CS 3710: Visual Recognition
Introduction

Adriana Kovashka
Department of Computer Science

January 6, 2015

Course Info

- **Course website:** http://people.cs.pitt.edu/~kovashka/cs3710_sp15/
- **Instructor**: Adriana Kovashka
- **Email**: kovashka@cs.pitt.edu (Please use "CS3710" at the beginning of the subject line.)
- **Office**: Sennott Square 5325
- **Office hours**: by appointment
- **Grades** on Blackboard / CourseWeb
- **Feedback** is welcome!

Plan for Today

- Introductions
- What is visual recognition?
  - How well does it work?
  - What are the challenges?
- Overview of topics
- Course structure and requirements
- Machine learning quiz (for self-evaluation)

Introductions

Introductions

- What is your name?
- What department are you at? Which year?
- What are your research interests?
- What one thing outside of research are you passionate about?
- What do you hope to get out of this class?

Visual Recognition

What is Computer Vision?

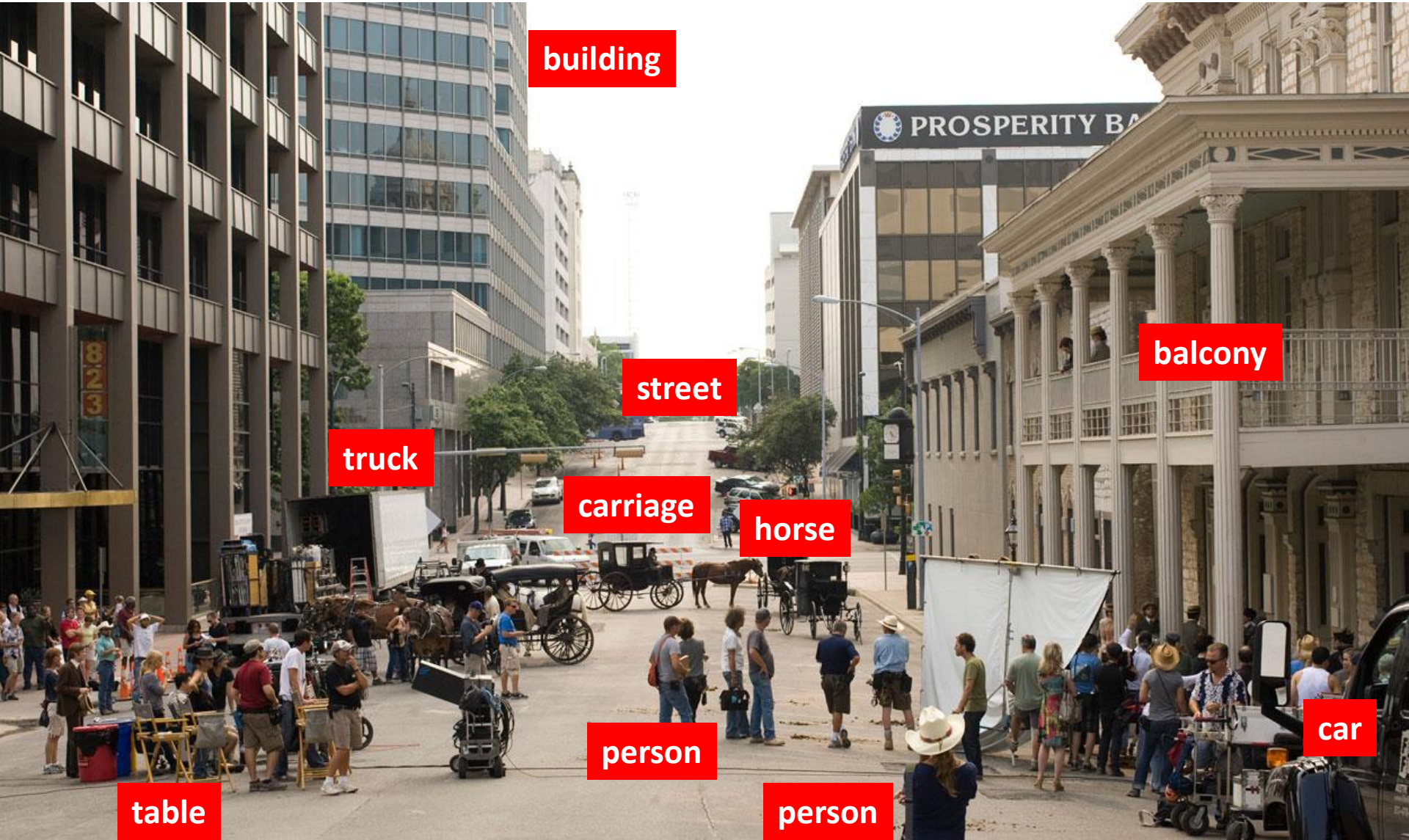
- Automatic understanding of images and video
  - Computing properties of the 3D world from visual data *(measurement)*
  - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities *(perception and interpretation)*
  - Algorithms to mine, search, and interact with visual data (*search and organization*)

Visual recognition


PROSPERITY BA
823

Recognition: What is this?

Recognition: What is this?



Detection: Where are the cars?

Activity: What is this person doing?

Scene: Is this an indoor scene?

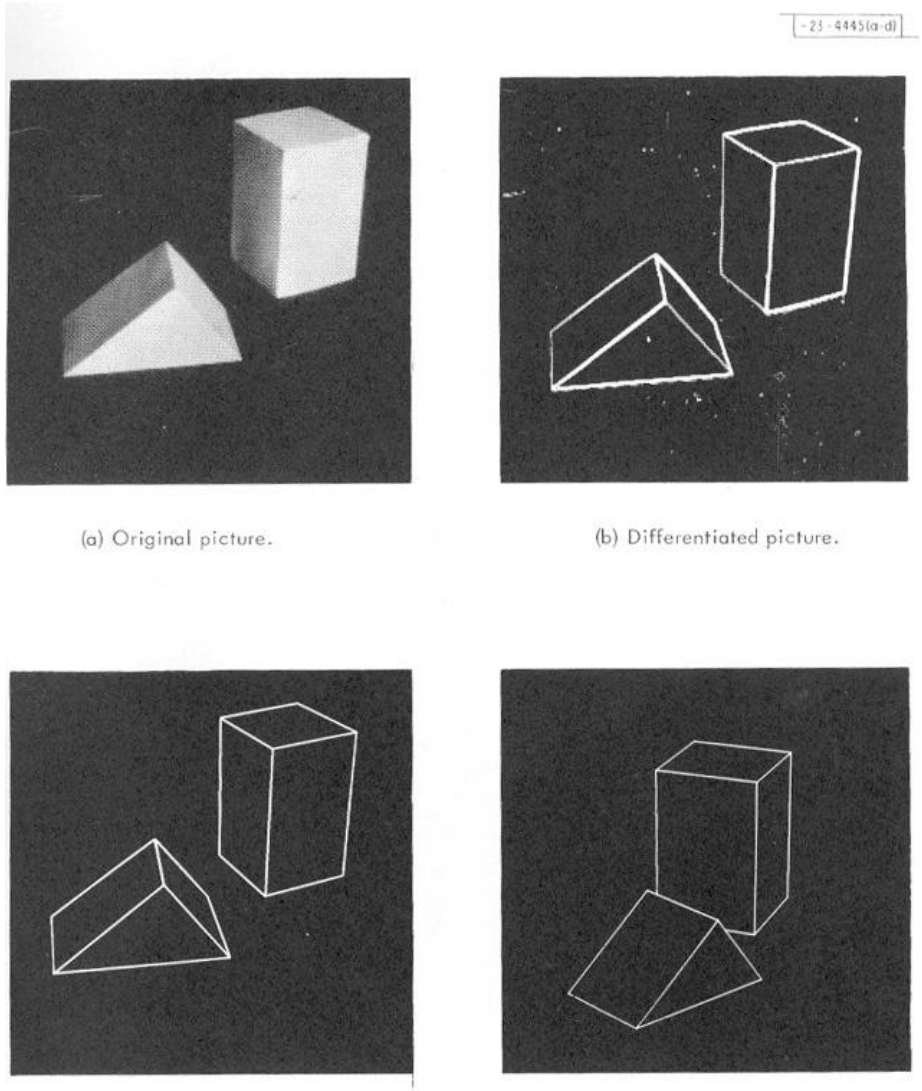
Instance: Which city? Which building?

