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

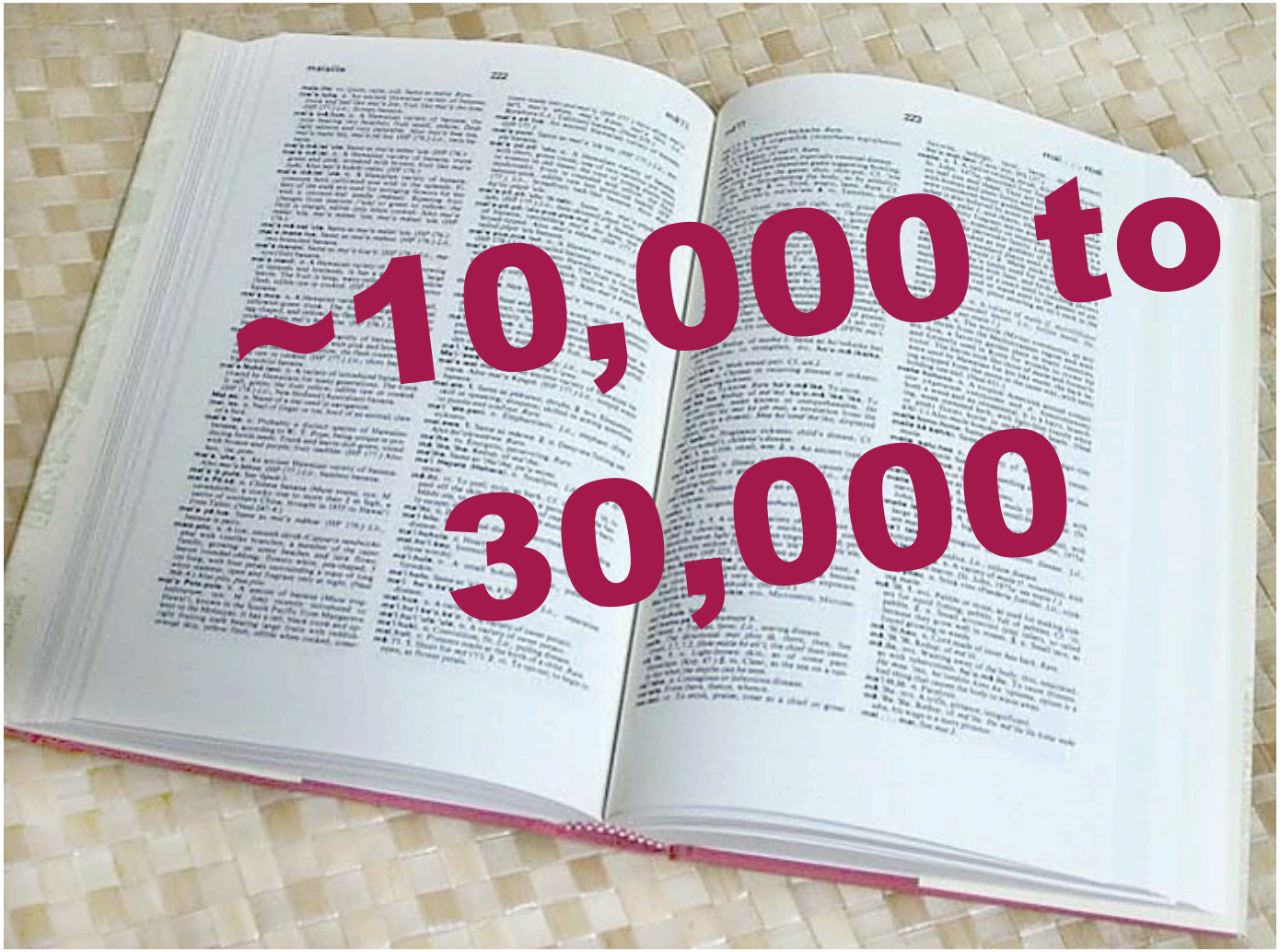
Does the approach to single object/category recognition scale?

How many categories are there?

