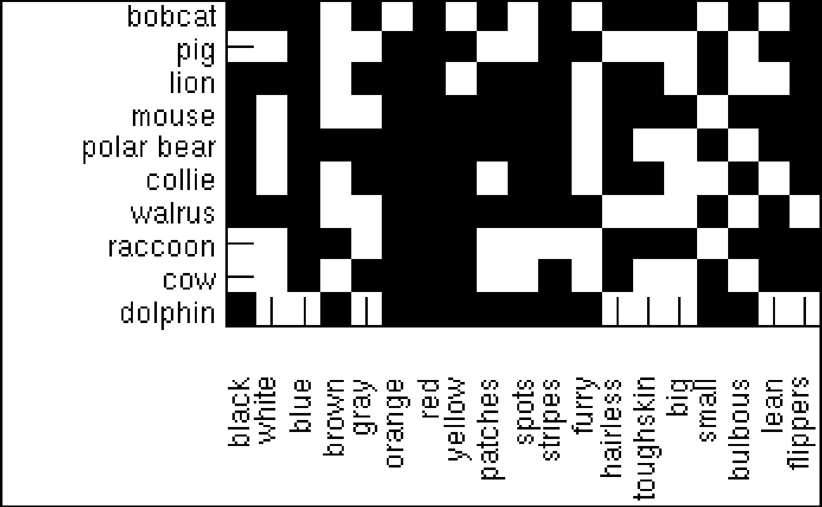
Attributes for Animal Classification

AwA dataset: 30K images, 50 classes, 85 attributes [Lampert CVPR'09]

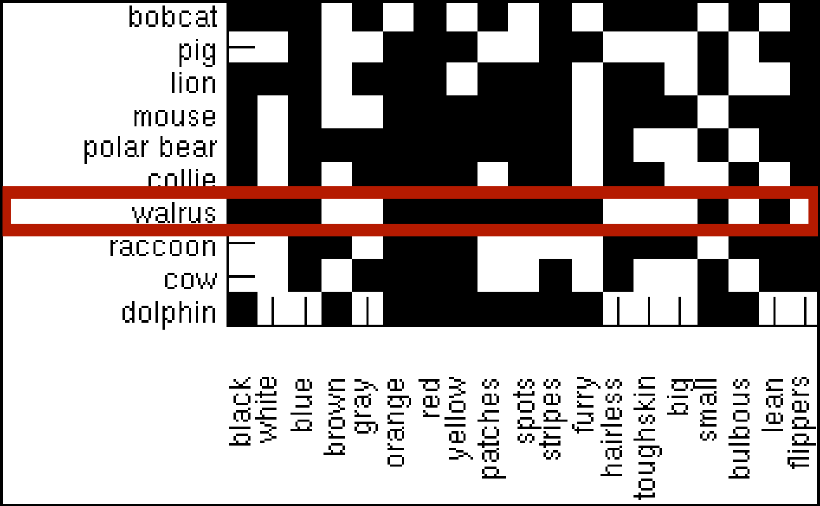
black	pads	strong	arctic
white	paws	weak	coastal
cyan	longleg	muscle	desert
brown	longneck	active	bush
gray	tail	inactive	plains
orange	chewteeth	nocturnal	forest
red	meatteeth	hibernate	fields
yellow	buckteeth	agility	jungle
patches	strainteeth	fish	mountains
spots	horns	meat	ocean
stripes	claws	plankton	ground
furry	tusks	vegetation	water
hairless	bipedal	insects	tree
toughskin	quadrapedal	forager	cave
big	flys	grazer	fierce
small	hops	hunter	timid
bulbous	swims	scavenger	smart
lean	tunnels	skimmer	group
flippers	walks	stalker	solitary
hands	fast	newworld	nestspot
hooves	slow	oldworld	domestic

Contain attributes about: color, texture, shape, body parts, behaviour, nutrition, activity, habitat, character

