



ECS 269: Visual Recognition Fall 2019

Yong Jae Lee

Department of Computer Science

Plan for today

- Topic overview
- Course requirements and administrivia
- Introductions
- Syllabus tour

Computer Vision

- Let computers see!
- Automatic understanding of visual data (i.e., images and video)
 - Computing properties of the 3D world from visual data (*measurement*)
 - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (*perception and interpretation*)

What does recognition involve?



Detection: are there people?



Activity: What are they doing?



Object categorization



mountain

tree

building

banner

street lamp

vendor

people

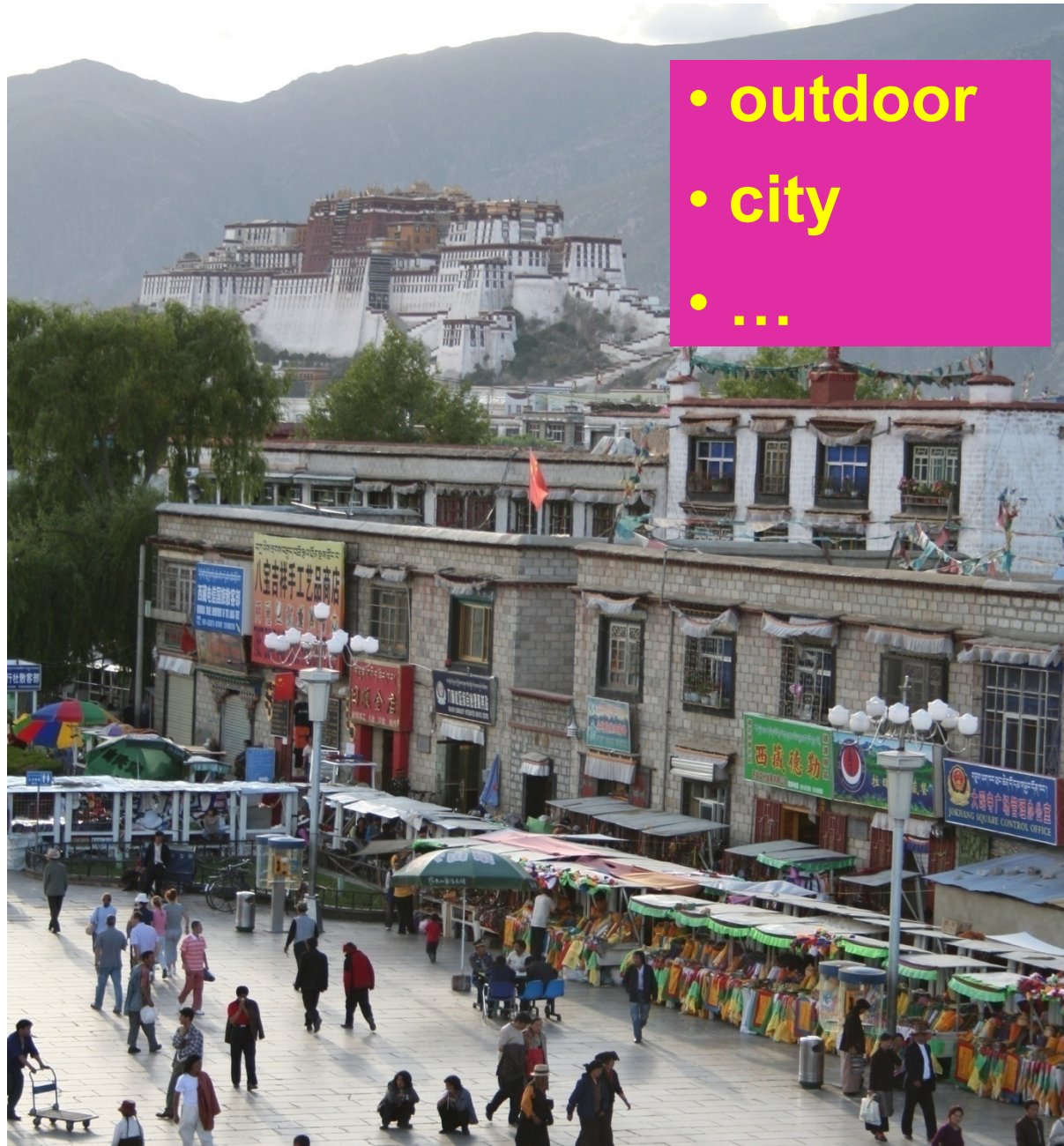
Instance recognition



**Potala
Palace**

**A particular
sign**

Scene categorization



Attribute recognition



gray

made of
fabric

crowded

flat

Why recognition?

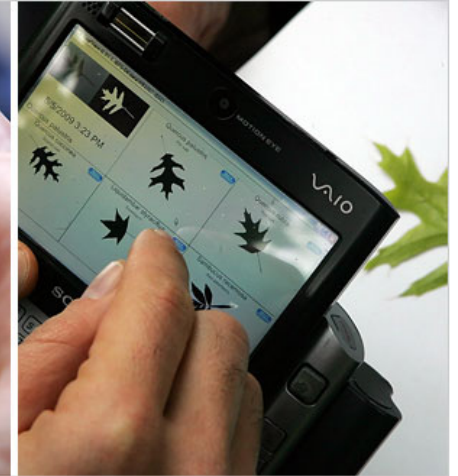
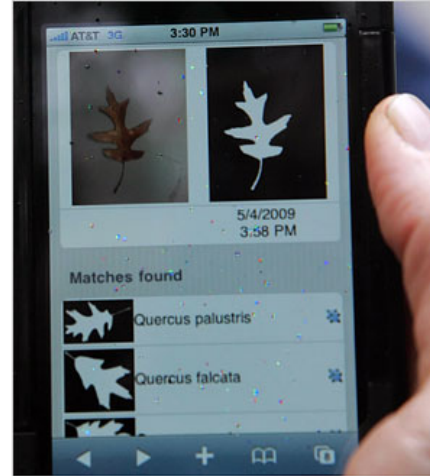
- Recognition a fundamental part of perception
 - e.g., robots, autonomous agents
- Organize and give access to visual content
 - Connect to information
 - Detect trends and themes

Posing visual queries



Yeh et al., MIT

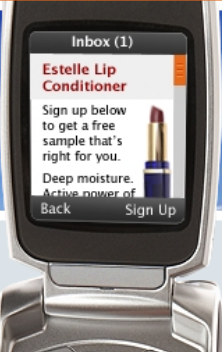
Digital Field Guides Eliminate the Guesswork



Belhumeur et al.

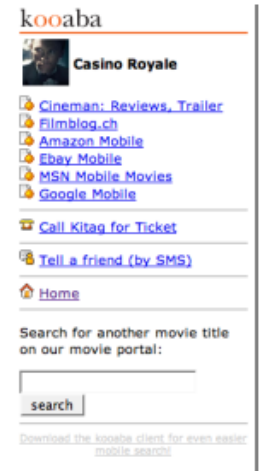
snaptell part of A9

Technology | News & Events



Get

Get back cool content on your phone. Videos, ringtones, WAP links and more!



Kooaba, Bay & Quack et al.

Finding visually similar objects

like visual shopping *alpha*

[My Like List](#) | [NewsLetter](#) | [Blog](#)

ALL SHOES BAGS WOMEN'S APPAREL MEN'S APPAREL KIDS ACCESSORIES JEWELRY & WATCHES HOLIDAY FOR THE HOME


IN Women's Shoes Search

Refine by Style Pumps Sandals Flats Pa
Refine by Color crimson taupe scarlet
Refine by Brand Clarks Sofft

Why is Like.com Different?
Like is a visual shopping engine that lets you find items by color, shape and pattern.
Click on [Likeness Search](#) to get started

Your Search Item



Which part of the image do you like?
Draw a box on the item to focus your search on that area.