“Muchas”

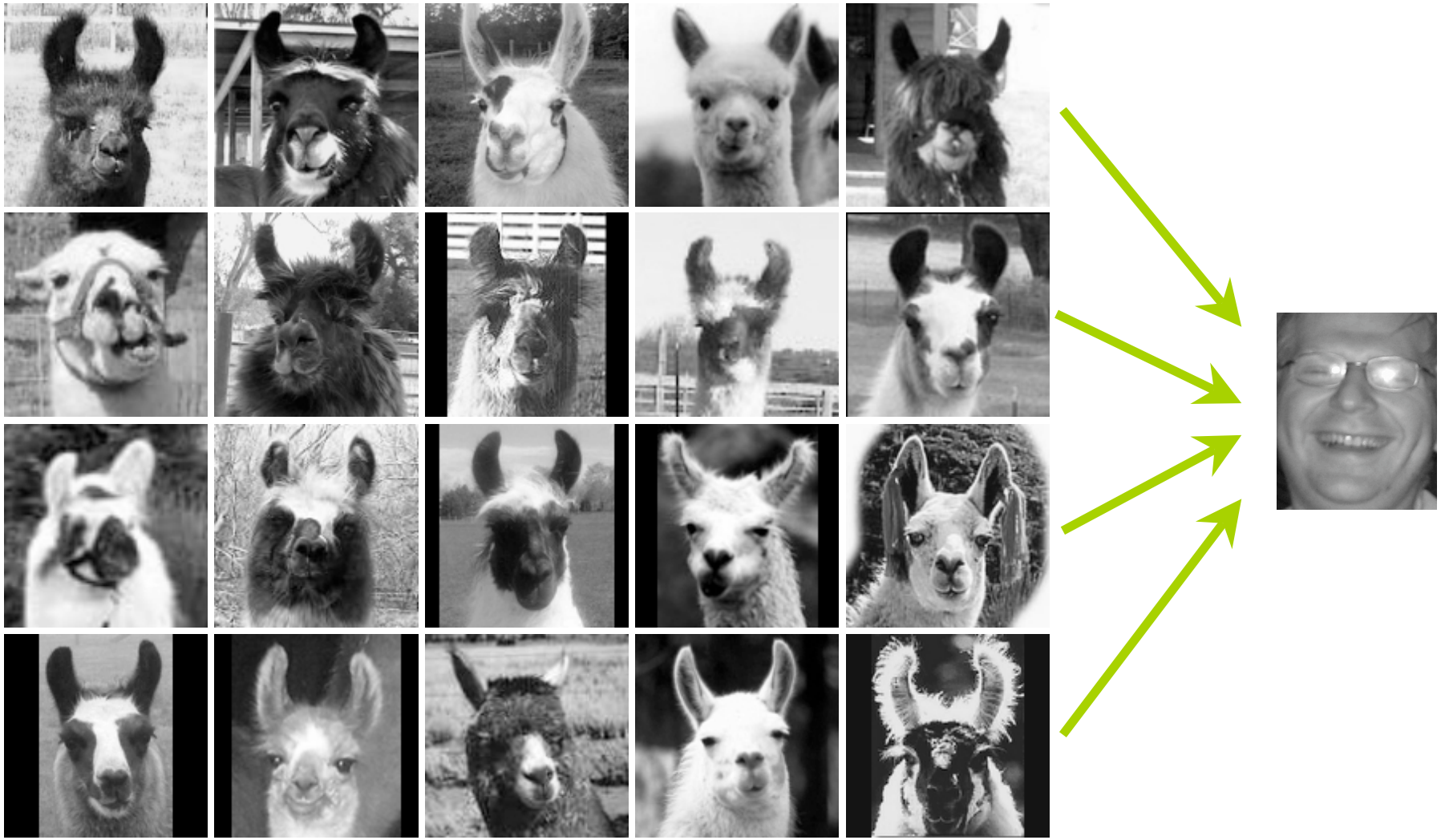
Slide by Aude Oliva

How many categories are there?



