CS5330: Pattern Recognition and Computer Vision

Northeastern University

Fall 2021

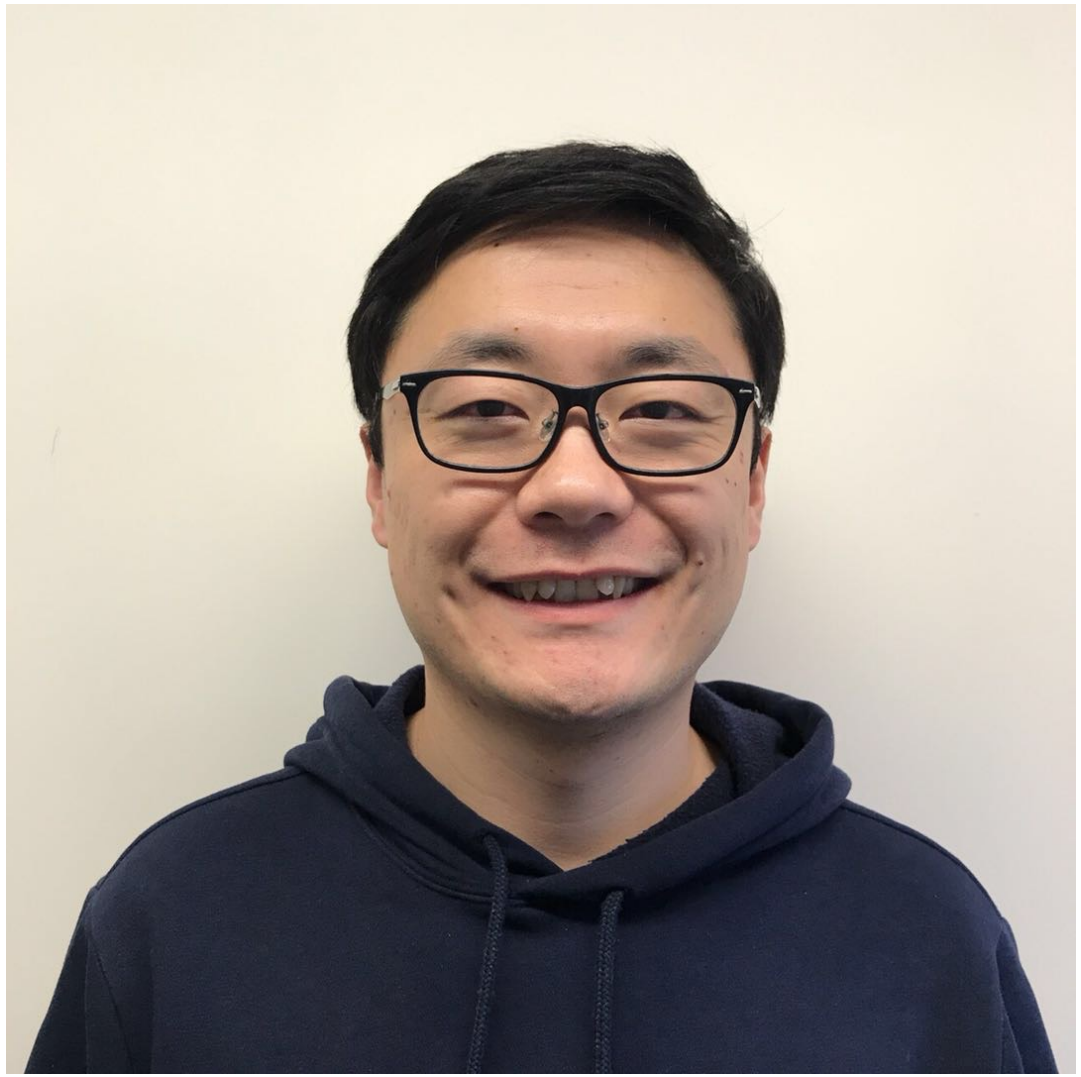
Instructor: Huaizu Jiang

TAs: Prajnan Goswami, Guangyuan Weng

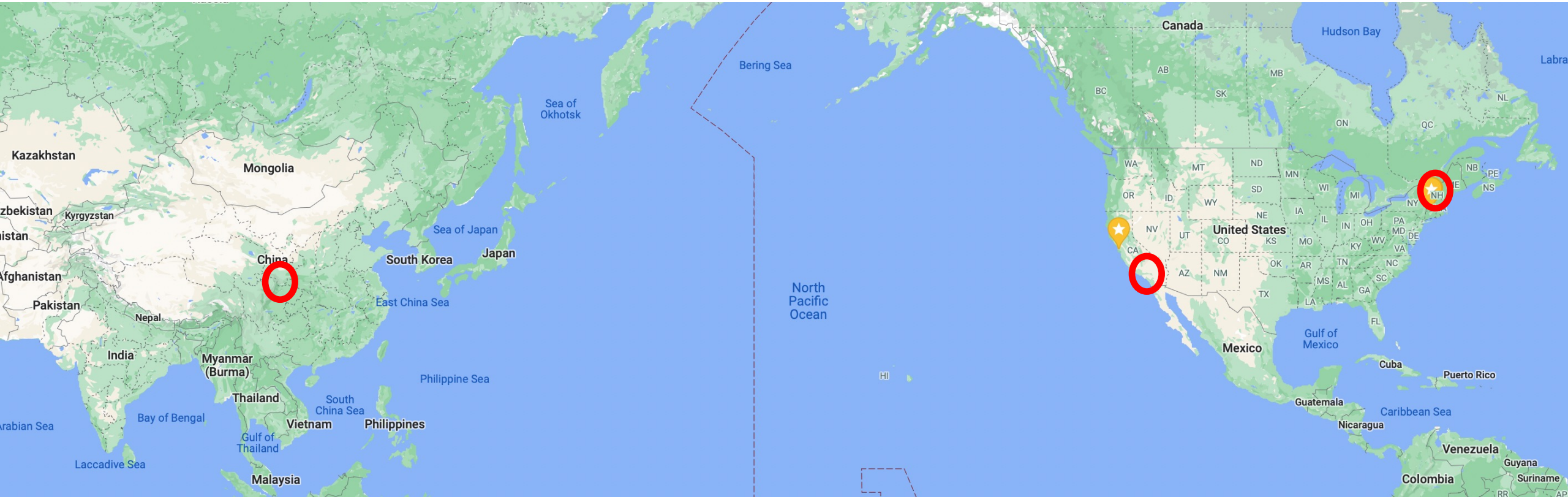
Today’s Class

- A little about me and the TAs
- Introduction to computer vision
- Course logistics
- Questions

About me



About me



2005-2012: Undergrad and master's student at Xi'an Jiaotong University, China
2015-2020: PhD student at University of Massachusetts, Amherst