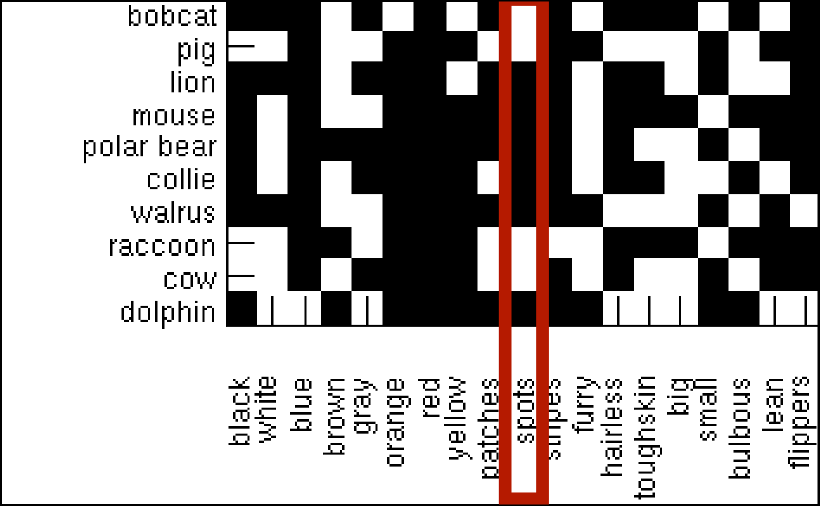
Binary Attribute-to-Class mapping



Binary Attribute-to-Class mapping



Binary Attribute-to-Class mapping



Deriving Attributes and Mappings

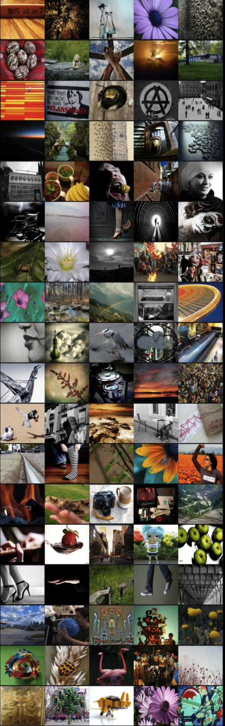
- Manual vocabulary, obtained from domain experts **[Lampert CVPR'09]**
- Tagged images of related classes **[Wah TR'11]**
- Automatic discovery from language resources **[Rohrbach CVPR'10]**
  - Such as: Experts descriptions, Ontologies, Wikipedia
- General classifiers / concepts **[Torresani ECCV'10]**
  - Such as Classemes or ImageNet
- Active Learning **[Parikh CVPR'11]**

How many attributes?

- In theory k binary attributes can represent …
- In practice for c classes we need …

How many attributes?

- In theory k binary attributes can represent ...
  2^k classes
- In practice for c classes we need ...
  Many attributes



3. Attribute predictors

Getting training examples

- Attribute names, without images
  - Search for attribute names on the Internet [Ferrari NIPS'07]
- Image labelled with attributes [Ferhadi CVPR'09]
- Class-specific descriptions [Lampert CVPR'09]
  - Use all images of class either as positive or as negative

