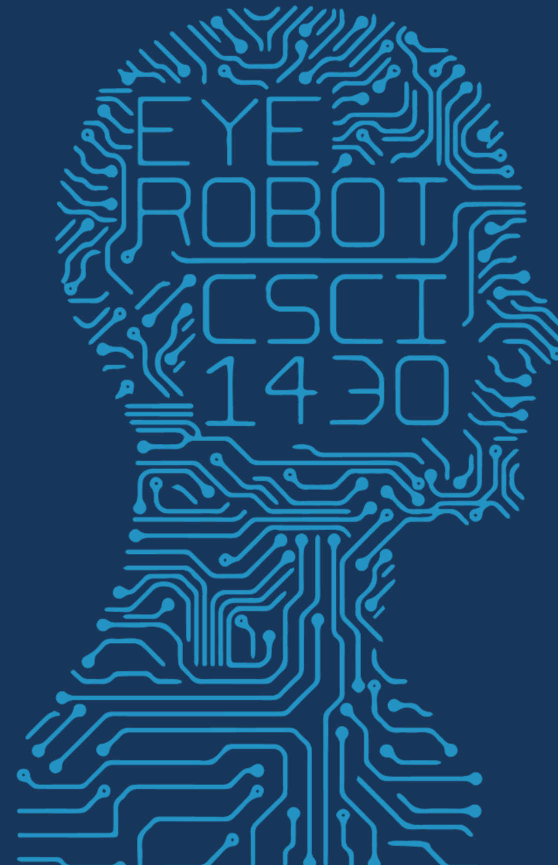
1950

Future Vision



2017 MWF 1pm

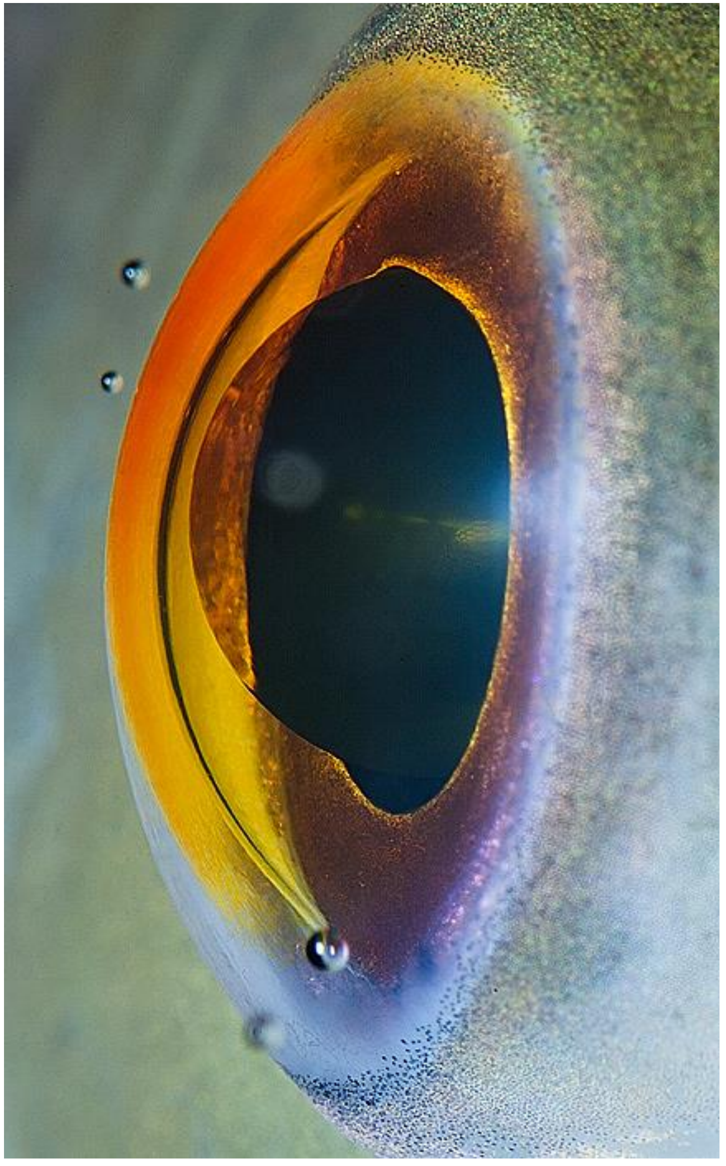
Computer Vision

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116



By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

Indy

Heterochromia iridum

From Wikipedia, the free encyclopedia

Not to be confused with Heterochromatin or Dichromatic (disambiguation).

In anatomy, **heterochromia** (ancient Greek: ἕτερος, *héteros*, different + χρῶμα, *chróma*, color[1]) is a difference in coloration, usually of the iris but also of hair or skin. Heterochromia is a result of the relative excess or lack of melanin (a pigment). It may be inherited, or caused by genetic mosaicism, chimerism, disease, or injury.[2]

Heterochromia of the eye (***heterochromia iridis*** or ***heterochromia iridum***) is of three kinds. In *complete heterochromia*, one iris is a different color from the other. In *sectoral heterochromia*, part of one iris is a different color from its remainder and finally in "central heterochromia" there are spikes of different colours radiating from the pupil.

Heterochromia

Complete heterochromia in human eyes: one brown and one green/hazel

Classification and external resources

Specialty	ophthalmology
ICD-10	Q13.2, H20.8, L67.1
ICD-9-CM	364.53
OMIM	142500
DiseasesDB	31289

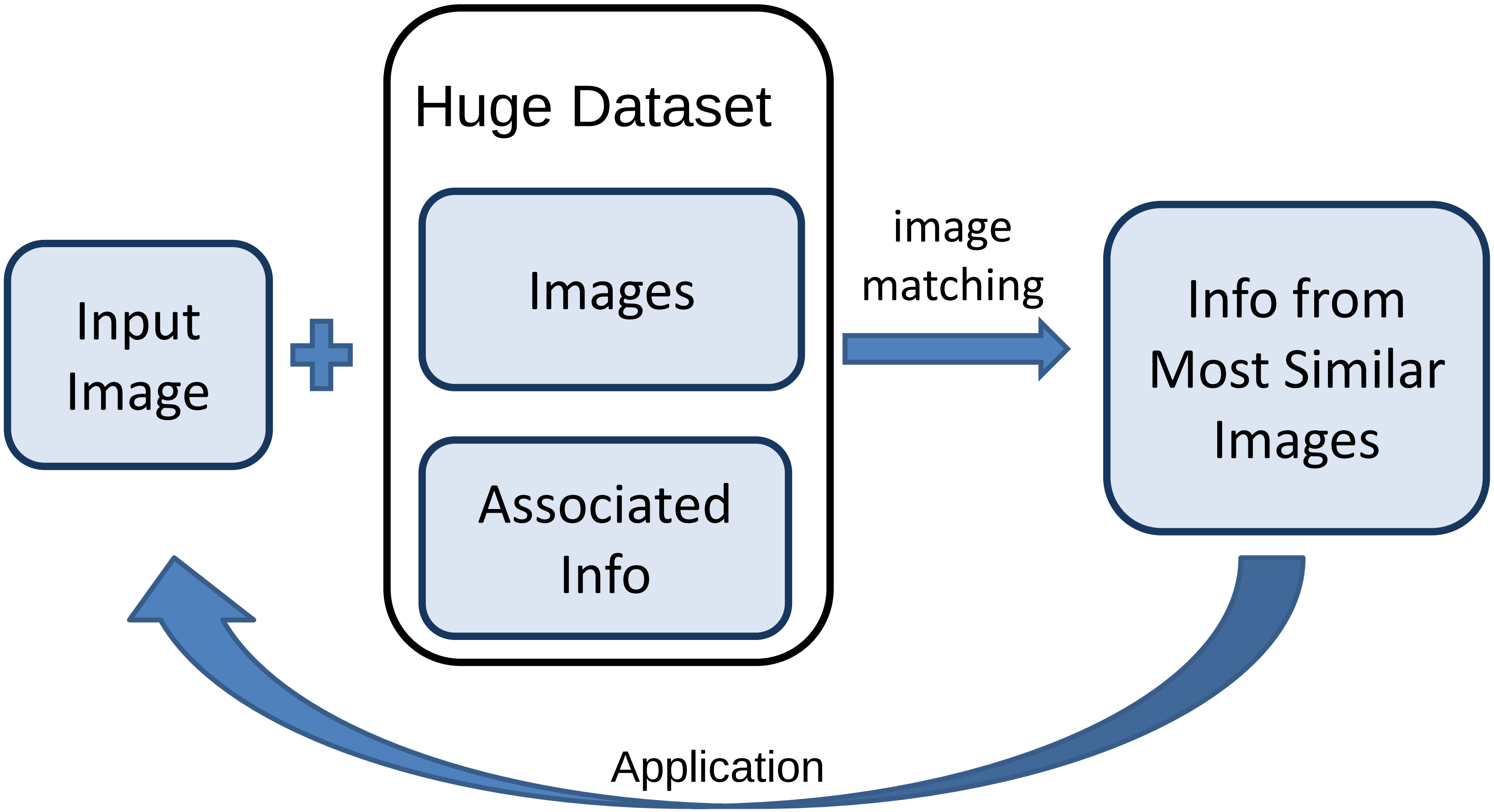
Project 4 is training and fine-tuning CNNs.

MATLAB can do this, but it’s not a common practice.

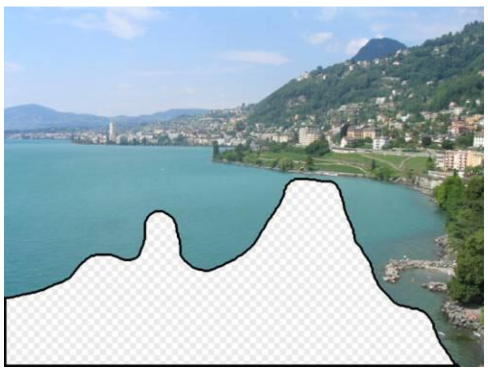
For more real-world experience, we will use TensorFlow – a Python package.

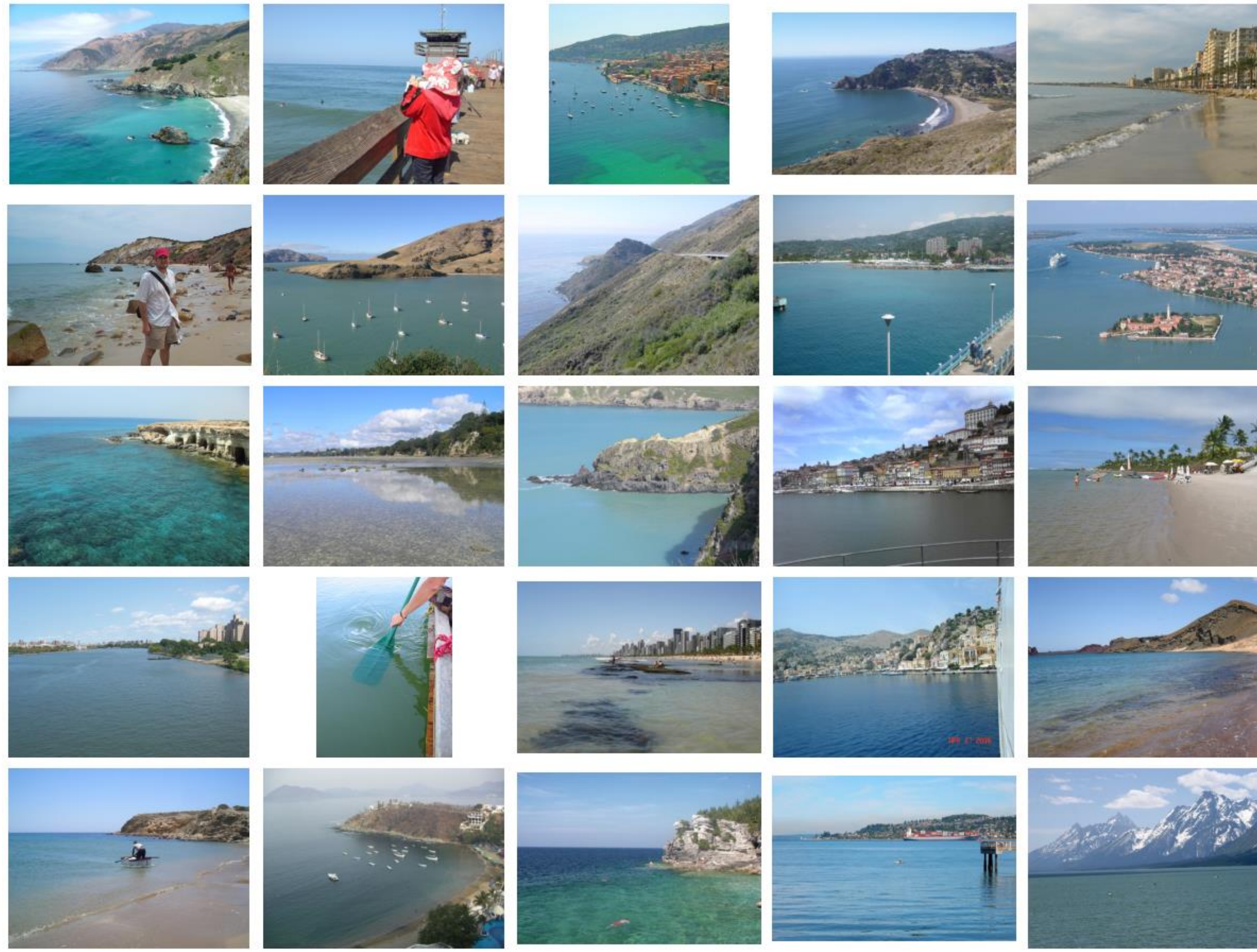
So…how’s your Python?

General Principal



Hopefully, If you have enough images, the dataset will contain very similar images that you can find with simple matching methods.





… 200 total

Graph cut + Poisson blending

How much can an image tell about its geographic location?

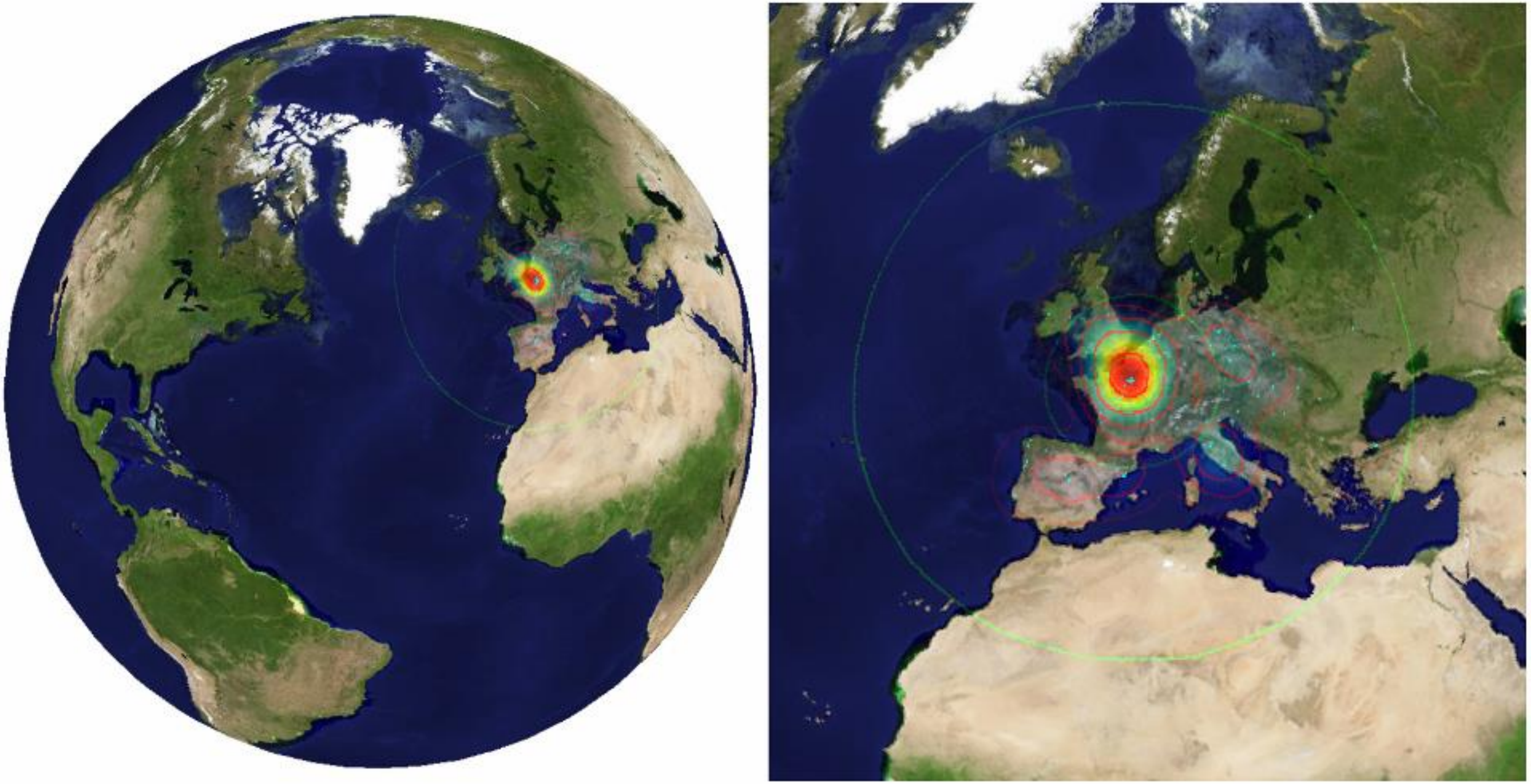
How much can an image tell about its geographic location?

