EECS 442 – Computer vision

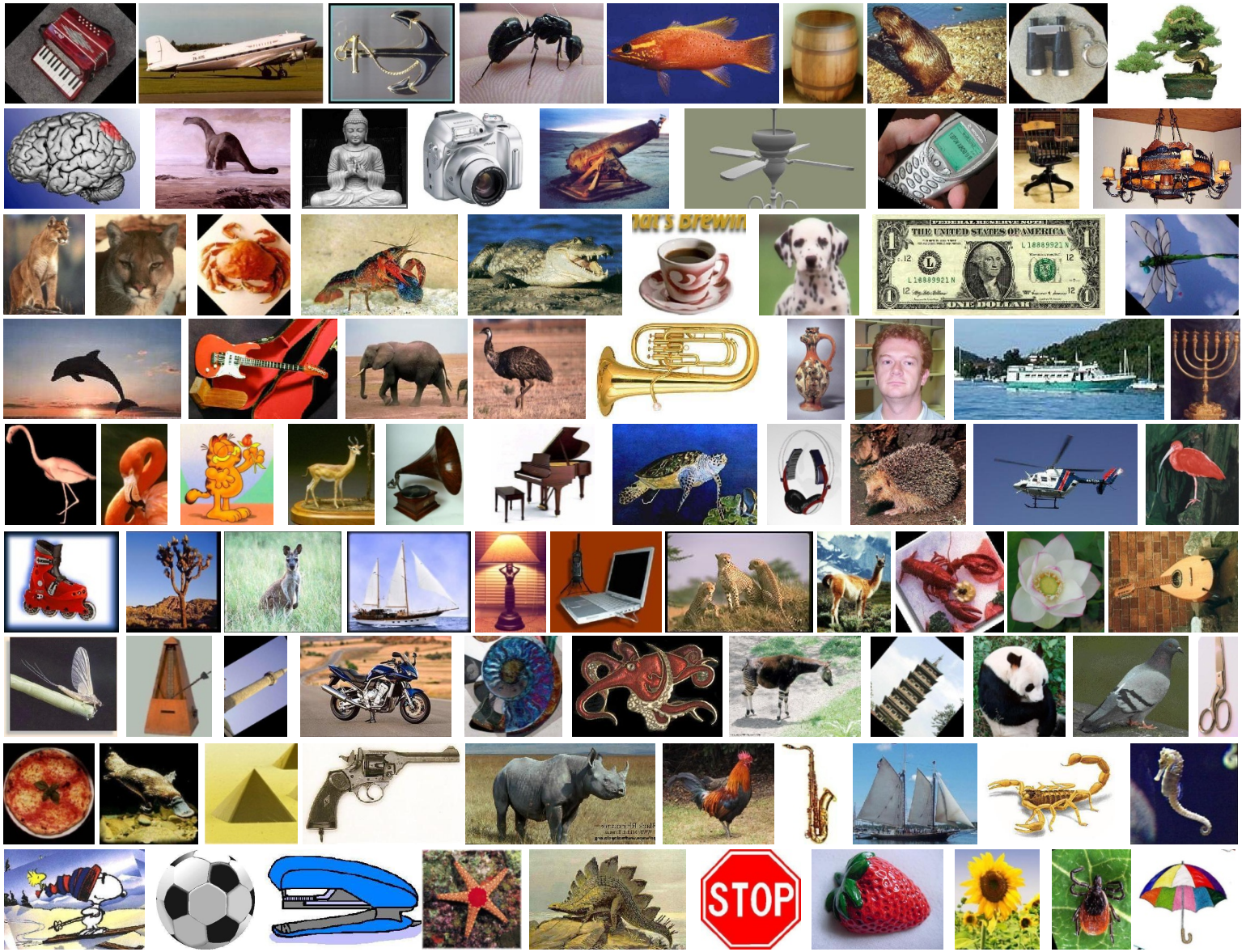
Databases for object recognition and beyond

Segments of this lectures are courtesy of Prof F. Li, R. Fergus and A. Zisserman

Caltech 101

- Pictures of objects belonging to 101 categories.
- About 40 to 800 images per category. Most categories have about 50 images.
- The size of each image is roughly 300 x 200 pixels.