Biederman 1987

Shared features and transfer learning



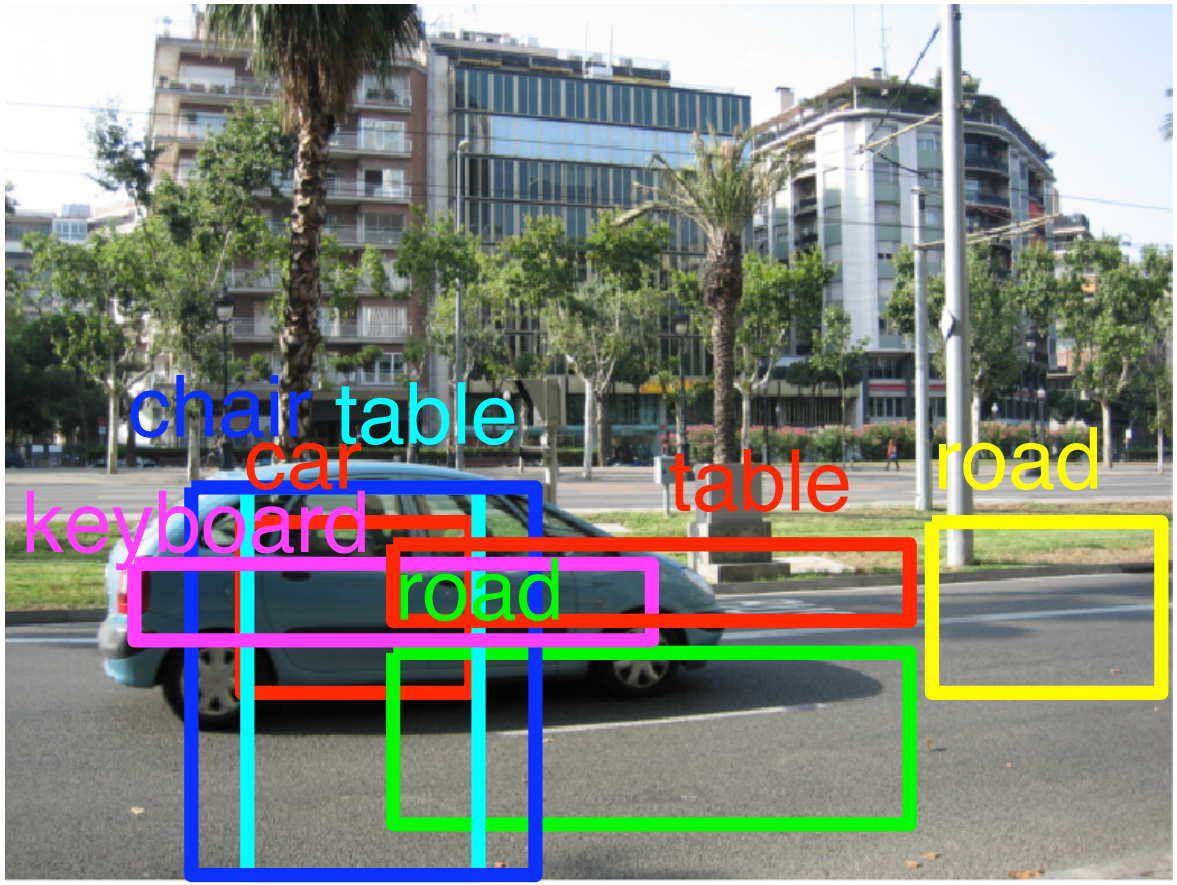
Can we transfer knowledge from one object category to another?