6 million geo-tagged Flickr images

http://graphics.cs.cmu.edu/projects/im2gps/

im2gps (Hays & Efros, CVPR 2008)

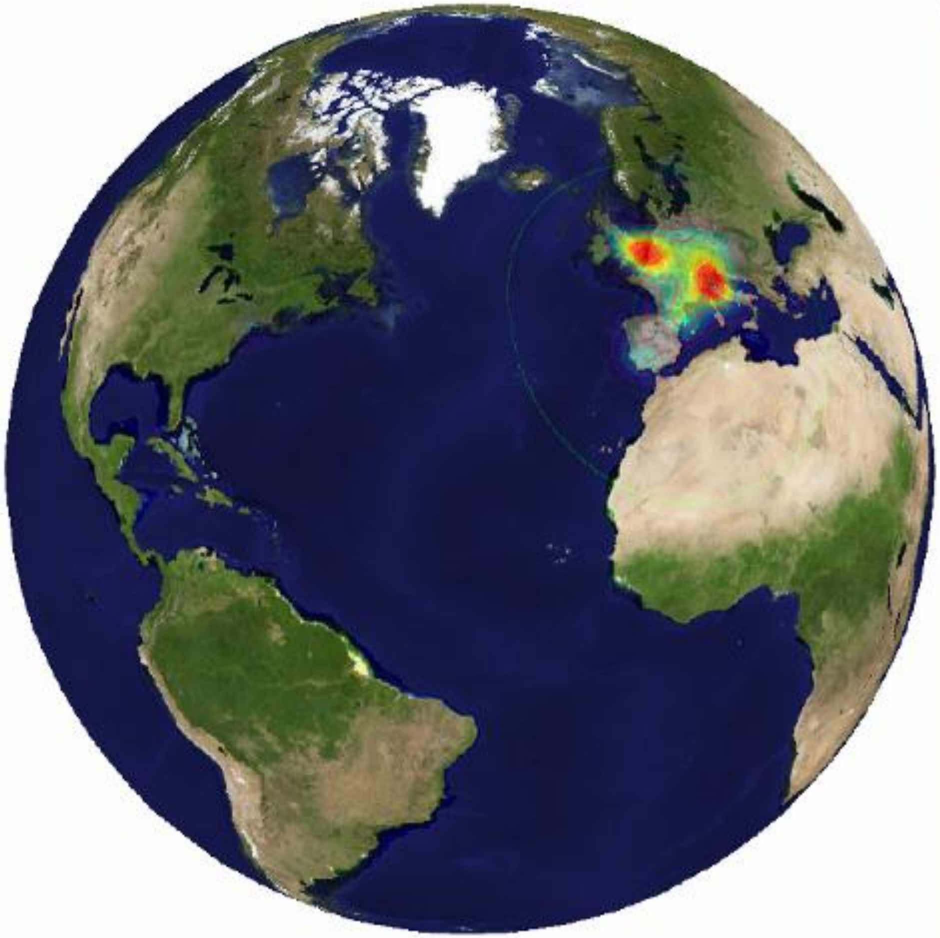
Nearest Neighbors according to gist + bag of SIFT + color histogram + a few others

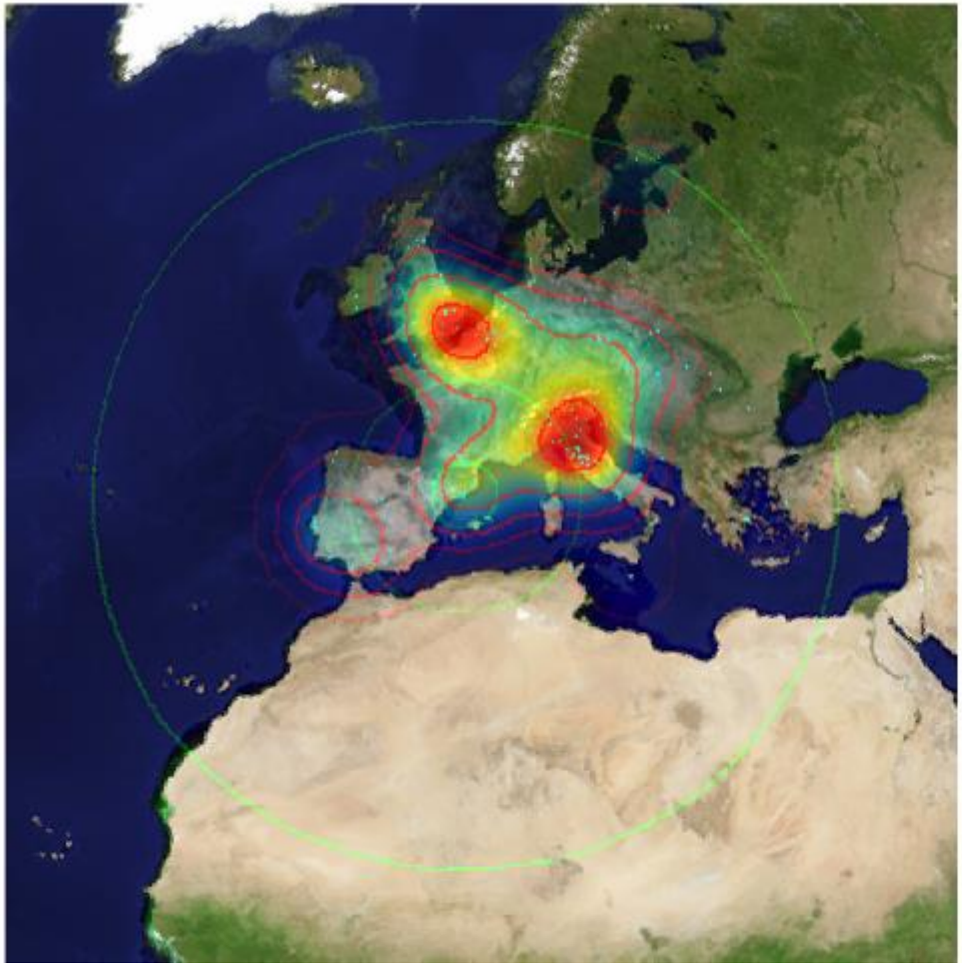


Im2gps

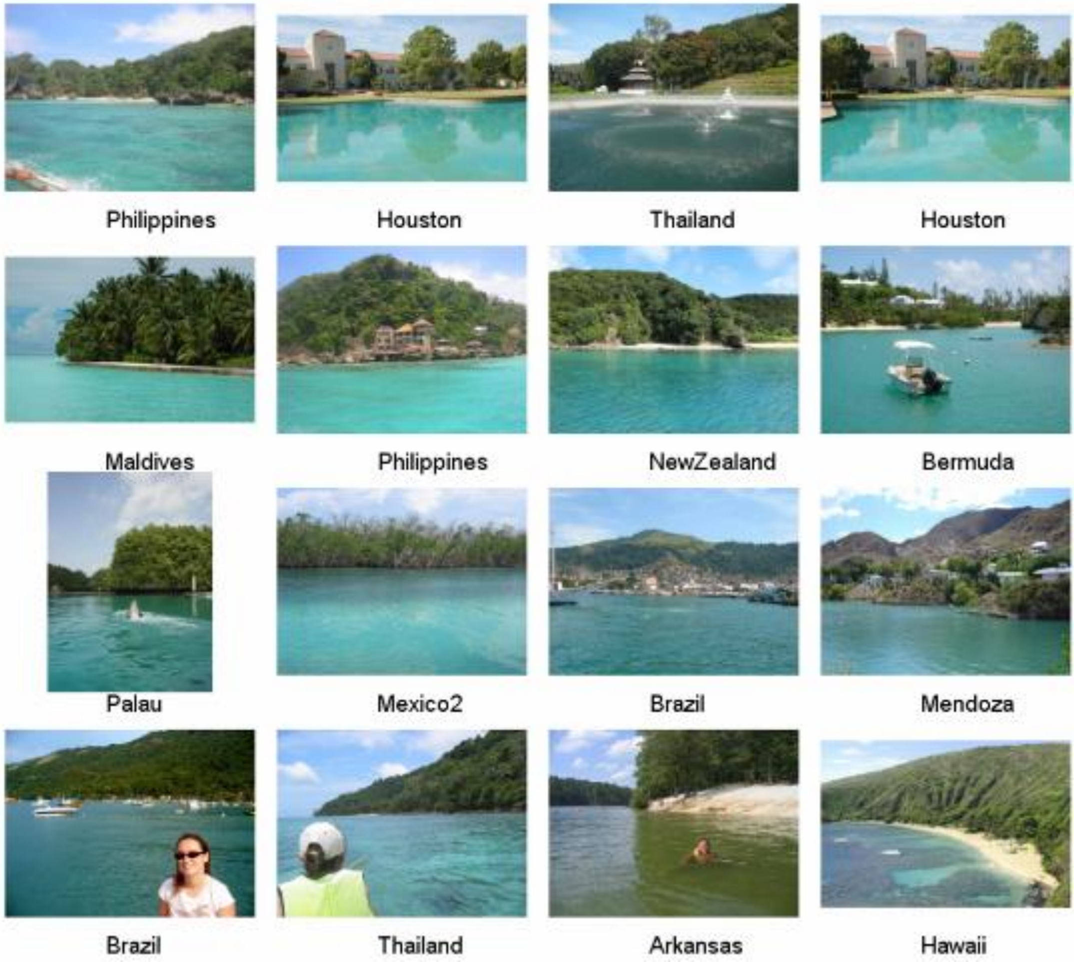
Example Scene Matches

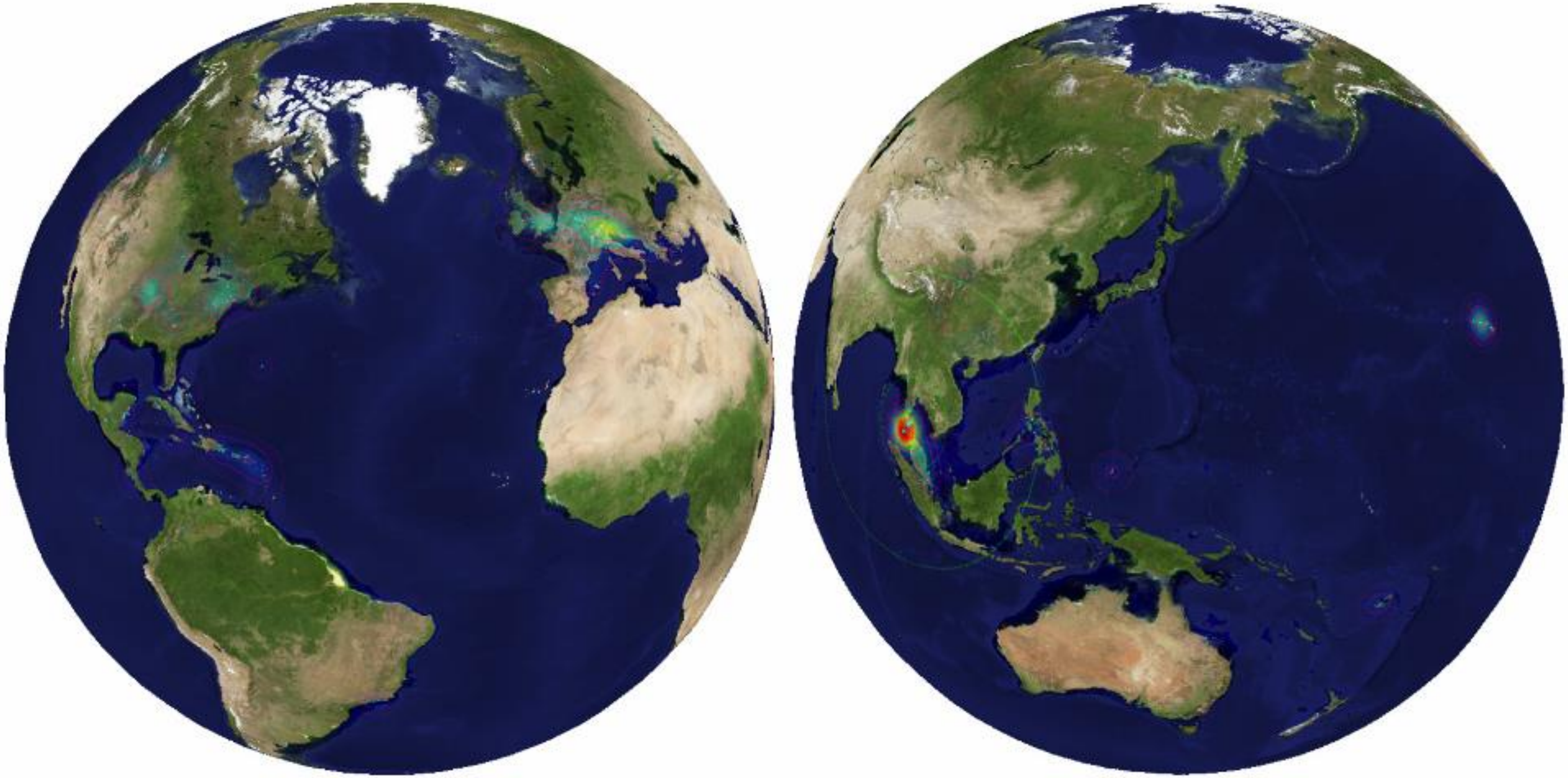
Voting Scheme



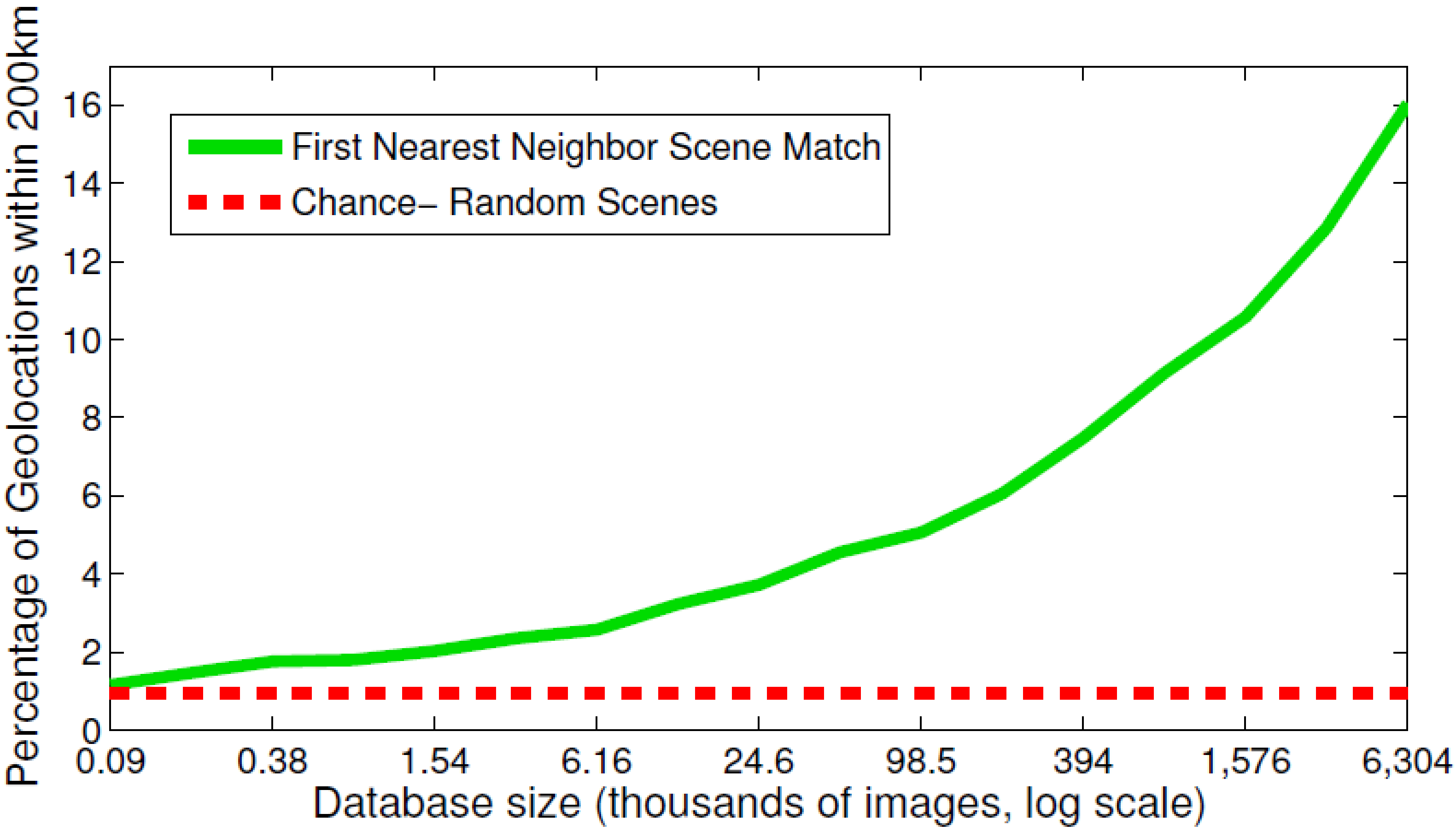


im2gps


Philippines
Houston
Thailand
Houston
Maldives
Philippines
NewZealand
Bermuda
Palau
Mexico2
Brazil
Mendoza
Brazil
Thailand
Arkansas
Hawaii