Caltech 101 images



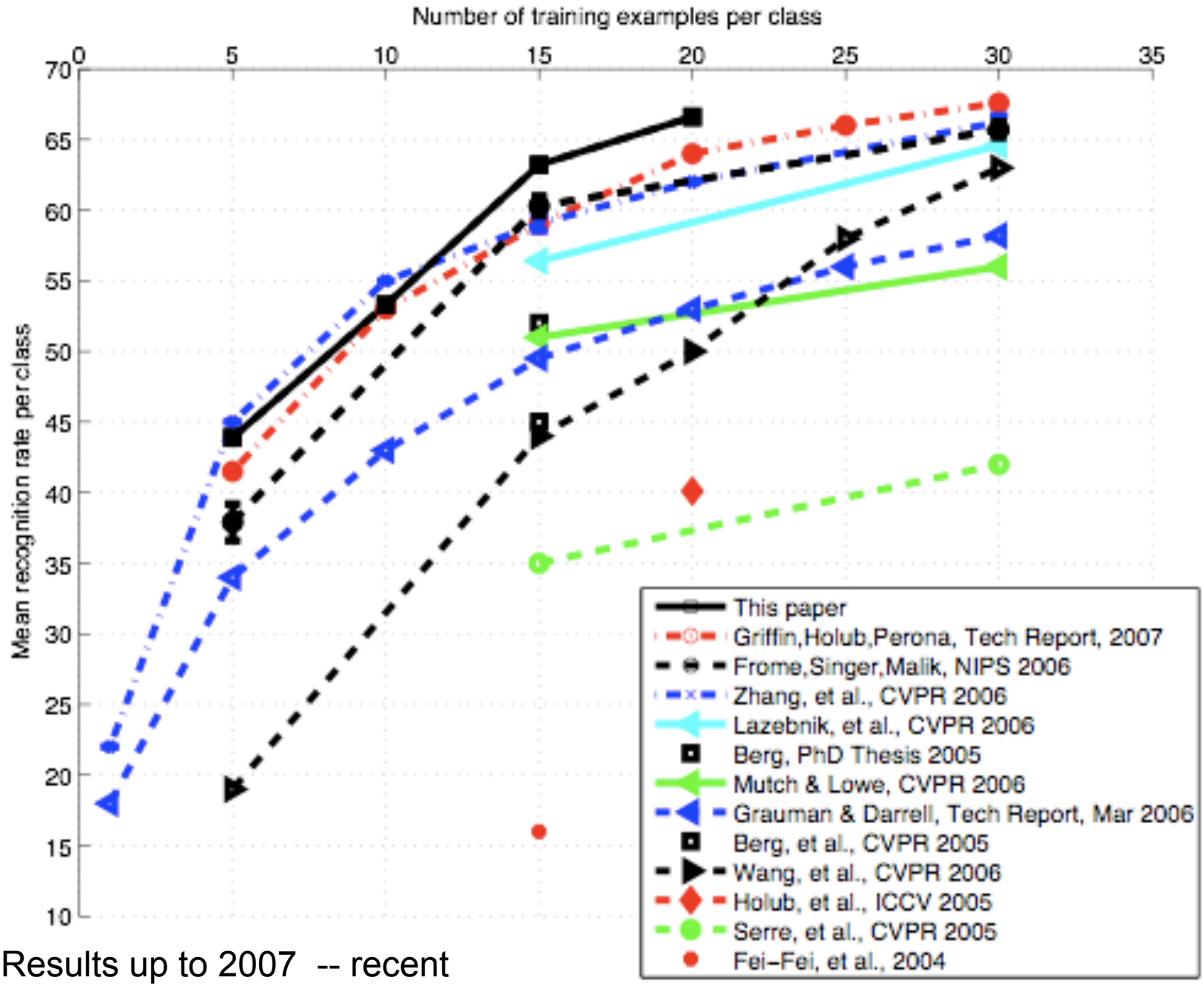
Caltech-101: Drawbacks

- Smallest category size is 31 images: $N_{train} \leq 30$

- Too easy?
  - left-right aligned
  - Rotation artifacts
  - Saturated performance







Results up to 2007 -- recent methods obtain almost 100%

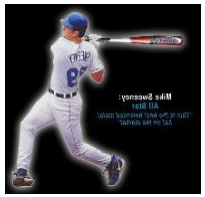
Caltech-256

- Smallest category size now 80 images
- About 30K images
- Harder
  - Not left-right aligned
  - No artifacts
  - Performance is halved
  - More categories
- New and larger clutter category