Cole Haan - Carma OT Air Pump
\$278.95
[More Details](#) + [Save to LikeList](#)
[Shop at Zappos.com](#)

[All Products](#) > [Shoes](#) > [Women's Shoes](#) > [Cole Haan](#) > Cole Haan - Carma OT Air Pump

Search Results Results 1 - 20 of 140,207

Sort By [LikenessSM](#) [Price](#) Change Your View:  

1 2 3 4 5 6 7 [NEXT >>](#)

Natural Comfort - LV58
a sexy classic pump with a pillow-like footbed to keep your feet happy. leather or patent leather upper. wrapped memory-foam footbed. covered heel. leather sole.
[Compare Prices](#) [More Details](#) [Save to LikeList](#)
\$99.95
[Shop at Zappos.com](#)
Free Shipping Available
Shop for more items like this:
[Likeness Search](#)

Cole Haan 'Carma Air' Patent Leather Open Toe Pump
Open toe styles a sleek, cushioned pump with a wrapped heel and a mini platform. Color(s): black patent, dark chocolate suede, wine patent, black python, natural python, beige leather. Brand: Cole Haan.
[Compare Prices](#) [More Details](#) [Save to LikeList](#)
\$275.00
[Shop at NORDSTROM.com](#)
Shop for more items like this:
[Likeness Search](#)

rsvp - Caitlyn
an easy on the eyes pump features craftsmanship to make it easy on your feet too. patent leather uppers. almond shaped toe. cushioned footbed. covered heel. leather outsole. made in brazil. 7 oz.
[More Details](#) [Save to LikeList](#)
\$89.95
[Shop at Zappos.com](#)
Free Shipping Available
Shop for more items like this:
[Likeness Search](#)

Interactive systems

KINECT
for XBOX 360.



Shotton et al.



Autonomous agents



Mars rover



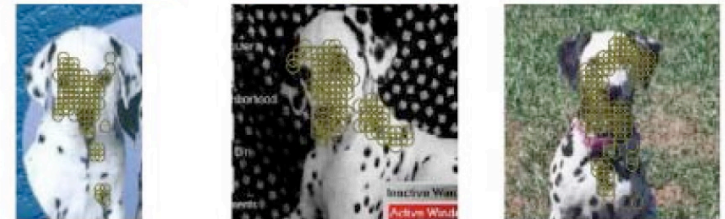
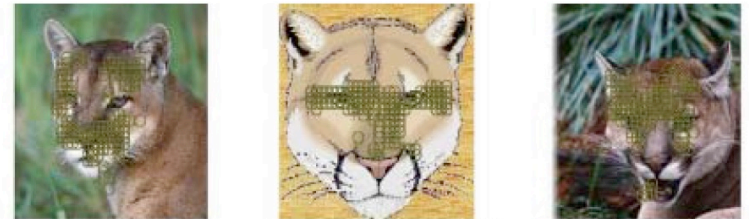
Google self-driving car

Discovering visual patterns



Actions

Wang et al.



Lee & Grauman

Categories



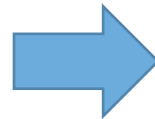
Doersch et al.

Characteristic elements

What are the challenges?



What we see



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

What a computer sees

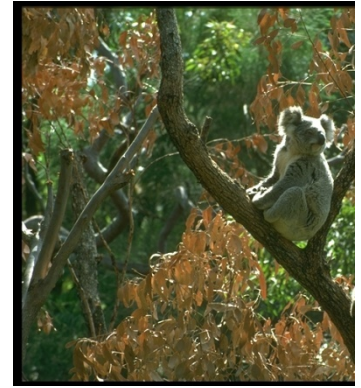
Challenges: robustness



Illumination



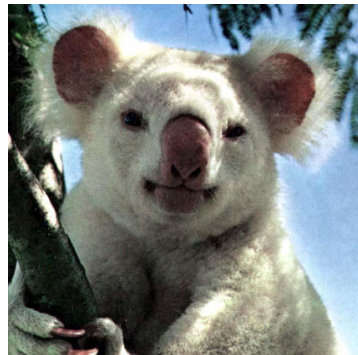
Object pose



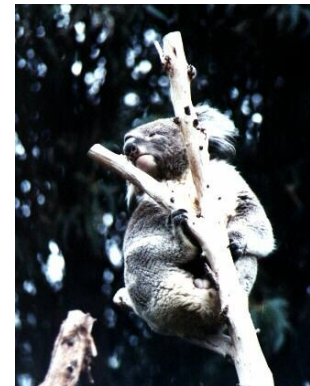
Clutter



Occlusions



Intra-class
appearance



Viewpoint

Challenges: context and human experience



Context cues

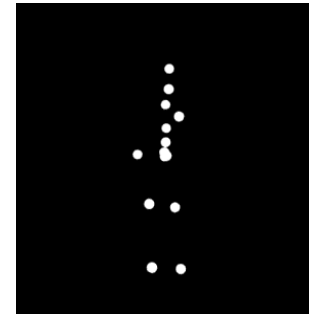
Challenges: context and human experience



Context cues

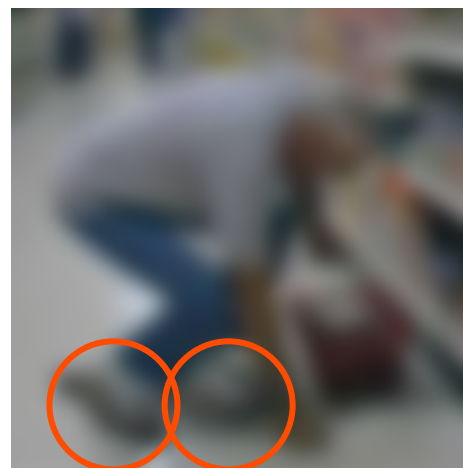
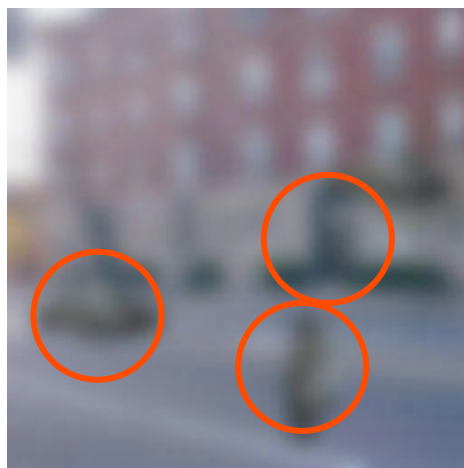
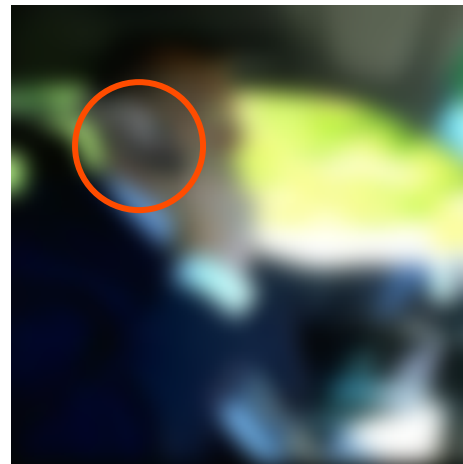
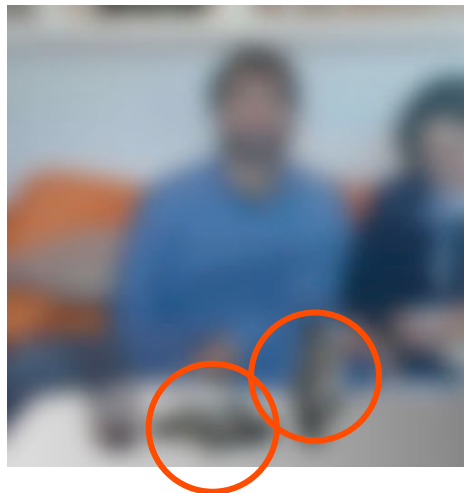
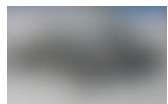


Function



Dynamics

Challenges: context and human experience



Challenges: context and human experience



Challenges: context and human experience



Challenges: scale, efficiency

- Half of the cerebral cortex in primates is devoted to processing visual information

Challenges: scale, efficiency

The Flickr logo, with 'flickr' in blue and 'r' in pink.