2020-2021: Postdoc at Caltech and Visiting Researcher at NVIDIA
2021-present: CS Professor at Northeastern University

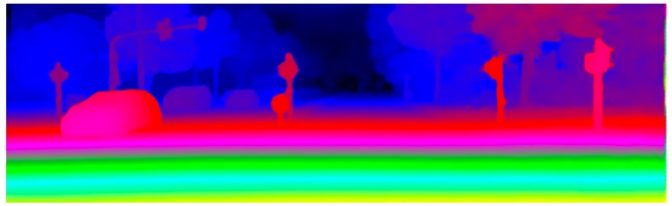
My Research: Dynamic Scene Understanding

(a) Image 1 of left camera

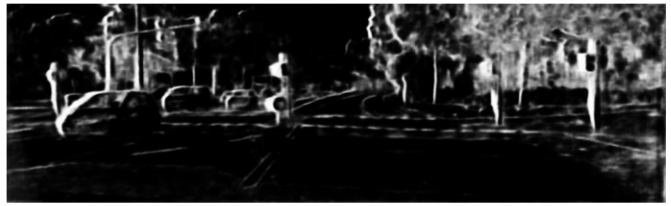
(b) Image 2 of left camera

(c) Optical flow

(d) Occlusions for flow



(e) Stereo disparity



(f) Occlusions for disparity

(g) Segmentation of (a)

(h) Segmentation of (b)

My Research: Use Natural Language as a Scaffold for Visual Understanding

Q: What are the ladies doing?