Effect of Dataset Size

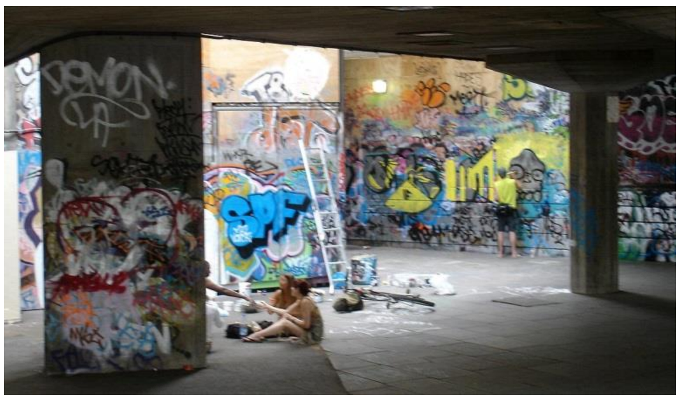


Where is This?

[Vesselova, Kalogerakis, Hertzmann, Hays, Efros. Image Sequence Geolocation. ICCV’09]

Where is This?

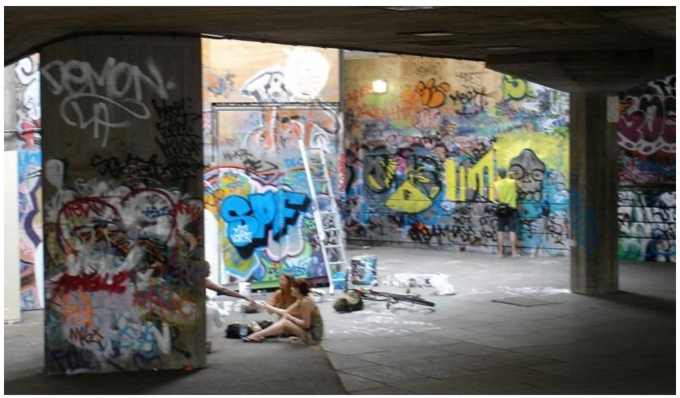
Where are These?



15:14,
June 18th, 2006

16:31,
June 18th, 2006

Where are These?



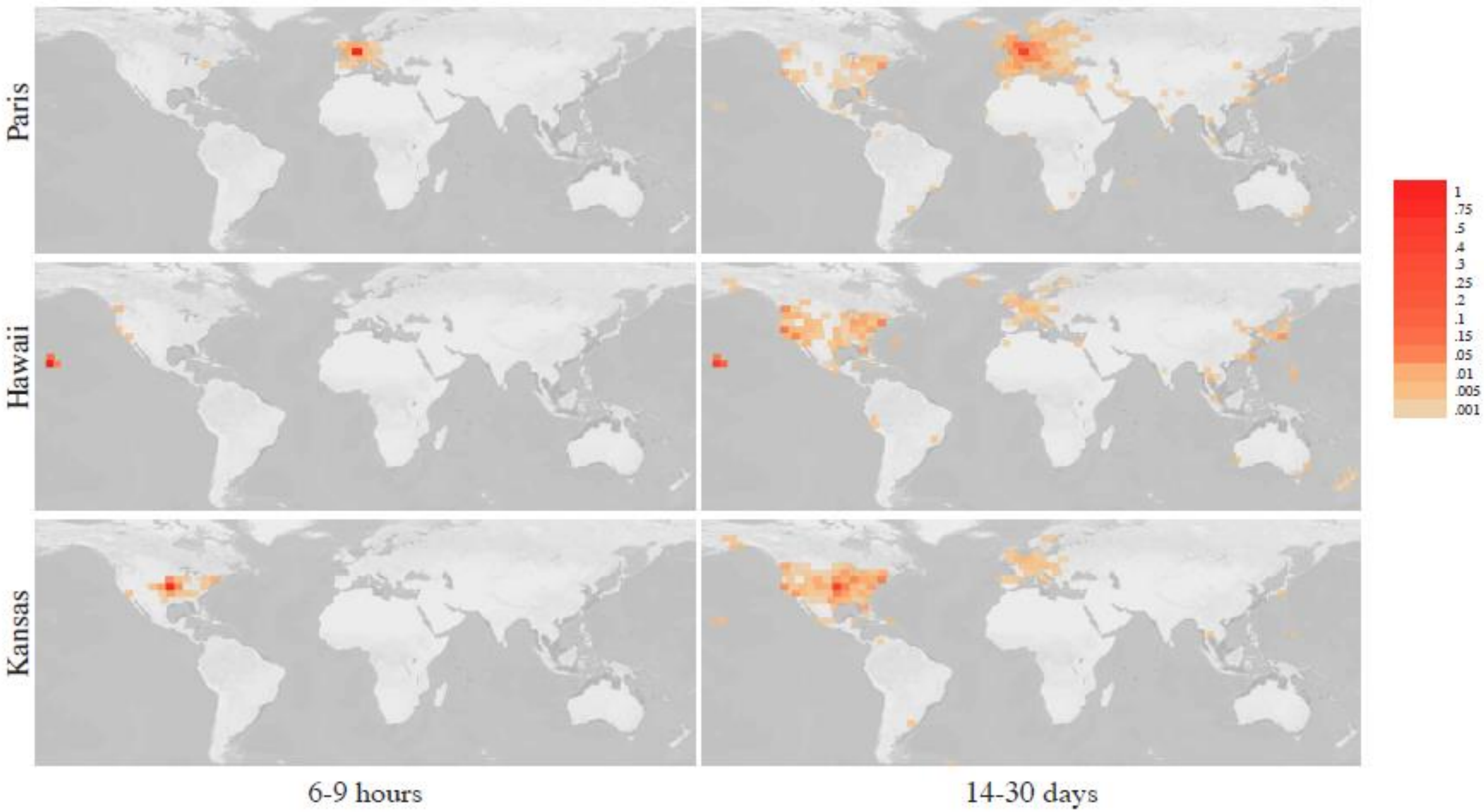
15:14,
June 18th, 2006

16:31,
June 18th, 2006

17:24,
June 19th, 2006

Results

- im2gps – 10% (geo-loc within 400 km)
- temporal im2gps – 56%



Tiny Images

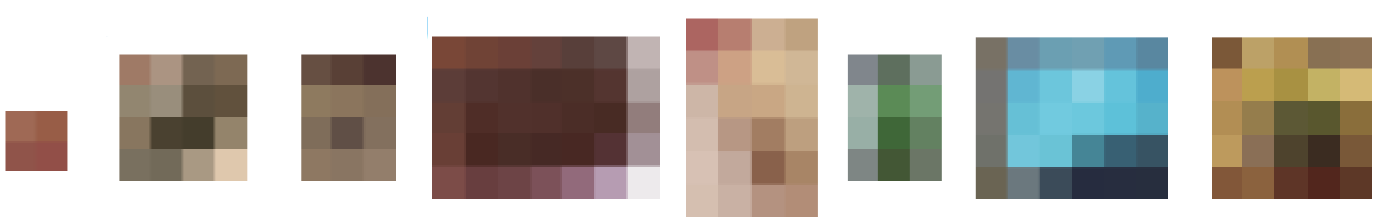
80 million tiny images: a large dataset for non-parametric object and scene recognition

Antonio Torralba, Rob Fergus and William T. Freeman. PAMI 2008.
http://groups.csail.mit.edu/vision/TinyImages/

32x32

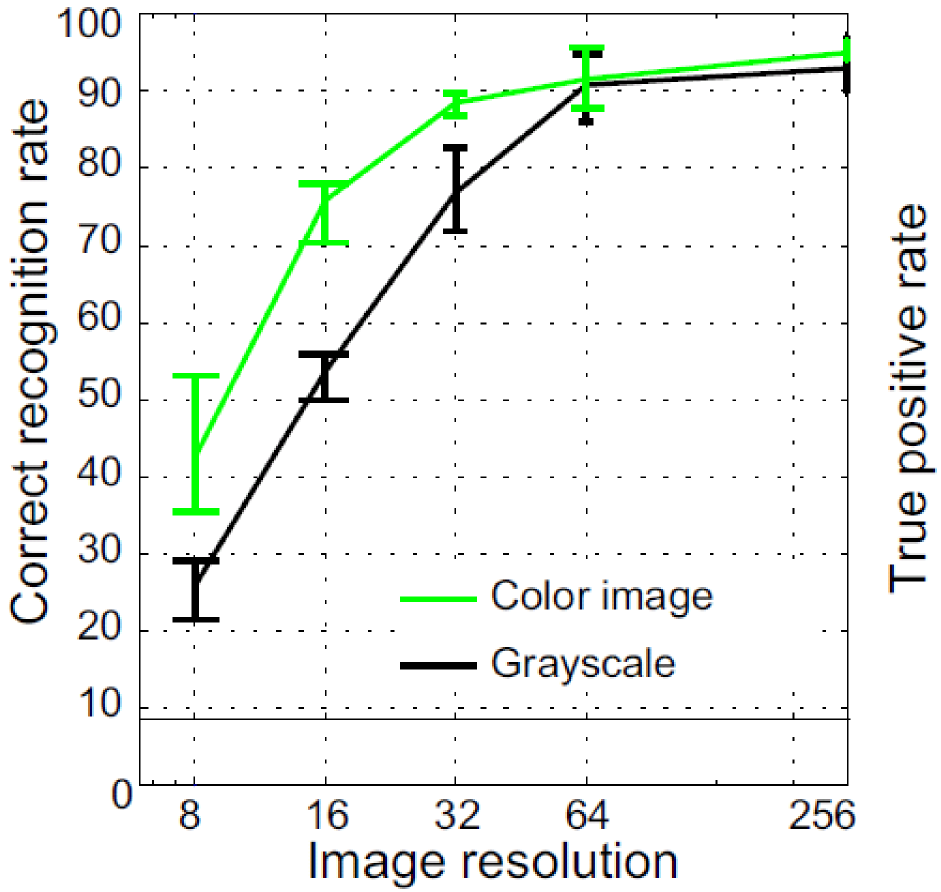

256x256
32x32


256x256
32x32



c) Segmentation of 32x32 images

Given a benchmark, resolution and human scene recognition accuracy increase to a limit

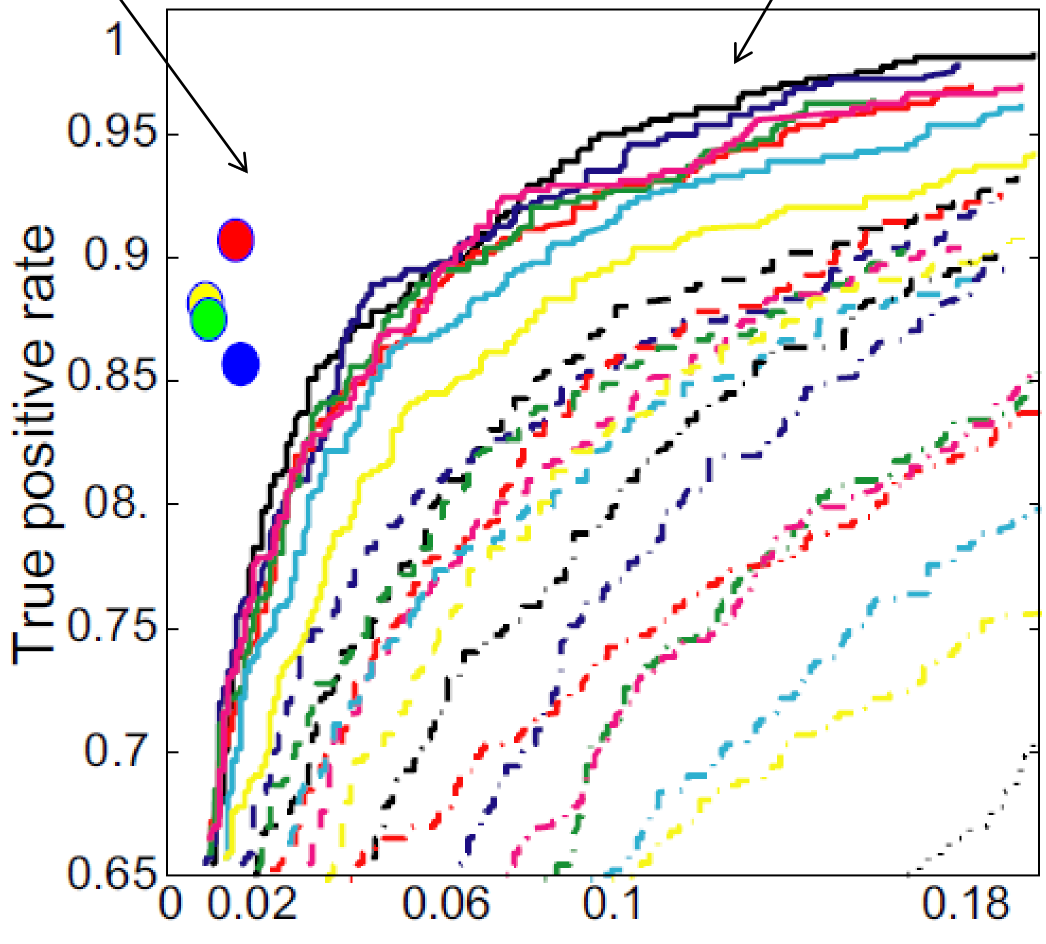


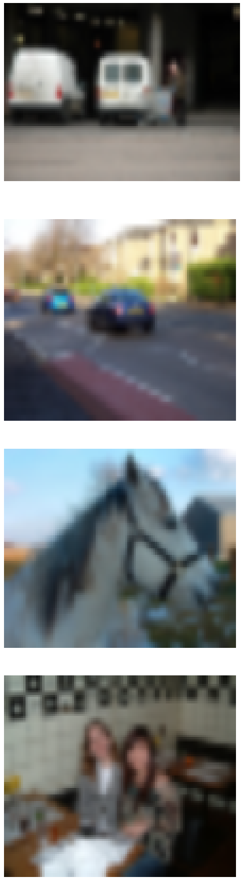
Torralba et al.

Humans vs. Computers: Car Classification

Humans for 32 pixel tall images

Various computer vision algorithms
for *full* resolution images





Example
test images

Torralba et al.