10 billion images

The Facebook logo, consisting of the word 'facebook' in white on a blue rounded rectangle background.

250 billion images

The Imgur logo, featuring the text 'the simple image sharer' in green and white above the word 'imgur' in white on a black background.

1 billion images
served daily

The YouTube logo, with 'You' in black and 'Tube' in white on a red rounded rectangle background.

300 hours uploaded per
minute

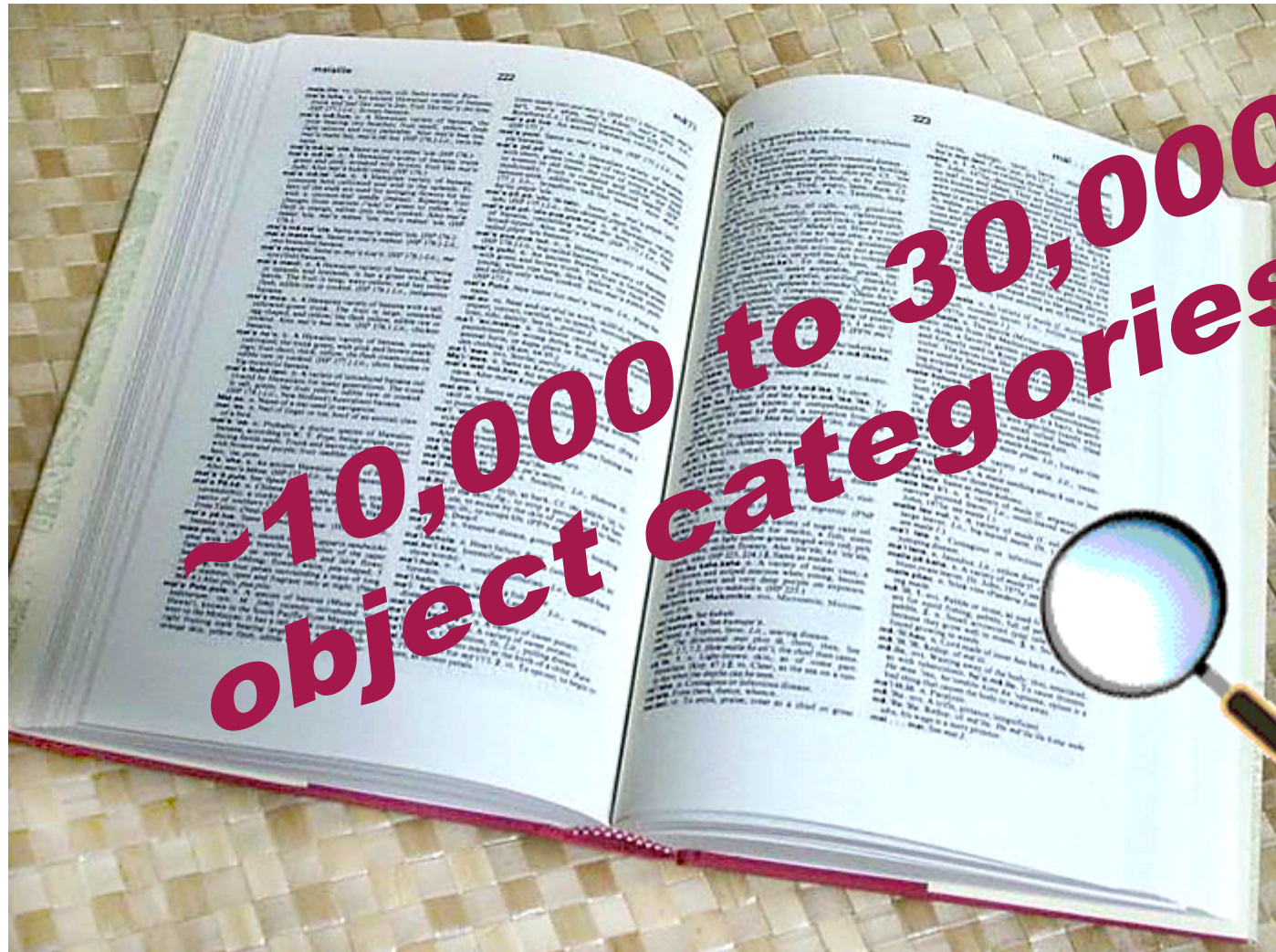
The Photobucket logo, featuring a camera icon and the word 'photobucket' in blue.

10 billion images

From  CISCO™

Almost **90%** of web traffic is visual!