A: Texting.

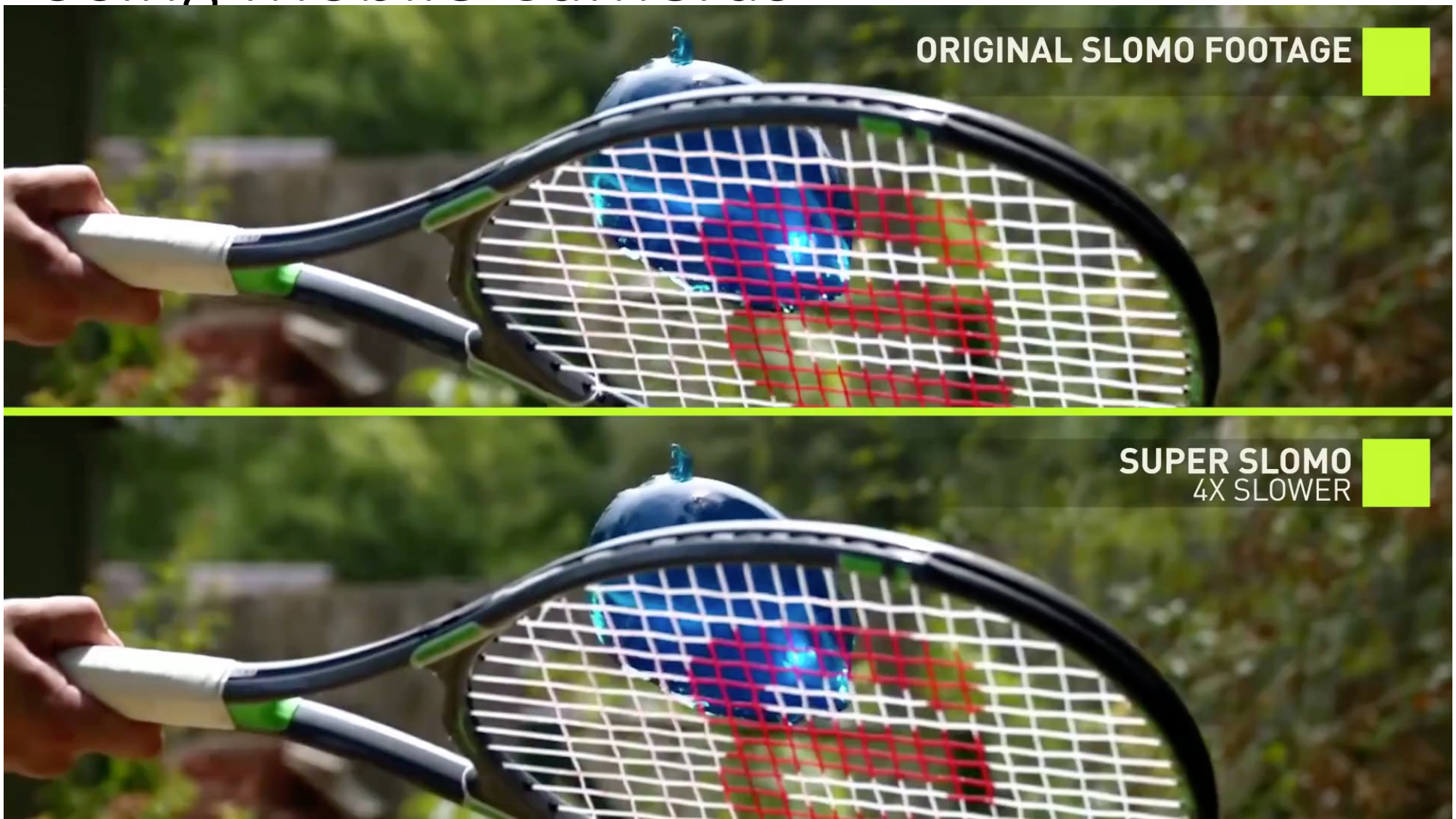
Q: Is there a bed?