Caltech 256 images

baseball-bat

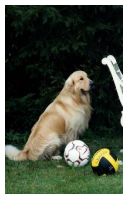



dog







basketball-hoop







kayac



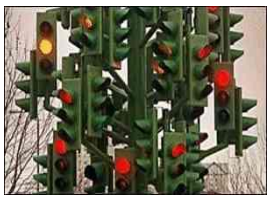




traffic light

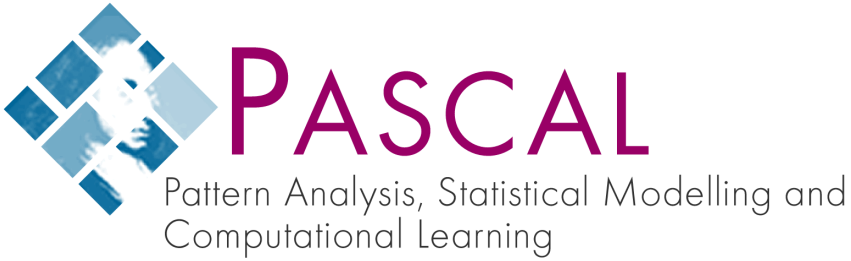






The PASCAL Visual Object Classes (VOC) Dataset and Challenge

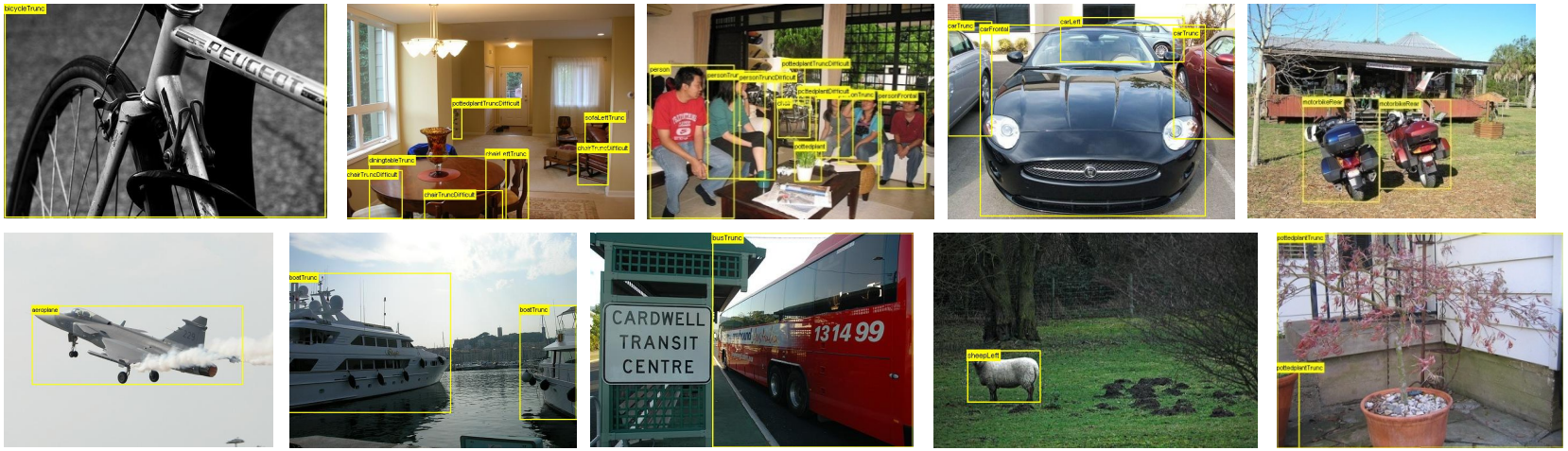
Mark Everingham
Luc Van Gool
Chris Williams
John Winn
Andrew Zisserman



Dataset Content

- 20 classes: aeroplane, bicycle, boat, bottle, bus, car, cat, chair, cow, dining table, dog, horse, motorbike, person, potted plant, sheep, train, TV

