

CSC2503: Foundations of Computer Vision

Object Recognition

Slides are modified from the excellent course notes and tutorials by Antonio Torralba, Fei-Fei Li and Rob Fergus.

<http://people.csail.mit.edu/torralba/cvpr2007/>

Where do we go from here?

- ▶ Single class recognition
- ▶ Multi-class recognition
- ▶ Scene Recognition and Context
- ▶ Parsing, Recognition and Segmentation

Multi-class category recognition

Does the approach to single object/category recognition scale?

How many categories are there?

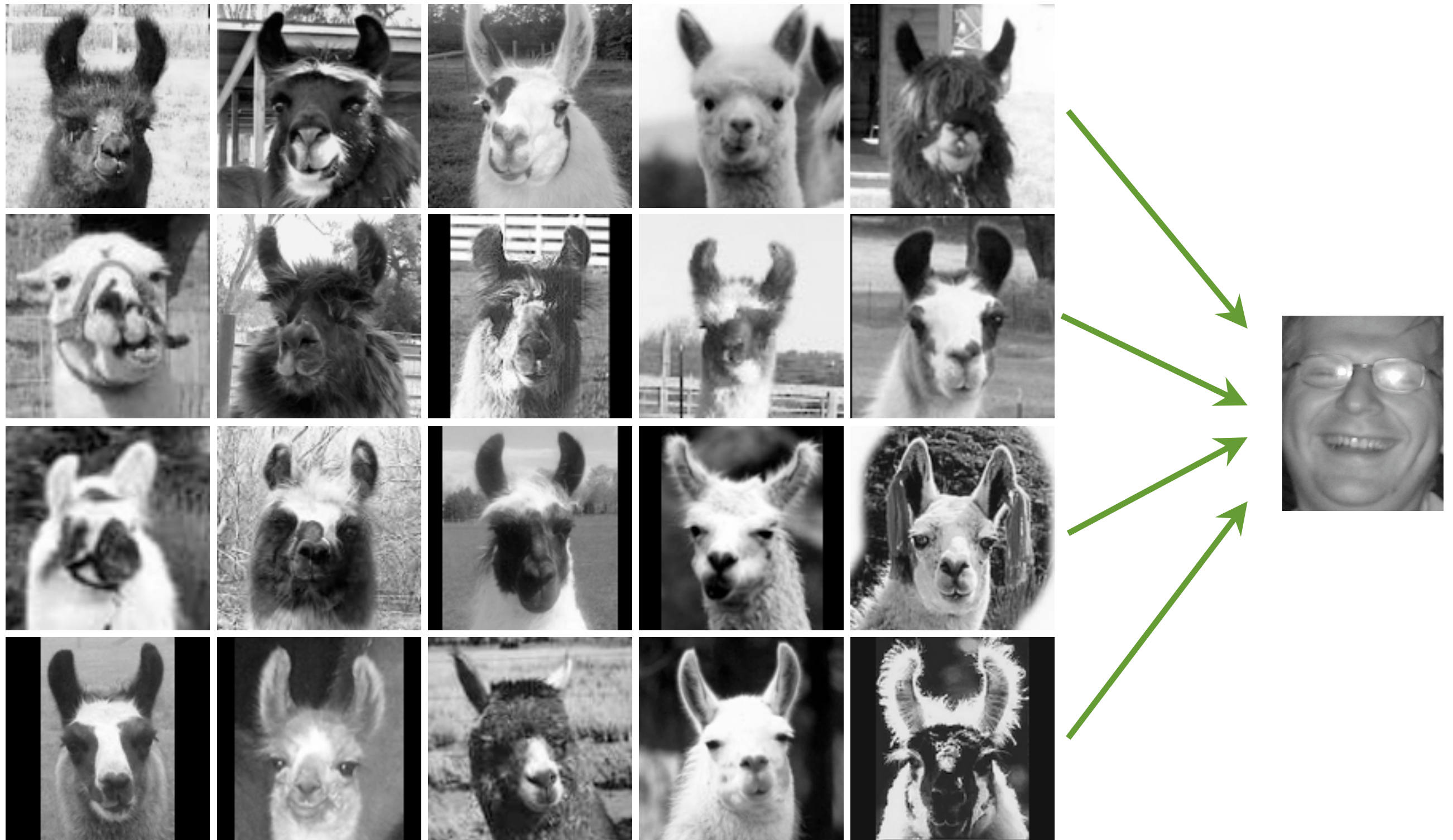
“Muchas”



How many categories are there?