Slide by Erik Sudderth

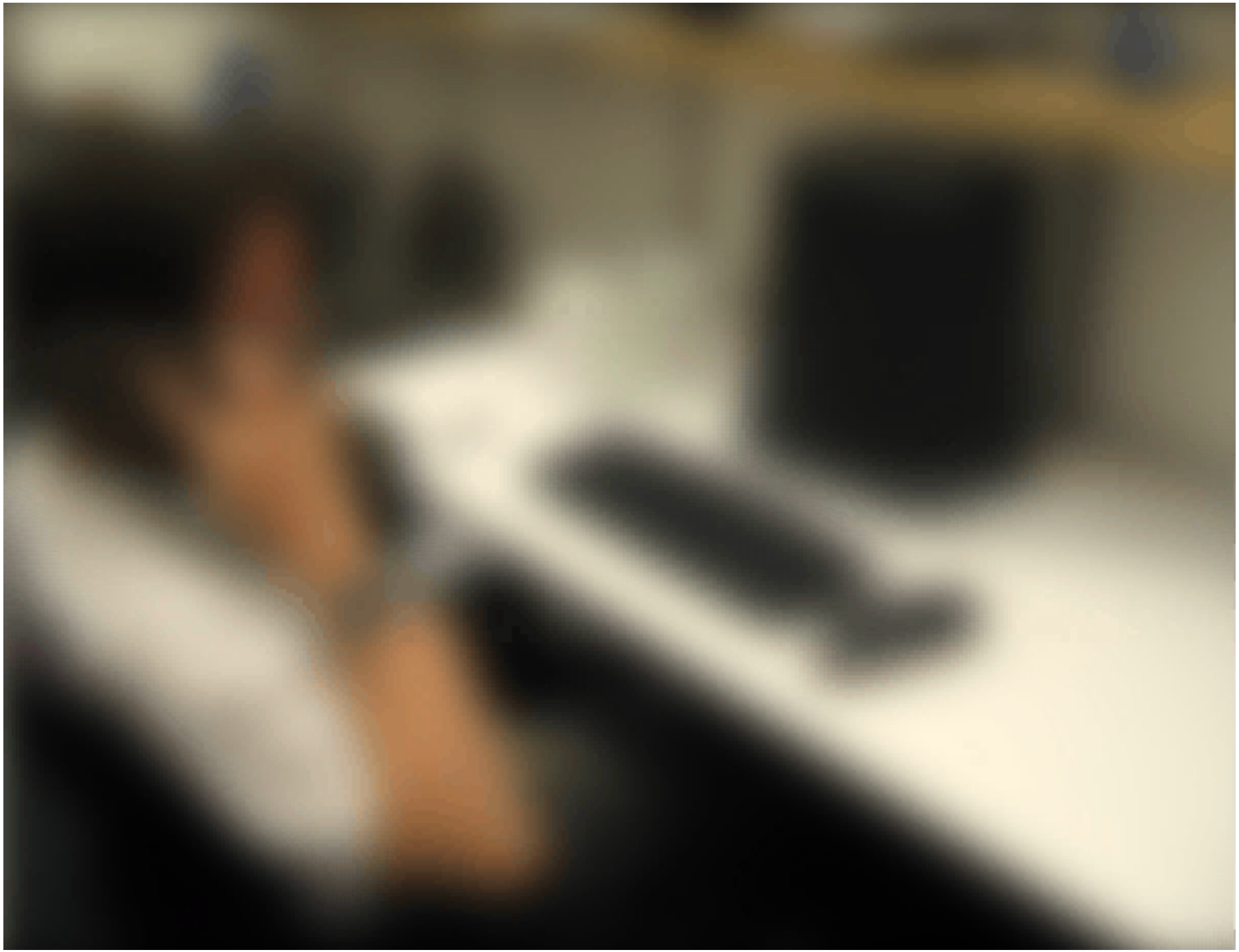
Scene recognition and context

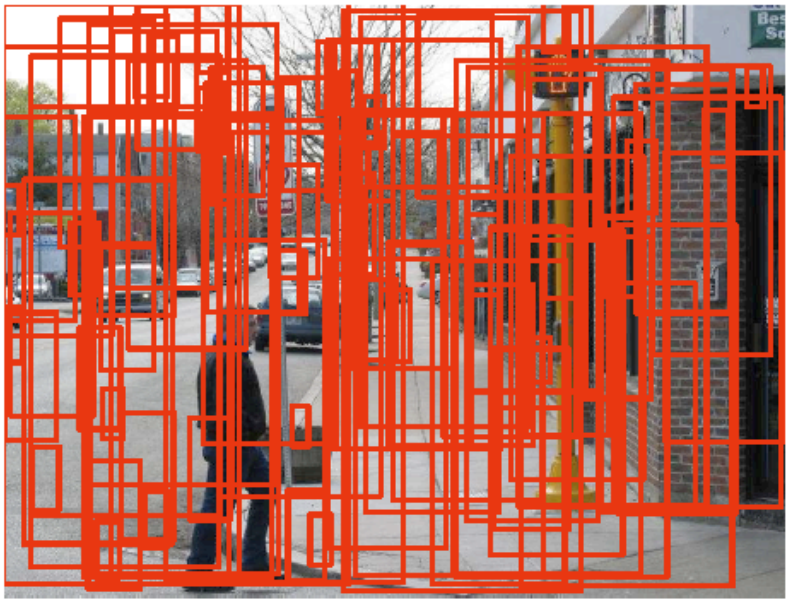
Is local information enough?