Powers of 10

Number of images on my hard drive: 10^{6}

Number of images seen during my first 10 years: 10^{8}

(3 images/second * 60 * 60 * 16 * 365 * 10 = 630,720,000)

Number of images seen by all humanity: 10^{20}

106,456,367,669 humans[1] * 60 years * 3 images/second * 60 * 60 * 16 * 365 =

1 from http://www.prb.org/Articles/2002/HowManyPeopleHaveEverLivedonEarth.aspx

Number of photons in the universe: 10^{88}

Number of all 32x32 images: 10^{7373}

$256^{32*32*3} \sim 10^{7373}$

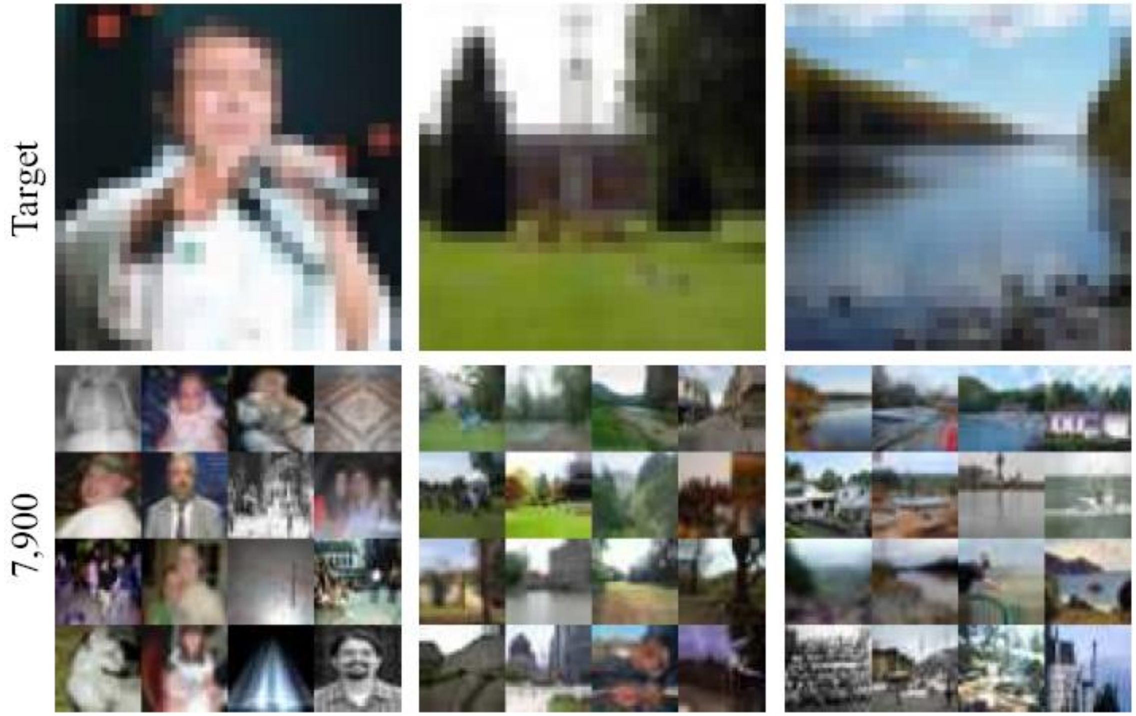
Understanding scenes encompasses all kinds of knowledge

But not all scenes are so original

Lots

Of

Images

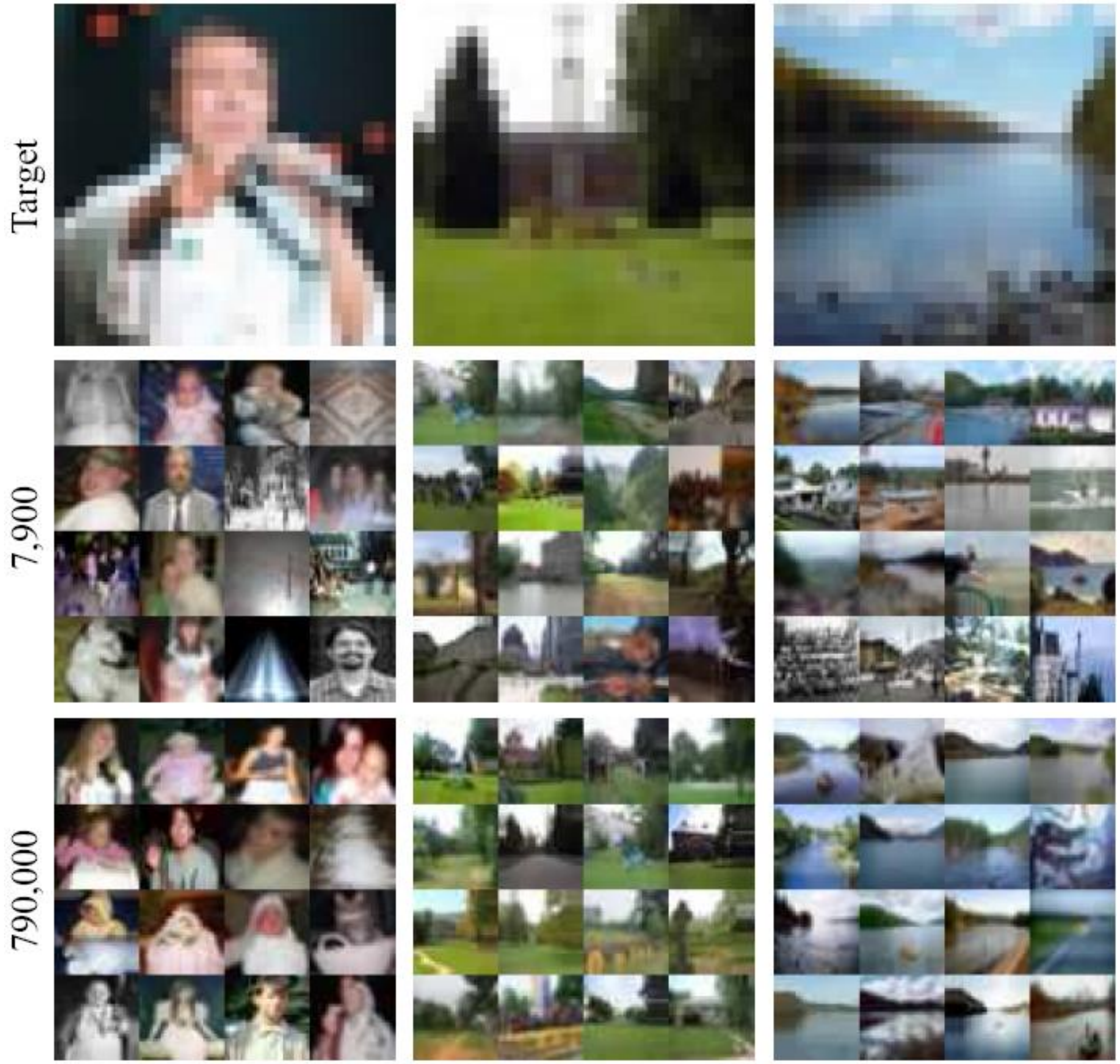


A. Torralba, R. Fergus, W.T.Freeman. PAMI 2008

Lots

Of

Images

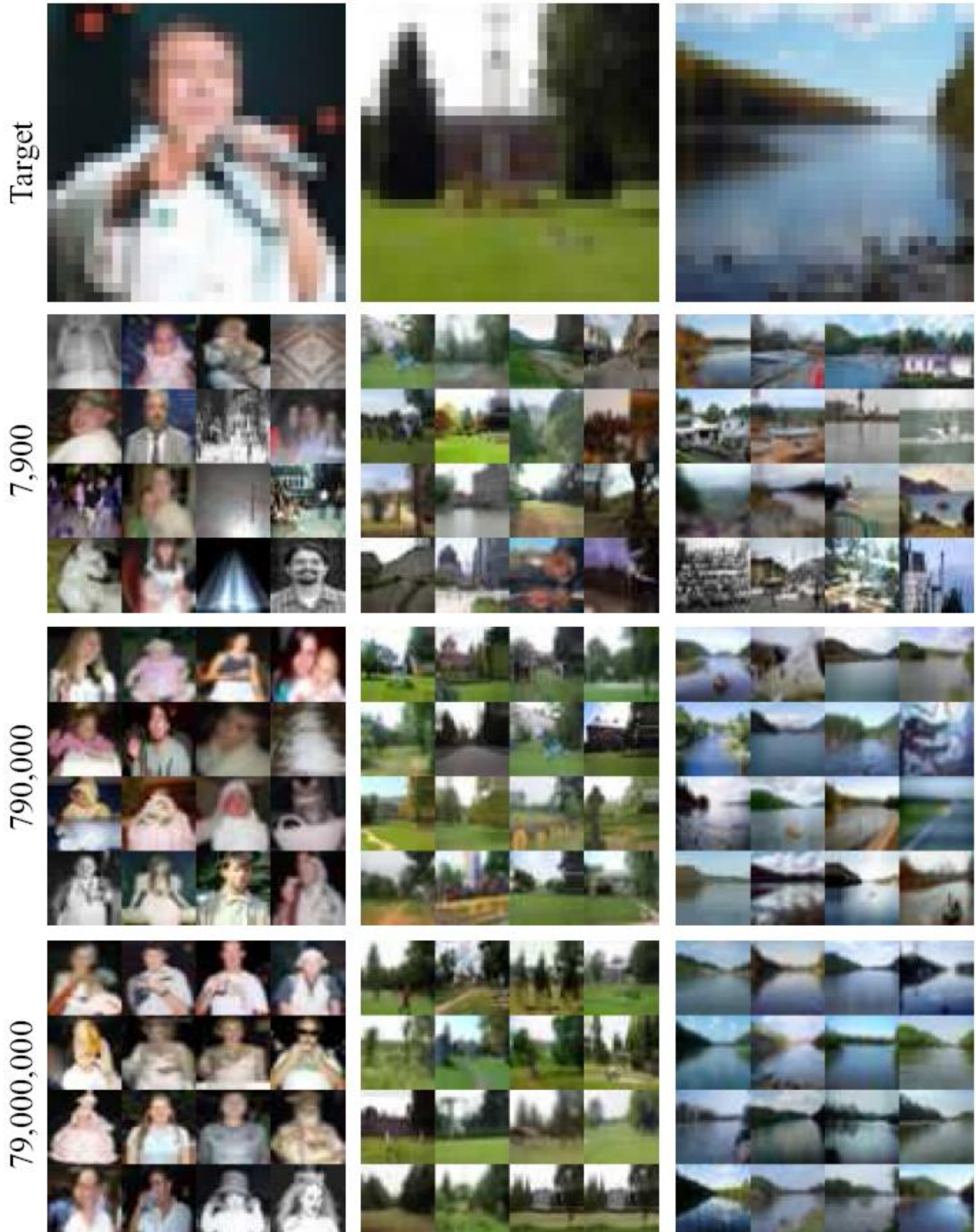


A. Torralba, R. Fergus, W.T.Freeman. PAMI 2008

Lots

Of

Images



Application: Automatic Colorization

Input

Color Transfer

Color Transfer

Matches (gray)

Matches (w/ color)

Avg Color of Match

Application: Automatic Colorization

Input

Color Transfer

Color Transfer

Matches (gray)

Matches (w/ color)

Avg Color of Match

Short cuts to AI

With billions of images on the web, it's often possible to find a close nearest neighbor.

We can shortcut hard problems by "looking up" the answer, stealing the labels from our nearest neighbor.

So what is intelligence?

Weak AI: The simulation of a ‘mind’ is a model for the ‘mind’.

Strong AI: The simulation of a ‘mind’ is a ‘mind’.

Chinese Room experiment, John Searle (1980)

Chinese Room experiment, John Searle (1980)

If a machine can convincingly simulate an intelligent conversation, does understand?

Searle imagines himself in a room, acting as a computer by manually executing a program that convincingly simulates the behavior of a native Chinese speaker.

Most of the discussion consists of attempts to refute it.
"The overwhelming majority," notes *BBS* editor Stevan Harnad,“ still think that the Chinese Room Argument is dead wrong.“

The sheer volume of the literature that has grown up around it inspired Pat Hayes to quip that the field of cognitive science ought to be redefined as "the ongoing research program of showing Searle's Chinese Room Argument to be false.

Yann LeCun

October 23 at 9:58pm ·

Questions from the piece:

Q1. Does the Chinese Room argument prove the impossibility of machine consciousness?
A1: Hell no. ... See More

Can Machines Become Moral?

The question is heard more and more often, both from those who think that machines cannot become moral, and who think that to believe otherwise is a dangerous illusion, and from those who think that machines must become moral,...

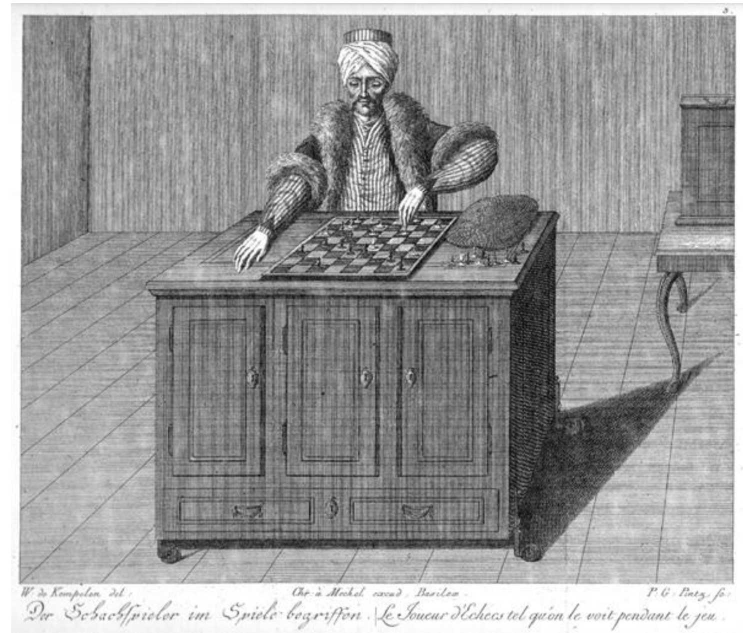
BIGQUESTIONSONLINE.COM | BY DON HOWARD

You and 156 others   30 Comments 20 Shares

Like   Comment   Share

Mechanical Turk

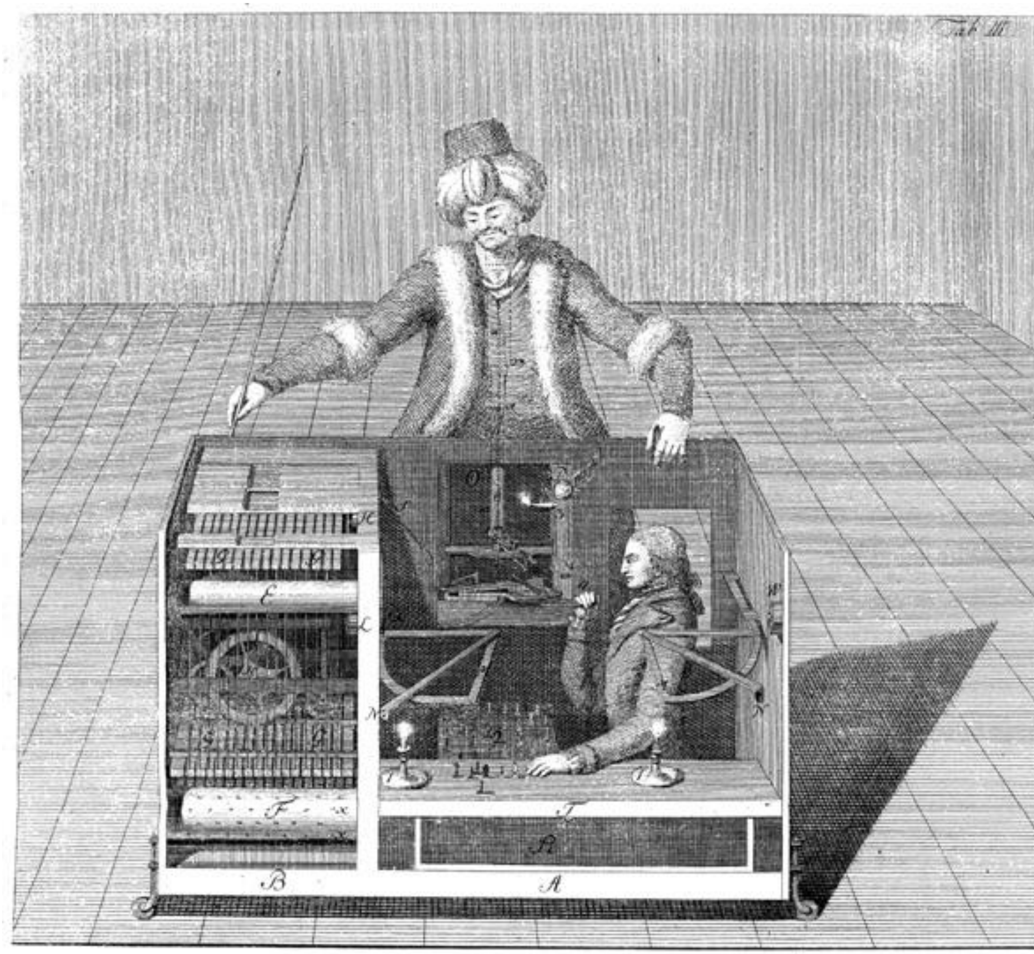
- von Kempelen, 1770.
- Robotic chess player.
- Clockwork routines.
- Magnetic induction (not vision)



- Toured the world; played
  Napoleon Bonaparte
  and Benjamin Franklin.