Shared features and transfer learning



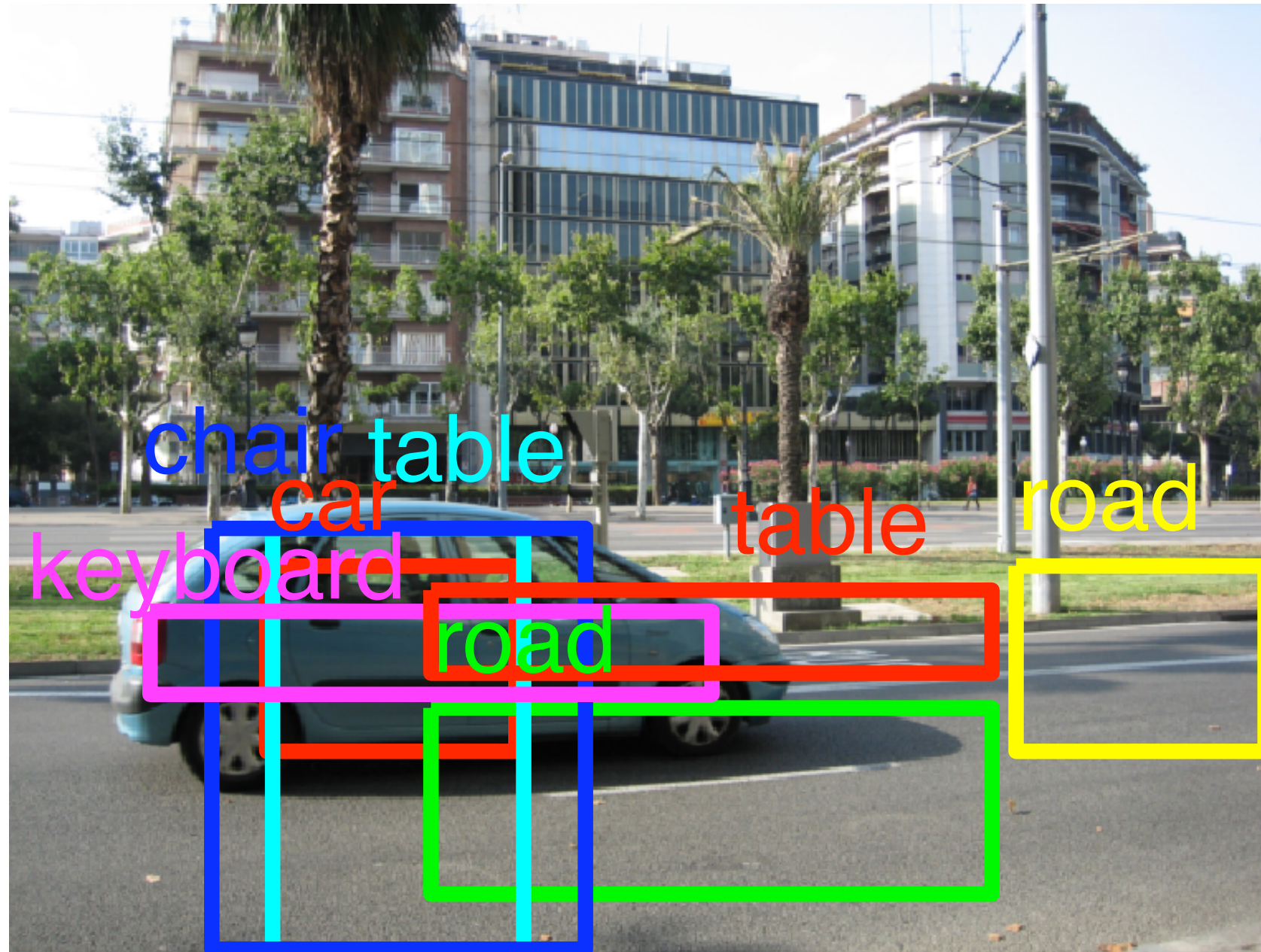
Can we transfer knowledge from one object category to another?

Scene recognition and context



Is local information enough?