Scene recognition and context

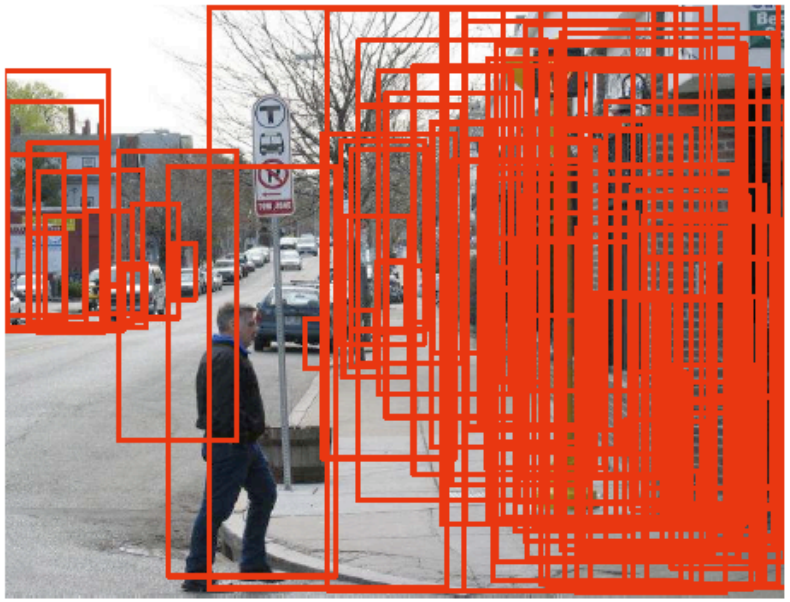


If we have 1000 categories (detectors), and each detector produces 1 false alarm every 10 images, we will have 100 false alarms per image… pretty much garbage…



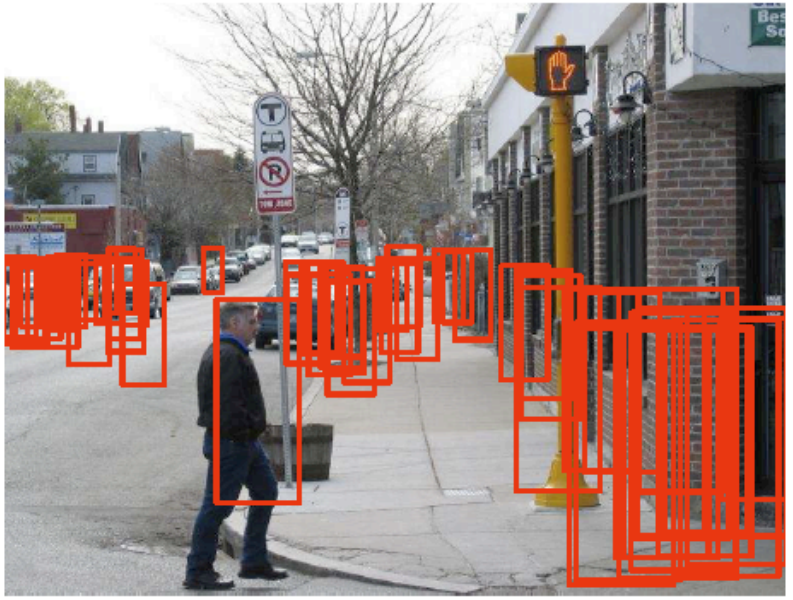


(b) P(person) = uniform



(d) P(person | geometry)