Challenges: scale, efficiency

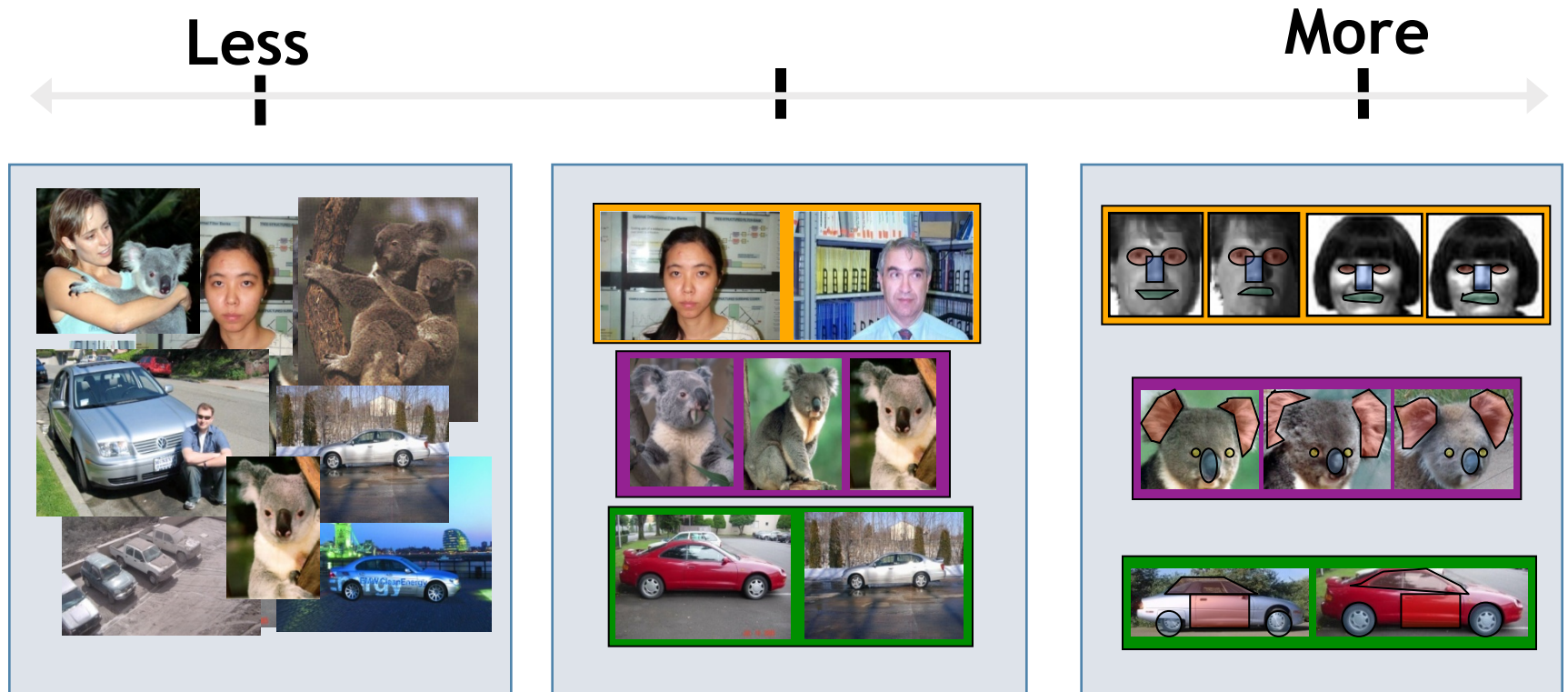




~10,000 to 30,000



Challenges: learning with minimal supervision



Unlabeled,
multiple objects

Classes labeled,
some clutter

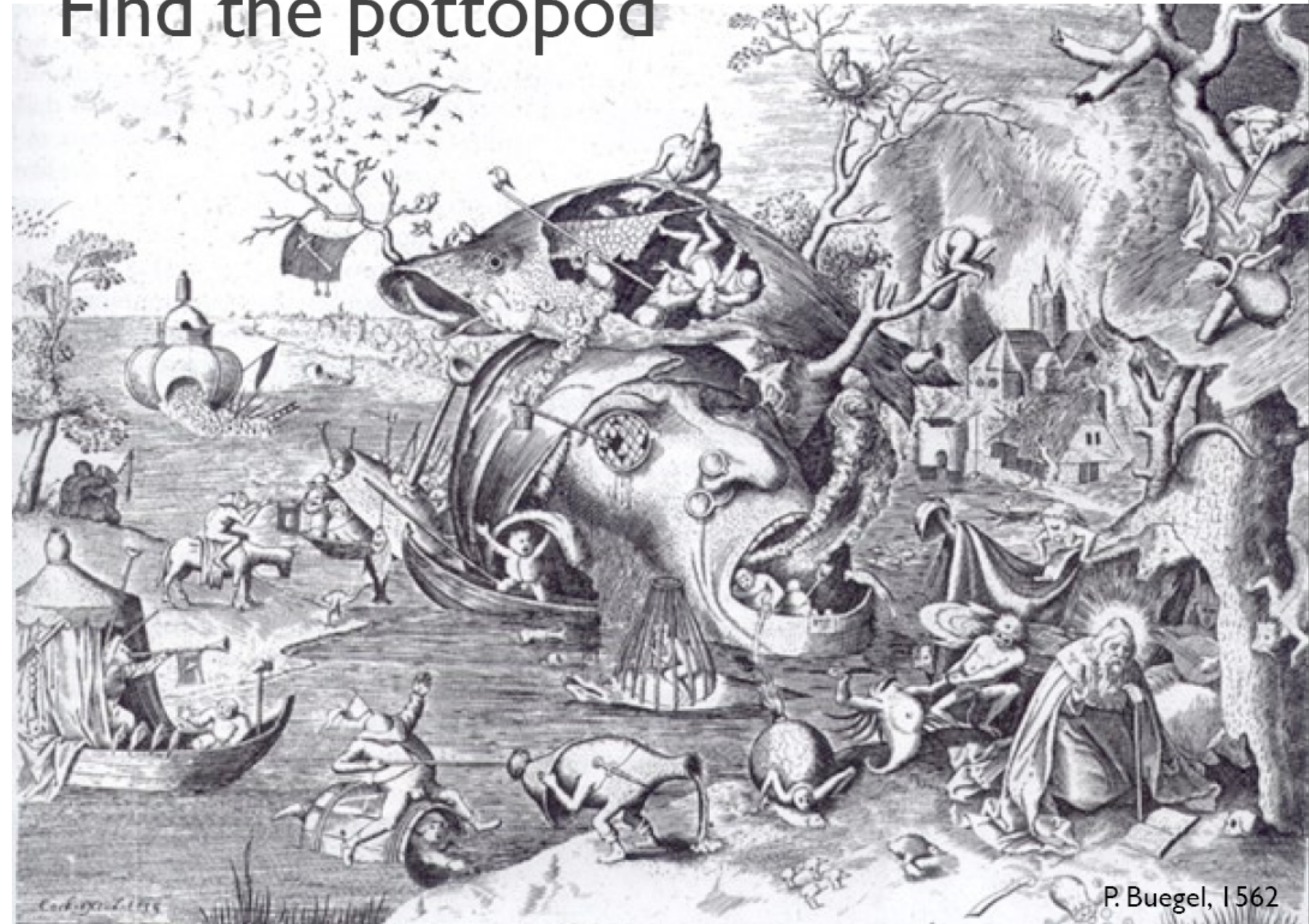
Cropped to object,
parts and classes
labeled