Why do visual recognition?

- Allows us to automate various processes
  - Checkout at a store
  - Surveillance
  - Annotating football videos
  - … etc.
- Enables content-based image search and organization
- Perception is an integral part of robotics

How has it evolved?
How well does it work?

Inputs in 1963…



(a) Original picture.

(b) Differentiated picture.

(c) Line drawing.

(d) Rotated view.

L. G. Roberts, *Machine Perception of Three Dimensional Solids*, Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

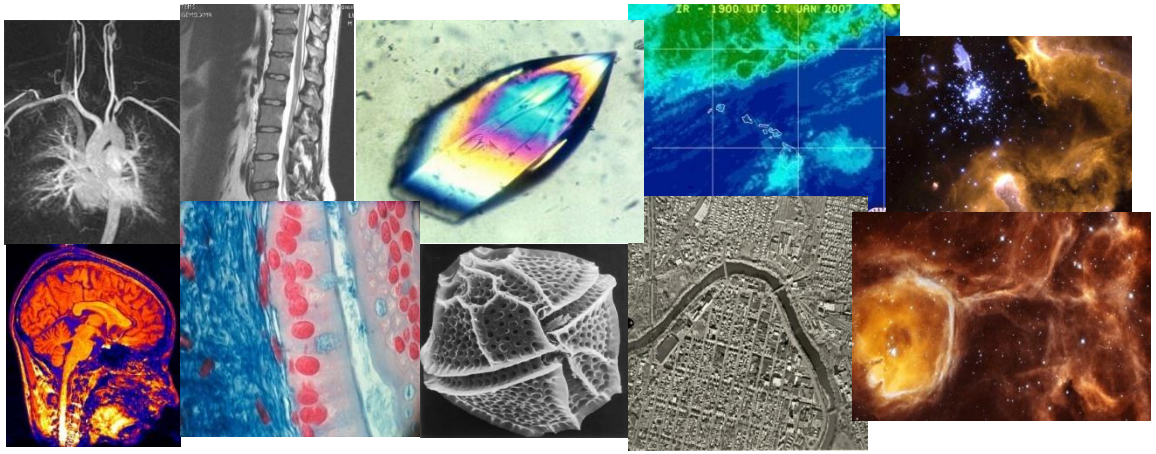
… and inputs today

Personal photo albums

Movies, news, sports

Google Image Search Picasa flickr webshots picsearch YouTube Broadcast Yourself

Surveillance and security

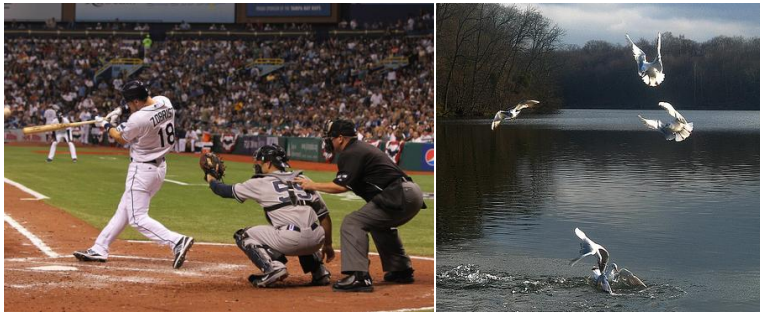


Medical and scientific images

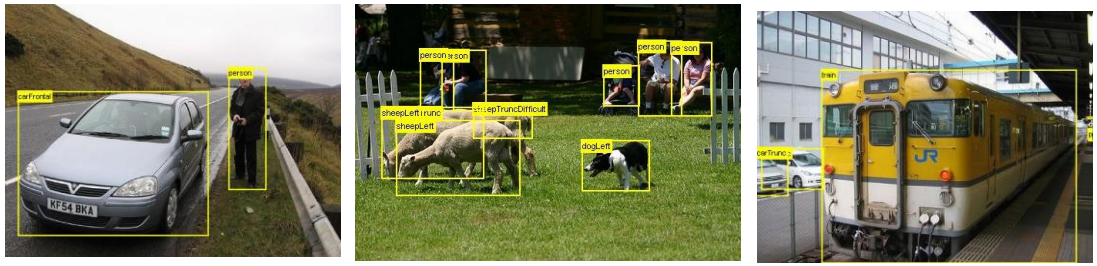
Svetlana Lazebnik

Datasets today

ImageNet:
22k categories, 14mil images



Microsoft COCO:
70 categories, 300k images



PASCAL:
20 categories, 12k images

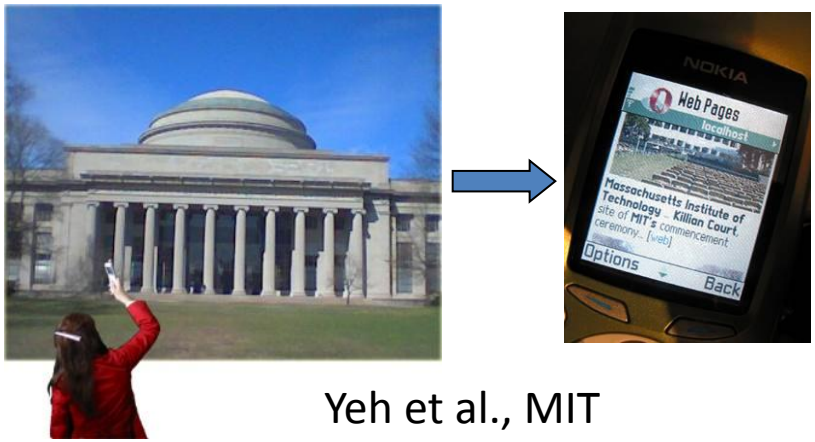
SUN:
5k categories, 130k images

Detecting faces

Interactive systems

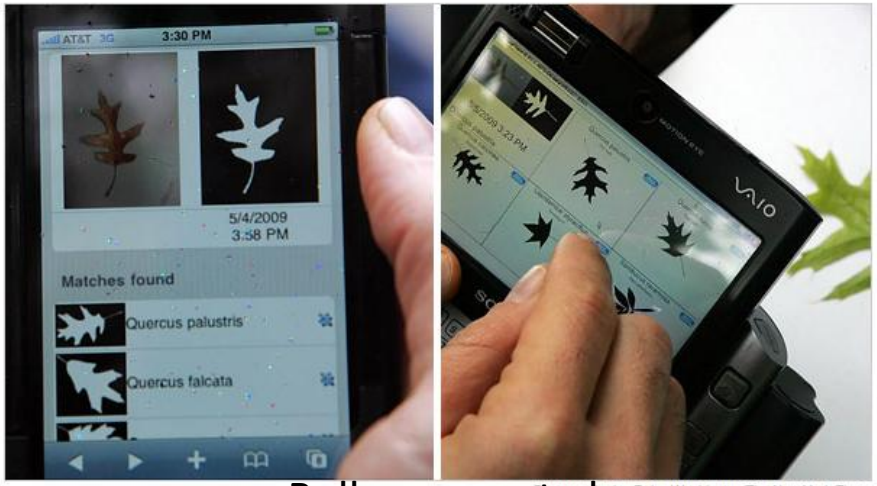
Shotton et al.

Posing visual queries



Yeh et al., MIT

Digital Field Guides Eliminate the Guesswork



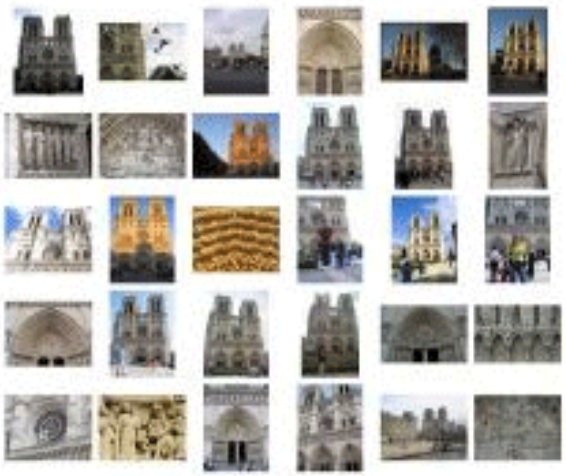
Belhumeur et al.