Scene recognition and context



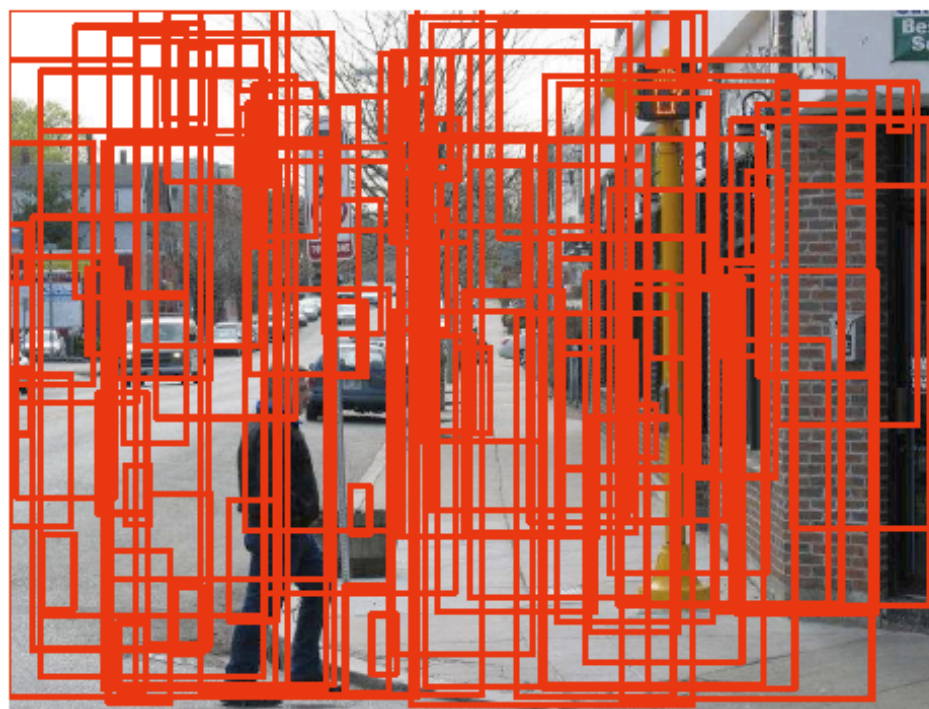
If we have 1000 categories (detectors), and each detector produces 1 false alarm every 10 images, we will have 100 false alarms per image... pretty much garbage...



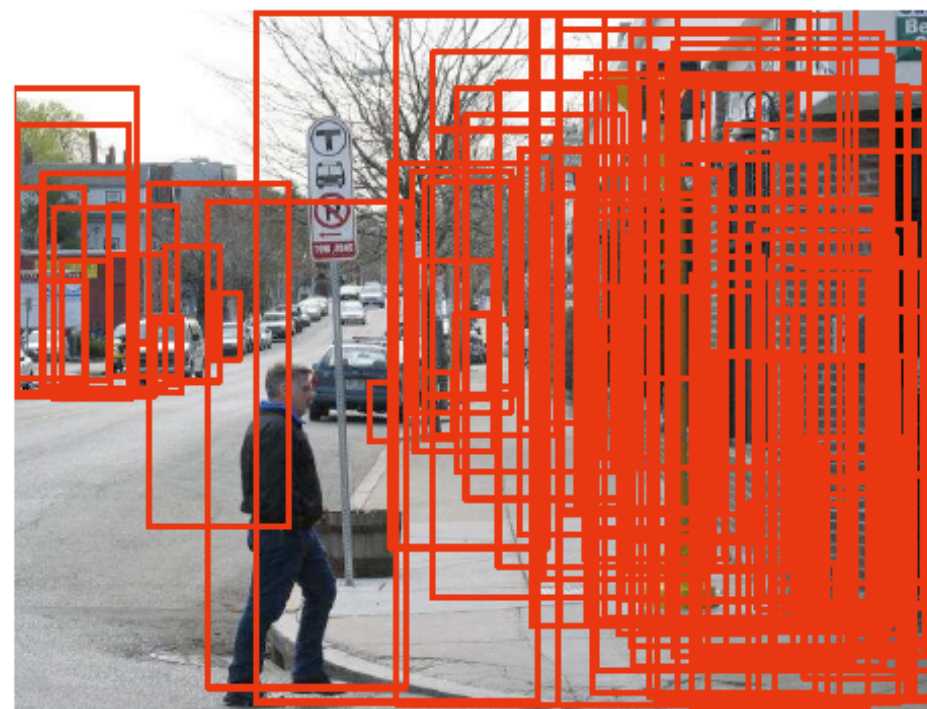


Geometric context

[Hoiem, Efros and Hebert, 2006]