(f) P(person | viewpoint)

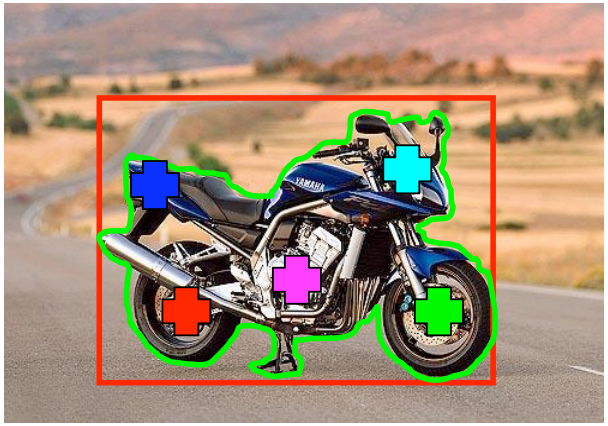


(g) P(person|viewpoint,geometry)

Hoiem, Efros, Herbert, 2006

Parsing, recognition and segmentation

Contains a motorbike



Datasets

Language 10^6 samples

Character Recognition (MNIST) 10^4 samples

Visual Objec 0^3 samples

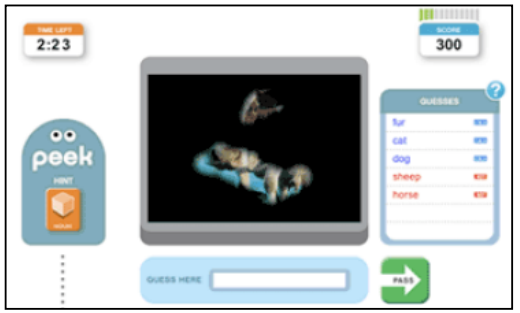
Coil

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: CUCS-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

- 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.*

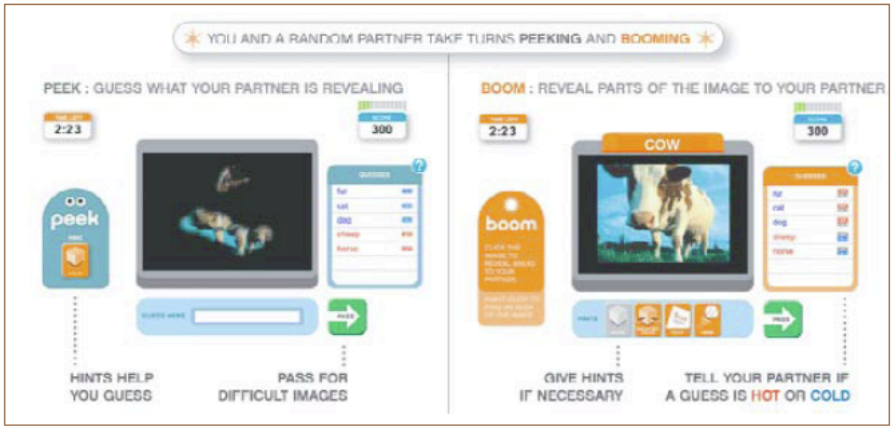


Figure 2. Peekaboom. *"Peek" tries to guess the word associated with an image slowly revealed by "Boom."*