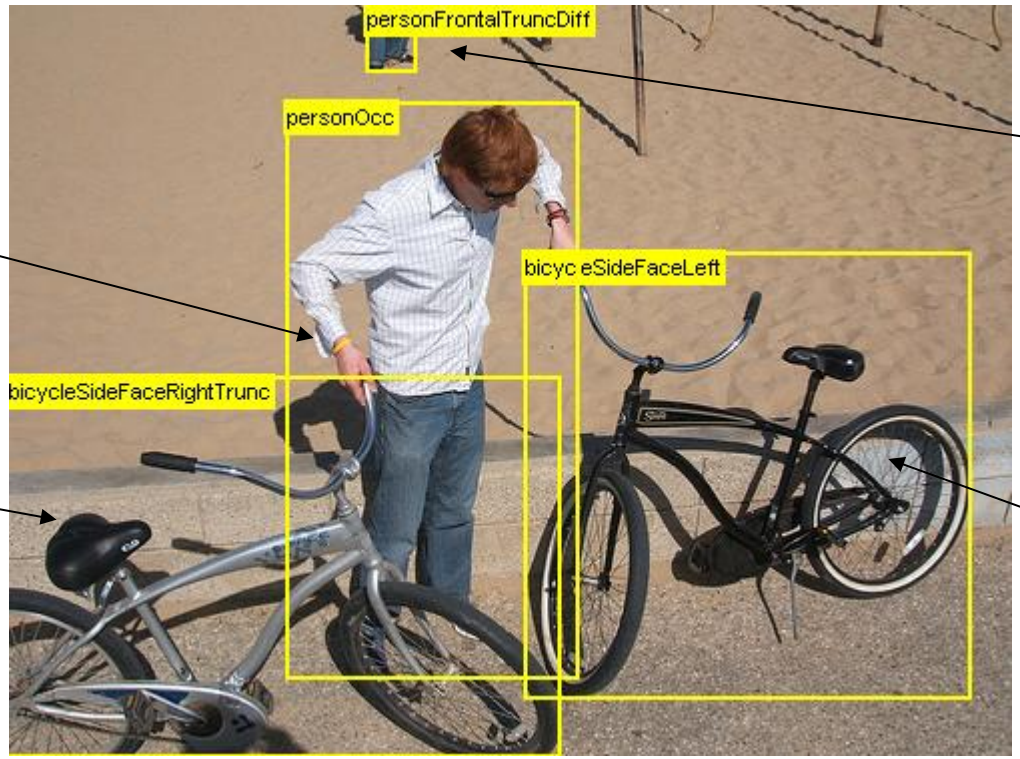
- Real images downloaded from flickr, not filtered for “quality”



- Complex scenes, scale, pose, lighting, occlusion, ...

Annotation

- Complete annotation of all objects
- Annotated in one session with written guidelines