Kooaba, Bay & Quack et al.

Exploring photo collections

Photo Tourism

Exploring photo collections in 3D



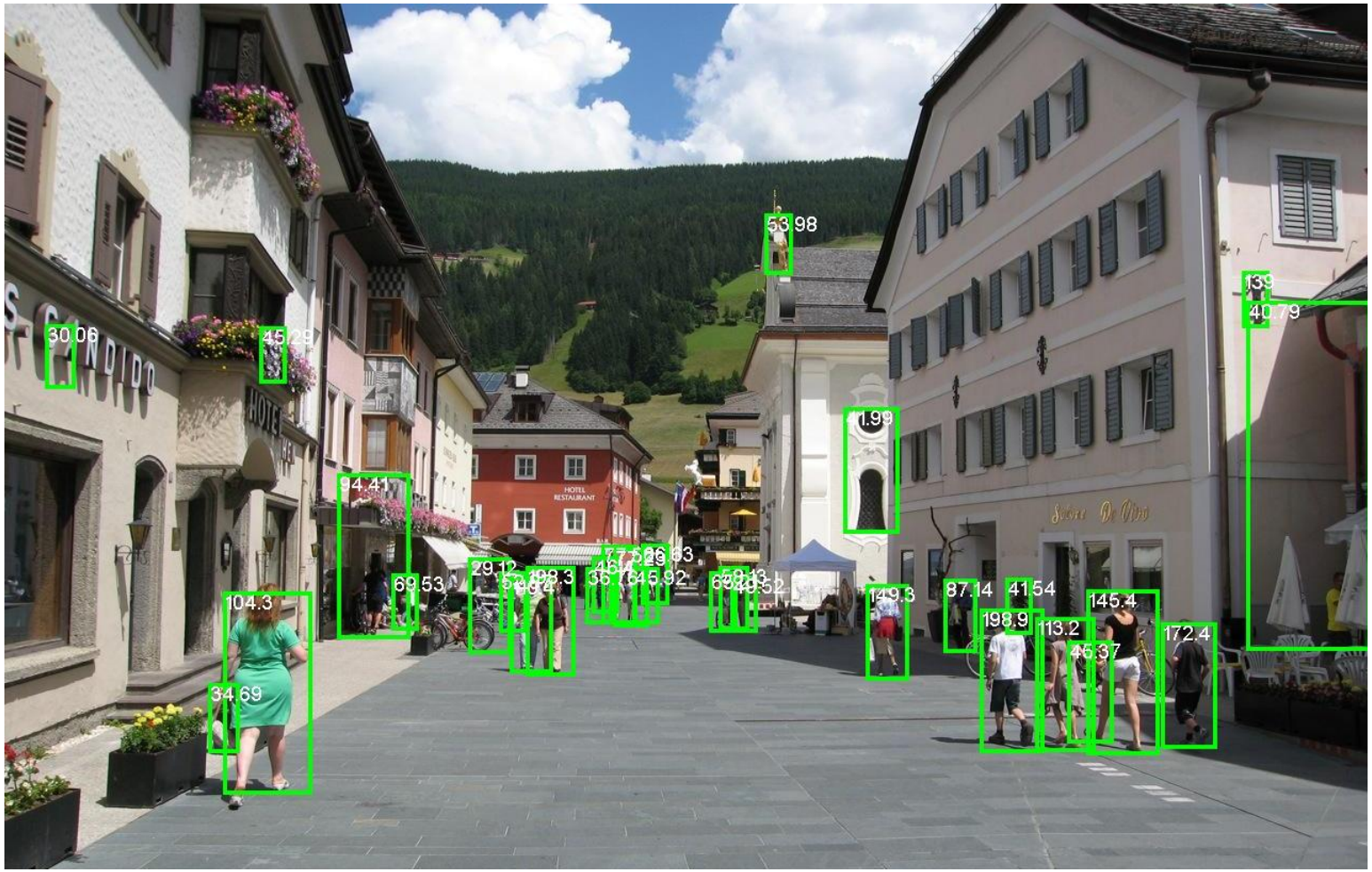
(a)

(b)

(c)

Snavely et al.

Detecting pedestrians



Devi Parikh

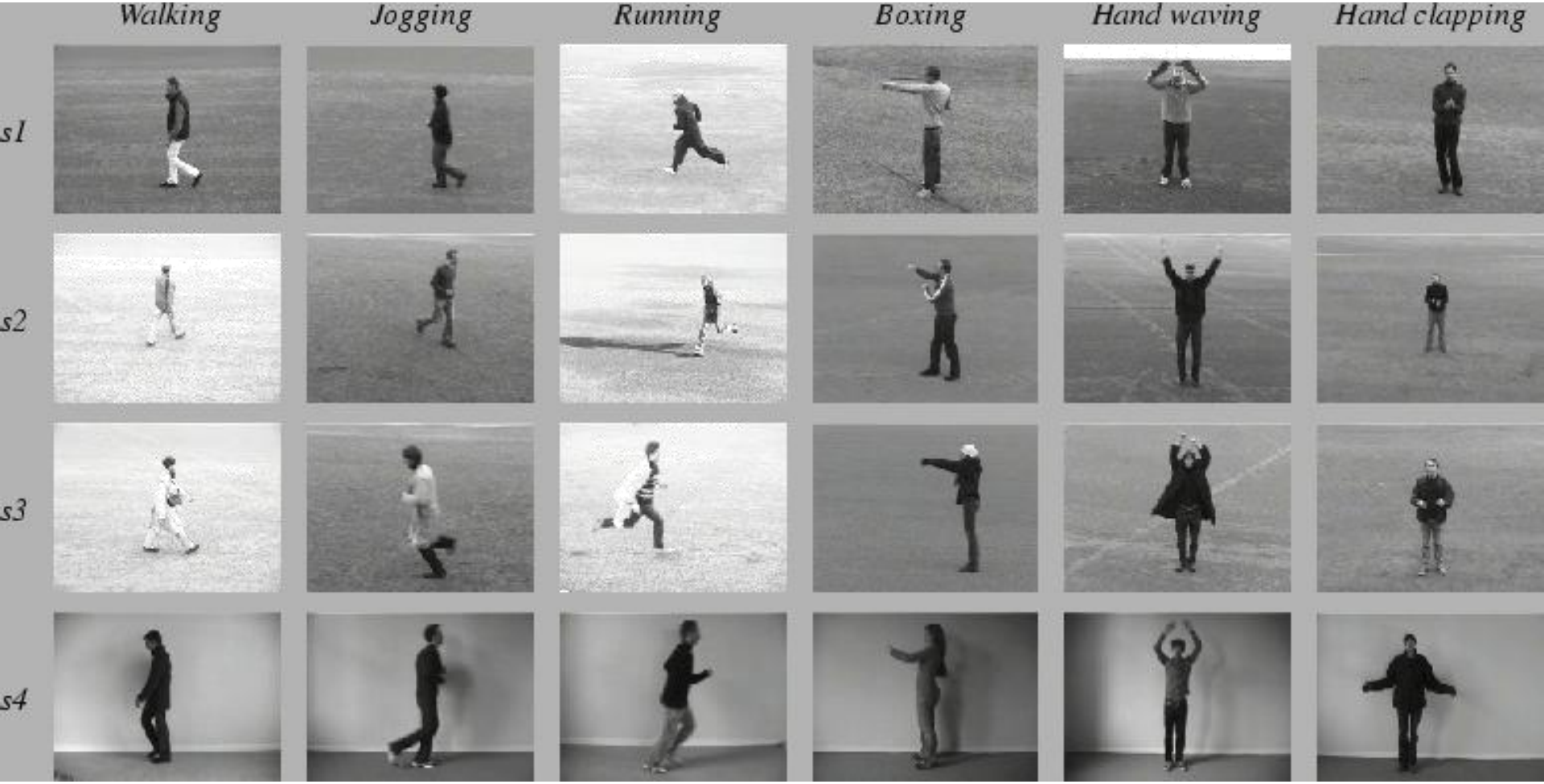
Dollar et al., BMVC 2009

Autonomous agents detecting objects

Google self-driving car

Mars rover

Classifying simple actions



Schuldt et al., ICPR 2004

What are the challenges?