A: Yes.

My Research: Enhance People's Experiences of Using Mobile Cameras

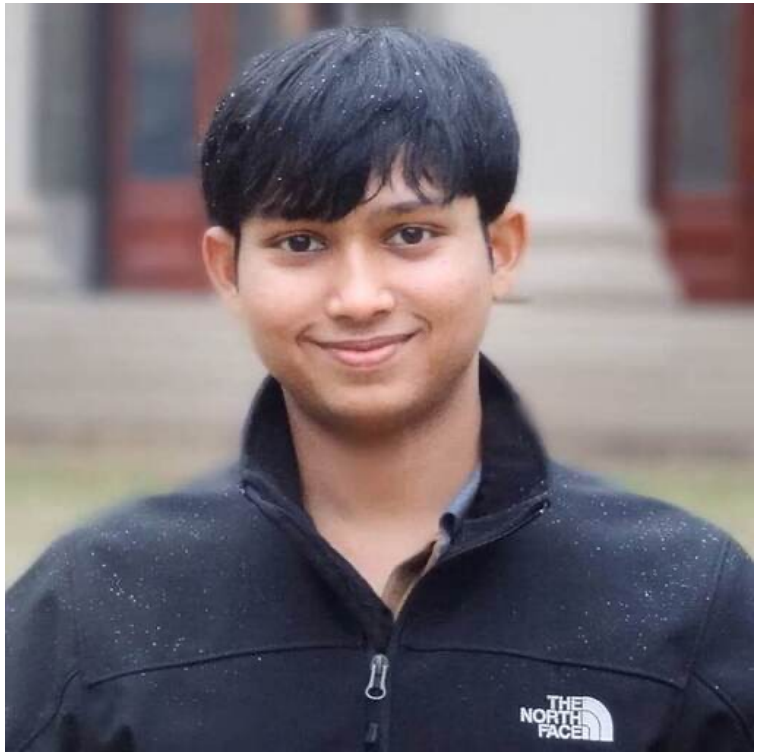

1x

My Research: Enhance People's Experiences of Using Mobile Cameras



My Research: Cognitive Reasoning

About TAs



Prajnan Goswami
Master's student in AI at Khoury College
Bachelor from
Sikkim Manipal Institute of Technology, Sikkim, India

Guangyuan Weng
PhD student at Khoury College
Bachelor from
ShanghaiTech University, Shanghai, China

~~Pattern Recognition and~~ Computer Vision

- Make (teach) computers to see and understand the world.

What kind of scene?
What are in the scene?
How far away is the building?
Can I walk on the grass?
…

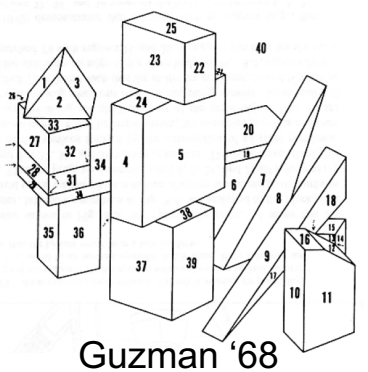
Why is this hard?