Occluded
Object is significantly occluded within BB

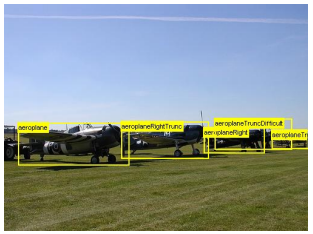
Truncated
Object extends beyond BB

Difficult
Not scored in evaluation

Pose
Facing left

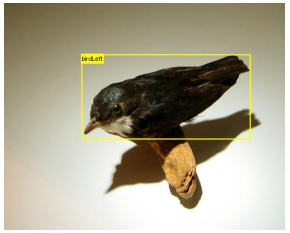
Examples

Aeroplane

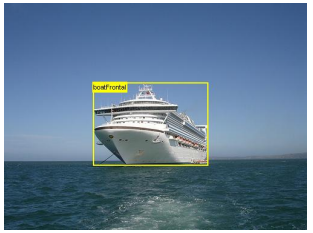


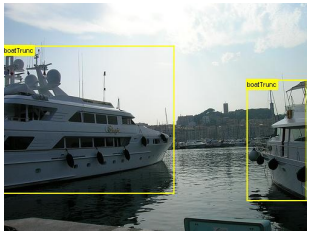
Bicycle

Bird

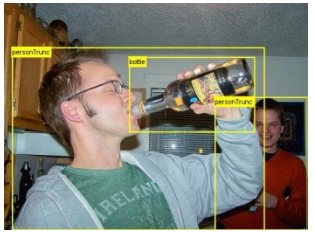


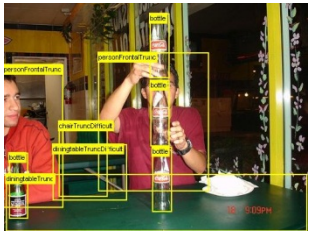
Boat



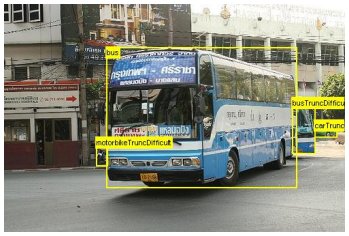


Bottle

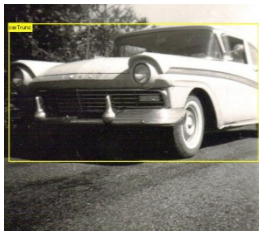




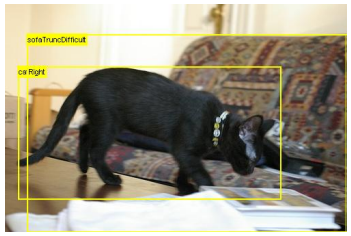
Bus



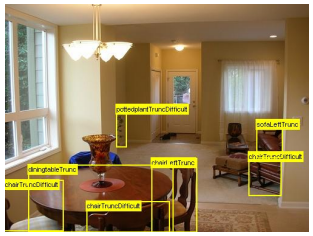
Car

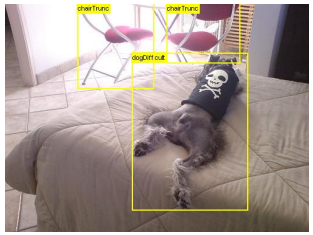


Cat

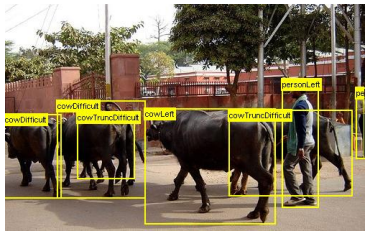


Chair





Cow



History

	Images	Objects	Classes	Entries	
2005	2,232	2,871	4	12	*Collection of existing and some new data.*
2006	5,304	9,507	10	25	*Completely new dataset from flickr (+MSRC)*
2007	9,963	24,640	20	28	*Increased classes to 20. Introduced tasters.*
2008	**8,776**	**20,739**	20		*Added "occlusion" flag. Reuse of taster data. Release detailed results to support "meta-analysis"*

- New dataset annotated annually
  - Annotation of test set is withheld until after challenge

Other recent datasets

ESP
[Ahn et al, 2006]

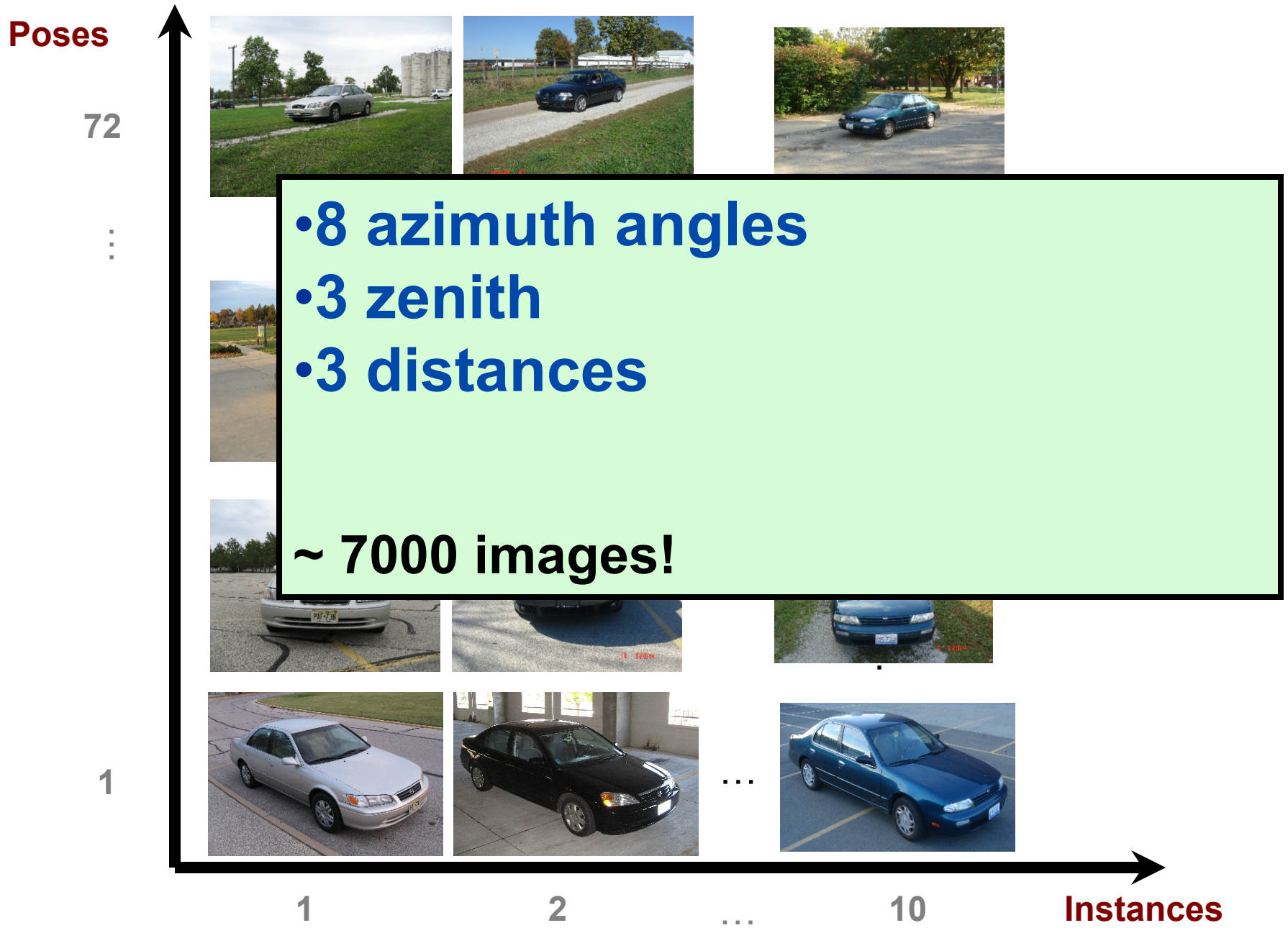
LabelMe
[Russell et al, 2005]