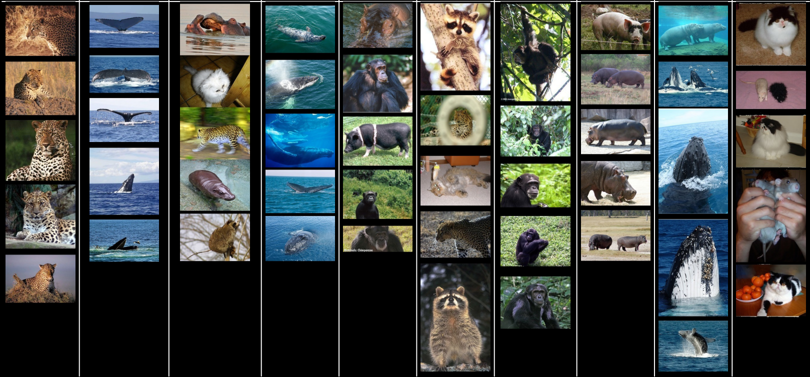
Use your favourite algorithm

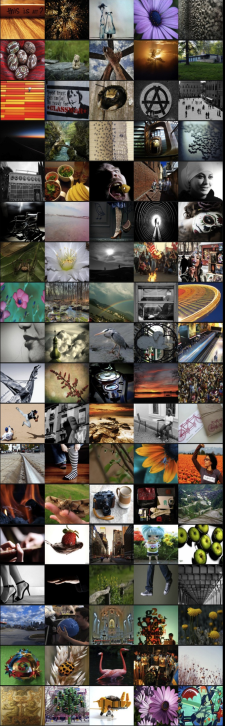
- SVM
- Logistic Regression
- DeepNet
- ...

Attributes for Animal Classification

AwA dataset: 30K images, 50 classes, 85 attributes [Lampert CVPR'09]

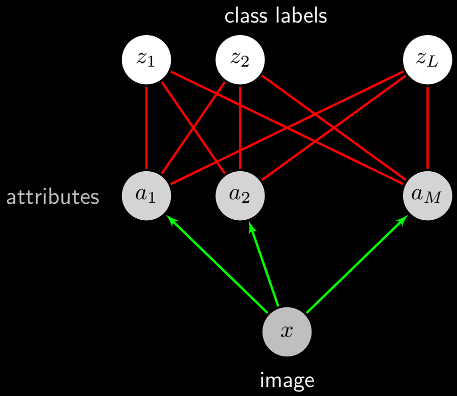
is yellow	*eats plankton*	*has buckteeth*	*is blue*	*is brown*	*has paws*	*lives in trees*	*is smelly*	*is big*	*is small*
(AUC 92.9)	(AUC 99.1)	(AUC 40.4)	(AUC 78.2)	(AUC 62.1)	(AUC 82.5)	(AUC 78.8)	(AUC 70.0)	(AUC 79.7)	(AUC 69.4)





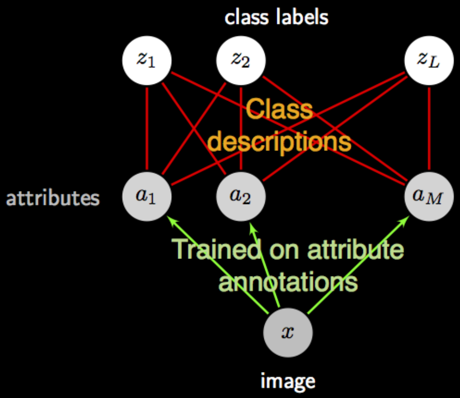
4. Attribute-based classification

Direct Attribute Prediction (DAP)



- Learn attribute classifiers from related classes **[Lampert CVPR'09]**
- Train and test classes are disjoint
- Use Attribute-to-class mapping for prediction

Direct Attribute Prediction (DAP)



- Learn attribute classifiers from related classes [Lampert CVPR'09]
- Train and test classes are disjoint
- Use Attribute-to-class mapping for prediction

DAP: Probabilistic model

- Class probability:

$$p(z|\boldsymbol{x}) = \frac{p(z)}{p(\boldsymbol{a}^z)} \prod_m p(a_m = a_m^z|\boldsymbol{x})$$

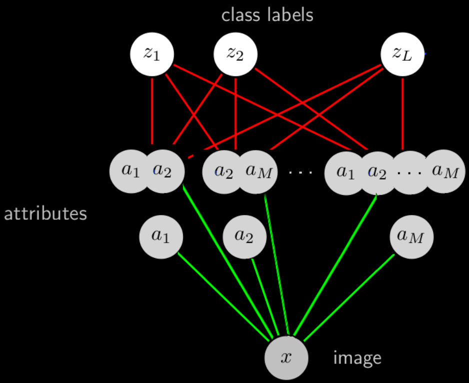
- Define attribute probability:

$$p(a_m = a_m^z|\boldsymbol{x}) = \begin{cases} p(a_m|\boldsymbol{x}) & \text{if } a_m^z = 1 \\ 1 - p(a_m|\boldsymbol{x}) & \text{otherwise} \end{cases}$$