Problem: Semantic Gap

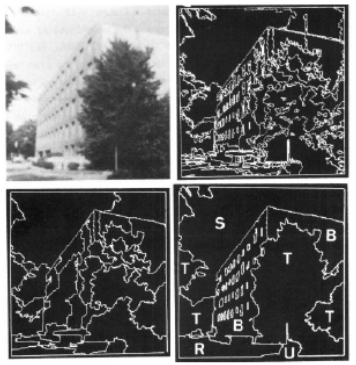
097	097	097	097	097	097	097	097	096	097	097	096	096	096
100	100	100	100	100	100	101	101	102	101	100	100	100	099
105	105	105	105	105	105	105	103	102	102	101	103	104	105
109	109	109	109	109	110	107	118	145	132	120	112	106	103
113	113	113	112	112	113	110	129	160	160	164	162	157	151
118	117	118	123	119	118	112	125	142	134	135	139	139	175
123	121	125	162	166	157	149	153	160	151	150	146	137	168
127	127	125	168	147	117	139	135	126	147	147	149	156	160
133	130	150	179	145	132	160	134	150	150	111	145	126	121
138	134	179	185	141	090	166	117	120	153	111	153	114	126
144	151	188	178	159	154	172	147	159	170	147	185	105	122
152	157	184	183	142	127	141	133	137	141	131	147	144	147
130	147	185	180	139	131	154	121	140	147	107	147	120	128
035	102	194	175	149	140	179	128	146	168	096	163	101	125

Ridiculously brief history of computer visior

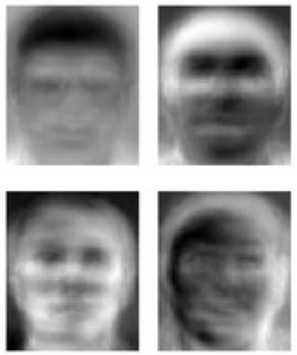
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: ANNs are back for big improvements in recognition; likely large increase in deployed vision systems
- 2020's: autonomous vehicles, the great robot rebellion?



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

Slide from Derek Hoiem

Why computer vision matters?

Why computer vision matters

Why computer vision matters?

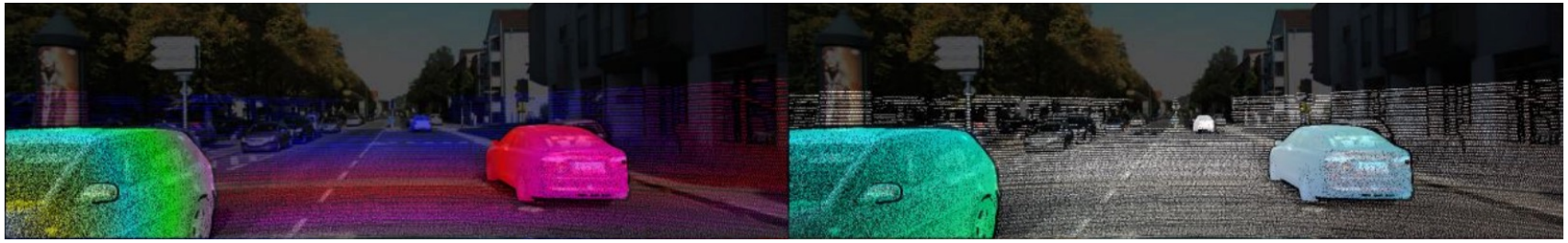
The KITTI Vision Benchmark Suite

[Geiger et al. CVPR 2012]