This is a
pottopod

S. Savarese, 2003

Find the pottopod



P. Buegel, 1562

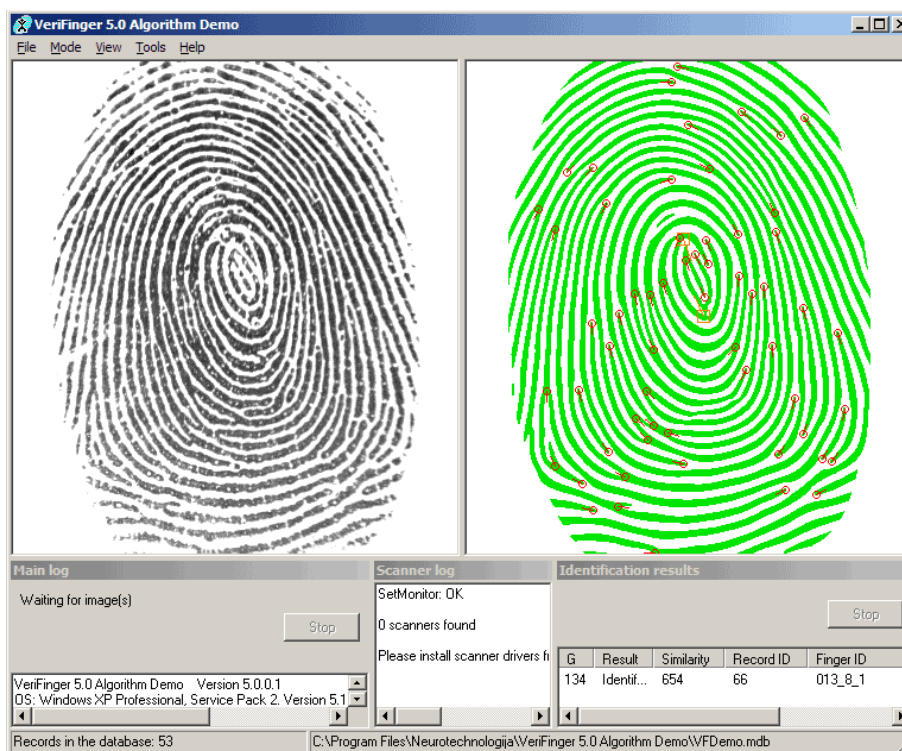
What works today

- Reading license plates, zip codes, checks

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

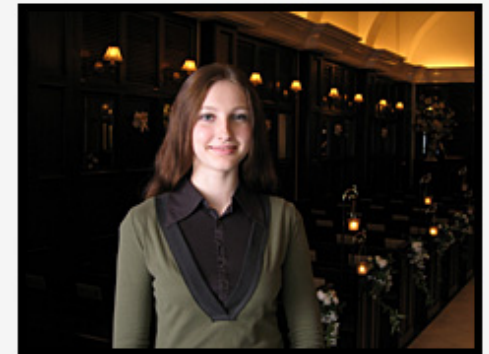
What works today

- Reading license plates, zip codes, checks
- Fingerprint recognition



What works today

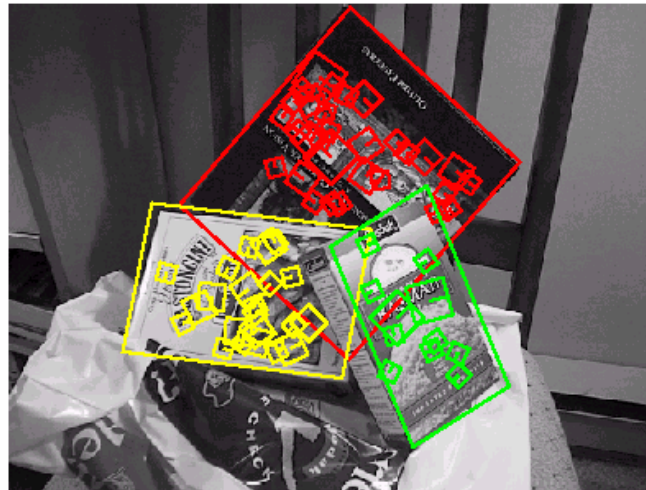
- Reading license plates, zip codes, checks
- Fingerprint recognition
- Face detection



[Face priority AE] When a bright part of the face is too bright

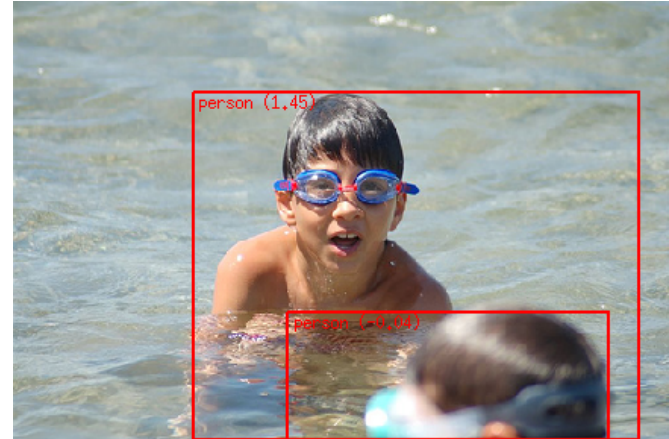
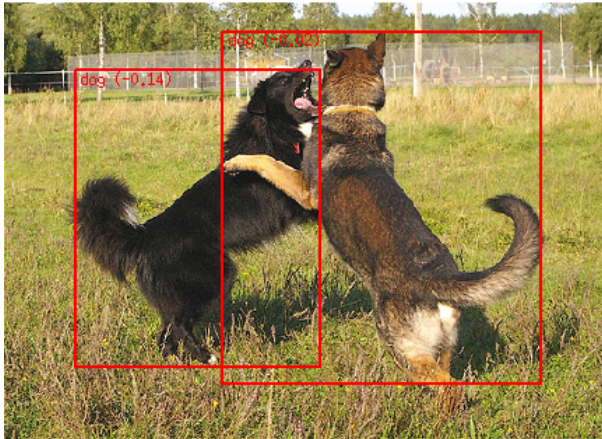
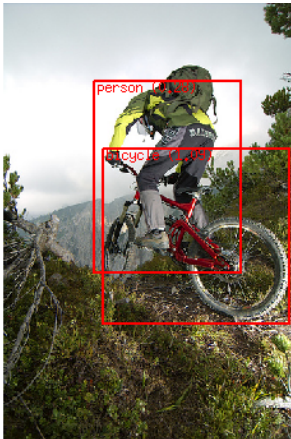
What works today

- Reading license plates, zip codes, checks
- Fingerprint recognition
- Face detection
- Recognition of flat textured objects (CD covers, book covers, etc.)

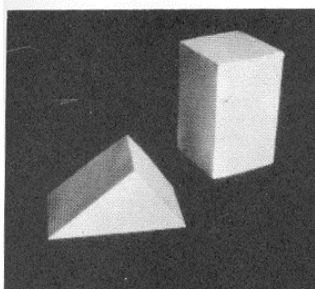


What works today

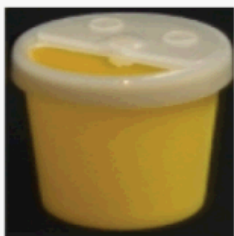
- Reading license plates, zip codes, checks
- Fingerprint recognition
- Face detection
- Recognition of flat textured objects (CD covers, book covers, etc.)
- Recognition of generic categories(*)!