Mechanical Turk

- It was all a ruse!
- Ho ho ho.



Amazon Mechanical Turk

Artificial artificial intelligence.

Launched 2005.
Small tasks, small pay.
Used extensively in data collection.

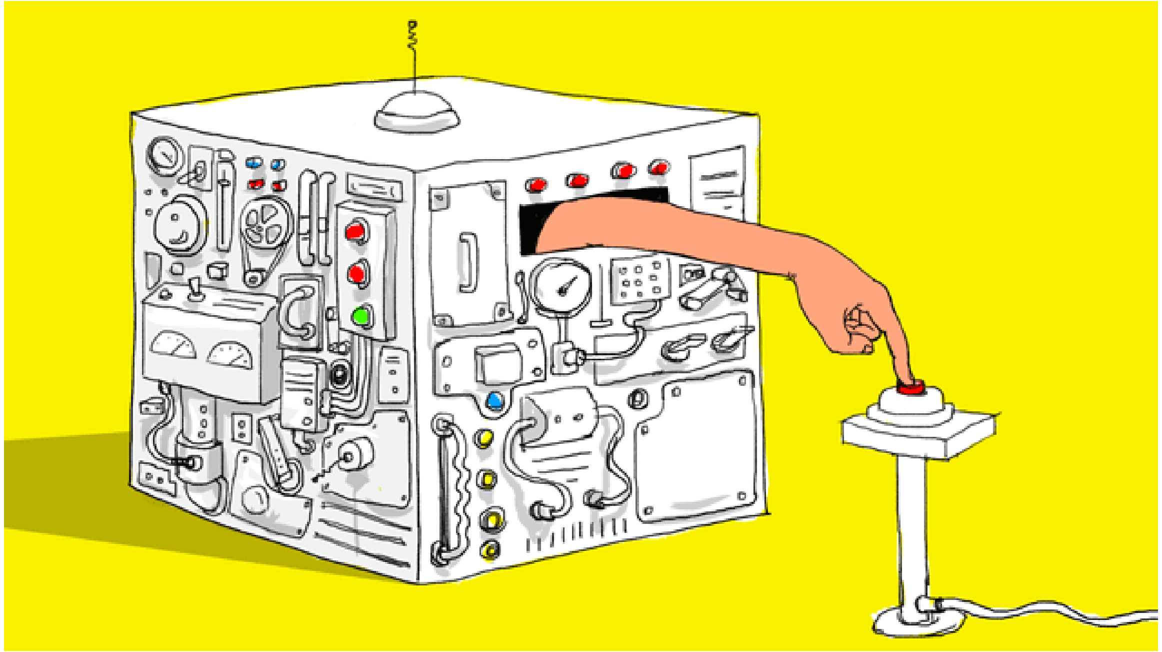
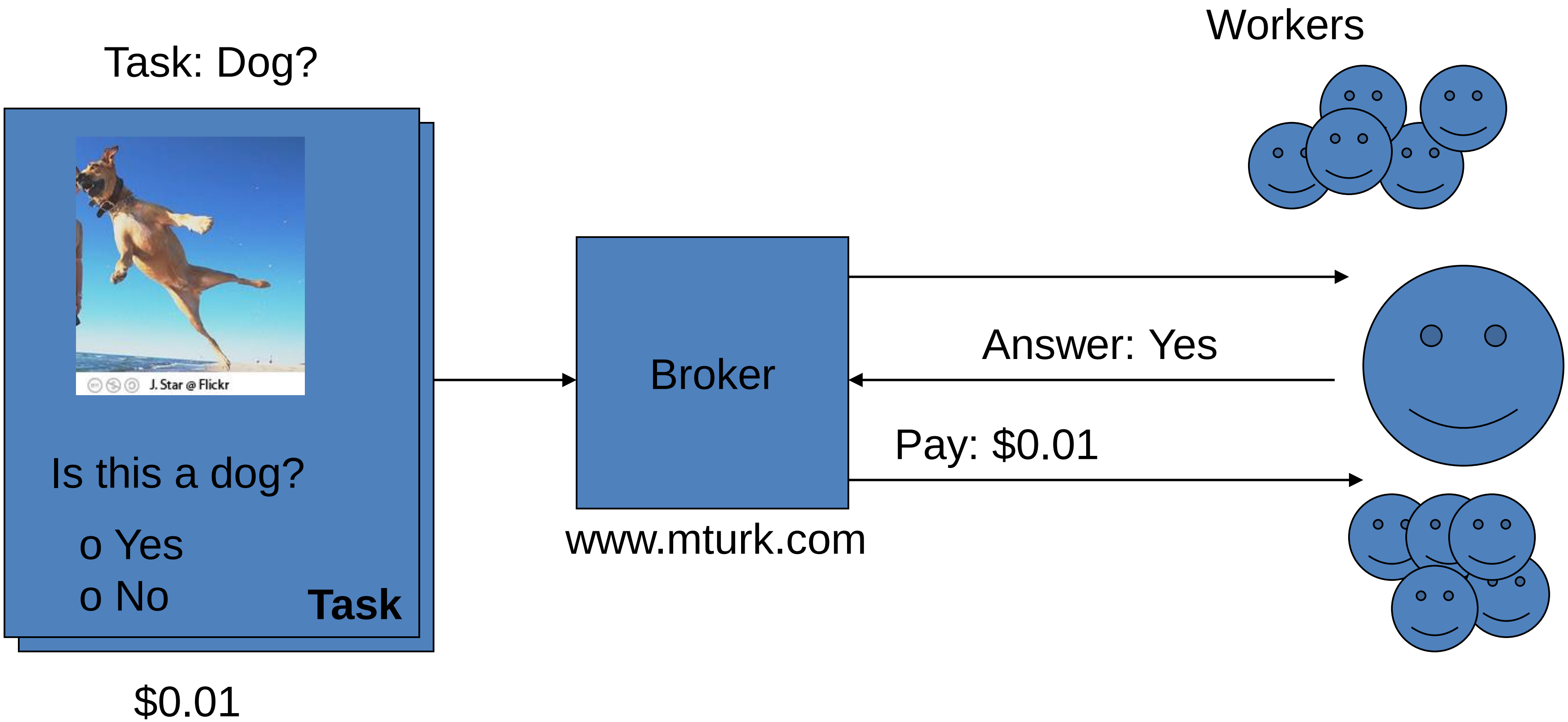
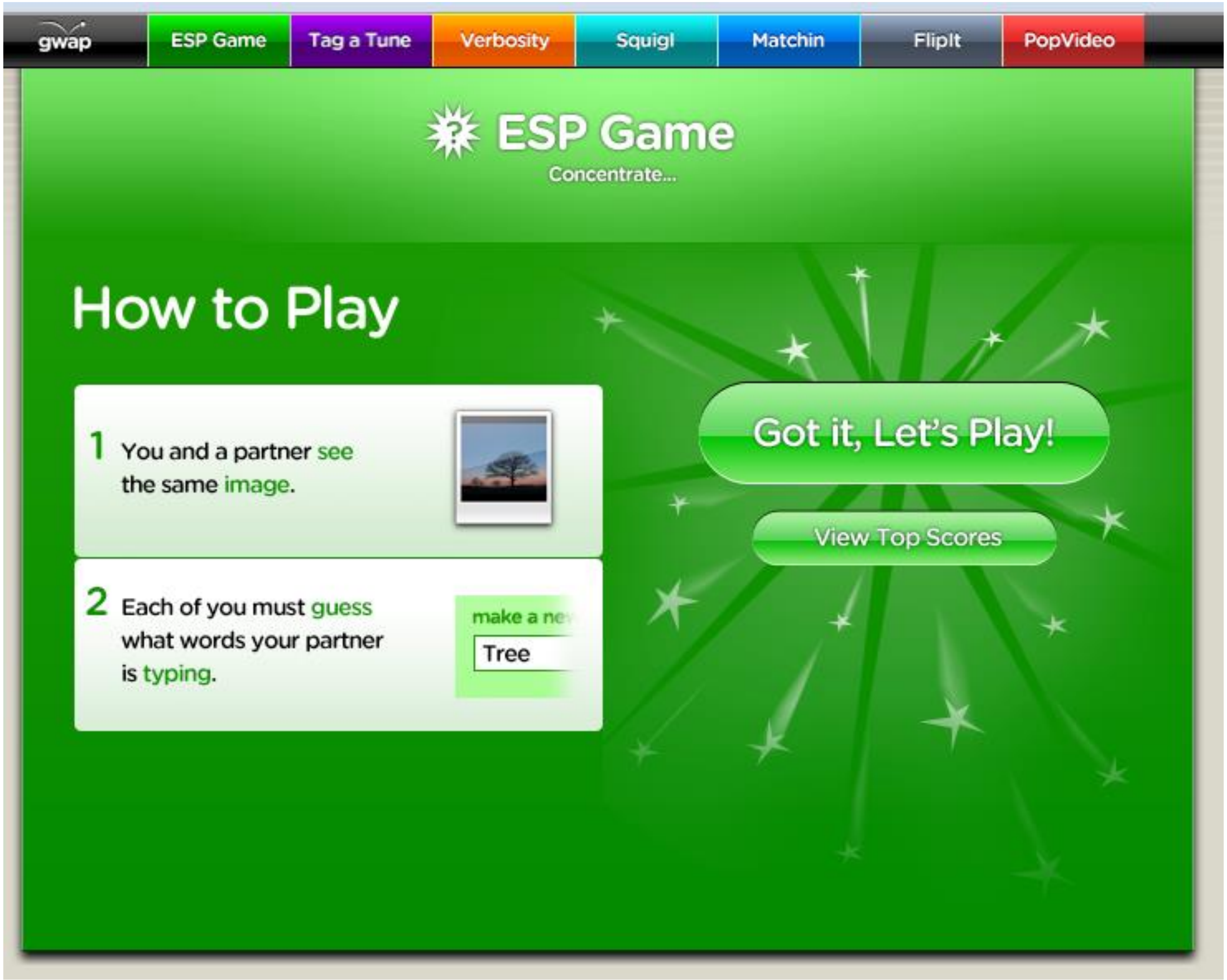


Image: Gizmodo

Amazon Mechanical Turk





Luis von Ahn and Laura Dabbish. [Labeling Images with a Computer Game](). ACM Conf. on Human Factors in Computing Systems, CHI 2004


ESP Game
Tag a Tune
Verbosity
Squigl
Matchin
logged
score
0
ESP Game
Concentrate...
time
2:56
What do you see?
taboo words
student
guesses
submit
pass
Play Anonymously
Screencast-O-Matic.com

Vision (Segmentation): LabelMe

http://labelme.csail.mit.edu

"Open world" database annotated by the community*

Notes on Image Annotation, Barriuso and Torralba 2012. http://arxiv.org/abs/1210.3448

Utility data annotation via Amazon Mechanical Turk

X 100 000 = \$5000

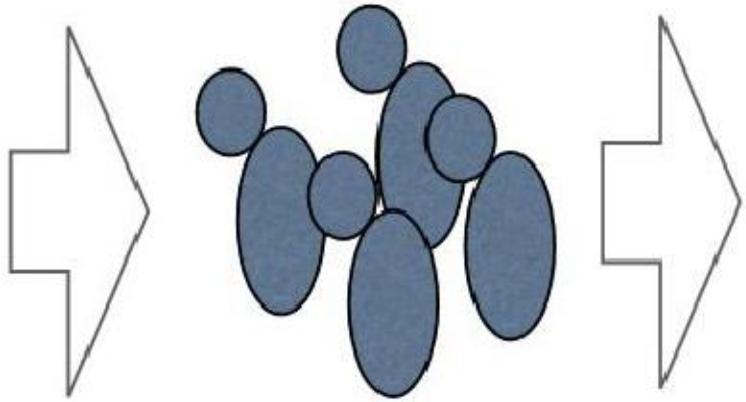
Alexander Sorokin
David Forsyth
CVPR Workshops 2008

Slides by Alexander Sorokin

Building datasets

6000 images
from flickr.com

Annotators



amazonmechanical turk
Artificial Artificial Intelligence

Is there an Indigo bunting in the image?

100s of
training images

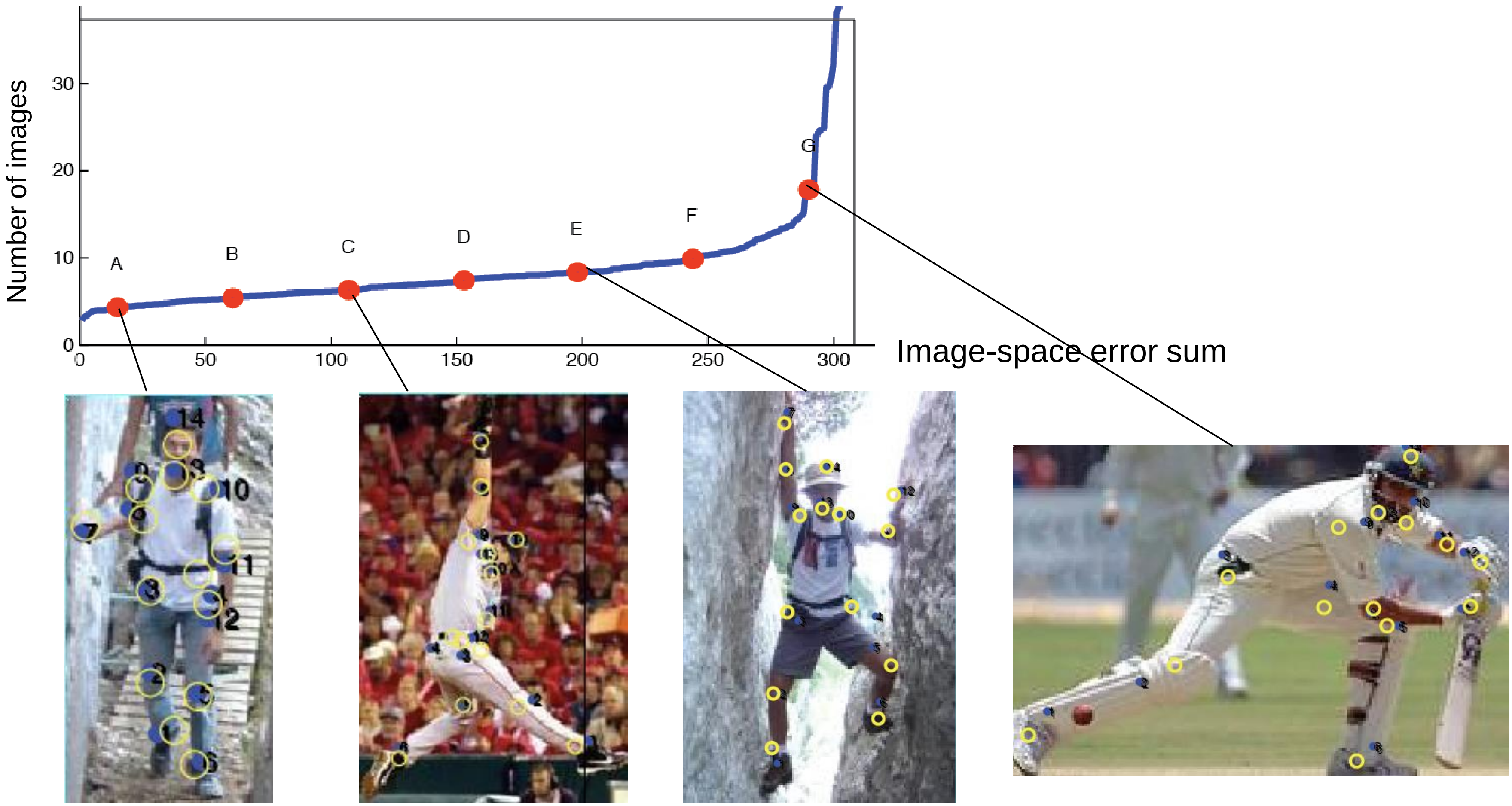
Slide credit: Welinder et al

Issues

- Quality?
  - How good is it?
  - How to be sure?
- Price?
  - Trade off between throughput and cost
    - *NOT* as much of a trade off with quality
  - Higher pay can actually attract scammers

Annotation quality

How much agreement is there on ‘ground truth’ and turker-labeled joint positions?
Points must agree within 5-10 pixels on 500x500 image.