L. von Ahn, L. Dabbish, 2004; L. von Ahn, R. Liu and M. Blum, 2006

Pascal Visual Objects Challenge

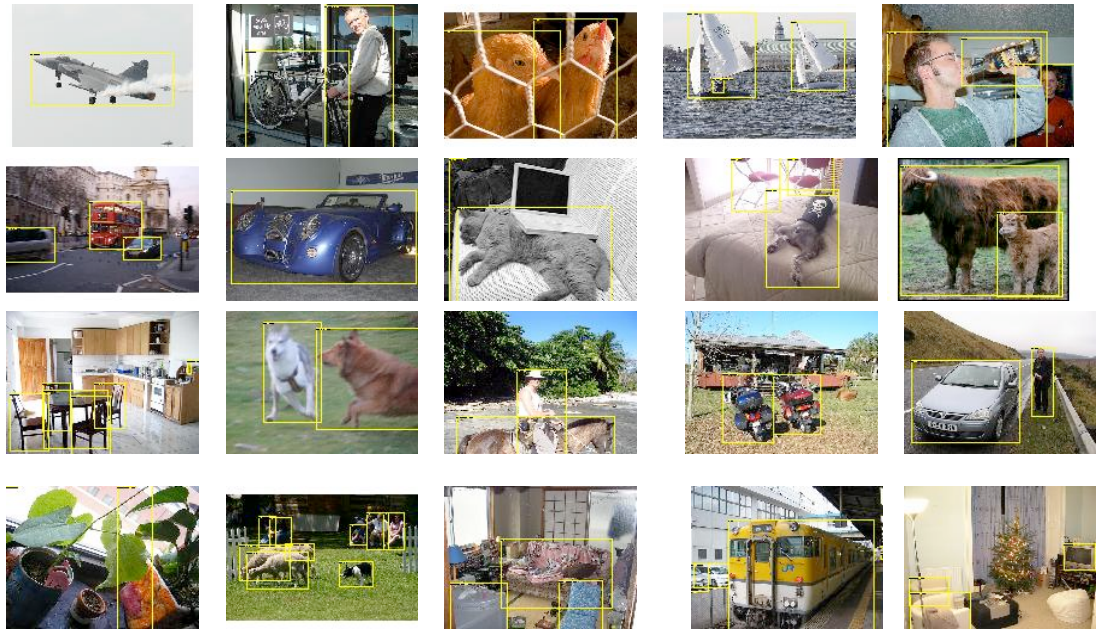
Twenty object classes selected are:

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

Went online July 1st, 2005 (290,000+ object annotations

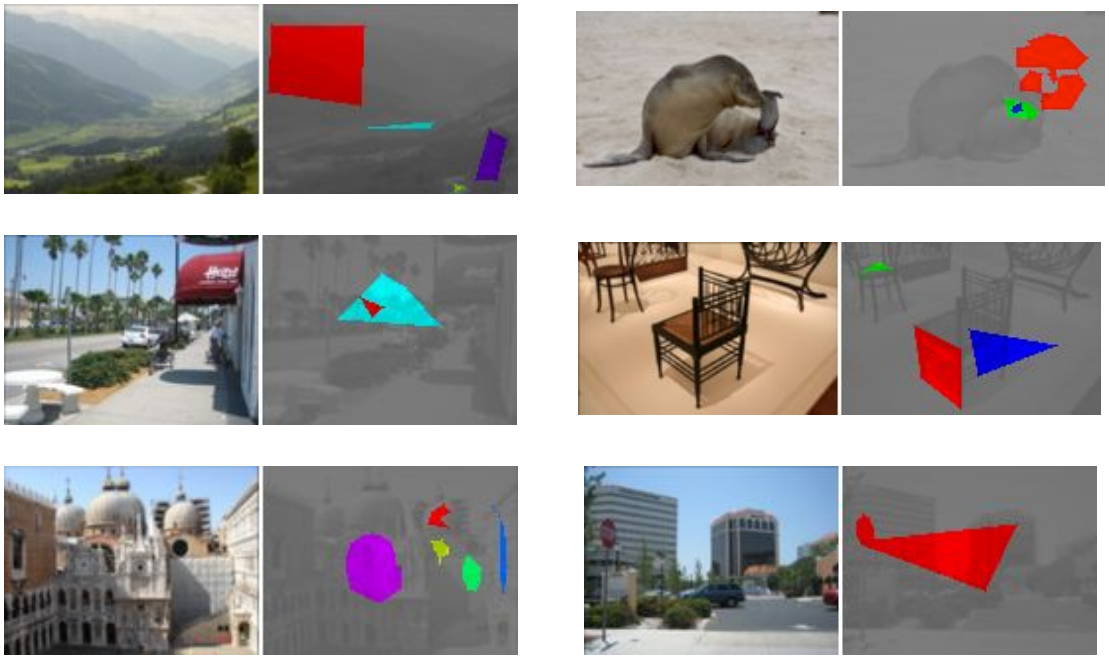
B. Russell, A. Torralba, K. Murphy, W.T. Freeman. IJCV ‘08