Progress charted by datasets



Roberts 1963

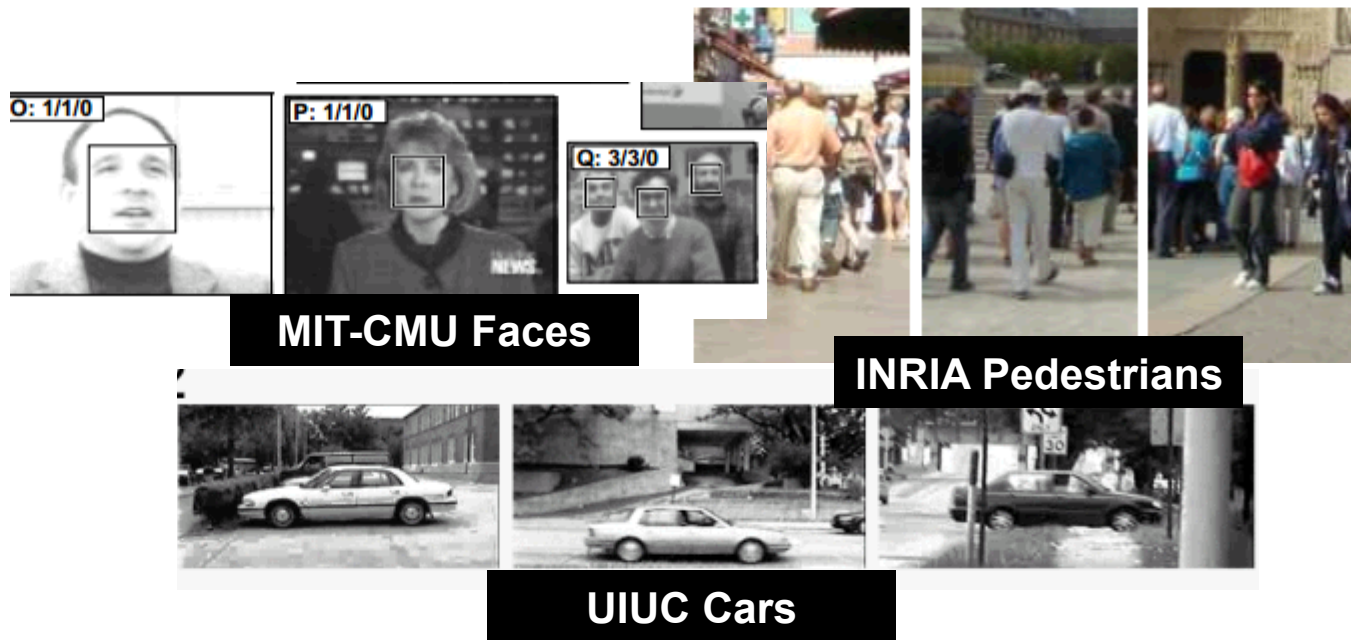


COIL



1963 ... 1996

Progress charted by datasets



1963 ... 1996 2000



Progress charted by datasets



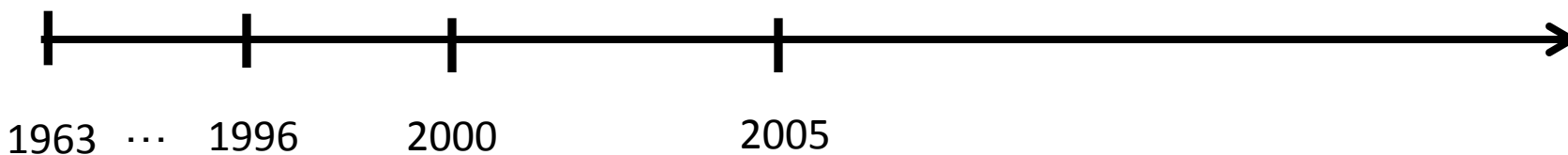
MSRC 21 Objects



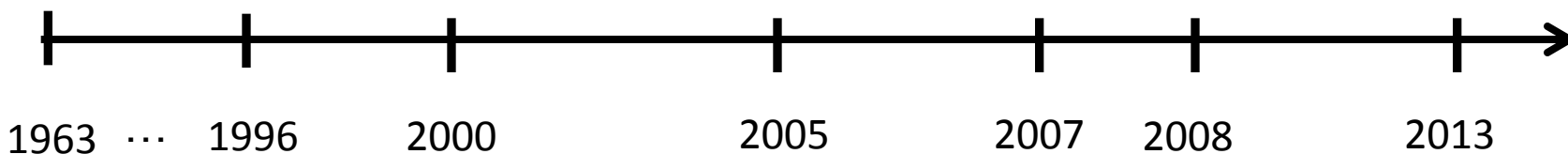
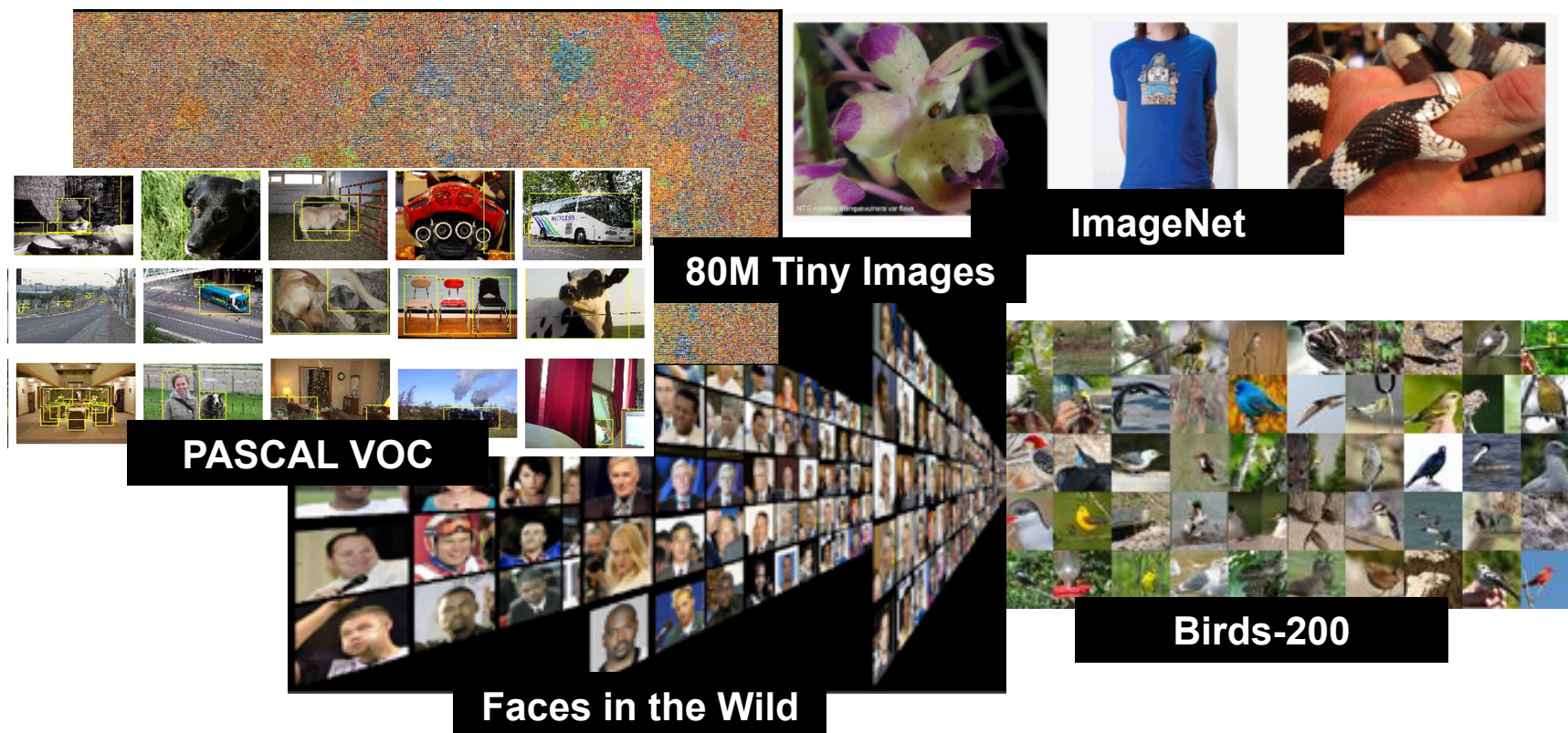
Caltech-101



Caltech-256



Progress charted by datasets



Evolution of methods

- Hand-crafted models
- 3D geometry
- Hypothesize and align
- Hand-crafted features
- Learned models
- Data-driven
- “End-to-end” learning of features and models*



* Labeled data, fast compute (GPUs).

Course Information

- Website:
<https://sites.google.com/a/ucdavis.edu/ecs-269-visual-recognition-fall-2019/>
- Canvas:
<https://canvas.ucdavis.edu/courses/361432>
- 4 Units
- Office hours: by appointment (email)
- Email: start subject with “[ECS269]”

TA

- Krishna Kumar Singh
- ecs269fall2019TA@gmail.com
- PhD student in CS



- Yuheng Li
- ecs269fall2019TA@gmail.com
- MS student in CS



This course

- Survey state-of-the-art research in computer vision
- High-level vision and learning problems with focus on recent deep learning techniques

Goals

- Understand, analyze, and discuss state-of-the-art techniques
- Identify interesting open questions and future directions

Prerequisites

- Course in computer vision (ECS 174) or machine learning (ECS 171)

Requirements

- Class Participation **[15%]**
- Paper Reviews **[25%]**
- Paper Presentations (1 time) **[30%]**
- Final Project **[30%]**

Class participation

- Actively participate in discussions
- **Read assigned paper** before each class
- Email TA after class if you contributed in any way to the discussions that day (1 sentence summary)

Paper reviews

- **Detailed review of paper** before each class
- *At most* one page in length
 - A summary of the paper (2-3 sentences)
 - Main contributions
 - Strengths and weaknesses
 - Experiments convincing?
 - Extensions?
 - Additional comments, including unclear points, open research questions, and applications
- Upload to Canvas by **11:59 pm the day before each class**