- Assume equal prior $p(z)$ and attribute prior $p(\boldsymbol{a}^z)$
- Assign a given image to class z^*

$$z^* = \arg\max_z \prod_m p(a_m^z|\boldsymbol{x})$$

Structured DAP



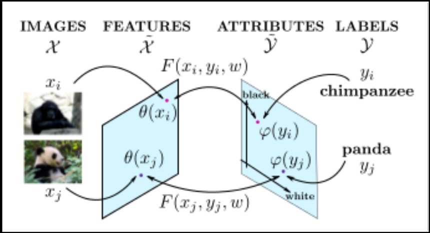
- Learn attributes jointly in a structured framework [Mensink PAMI'12]
- Train and test classes are disjoint
- Use Attribute-to-class mapping for prediction

Attribute Label Embedding (ALE)

- **Limitation of direct attribute prediction:**
  not optimized for the final classification objective!
- DAP uses two-stage learning / predicting:
  1. Learn Attribute Predictors
  2. Use for classification
- **Solution:**
  ALE learns for zero-shot classification [Akata CVPR'13]

Attribute Label Embedding (ALE)

- **Limitation of direct attribute prediction:**
  not optimized for the final classification objective!
- DAP uses two-stage learning / predicting:
  1. Learn Attribute Predictors
  2. Use for classification
- **Solution:**
  ALE learns for zero-shot classification **[Akata CVPR'13]**



ALE: Model

$$F(z) = \boldsymbol{x}^\top W \boldsymbol{a}_z = \sum_m a_{zm}\, \boldsymbol{x}^\top \boldsymbol{w}_a$$

- Image features $\boldsymbol{x}$
- Attribute vector $\boldsymbol{a}_z$
- Attribute predictors W
  - Each column is an attribute predictor
- **Trained to optimise zero-shot classification** z
  - When trained for attribute prediction $a \rightsquigarrow$ DAP

ALE: Model

$$
\begin{aligned}
F(z) &= \boldsymbol{x}^\top W \boldsymbol{a}_z \\
&= \sum_m a_{zm}\, \boldsymbol{x}^\top \boldsymbol{w}_a
\end{aligned}
$$

- Image features $\boldsymbol{x}$
- Attribute vector $\boldsymbol{a}_z$
- Attribute predictors W
  - Each column is an attribute predictor
- **Trained to optimise zero-shot classification** z
  - When trained for attribute prediction $a \rightsquigarrow$ DAP

ALE Results

- Zero-shot learning
  - Train and test classes are disjoint
- Evaluation of class prediction and attribute prediction

	Obj. pred.		Att. pred.	
	DAP	ALE	DAP	ALE
AWA	36.1	**37.4**	**71.9**	65.7
CUB	10.5	**18.0**	**61.8**	60.3

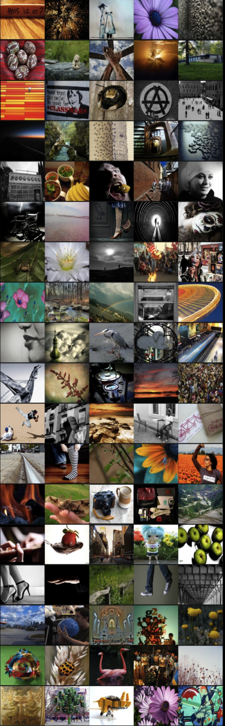
- ALE improves zero-shot recognition
- But, attribute prediction decreased!

ALE Results

- Zero-shot learning
  - Train and test classes are disjoint
- Evaluation of class prediction and attribute prediction

	Obj. pred.		Att. pred.	
	DAP	ALE	DAP	ALE
AWA	36.1	**37.4**	**71.9**	65.7
CUB	10.5	**18.0**	**61.8**	60.3

- ALE improves zero-shot recognition
- But, attribute prediction decreased!