TinyImage
Torralba et al. 2007

Lotus Hill
[Yao et al, 2007]

MSRC
[Shotton et al. 2006]

3D object dataset [Savarese & Fei-Fei 07]


Poses
72
1
1
2
10
Instances
•8 azimuth angles
•3 zenith
•3 distances
~ 7000 images!

Largest dataset for object categories up to date

J. Deng, H. Su. K. Li , L. Fei-Fei ,

- ~20K categories;
- 14 million images;
- ~700im/categ;
- free to public at **www.image-net.org**

http://www.image-net.org

IMAGENET

9,956,478 images, 14841 synsets indexed

Explore Download **Challenge**[New!] People Publication Sponsors About

Not logged in. Login | Signup

ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently we have an average of over five hundred images per node. We hope ImageNet will become a useful resource for researchers, educators, students and all of you who share our passion for pictures.
Click Here to learn more about ImageNet, Click Here to join ImageNet mailing list.

SEARCH

What do these images have in common? *Find out!*

Update Notice: ImageNet 2010 Spring Version will be released in April, 2010

IMAGENET is a knowledge ontology

- Taxonomy

- S: (n) Eskimo dog, **husky** (breed of heavy-coated Arctic sled dog)
  - *direct hypernym* / ***inherited hypernym*** / *sister term*
    - S: (n) working dog (any of several breeds of usually large powerful dogs bred to work as draft animals and guard and guide dogs)
      - S: (n) dog, domestic dog, Canis familiaris (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) *"the dog barked all night"*
        - S: (n) canine, canid (any of various fissiped mammals with nonretractile claws and typically long muzzles)
          - S: (n) carnivore (a terrestrial or aquatic flesh-eating mammal) *"terrestrial carnivores have four or five clawed digits on each limb"*
            - S: (n) placental, placental mammal, eutherian, eutherian mammal (mammals having a placenta; all mammals except monotremes and marsupials)
              - S: (n) mammal, mammalian (any warm-blooded vertebrate having the skin more or less covered with hair; young are born alive except for the small subclass of monotremes and nourished with milk)
                - S: (n) vertebrate, craniate (animals having a bony or cartilaginous skeleton with a segmented spinal column and a large brain enclosed in a skull or cranium)
                  - S: (n) chordate (any animal of the phylum Chordata having a notochord or spinal column)
                    - S: (n) animal, animate being, beast, brute, creature, fauna (a living organism characterized by voluntary movement)
                      - S: (n) organism, being (a living thing that has (or can develop) the ability to act or function independently)
                        - S: (n) living thing, animate thing (a living (or once living) entity)
                          - S: (n) whole, unit (an assemblage of parts that is regarded as a single entity) *"how big is that part compared to the whole?"; "the team is a unit"*
                            - S: (n) object, physical object (a tangible and visible entity; an entity that can cast a shadow) *"it was full of rackets, balls and other objects"*
                              - S: (n) physical entity (an entity that has physical existence)
                                - S: (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

More Datasets….