Challenges: Appearance variation

Illumination

Object pose

Viewpoint

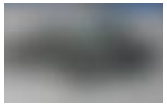
Intra-class appearance

Occlusions

Clutter

Kristen Grauman

Challenges: Context and function



Antonio Torralba, Kristen Grauman

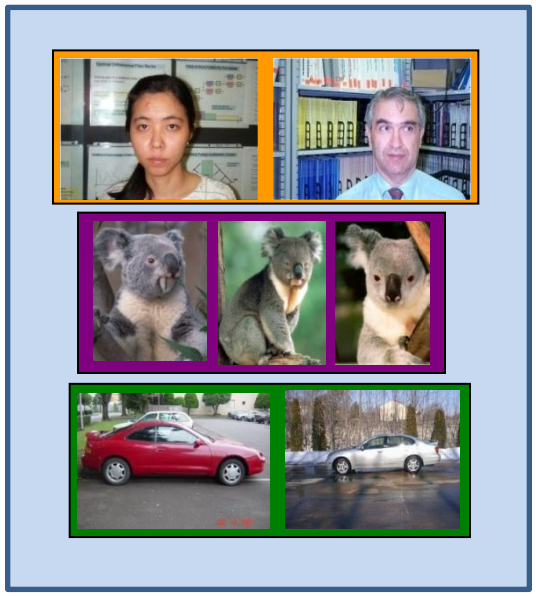
Challenges: Scale and efficiency

- ~20 hours of video added to YouTube per minute
- ~5,000 new tagged photos added to Flickr per minute
- Thousands to millions of pixels in an image
- 30+ degrees of freedom in the pose of articulated objects (humans)
- 3,000-30,000 human recognizable object categories
- Half of the human brain is devoted to processing visual information

Challenges: Limited supervision

Less — **More**

Unlabeled, multiple objects



Classes labeled, some clutter

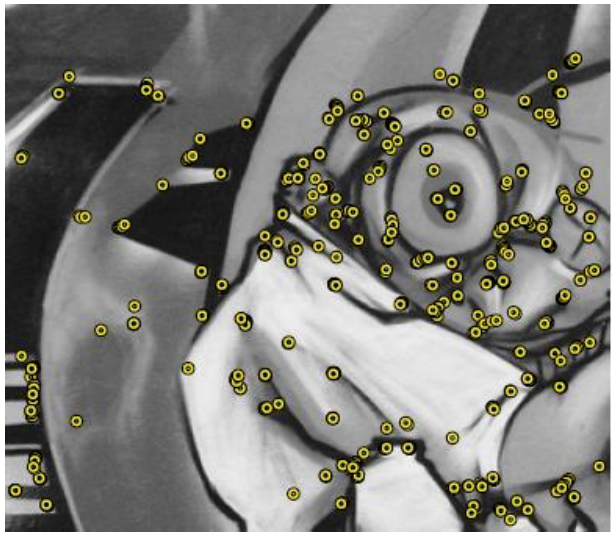
Cropped to object, parts and classes labeled

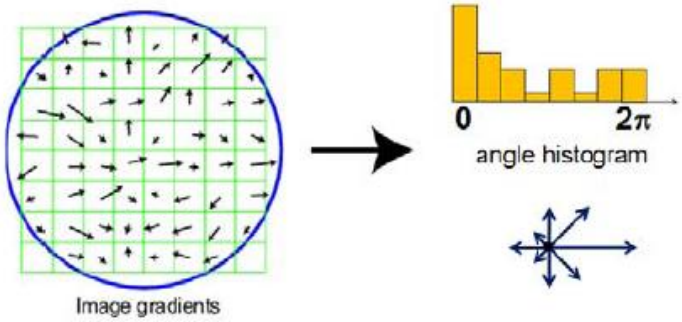
Kristen Grauman

Overview of Topics

Feature detection and matching

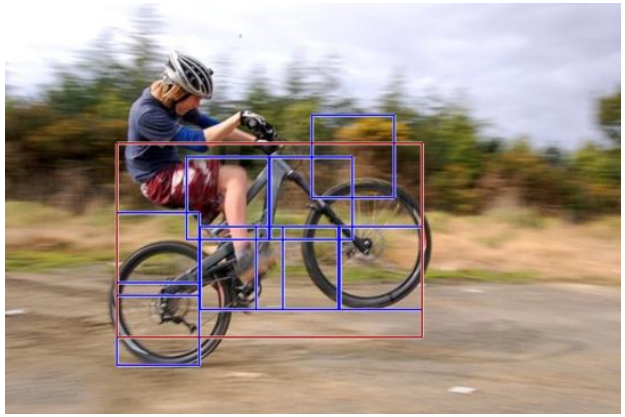
- Detecting repeatable features
- Describing images with local statistics
- Matching features across images

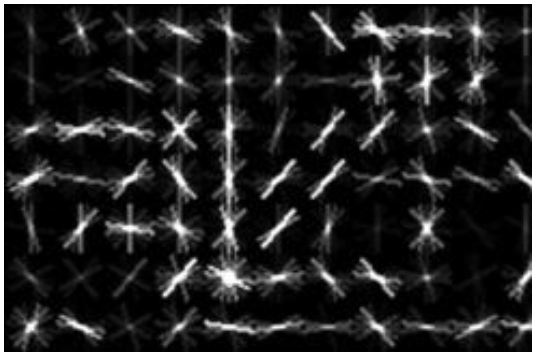


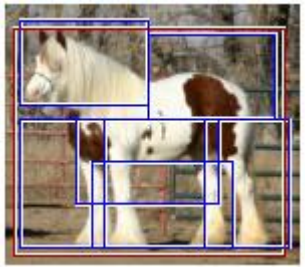


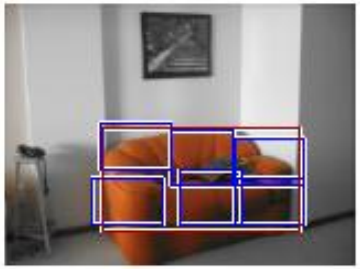
Detection and classification

- Detecting novel instances of objects
- Classifying regions as one of several categories