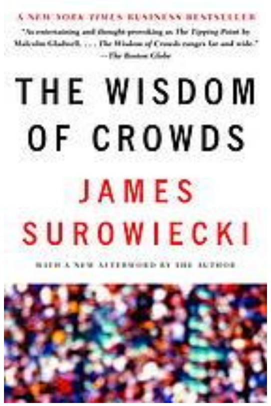
A C E G

Yellow ring = ‘ground truth’; Blue circle = human labeled

Ensuring Annotation Quality

- Consensus in multiple annotations "Wisdom of the Crowds"

  Not enough on its own, but widely used



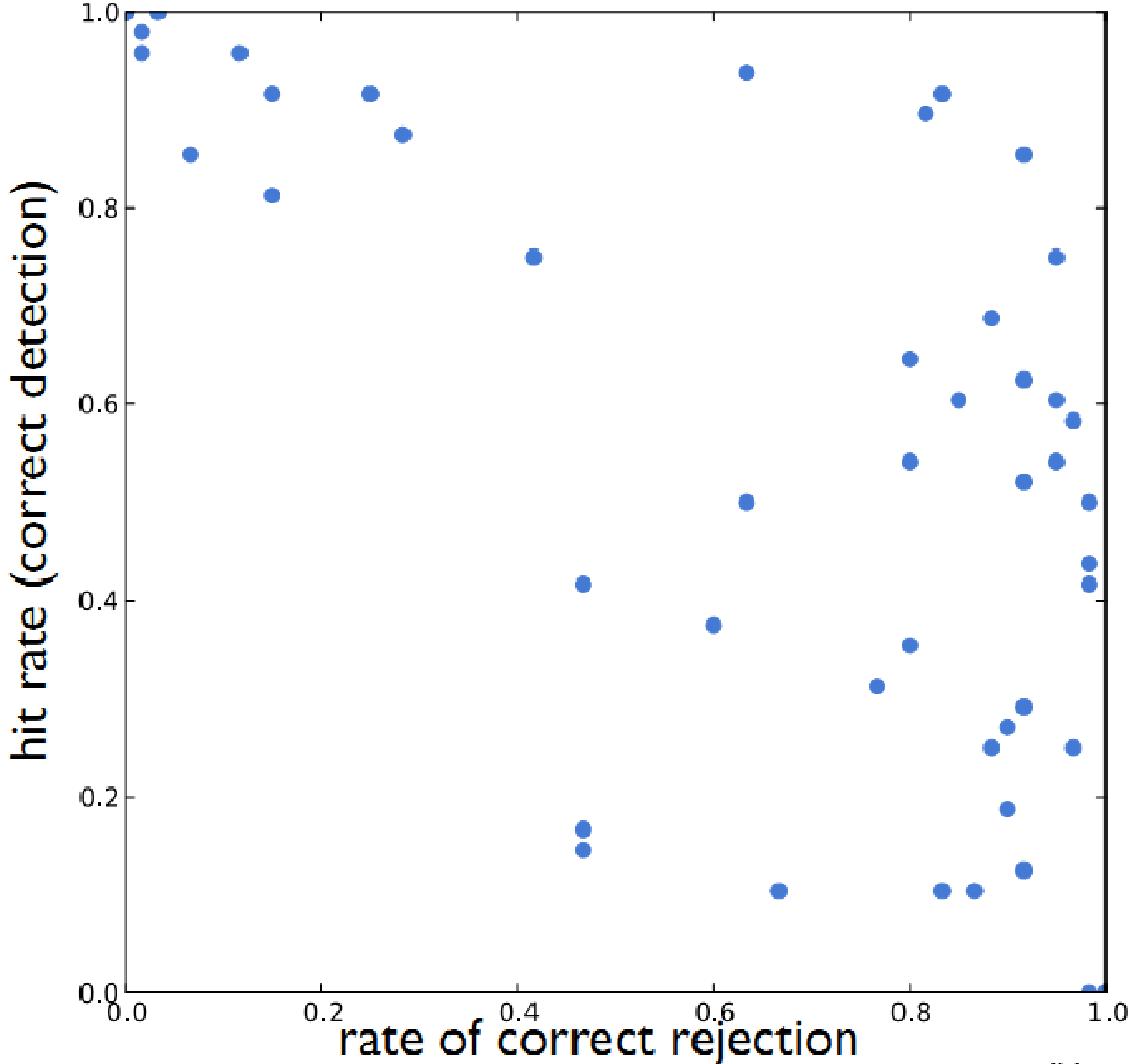
- Gold Standard / Sentinel
  – Special case: qualification exam

    Widely used & important. Find good annotators; keep them honest.