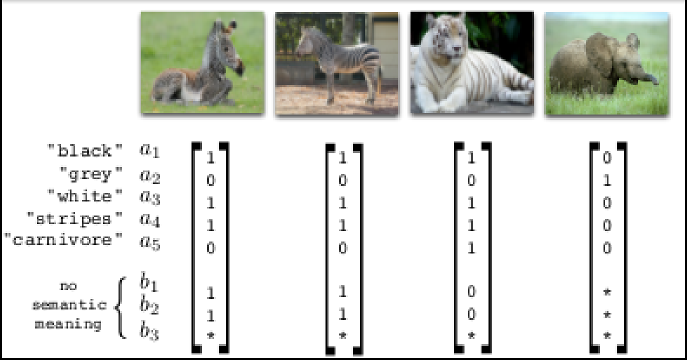
5. Fun with Attributes

Discriminative Attribute Representations

- Attributes are interpretable
- Can we learn discriminative attributes?
- Augmented Attributes [Sharmanska ECCV'12]
- Discriminative Binary Codes [Rastegari ECCV'12]

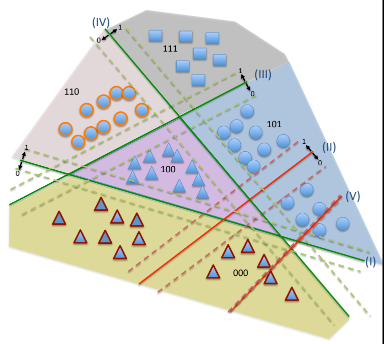
Discriminative Attribute Representations

- Attributes are interpretable
- Can we learn discriminative attributes?
- Augmented Attributes [Sharmanska ECCV'12]
- Discriminative Binary Codes [Rastegari ECCV'12]



Discriminative Attribute Representations

- Attributes are interpretable
- Can we learn discriminative attributes?
- Augmented Attributes **[Sharmanska ECCV'12]**
- Discriminative Binary Codes **[Rastegari ECCV'12]**



Relative Attributes

- **Problem:** Binary attributes are very crude
  - If mouse = small, then cat $\neq$ small
  - If elephant = large, then cat $\neq$ large
- **Solution:** Relative attributes [Parikh ICCV'11]

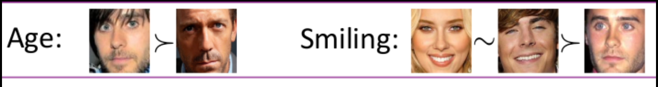
Relative Attributes

- **Problem:** Binary attributes are very crude
  - If mouse = small, then cat $\neq$ small
  - If elephant = large, then cat $\neq$ large
- **Solution:** Relative attributes **[Parikh ICCV'11]**
- Rank images to a level of *degree*

Relative Attributes

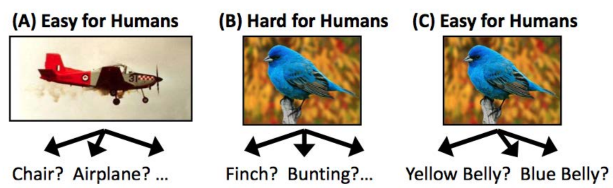
- **Problem:** Binary attributes are very crude
  - If mouse = small, then cat ≠ small
  - If elephant = large, then cat ≠ large
- **Solution:** Relative attributes **[Parikh ICCV'11]**
- Rank images to a level of *degree*

- Use distance in ranking for comparisons:



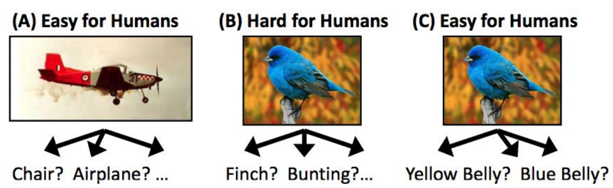
Humans in the Loop

- A computer should help the human
- Easy and hard classification problems for humans:

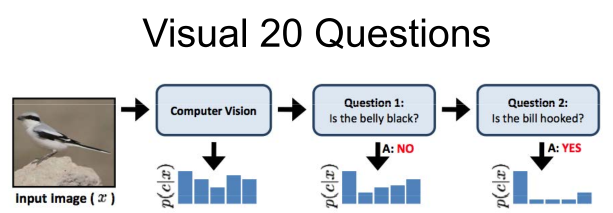


Humans in the Loop

- A computer should help the human
- Easy and hard classification problems for humans:



- Solve *hard for human* problems with interaction [Branson ECCV'10]

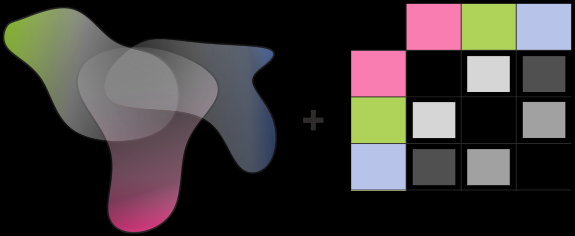


Labels as Attributes and Classes

- **Problem:** distinction between *classes* and *attributes*
- **Solution:** Use labels to predict unseen labels **[Mensink CVPR'14]**
- Predict unseen labels based on co-occurrence with other labels

Labels as Attributes and Classes

- **Problem:** distinction between *classes* and *attributes*
- **Solution:** Use labels to predict unseen labels [Mensink CVPR'14]
- Predict unseen labels based on co-occurrence with other labels



Existing classifiers

Co-occurences from **Multi-Labeled** Images

Labels as Attributes and Classes

- **Problem:** distinction between *classes* and *attributes*
- **Solution:** Use labels to predict unseen labels [Mensink CVPR'14]
- Predict unseen labels based on co-occurrence with other labels

Can attributes be used for known classes?

- And will it be any better than low-level features?

Fine-Grained Classification

- **Goal:** Classify similar objects into specific types

- **Normal classification**: Elephant or other animal?
- **Fine-grained classification**: Indian or African Elephant?

Fine-Grained Classification (2)

African

An African or Indian Elephant?

Indian

Fine-Grained Classification (3)

An African or Indian Elephant?

The **African Elephant** is described as the Loxodonta africana of Africa. They are very large, grey, four-legged herbivorous mammals. They have almost hairless skin, a distinctive long, flexible, prehensile trunk. Its upper incisors form long curved tusks of ivory. African elephants have large fan-shaped ears and two fingers at the tip of its trunk, compared to only one in the Asian species.

The **Indian Elephant** is described as Elephas maximus of south-central Asia. They are very large, grey, four-legged herbivorous mammals. They have almost hairless skin, a distinctive long, flexible, prehensile trunk. Its upper incisors form long curved tusks of ivory. The ears of Indian elephants are significantly smaller than African elephants.

1. Source: `http://www.findfast.org/animals-elephants.htm`

Fine-Grained Classification (3)

An African or Indian Elephant?

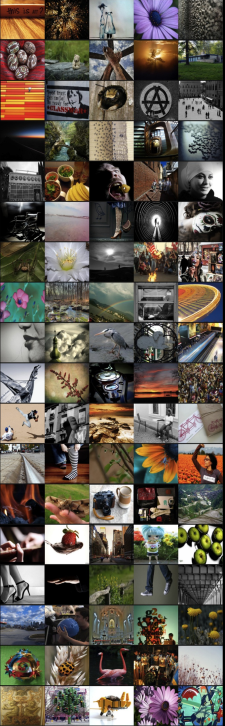
The **African Elephant** is described as the Loxodonta africana of Africa. They are very large, grey, four-legged herbivorous mammals. They have almost hairless skin, a distinctive long, flexible, prehensile trunk. Its upper incisors form long curved tusks of ivory. African elephants have large fan-shaped ears and two fingers at the tip of its trunk, compared to only one in the Asian species.