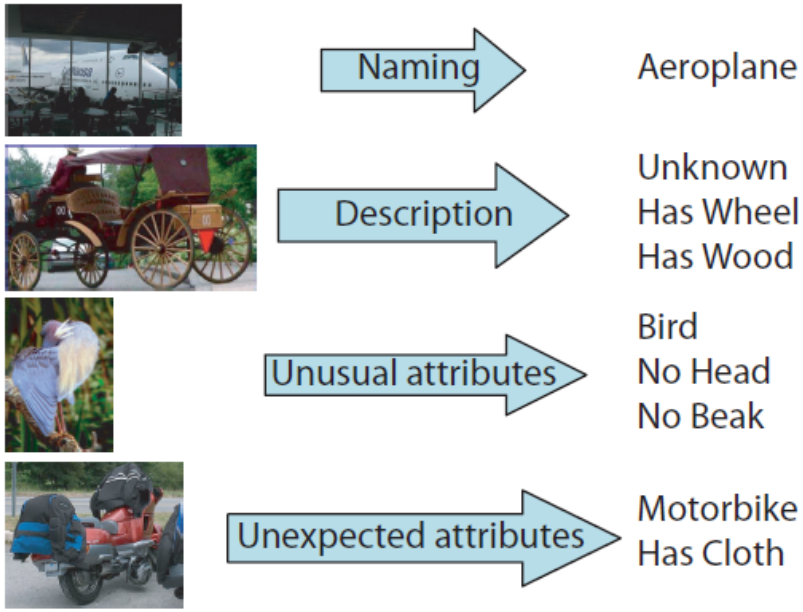






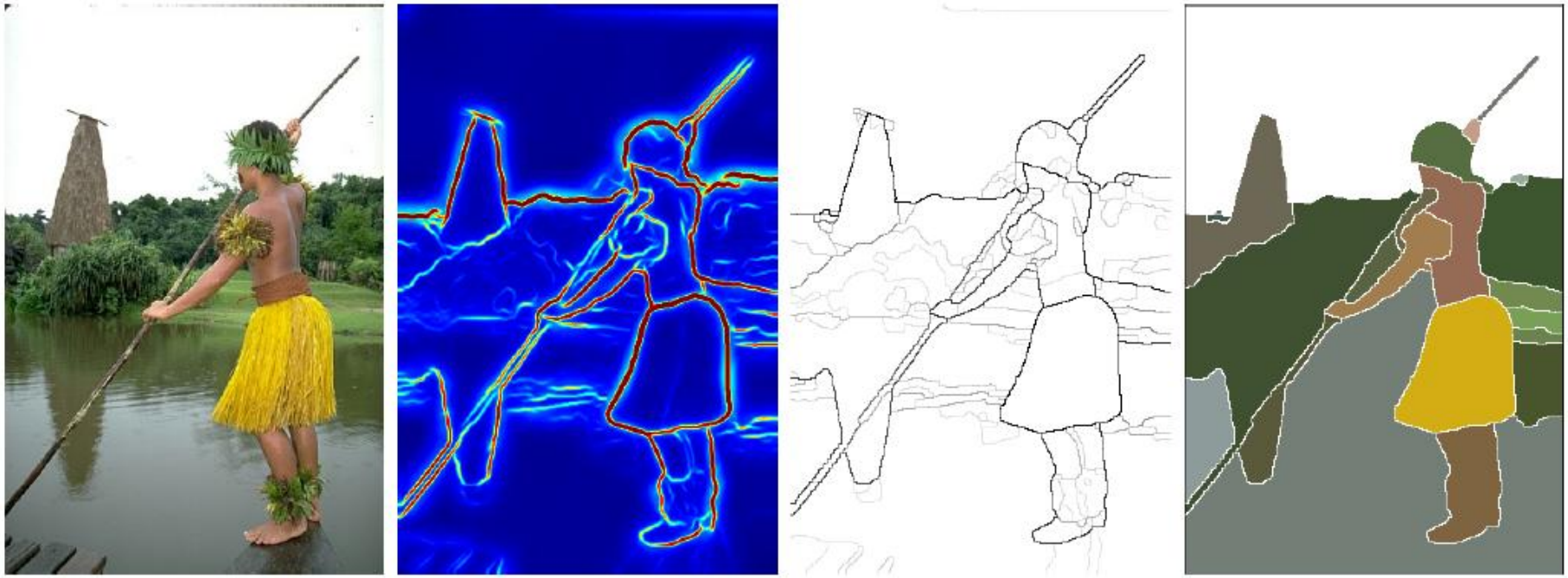
Attributes

- Describing the high-level properties of objects
- Measuring the degree of attribute presence



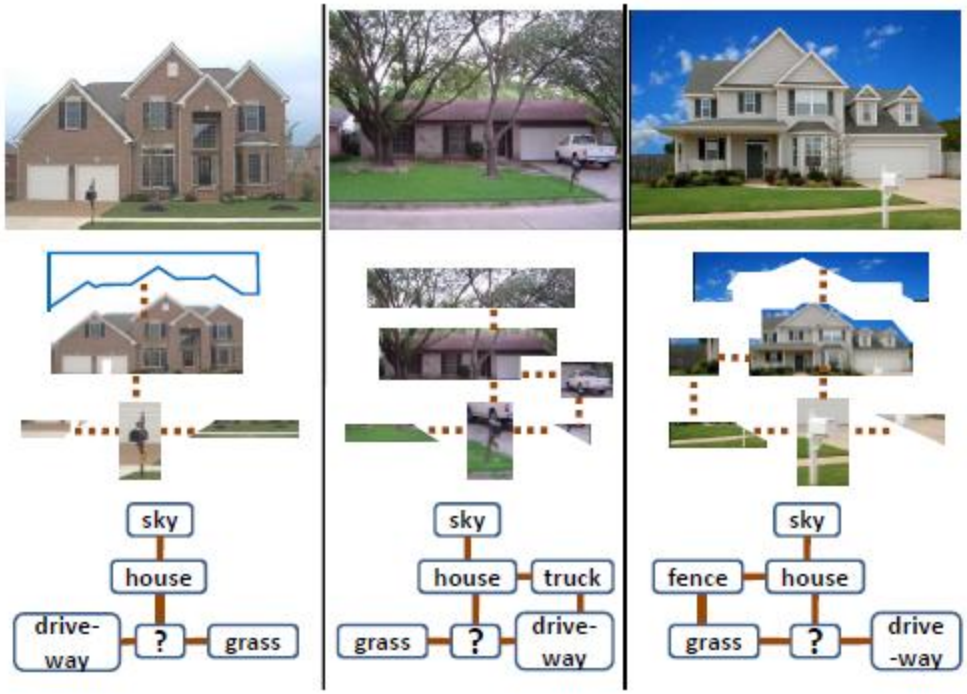
Segmentation

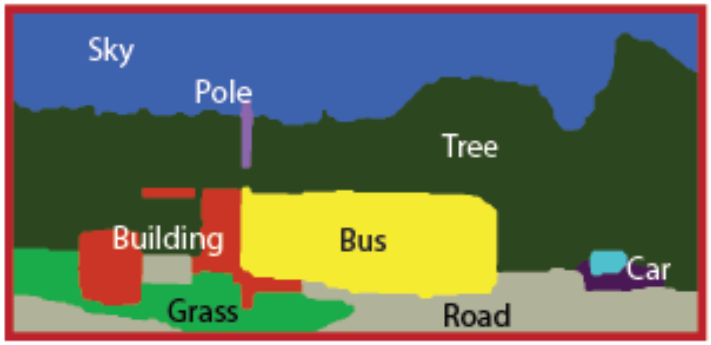
- Detecting contours
- Grouping pixels into semantic regions



Groups of objects, scenes, and context

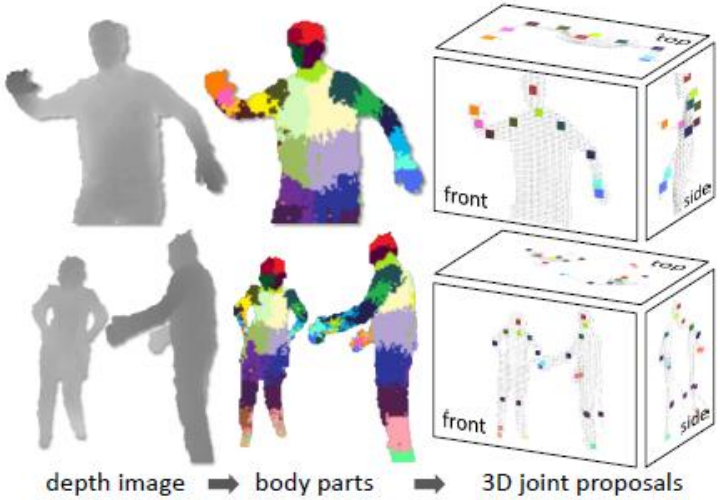
- Exploiting context to detect novel objects
- Parsing out the elements of scenes





Pose and actions

- Automatically annotating a human's pose
- Recognizing different human activities