- Grading Tasks
  – A second tier of workers who grade others

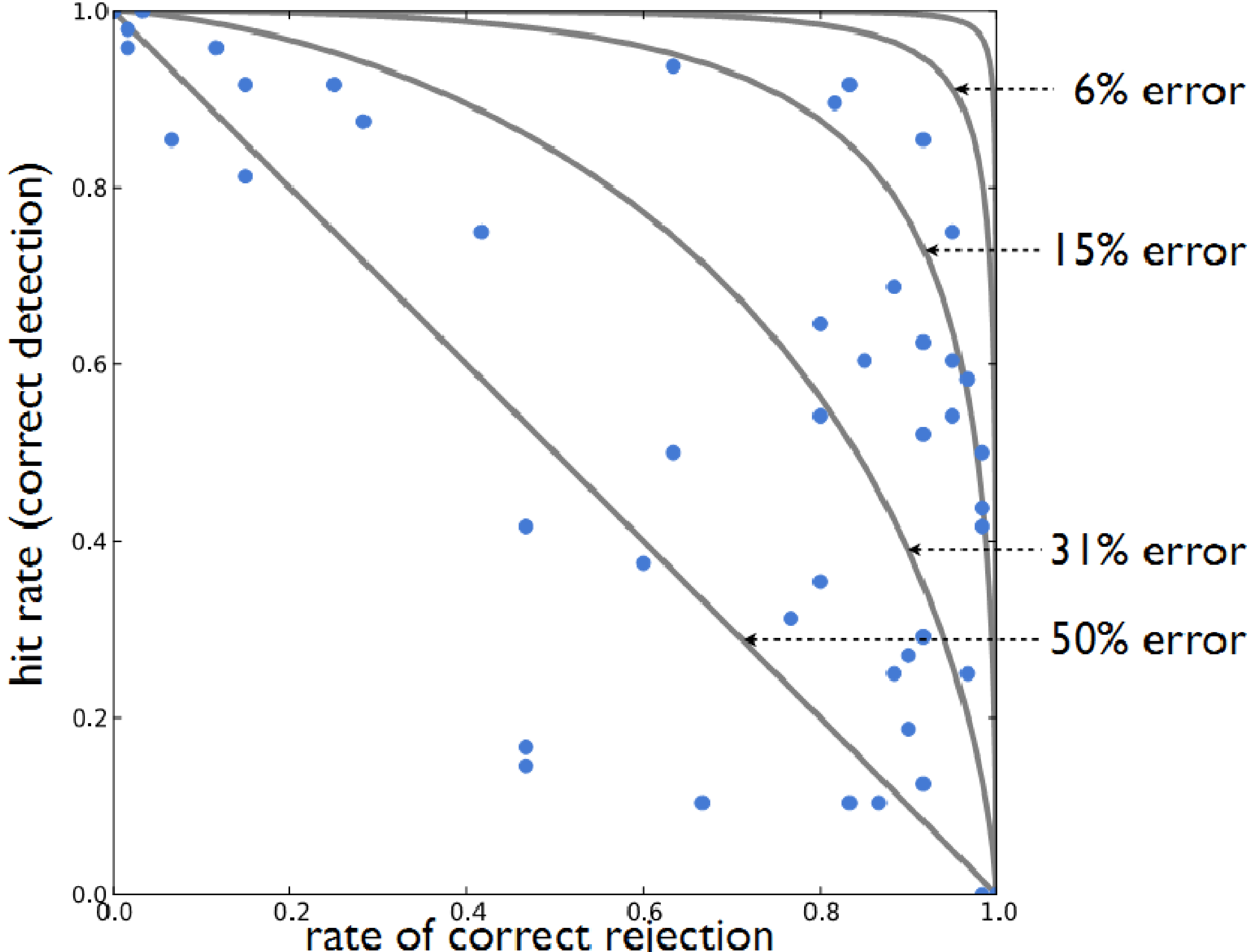
    Not widely used

Task: Find the Indigo Bunting



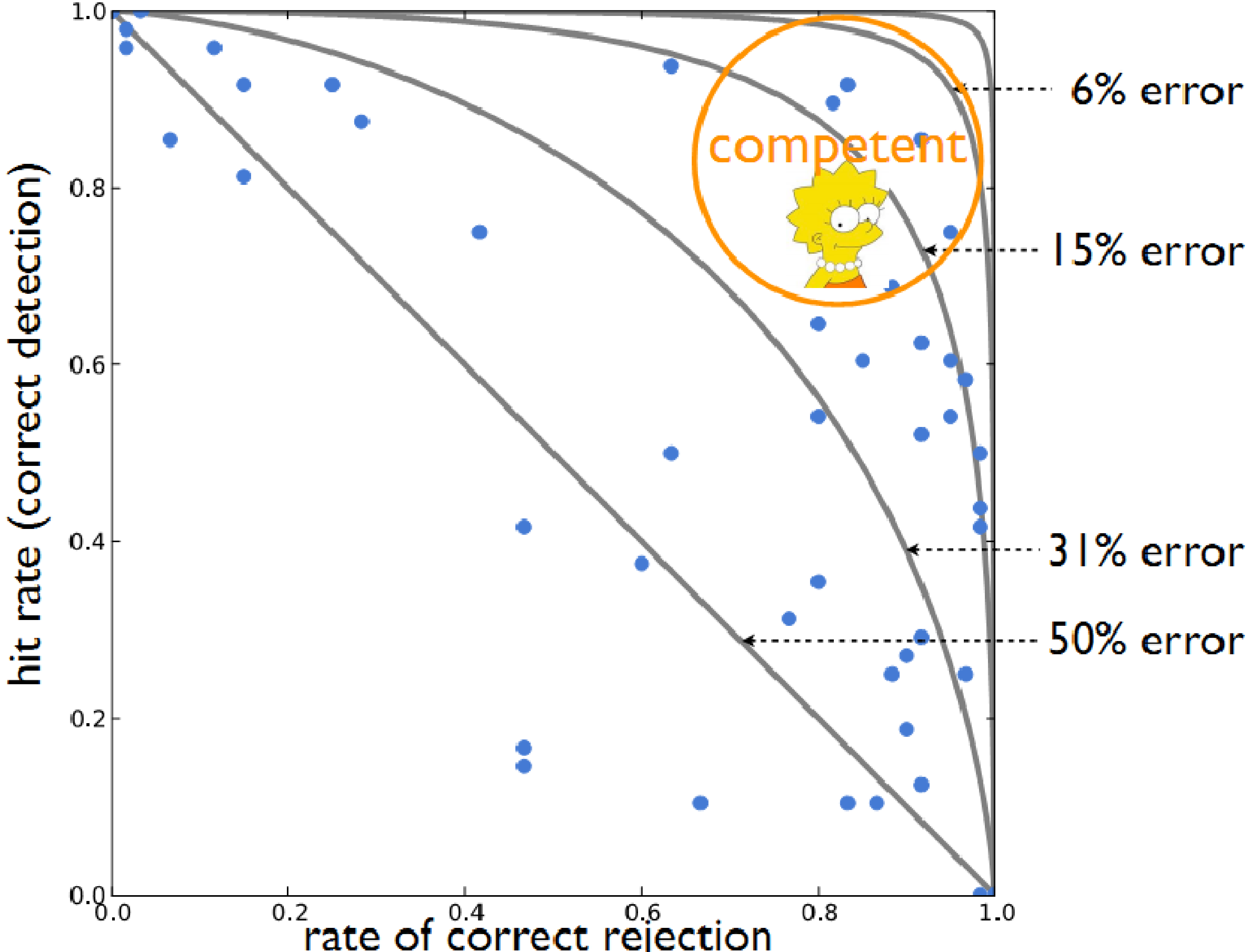
Slide credit: Welinder et al

Task: Find the Indigo Bunting



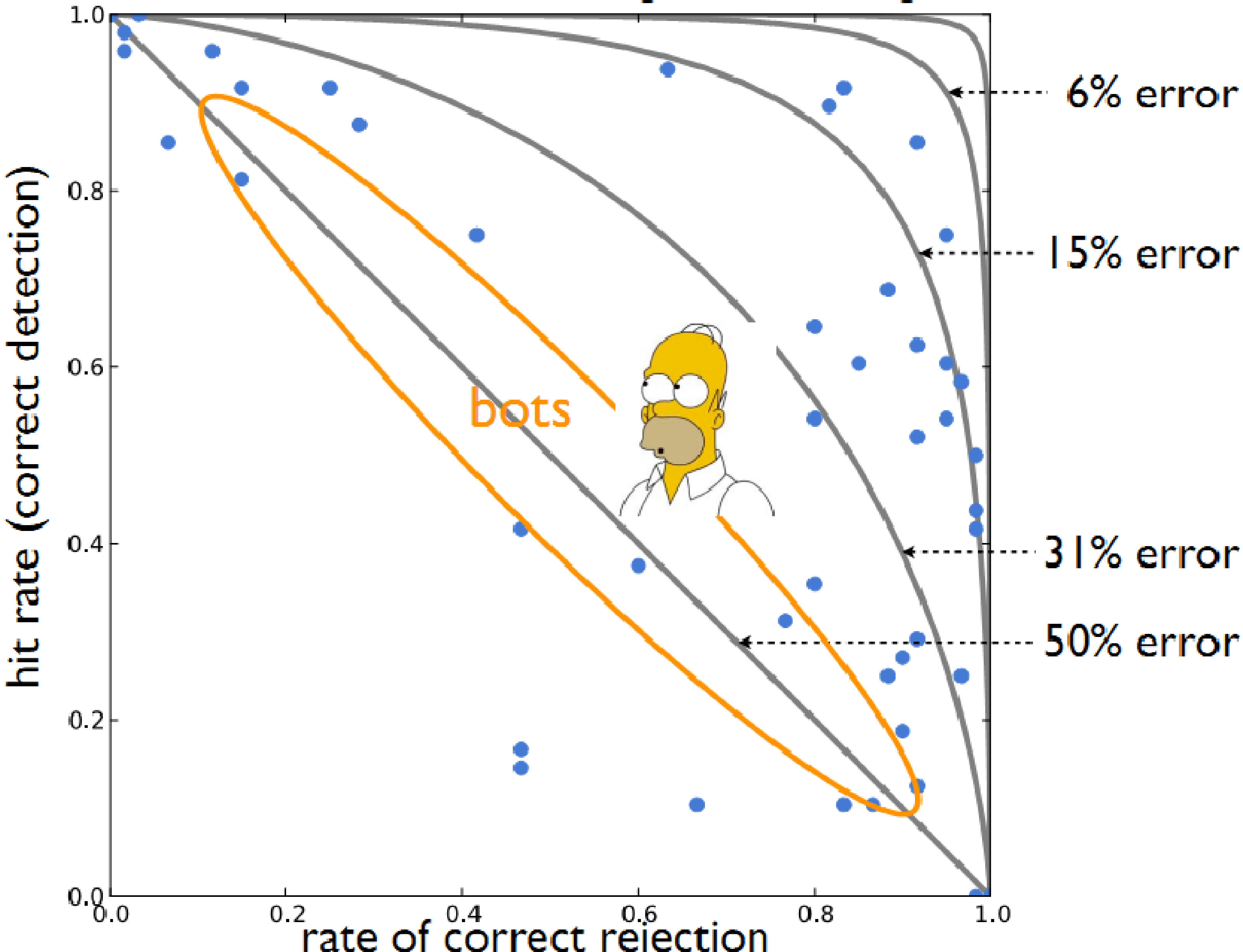
Slide credit: Welinder et al

Task: Find the Indigo Bunting



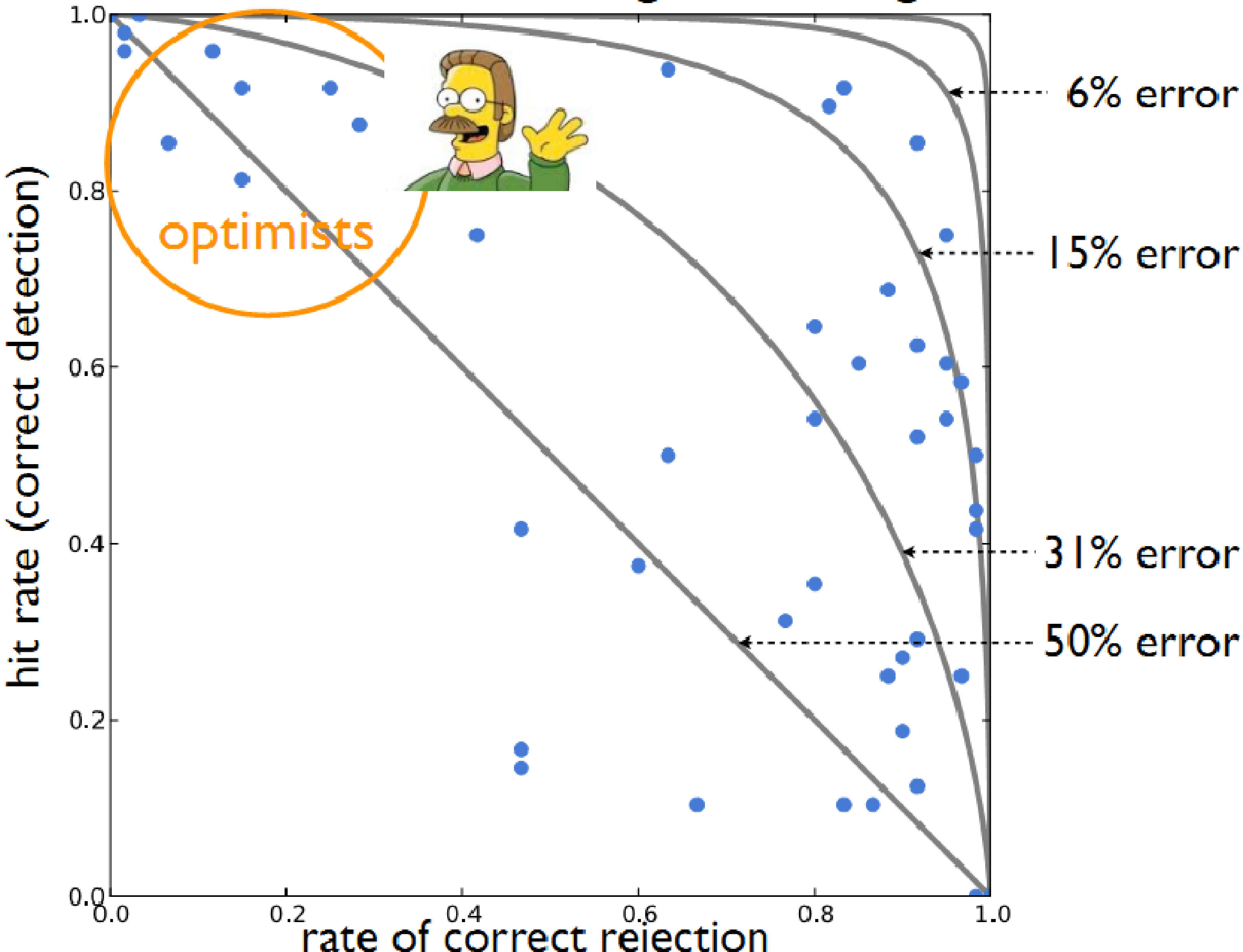
Slide credit: Welinder et al

Task: Find the Indigo Bunting



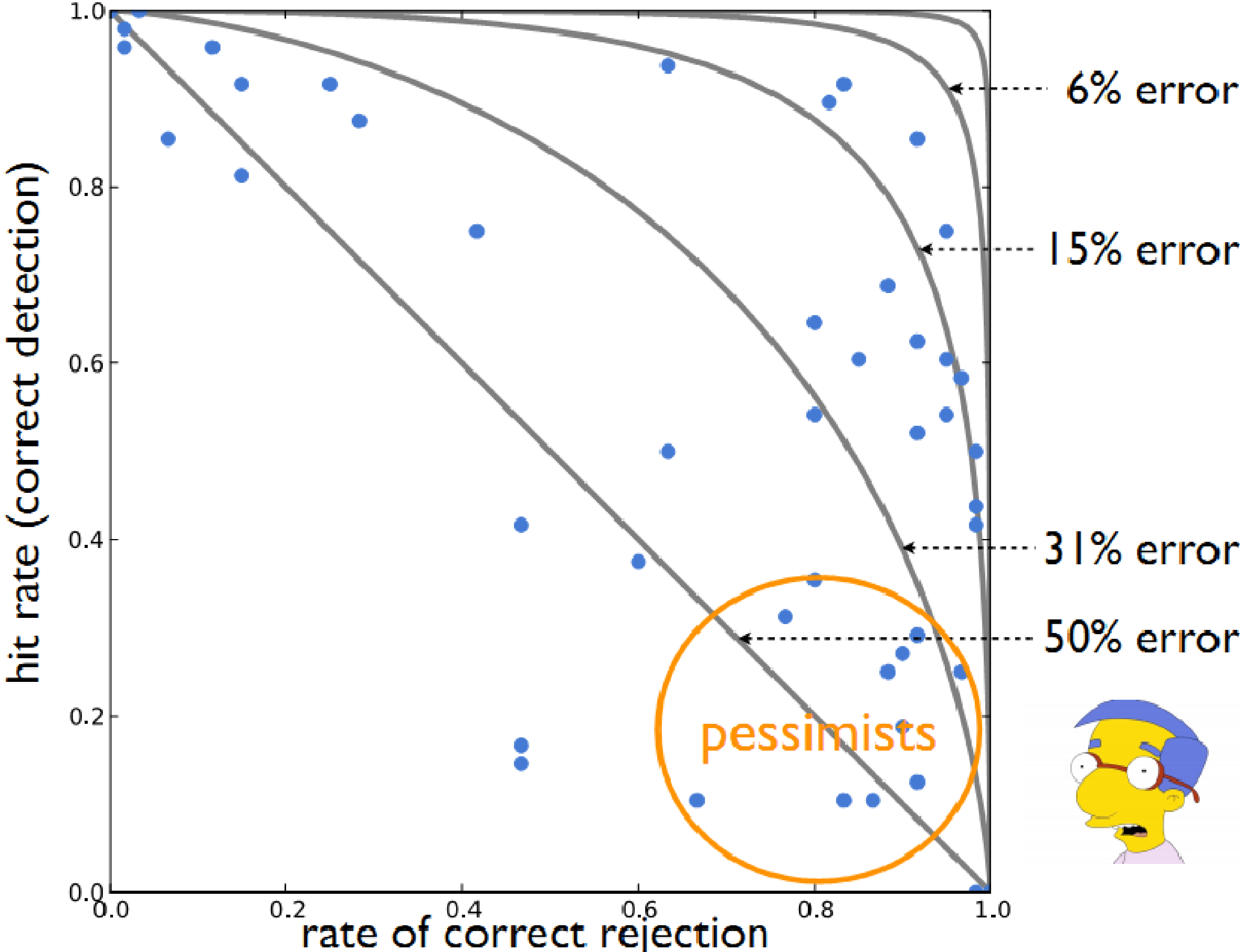
Slide credit: Welinder et al

Task: Find the Indigo Bunting



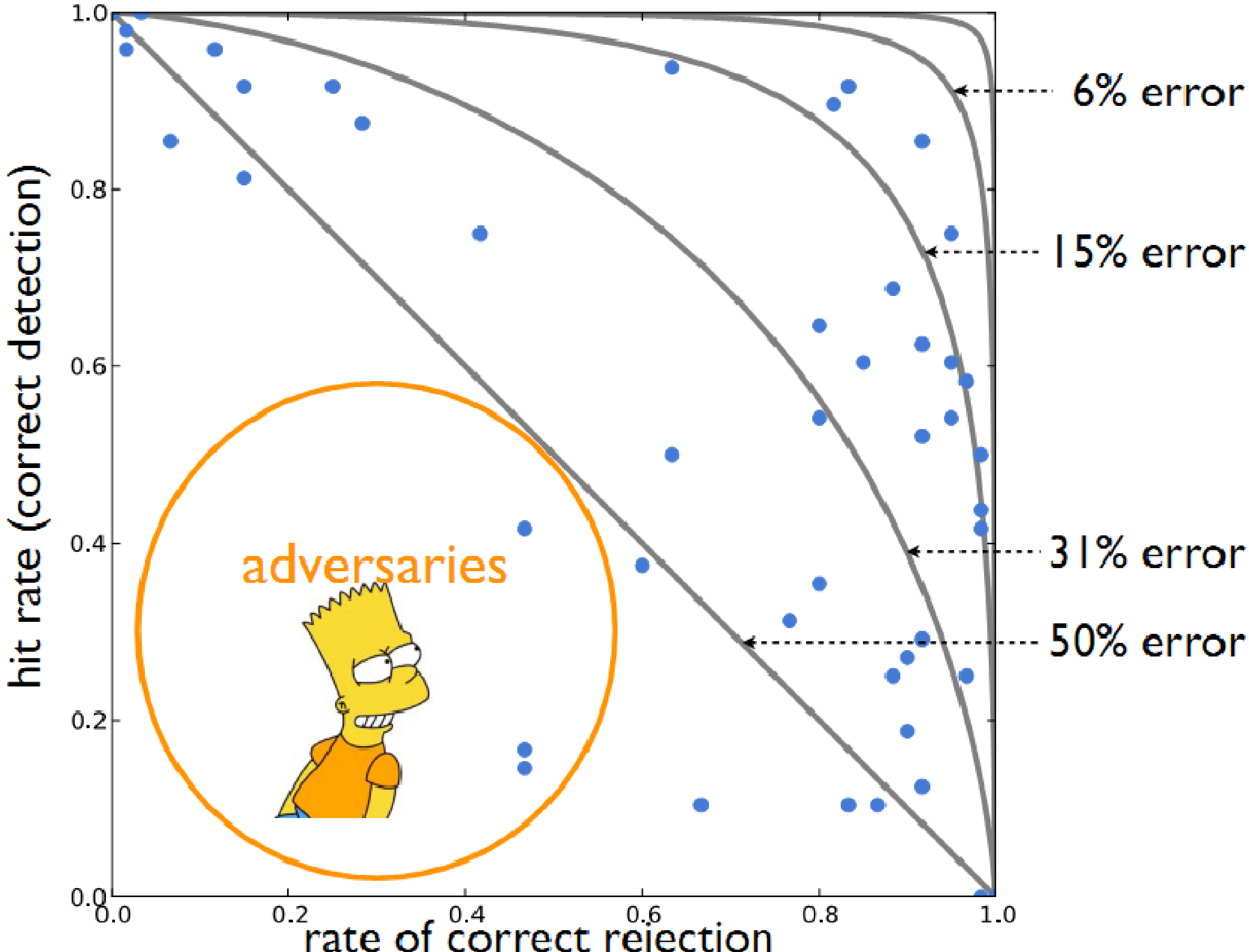
Slide credit: Welinder et al

Task: Find the Indigo Bunting



Slide credit: Welinder et al

Task: Find the Indigo Bunting



Slide credit: Welinder et al

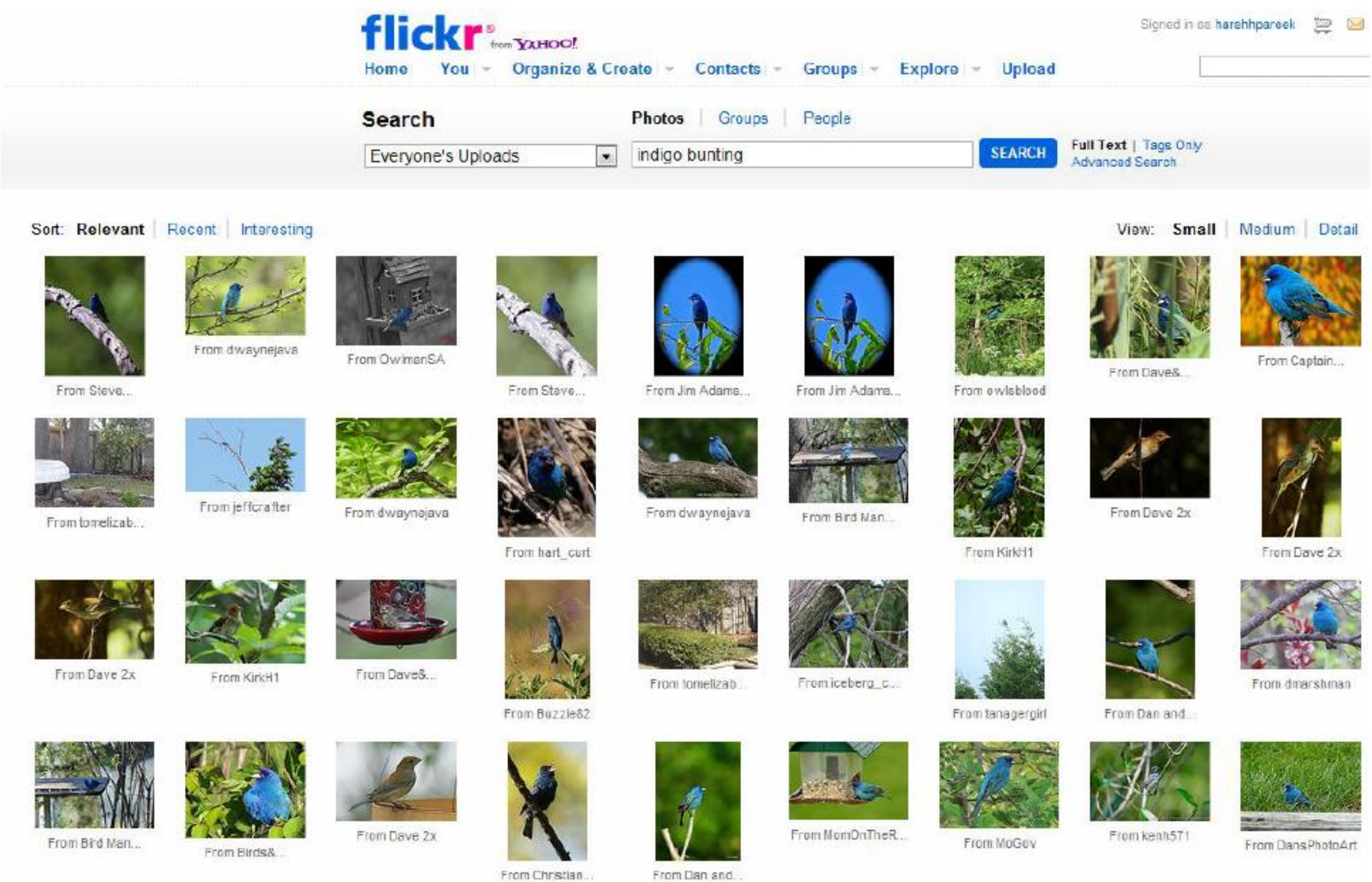


Image credit: Flickr.com

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

By Suren Manvelyan, http://www.surenmanvelyan.com/gallery/7116

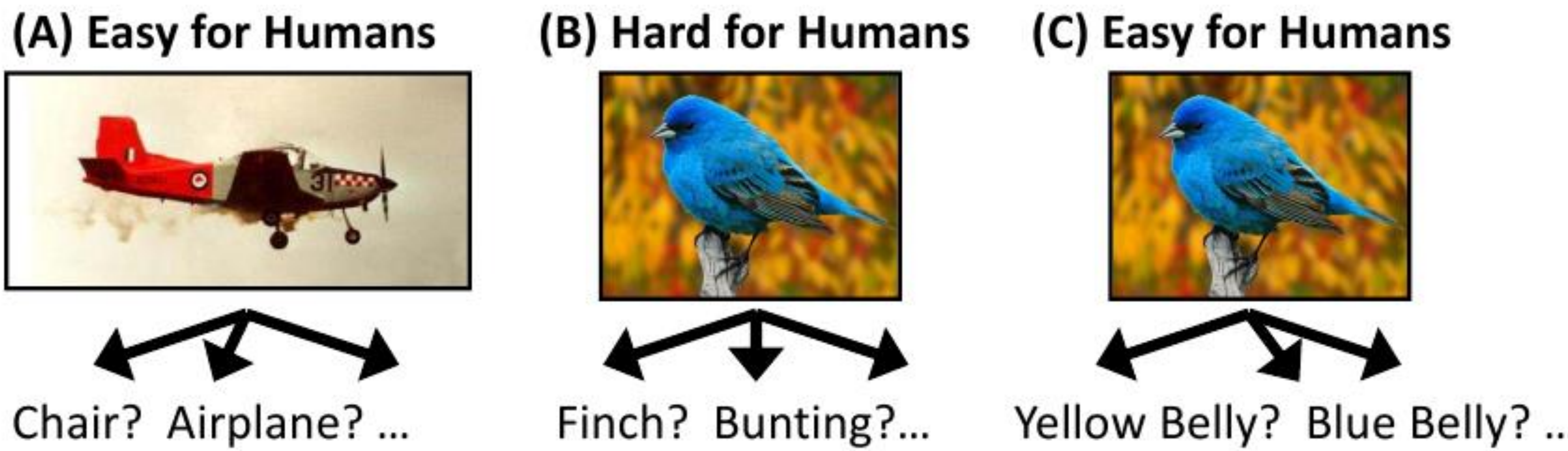
Visual Recognition with Humans in the Loop

Steve Branson, Catherine Wah, Florian Schroff, Boris Babenko, Peter Welinder, Pietro Perona, Serge Belongie

Part of the [Visipedia project]

Slides from Brian O'Neil

Introduction:



Computers starting to get good at this.