Labelme.csail.mit.edu

LabelMe: Polygon quality

… things do not always look good…

LabelMe: Polygon quality

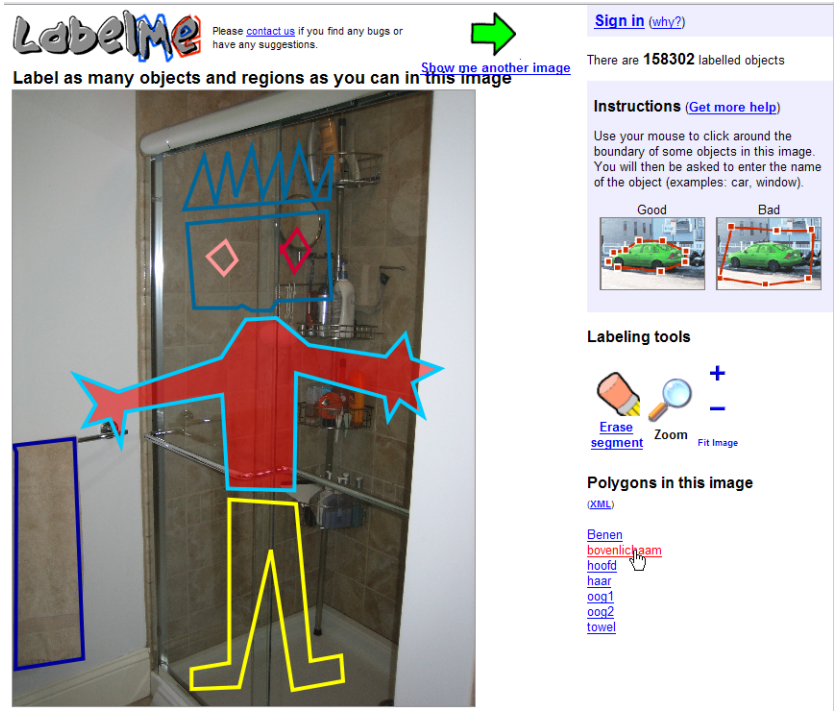
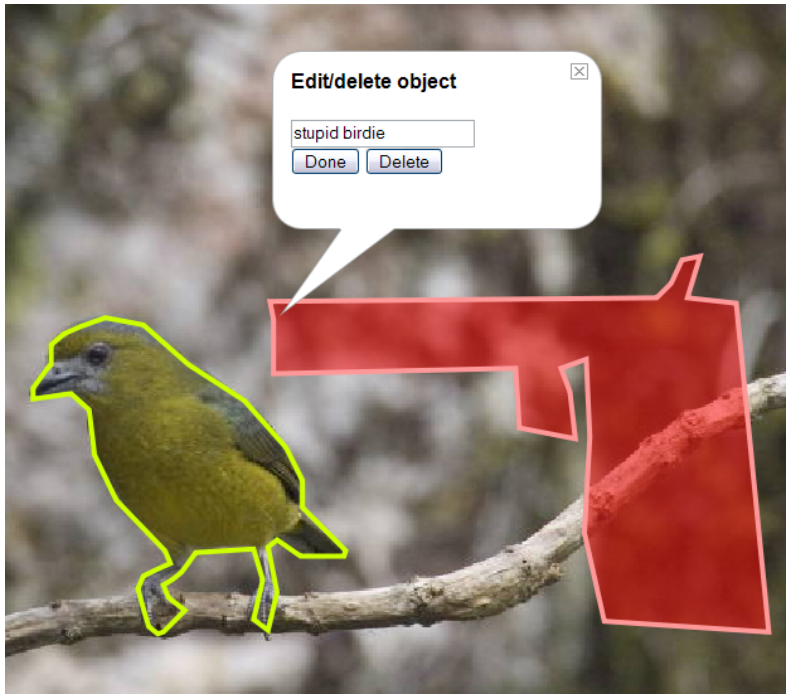


Most common labels:

test

adksdsa

woiieiie

…

LabelMe: Online hooligans