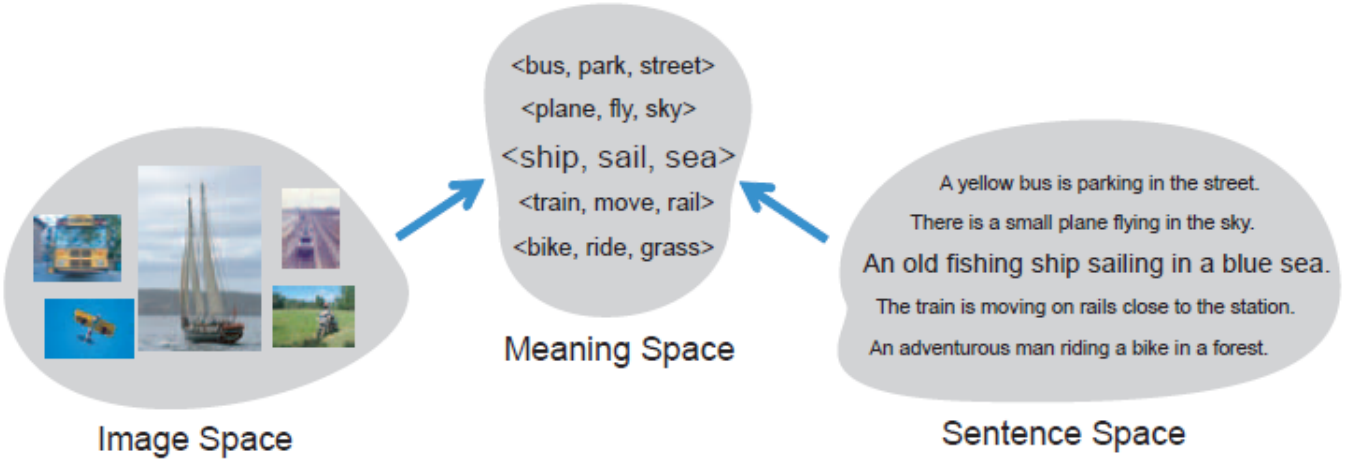


Unsupervised visual discovery

- Finding patterns in unannotated data

Vision and language

- Describing images with sentences

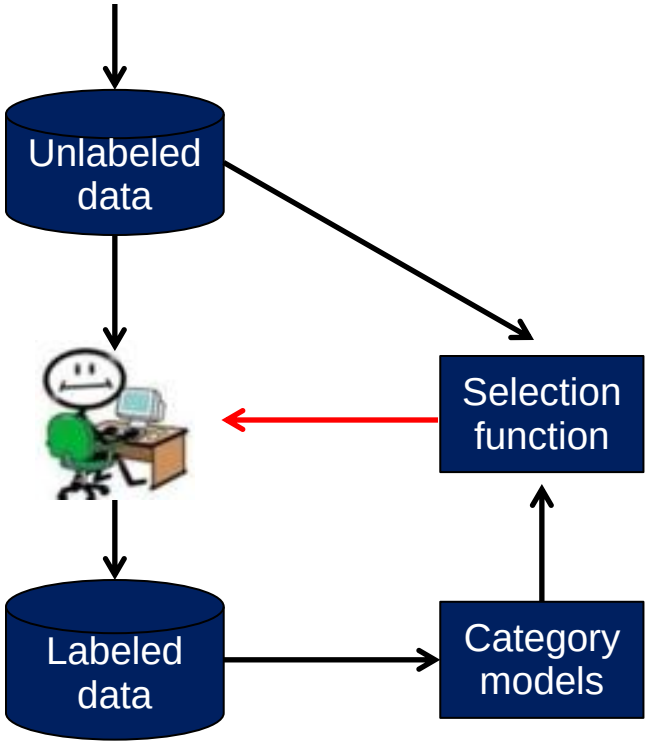


There are two aeroplanes. The first shiny aeroplane is near the second shiny aeroplane.

There are one cow and one sky. The golden cow is by the blue sky.

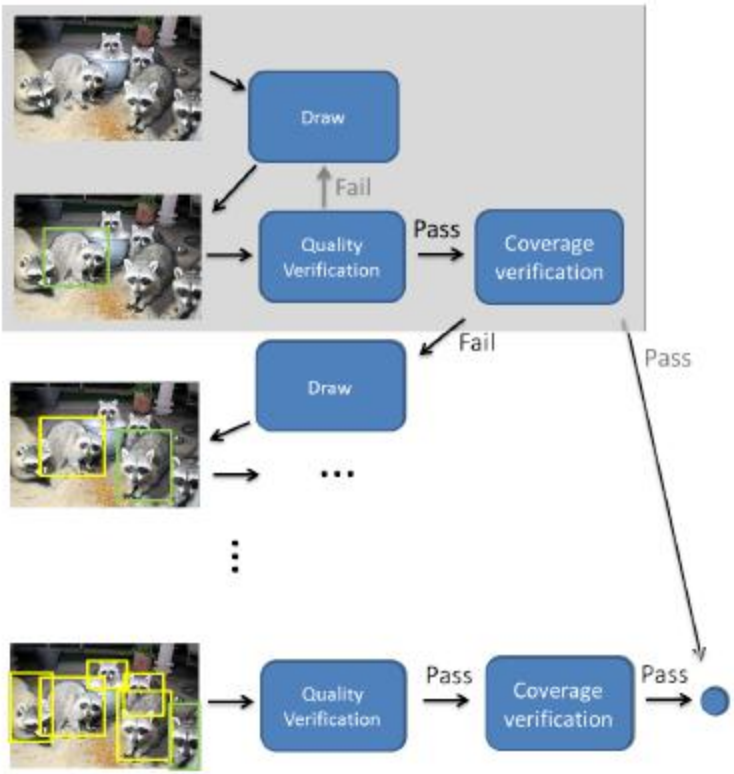
Active and transfer learning

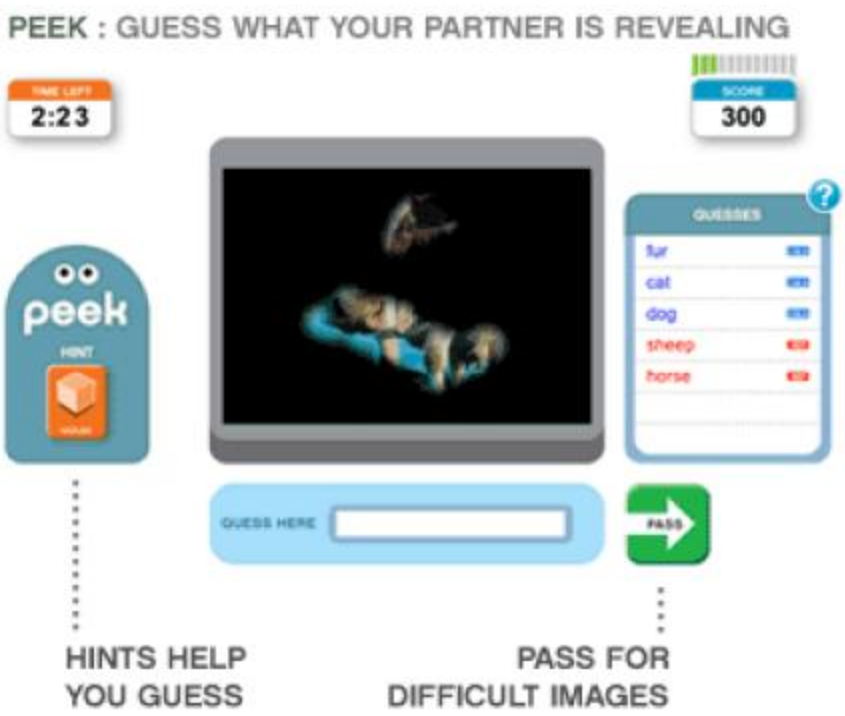
- Making better use of training data
- Using a known category to learn another