UIUC Cars (2004)
S. Agarwal, A. Awan, D. Roth

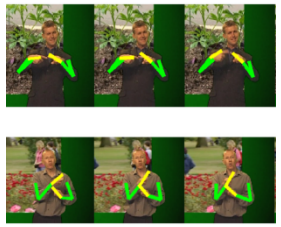
CMU/VASC Faces (1998)
H. Rowley, S. Baluja, T. Kanade

FERET Faces (1998)
P. Phillips, H. Wechsler, J. Huang, P. Raus

COIL Objects (1996)
S. Nene, S. Nayar, H. Murase

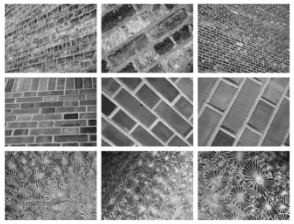
MNIST digits (1998-10)
Y LeCun & C. Cortes

KTH human action (2004)
I. Leptev & B. Caputo

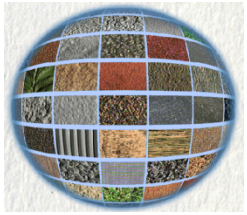


Sign Language (2008)
P. Buehler, M. Everingham, A. Zisserman

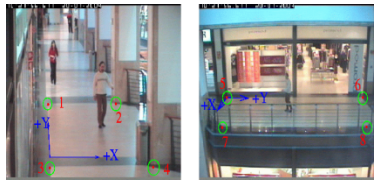
Segmentation (2001)
D. Martin, C. Fowlkes, D. Tal, J. Malik.



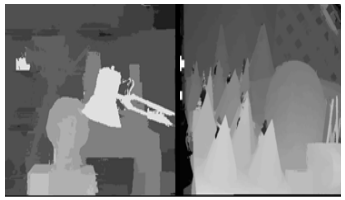
3D Textures (2005)
S. Lazebnik, C. Schmid, J. Ponce



CuRRET Textures (1999)
K. Dana B. Van Ginneken S. Nayar J. Koenderink



CAVIAR Tracking (2005)
R. Fisher, J. Santos-Victor J. Crowley



Middlebury Stereo (2002)
D. Scharstein R. Szeliski

Links to datasets

The next tables summarize some of the available datasets for training and testing object detection and recognition algorithms. These lists are far from exhaustive.

Databases for object localization

CMU/MIT frontal faces	vasc.ri.cmu.edu/idb/html/face/frontal_images cbcl.mit.edu/software-datasets/FaceData2.html	Patches	Frontal faces
Graz-02 Database	www.emt.tugraz.at/~pinz/data/GRAZ_02/	Segmentation masks	Bikes, cars, people
UIUC Image Database	l2r.cs.uiuc.edu/~cogcomp/Data/Car/	Bounding boxes	Cars
TU Darmstadt Database	www.vision.ethz.ch/leibe/data/	Segmentation masks	Motorbikes, cars, cows
LabelMe dataset	people.csail.mit.edu/brussell/research/LabelMe/intro.html	Polygonal boundary	>500 Categories

Databases for object recognition

Caltech 101	www.vision.caltech.edu/Image_Datasets/Caltech101/Caltech101.html	Segmentation masks	101 categories
Caltech 256	http://www.vision.caltech.edu/Image_Datasets/Caltech256/	Bounding Box	256 Categories
COIL-100	www1.cs.columbia.edu/CAVE/research/softlib/coil-100.html	Patches	100 instances
NORB	www.cs.nyu.edu/~ylclab/data/norb-v1.0/	Bounding box	50 toys

On-line annotation tools

ESP game	www.espgame.org	Global image descriptions	Web images
LabelMe	people.csail.mit.edu/brussell/research/LabelMe/intro.html	Polygonal boundary	High resolution images

Collections

PASCAL	http://www.pascal-network.org/challenges/VOC/	Segmentation, boxes	various