Stereo videos are used for precise localization

left camera

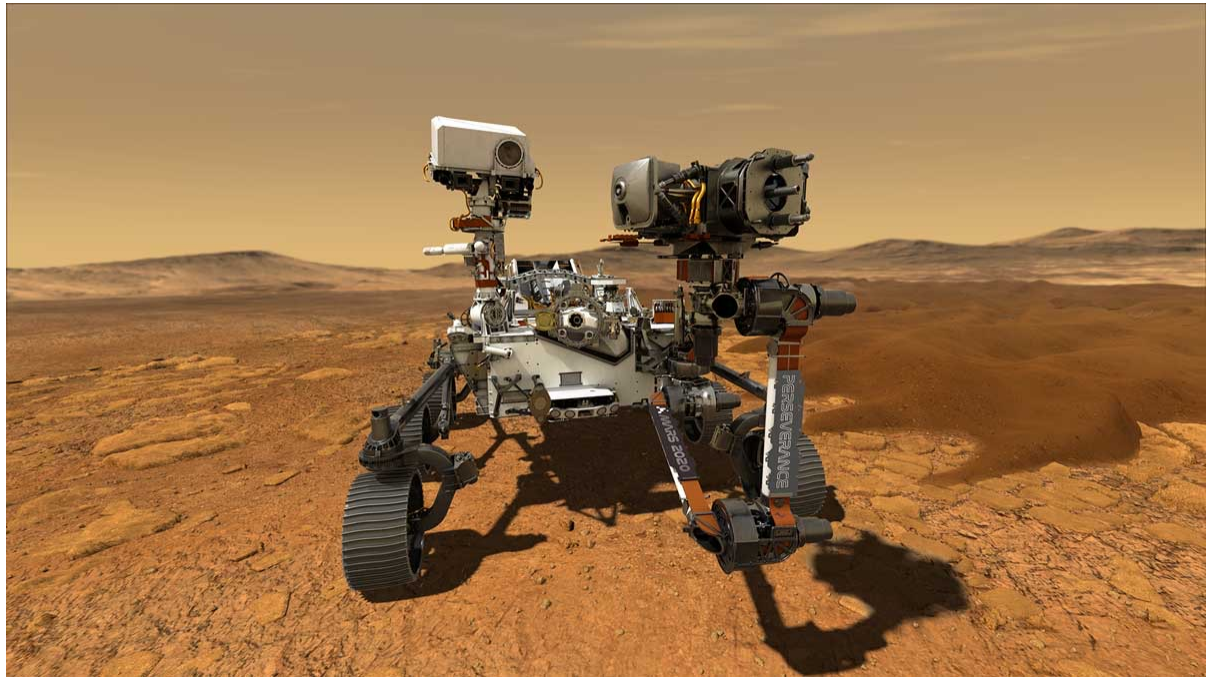
right camera



Autonomous driving


V53005
tusimple
357273
E001

Why computer vision matters?



Why computer vision matters

Why computer vision matters?

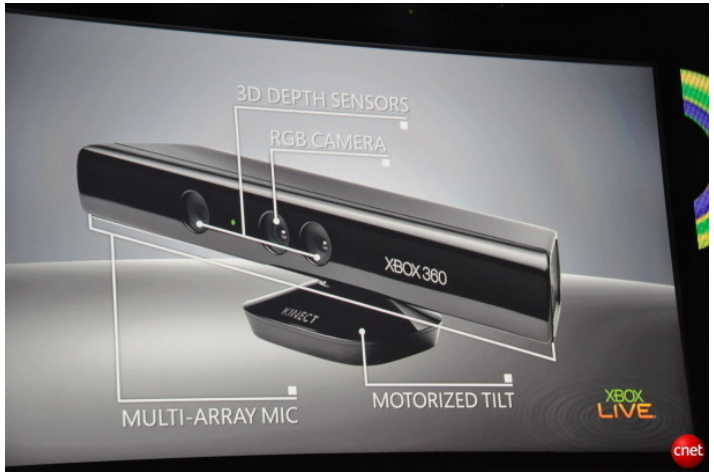
Pirates of the Carribean, Industrial Light and Magic

Slide from Derek Hoiem

Why computer vision matters?