Crowdsourcing annotations

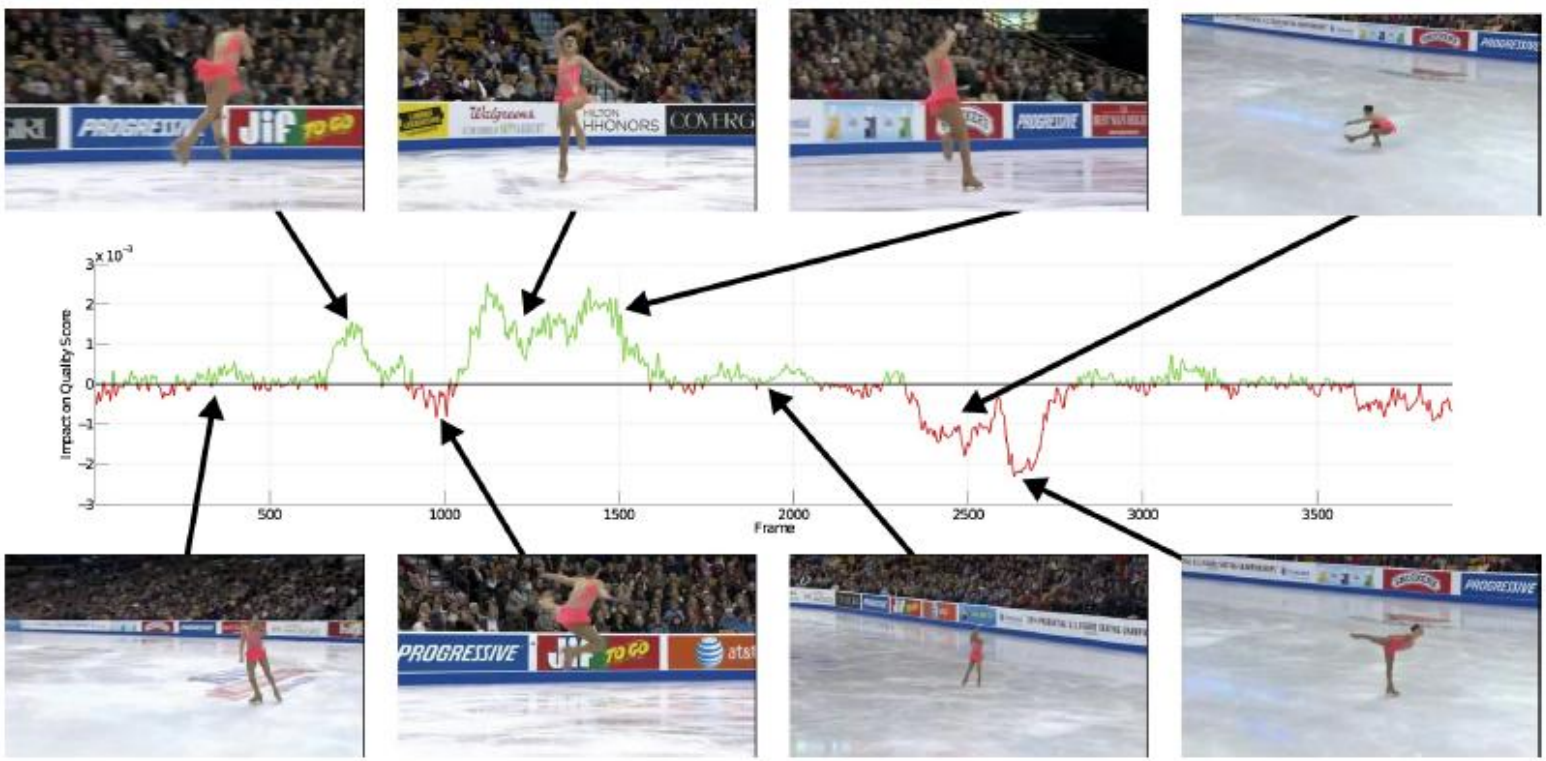
- Using non-expert labelers to collect data
- Ensuring the quality of this data





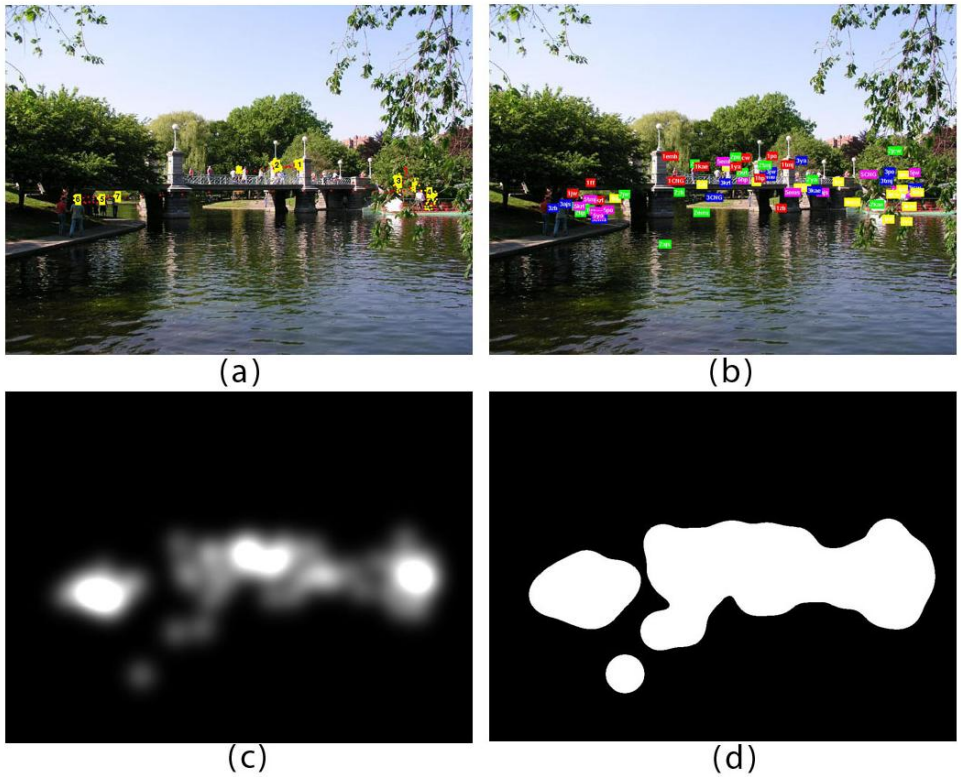
Fine-grained analysis

- Use rationales to learn faster
- Evaluate the quality of performed actions



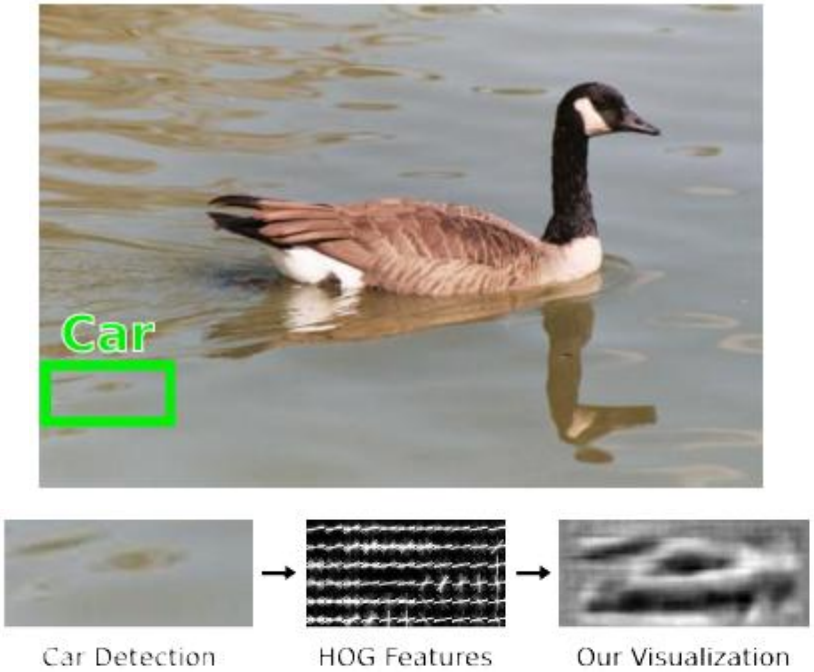
Saliency and importance

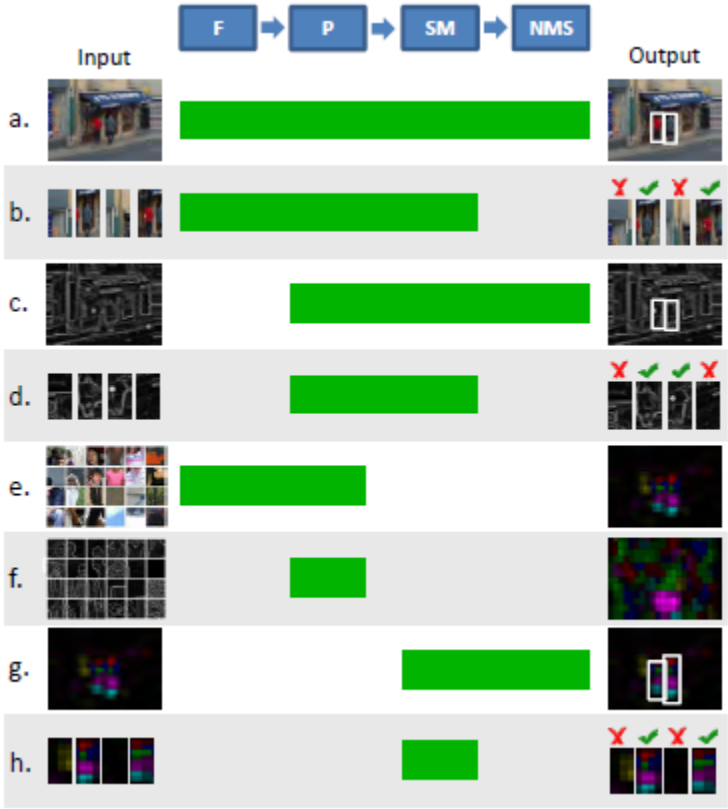
- Learning where in an image a human will look
- Predicting which content is worth mentioning