If it's hard for humans, it's probably too hard for computers.

Semantic feature extraction difficult for computers.

Combine strengths to solve this problem.

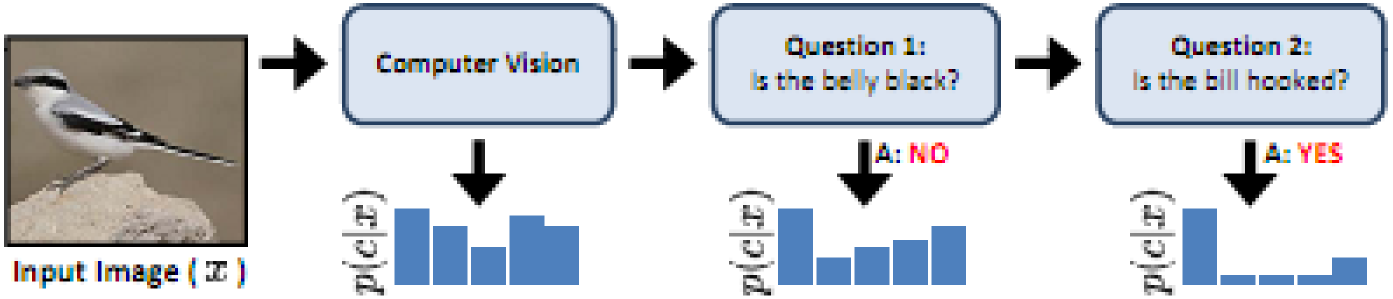
The Approach: What is progress?

- Supplement visual recognition with the human capacity for visual feature extraction to tackle difficult (fine-grained) recognition problems.
- Typical progress is viewed as increasing data difficulty while maintaining full autonomy

- Reduction in human effort on difficult data.

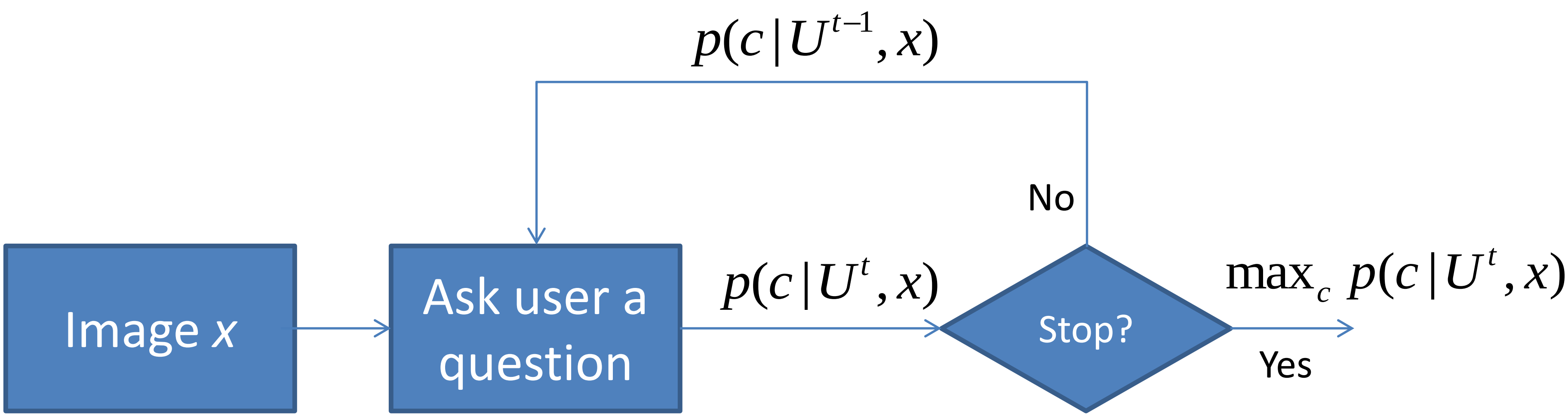
The Approach: 20 Questions

- Ask the user a series of discriminative visual questions to make the classification.



Which 20 questions?

- At each step, exploit the image itself and the user response history to select the most informative question to ask next.



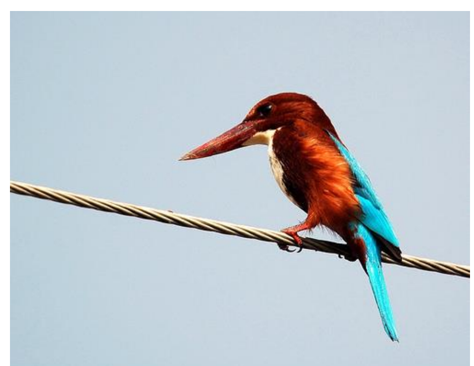
Which question to ask?

- The question that will reduce entropy the most, taking into consideration the computer vision classifier confidences for each category.

The Dataset: Birds-200

- 6033 images of 200 species




Implementation

- Assembled 25 visual questions encompassing 288 visual attributes extracted from www.whatbird.com
- Mechanical Turk users asked to answer questions and provide confidence scores.

User Responses.

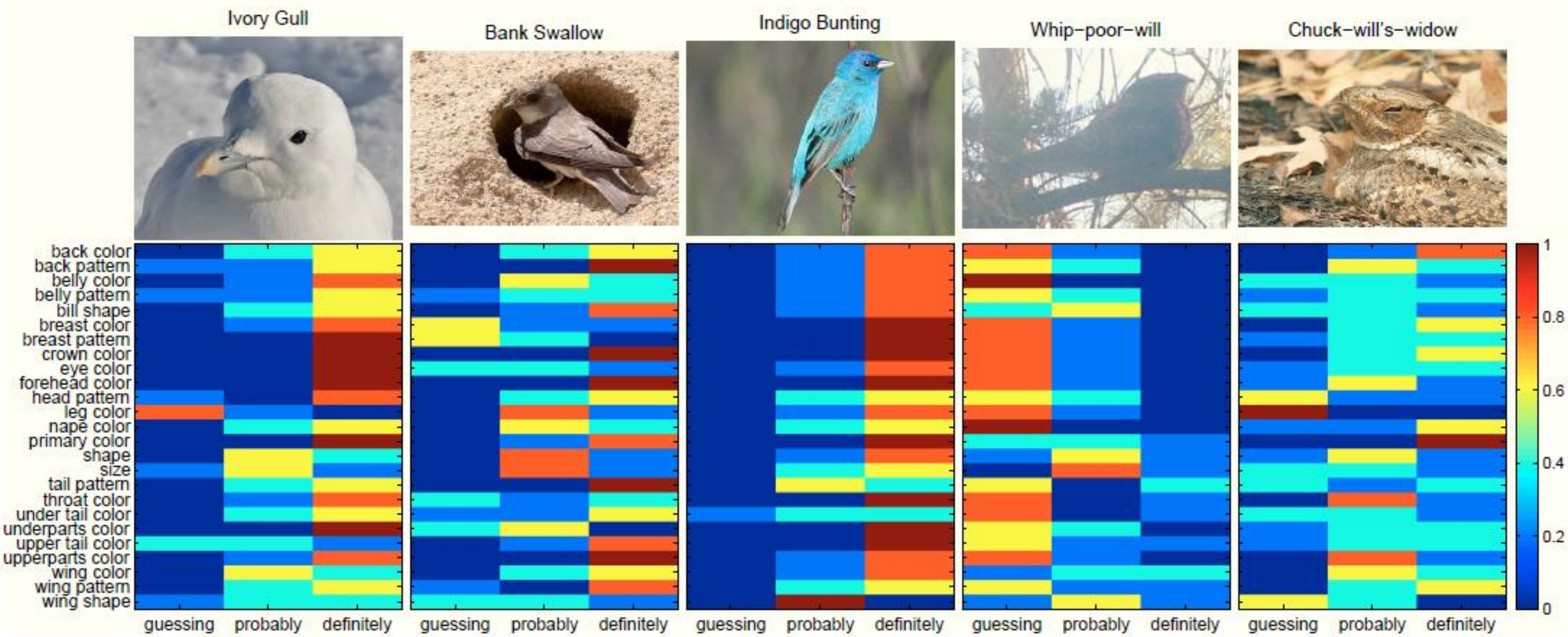
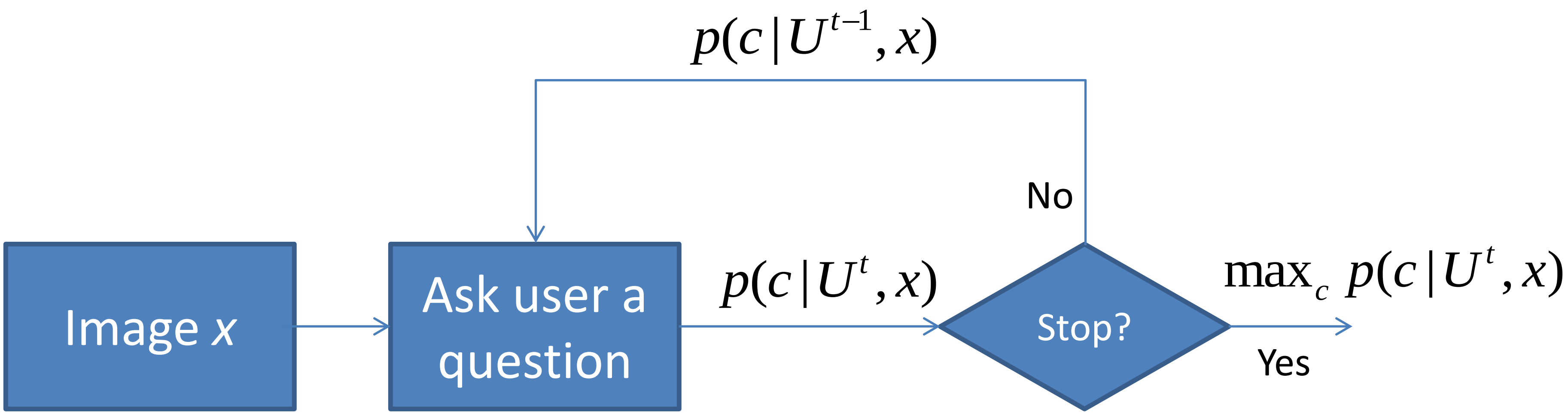


Fig. 4. Examples of user responses for each of the 25 attributes. The distribution over {*Guessing*,*Probably*,*Definitely*} is color coded with blue denoting 0% and red denoting 100% of the five answers per image attribute pair.

Visual recognition

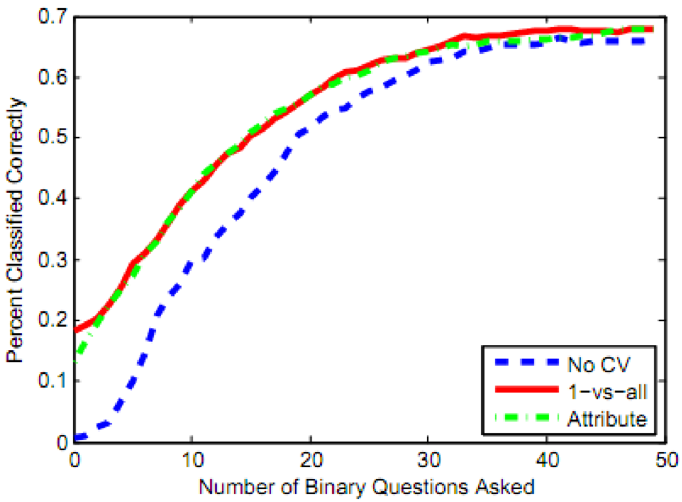
- Any vision system that can output a probability distribution across classes will work.
- Authors used Andrea Vedaldis's code.
  - Color/gray SIFT
  - VQ geometric blur
  - 1 v All SVM
- Authors added full image color histograms and VQ color histograms

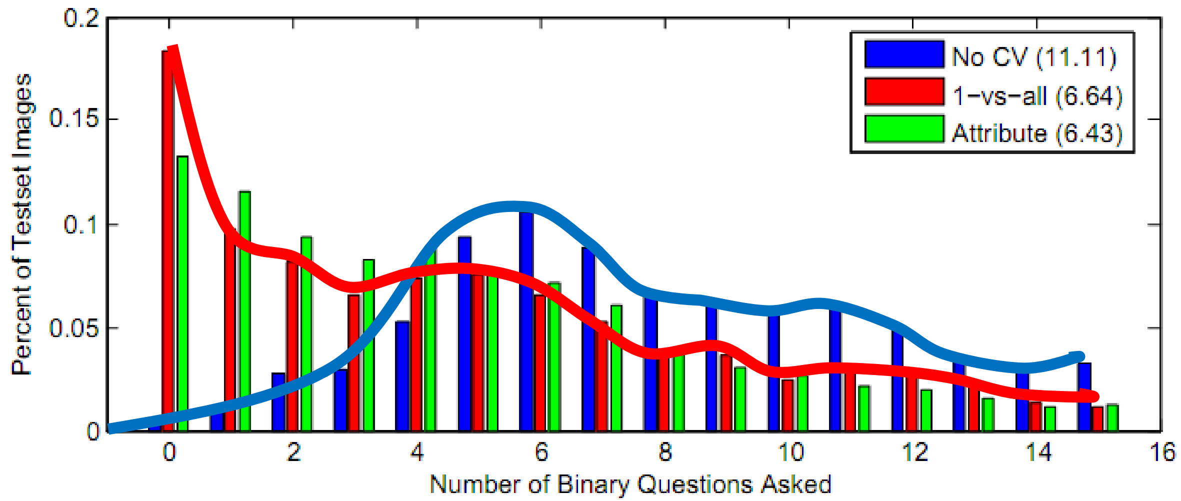
Experiments



- 2 Stop criteria:
  - Fixed number of questions – evaluate accuacy
  - User stops when bird identified – measure number of questions required.

Results





- Average number of questions to make ID reduced from 11.11 to 6.43
- Method allows CV to handle the easy cases, consulting with users only on the more difficult cases.

Key Observations

- Visual recognition reduces labor over a pure "20 Q" approach.
- Visual recognition improves performance over a pure "20 Q" approach. (69% vs 66%)
- User input dramatically improves recognition results. (66% vs 19%)

Strengths and weaknesses

- Handles very difficult data and yields excellent results.
- Plug-and-play with many recognition algorithms.
- Requires significant user assistance
- Reported results assume humans are perfect verifiers
- Is the reduction from 11 questions to 6 really that significant?