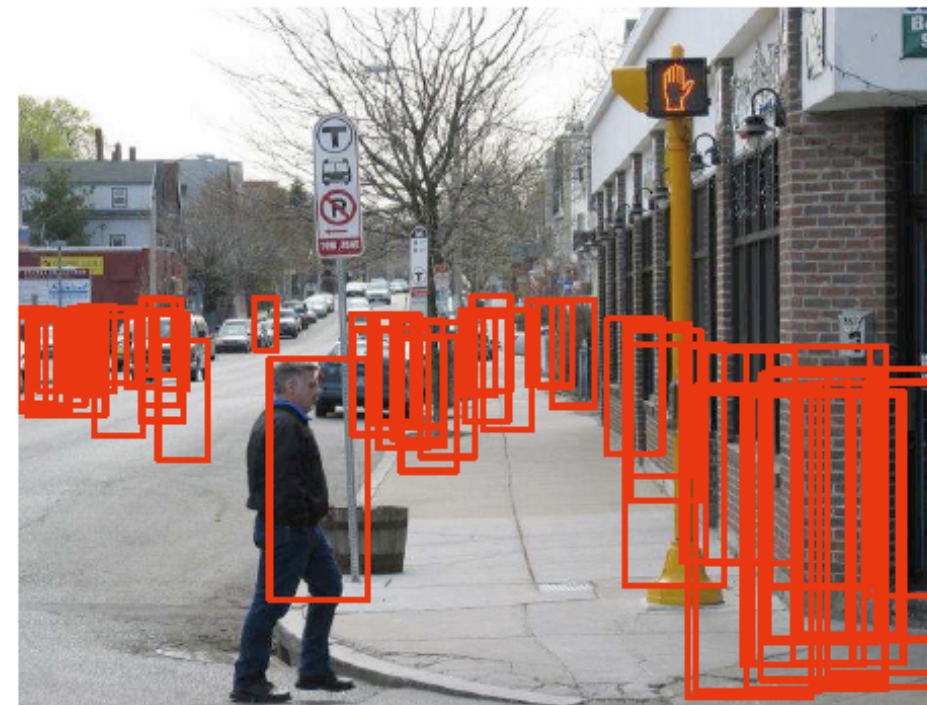
(b) $P(\text{person}) = \text{uniform}$



(d) $P(\text{person} \mid \text{geometry})$



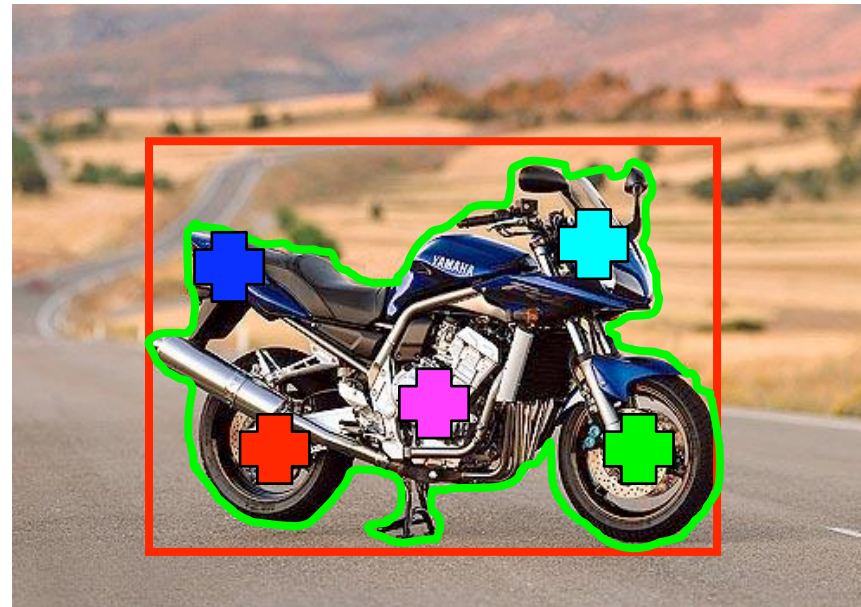
(f) $P(\text{person} \mid \text{viewpoint})$