The **Indian Elephant** is described as Elephas maximus of south-central Asia. They are very large, grey, four-legged herbivorous mammals. They have almost hairless skin, a distinctive long, flexible, prehensile trunk. Its upper incisors form long curved tusks of ivory. The ears of Indian elephants are significantly smaller than African elephants.

1. Source: http://www.findfast.org/animals-elephants.htm

Fine-Grained Classification (4)

- **Goal:** Classify similar objects into specific types
- **Observation:** Visual examples might not help to distinguish.
- **Attributes:** Could provide a way to use *expert knowledge* about the differences between visual similary types.



6. Conclusions

Take home messages

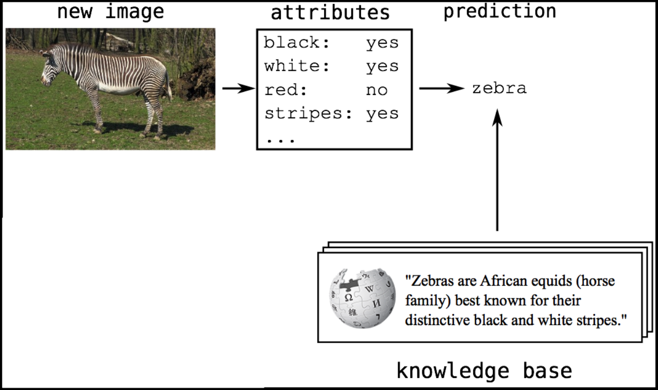
Attribute-based Classification

1. Vocabulary of attributes and class descriptions
   - Attributes are semantic and detectable object properties
2. Attribute Predictors
   - Attributes provide an intermediate semantic representation
     Often of lower dimensionality as low-level image features
3. Combining into decision
   - Allows to use expert (a priori) knowledge about classes

Take home messages: Illustration

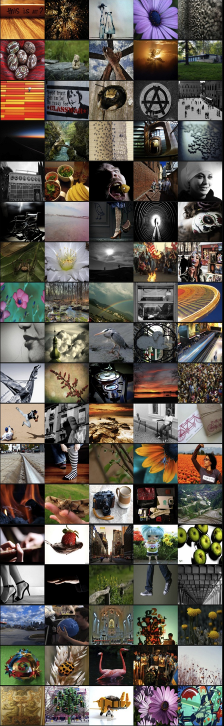
Attribute-based Classification

1. Vocabulary of attributes and class descriptions
2. Attribute Predictors
3. Combining into decision



Thanks to . . .

- Christoph Lampert for slides and inspiration
- The organizers (Arnold, Laurens and Cees, for asking me)
- My colleagues and former colleagues
- Authors of the papers I've used for this presentation



Learning using attributes

Questions?

References

- Akata et al., Label-Embedding for Attribute-Based Classification, CVPR’13
- Branson et al., Visual Recognition with Humans in the Loop, ECCV’10
- Ferrari and Zisserman, Learning Visual Attributes, NIPS’07
- Ferhadi et al, Describing Objects by Their Attributes, CVPR’09
- **Lampert et al., Learning To Detect Unseen Object Classes, CVPR’09**
- Li et al., Object Bank: A High-Level Image Representation, NIPS’10
- Mensink et al., Tree-structured CRF Models for Interactive Image Labeling, PAMI’12
- Mensink et al., Co-Occurrence Statistics for Zero-Shot Classification, CVPR’14
- Parikh and Grauman, Relative Attributes, ICCV’11
- Parikh and Grauman, Interactively Building a Vocabulary of Nameable Attributes, CVPR’11
- Rastegari et al., Attribute Discovery via Discriminative Binary Codes, ECCV’12
- Rohrbach et al., What Helps Where And Why? Semantic Knowledge Transfer, CVPR’10
- Sharmanska et al., Augmented Attribute Representations, ECCV’12
- Torresani et al., Efficient Object Category Recognition Using Classemes, ECCV’10
- Wah et al., The Caltech-UCSD Birds-200-2011 Dataset, TR’11