Paper presentations

- 1 time during the quarter (in groups of 2)
- ~20 minutes long, well-organized, and polished:
 - Clear statement of the problem
 - Motivate why the problem is interesting, important, and/or difficult
 - Describe key contributions and technical ideas
 - Experimental setup and results
 - Strengths and weaknesses
 - Interesting open research questions, possible extensions, and applications
- Slides should be clear and **mostly visual** (figures, animations, videos)
- Look at background papers as necessary

Paper presentations

- Search for relevant material, e.g., from the authors' webpage, project pages, etc.
- **Each slide that is not your own must be clearly cited**
- Meet with TA **one day before your presentation** with prepared slides; your responsibility to email TA to schedule a time
- Upload slides to Canvas on day of presentation (after class is fine)
- Skip paper review the day you are presenting

Final project

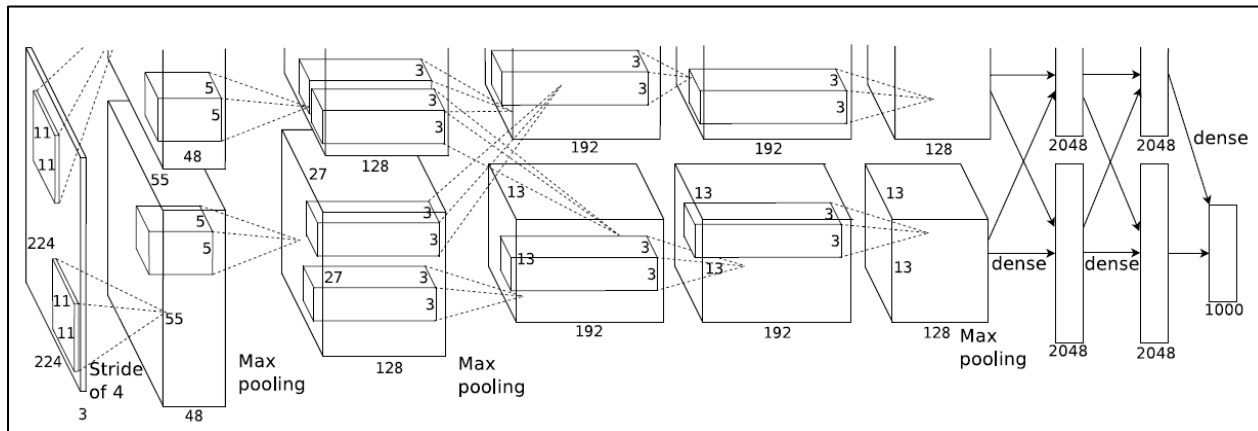
- One of:
 - Design and evaluation of a novel approach
 - An extension of an approach studied in class
 - In-depth analysis of an existing technique
 - In-depth comparison of two related existing techniques
- Work in groups of 4 or 5
- Talk to me/TA if you need ideas
- Google Cloud Credits (will be explained in a future class)

Final project

- Final Project Proposal (5%) due **10/23**
 - 1 page
- Final Project Presentation (10%) due **12/2, 12/4, and 12/6**
 - 10 minutes
- Final Project Paper (15%) due **12/9**
 - 6-8 pages
- Upload pdf/slides to Canvas