(g) $P(\text{person} \mid \text{viewpoint, geometry})$

Parsing, recognition and segmentation

Contains a motorbike



Datasets

Language 10^6 samples

Character Recognition (MNIST) 10^4 samples

Visual Objec 10^3 samples



3	6	8	1	7	9	6	6	9	1
6	7	5	7	8	6	3	4	8	5
2	1	7	9	7	1	2	8	4	5
4	8	1	9	0	1	8	8	9	4
7	6	1	8	6	4	1	5	6	0
7	5	9	2	6	5	8	1	9	7
2	2	2	2	2	3	4	4	8	0
0	2	3	8	0	7	3	8	5	7
0	1	4	6	4	6	0	2	4	3
7	1	2	8	7	6	9	8	6	1

Coil

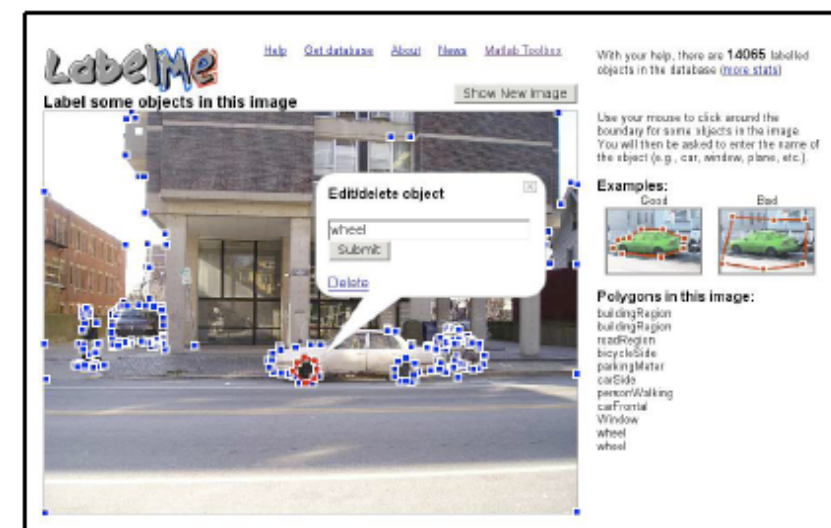
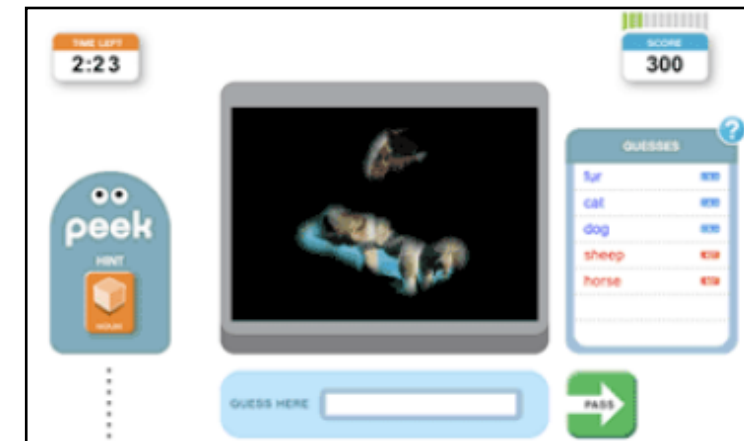


The Columbia Object Image Library (COIL-100): colour images of 100 objects taken at pose intervals of 5 degrees (72 poses per object).

[S. A. Nene, S. K. Nayar & H. Murase, TR: CUUCS-006-96, 1996]

Collecting 10^{6-7} Examples

- ▶ ESP game (CMU)
Luis Von Ahn and Laura Dabbish 2004
- ▶ LabelMe (MIT)
Russell, Torralba, Freeman, 2005
- ▶ StreetScenes (CBCL-MIT)
Bileschi, Poggio, 2006
- ▶ WhatWhere (Caltech)
Perona et al, 2007
- ▶ PASCAL challenge
2006, 2007, 2008, 2009, ...
- ▶ Lotus Hill Institute
Song-Chun Zhu et al 2007



Labeling with games



Figure 1. Partners agreeing on an image in the ESP Game. Neither player can see the other's guesses.