Debugging vision systems

- Visualizing automatic predictions and outputs
- Finding the “weakest link” in systems



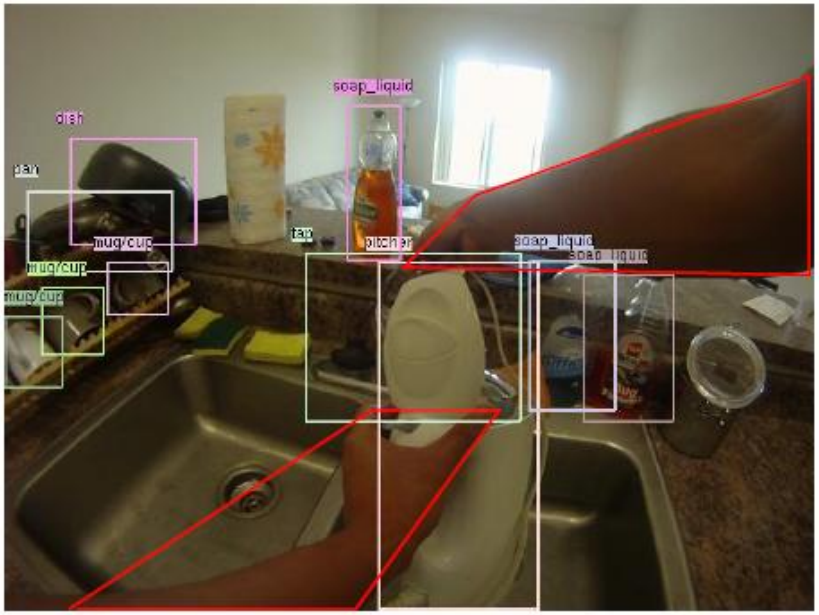


Exploiting big data

- Letting the data solve the problem
- Life-long learning and mining of patterns
- Dataset bias

Video analysis

- Recognizing actions in first-person video
- Summarizing long videos



Vision and art

- Matching queries across domains
- Modeling painterly styles

Course Structure and Requirements

Course Goals

- To learn about the state of the art approaches in visual recognition
- To think critically about vision approaches, and to see connections between works and potential for improvement
- To practice critical reading, clear writing, and engaging presentation skills

Pre-requisites

- Expected basic knowledge of probability and linear algebra
- Experience or familiarity with machine learning is recommended

Course Components

- 20% paper reviews
- 20% participation/discussion
- 20% paper presentations
- 10% experiment presentation
- 30% course project
  - 5% proposal
  - 5% mid-semester progress/status report
  - 10% final presentation
  - 10% final report

Paper Reviews

- 1-page review for 1 of the 2 papers discussed each class (usually the Primary paper)
- If any paper marked with *, read that first
- Due 10pm the day before each class
- Send email to instructor with Subject: “CS3710 Paper Review”
- Name your file: [first name]_[last name]_ review_[month]_[day].pdf (or .doc, .docx)
- Skip paper reviews for papers you present

Paper Reviews

- Answer the following questions:
  1. Summarize what this paper aims to do, and what its main contribution is.
  2. Summarize the proposed approach.
  3. Summarize the experimental validation of the approach.
  4. What are three advantages of the proposed approach?
  5. What are three disadvantages or weaknesses of the approach or experimental validation?
  6. Suggest one possible extension of this approach, i.e. one idea for future work.
  7. Any other thoughts, comments, or questions on this paper?

Participation

- Carefully read the assigned papers
- Ask meaningful questions
- Make meaningful comments about the paper’s strengths and weaknesses
- Answer questions asked by others

Paper Presentations

- Each student will give about 2 presentations
- Each presentation will cover 2 papers
- Presentations should be 30 min long (40 min if no experiment presentation), and should be clear and well rehearsed
- Presentations will be followed by a discussion moderated by the presenter
- See the course website for questions to address in presentation

Paper Presentations

- Often you can find slides on the ***authors' websites***
- Cite all sources and slide credits, and use your own words!
- Slides should use text sparingly
- Slides will be uploaded to course website

Paper Presentations

- **Important:** Presenters should ***meet with the instructor*** on Friday (for Tuesday presentations) or Monday (for Thursday presentations) at the latest!
- **Important:** Presenters should ***email their draft slides*** (and meeting time availability) to the instructor 2 days before desired meeting date, with Subject "CS3710 Presentation Slides“
- Exception: 1 day before meeting if presentation is on 1/15

Experiment Presentations

- Present experimental evaluation of 1 paper
- Students can volunteer to present an extra paper later (better grade will be used)
- Presentations will be about 15 min
- Pick one aspect of the paper to evaluate
- Cite any code you used
- Explain what/why you did, and what you found out in your experiments
- Ask for guidance if needed

Final Project

- Students will complete in-depth study of one topic covered in class
- These projects can become conference publications!
- For most types of projects, students can work in pairs
- Most project types require experimental evaluation

Final Project

- Projects can be one of the following:
  - An extension of one or more papers covered in class
  - A novel approach (with evaluation)
  - A definition of a new problem, along with detailed argumentation of why this problem is important and challenging, and an approach to solve this problem
  - Extensive analysis and experimental evaluation of one or more of the approaches covered in class
  - Extensive literature review and analysis on one of the topics covered in class *(this can only be done by students working individually!!!)*

Final Project

- Timeline
  - **March 6**, 5pm – project proposals due (5% of course grade)
  - **April 3**, 5pm – project status reports due (5%)
  - **April 16, 21, 23** – project presentations (10%)
  - **April 24**, 5pm – project final report due (10%)
- Email these to the instructor with Subject: "CS3710 Project“
- See course website for details

Final Project

- Proposal
  - 3-5 pages in length
  - Include: clear problem statement, extensive literature review, detailed outline of the approach, and planned experimental setup
  - Students are encouraged to meet with the instructor to discuss proposal before proposals are due

Final Project

- Progress report
  - Describe progress, identify problem areas
  - Use the CVPR LaTeX template
  - Include sections: Introduction, Related Work, Approach, and Results
- Presentation
  - 20 min long
- Final report

Homework for Next Time

- By **10pm on 1/07 (tomorrow),** email the instructor (kovashka@cs.pitt.edu):
  - A list of the ten topics from the schedule on the course website that you are most interested in presenting as a paper presentation
  - A list of ten papers that you might like to present as an experiment presentation
  - Sort these from 1 to 10, where 1 denotes "most interested”
  - Use Subject: "CS3710 Topic Preferences”

Questions?

Quiz

Question # 1

- How would you learn to predict whether author A or author B wrote a given piece of text?

Question # 2

- What learning problem corresponds to predicting the price of a car based on its features (e.g., engine specs) and the prices of other cars?

Question # 3

- Other than learning an individual model, what else can you use training data for?

Question # 4

- Given a model that can predict whether an image has a cat in it or not, what do you usually need to do to a new image in order to be able to apply this model to it?

Question # 5

- What does “overfitting” mean?
- Why is it a problem?
- How can you overcome this problem?

Question # 6

- What two goals does a soft margin linear SVM optimization objective capture?

Question # 7

- How can you use maximum likelihood to select model parameters?

Question # 8

- State Bayes' theorem.

Question # 9

- What is the goal of performing *k*-means?
- How is this goal accomplished?

Question # 10

- What are λ and v in relationship to A, according to the following equation?

$$A\ v = \lambda\ v$$

- If the dimensionality of v is mx1, what is the dimensionality of A?

Next Two Weeks

- This Thursday: Describing images with features
- Next Tuesday: Recognition basics; Adriana's research
- Next Thursday: First presentation (features)