Introductions

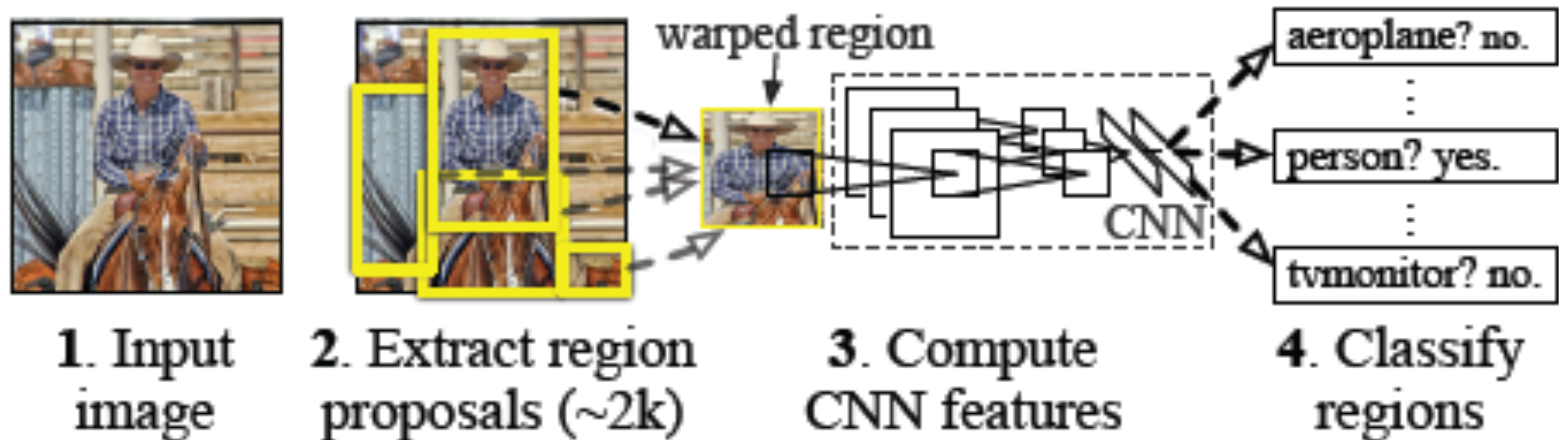
Image classification



- Convolutional neural networks AlexNet, ResNet

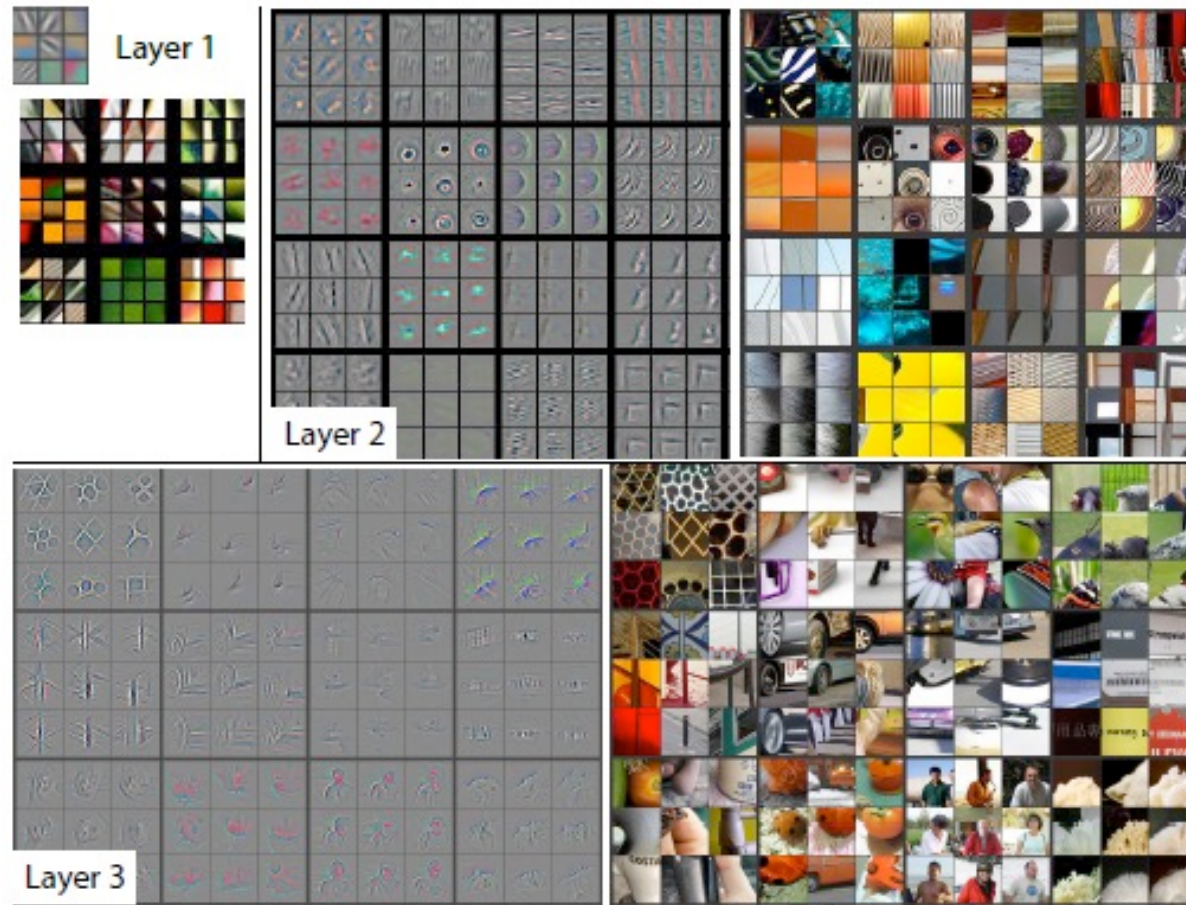
Object detection

R-CNN: *Regions with CNN features*



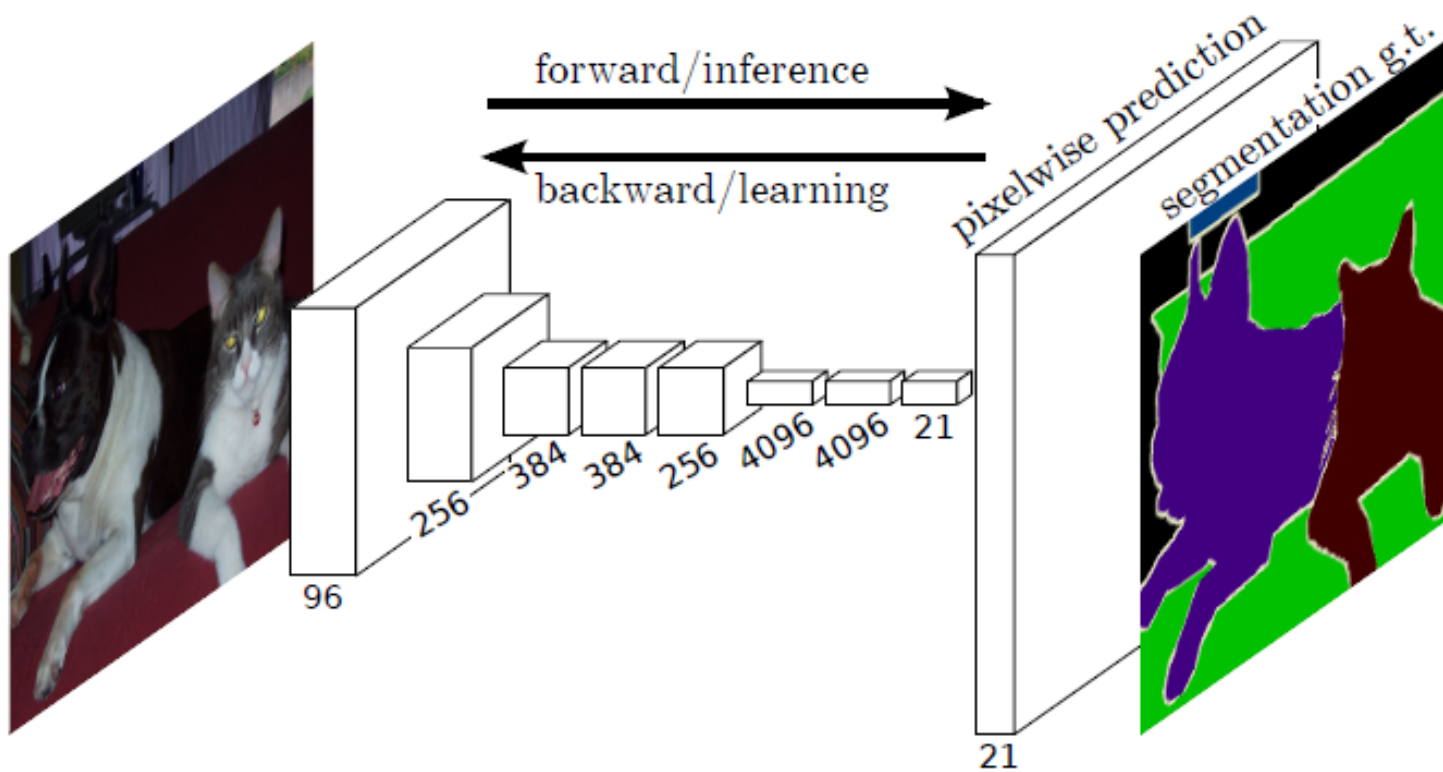
- Two-stage and single-stage frameworks

CNN Visualization and Analysis



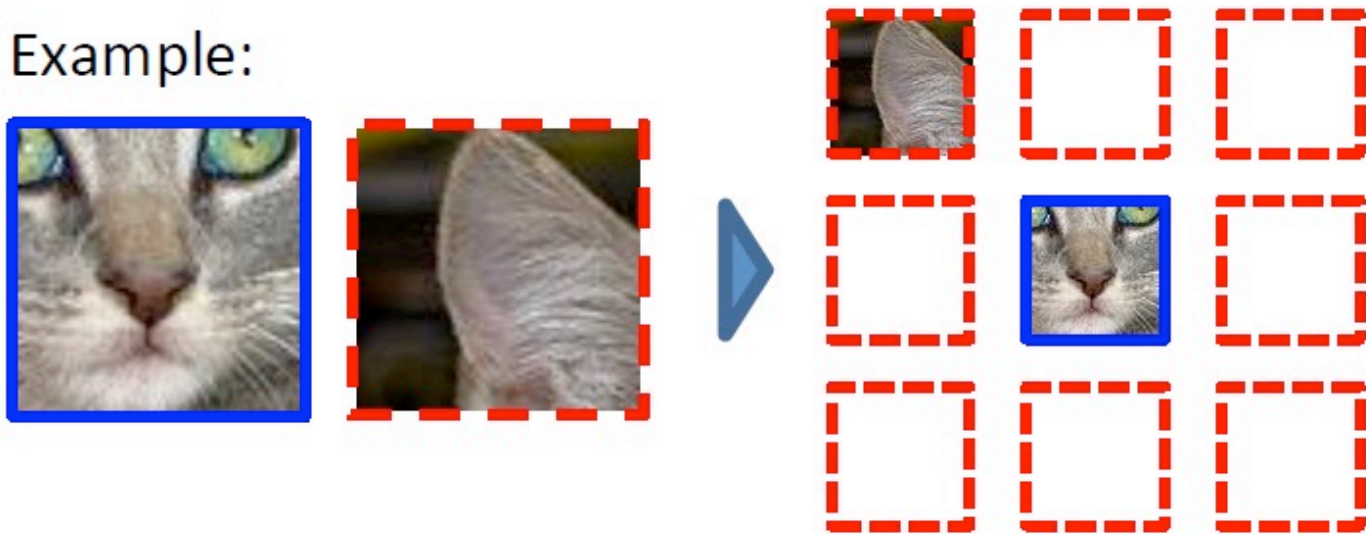
- What is encoded by a CNN?

Semantic Segmentation



Self-supervision

Example:



Question 1:



Question 2:



Neural Network Art



Coming up

- Carefully read class webpage
- Sign-up for papers you want to present (spreadsheet will be emailed shortly)
- Next class: CNN basics