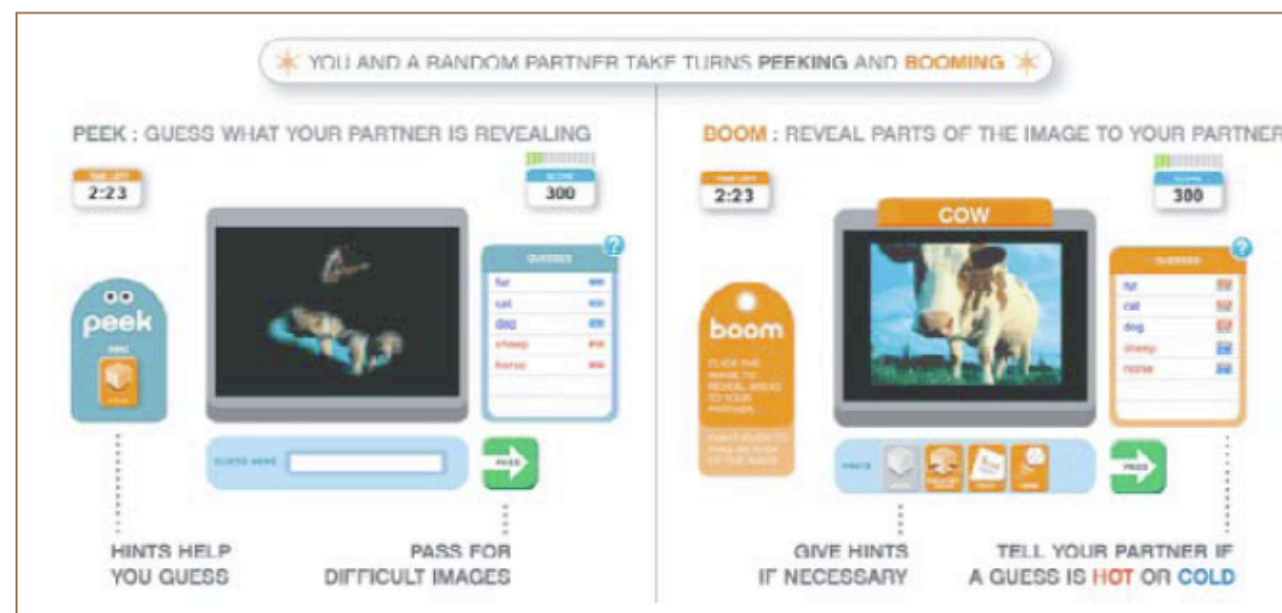


Figure 2. Peekaboom. "Peek" tries to guess the word associated with an image slowly revealed by "Boom."

Pascal Visual Objects Challenge

20 object classes selected are:

Person: person

Animal: bird, cat, cow, dog, horse, sheep

Vehicle: aeroplane, bicycle, boat, bus, car, motorbike, train

Indoor: bottle, chair, dining table, potted plant, sofa, tv/monitor



M. Everingham, Luc van Gool , C. Williams, J. Winn, A. Zisserman 2007



Please [contact us](#) if you find any bugs or have any suggestions.



[Show me another image](#)

Label as many objects and regions as you can in this image



[Sign in](#) ([why?](#))

With your help, there are **91348** labelled objects in the database ([more stats](#))

Instructions ([Get more help](#))

Use your mouse to click around the boundary of some objects in this image. You will then be asked to enter the name of the object (examples: car, window).



Labeling tools



Polygons in this image ([XML](#))

- [door](#)
- [door](#)
- [road](#)
- [stair](#)
- [window](#)
- [window](#)
- [sidewalk](#)
- [building region](#)
- [house](#)
- [window](#)
- [window](#)
- [window](#)

Went online July 1st, 2005 (290,000+ object annotations)
B. Russell, A. Torralba, K. Murphy, W.T. Freeman. IJCV '08
Labelme.csail.mit.edu

LabelMe: Polygon quality



LabelMe: Not all data is reliable



Most common labels:

test

adksdsa

woiieie

...

LabelMe: Online hooligans

LabelMe Please [contact us](#) if you find any bugs or have any suggestions. [Show me another image](#)

Label as many objects and regions as you can in this image

Sign in (why?)

There are **158302** labelled objects

Instructions ([Get more help](#))

Use your mouse to click around the boundary of some objects in this image. You will then be asked to enter the name of the object (examples: car, window).

Good Bad

Labeling tools

Erase segment Zoom Fit Image

Polygons in this image (XML)

[Benen](#)
[bovenlichaam](#)
[hoofd](#)
[haar](#)
[oog1](#)
[oog2](#)
[towel](#)