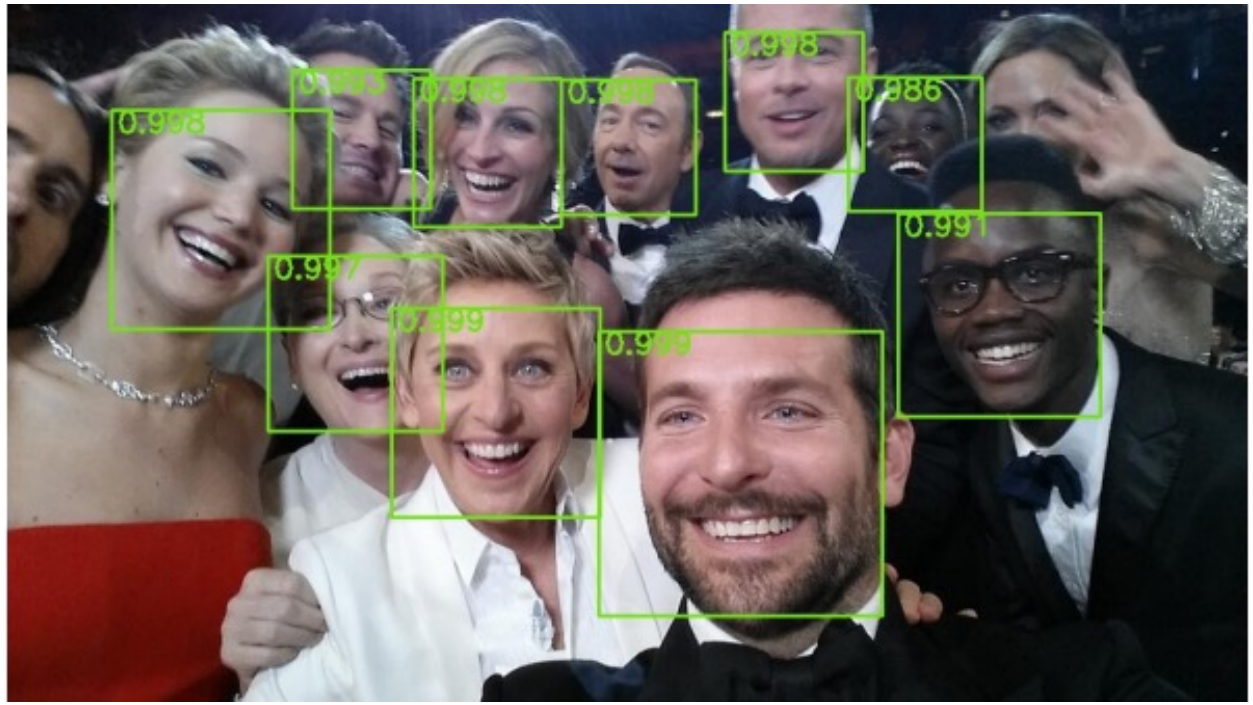
Interactive Games: Kinect

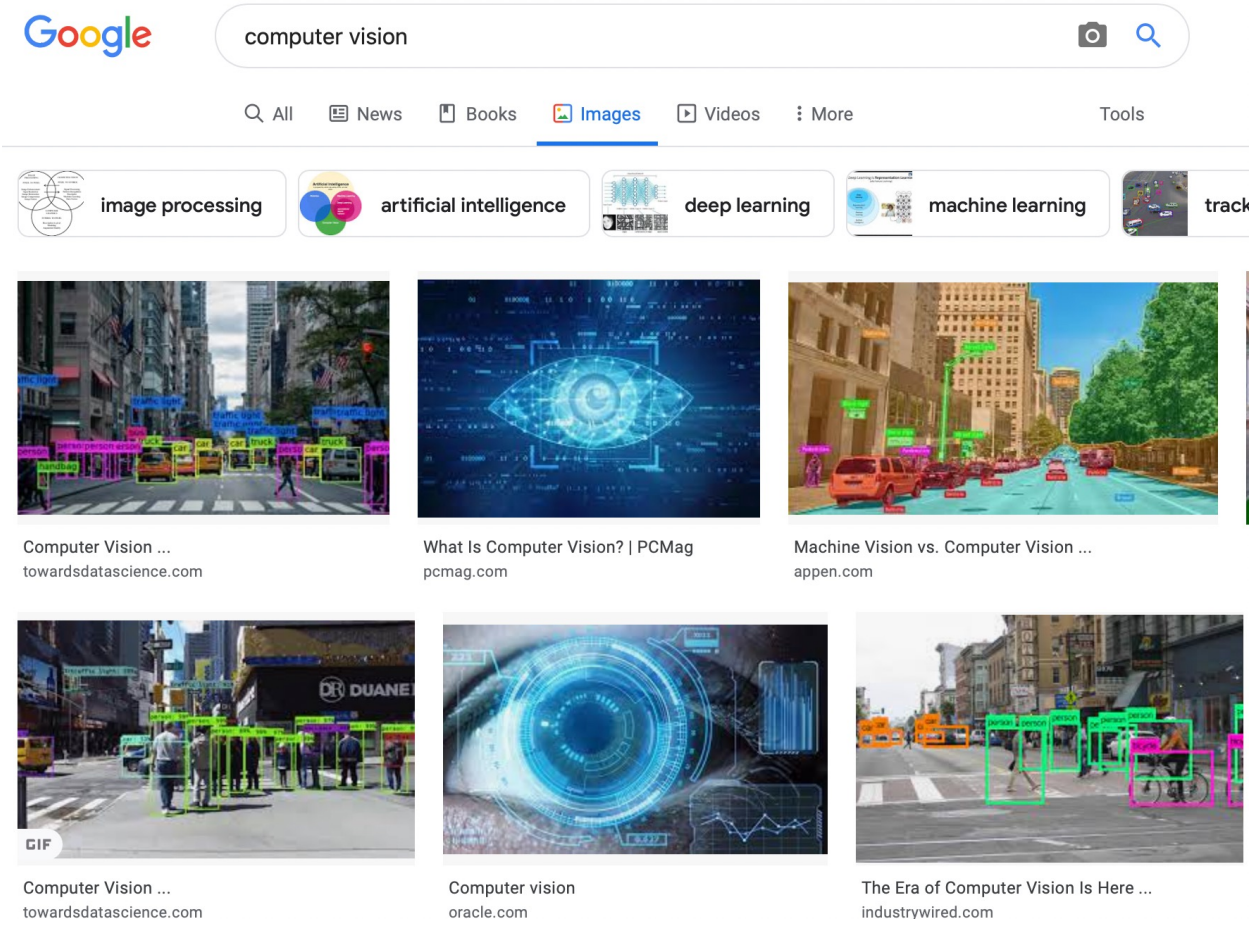
- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



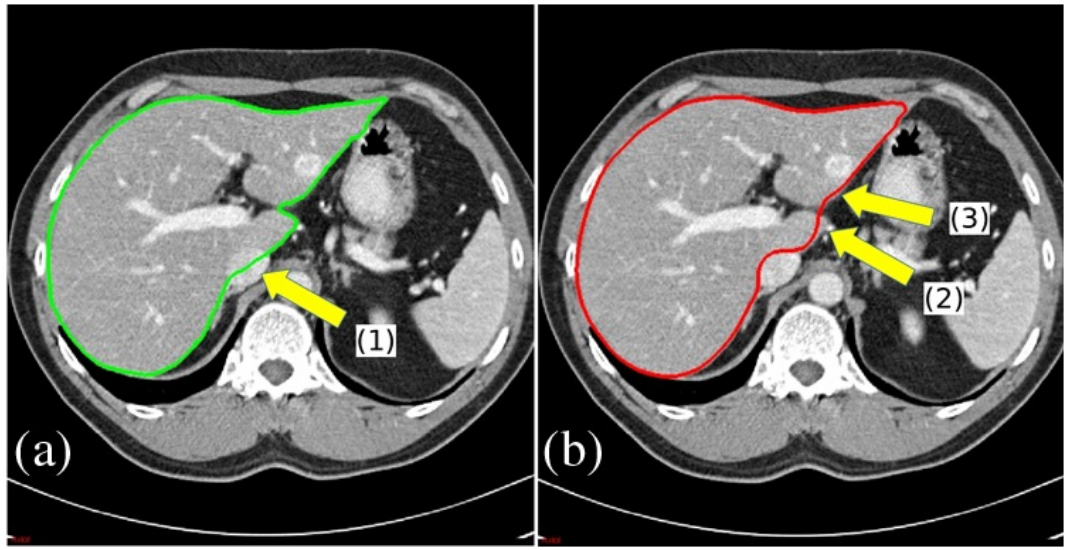
Slide from Derek Hoiem

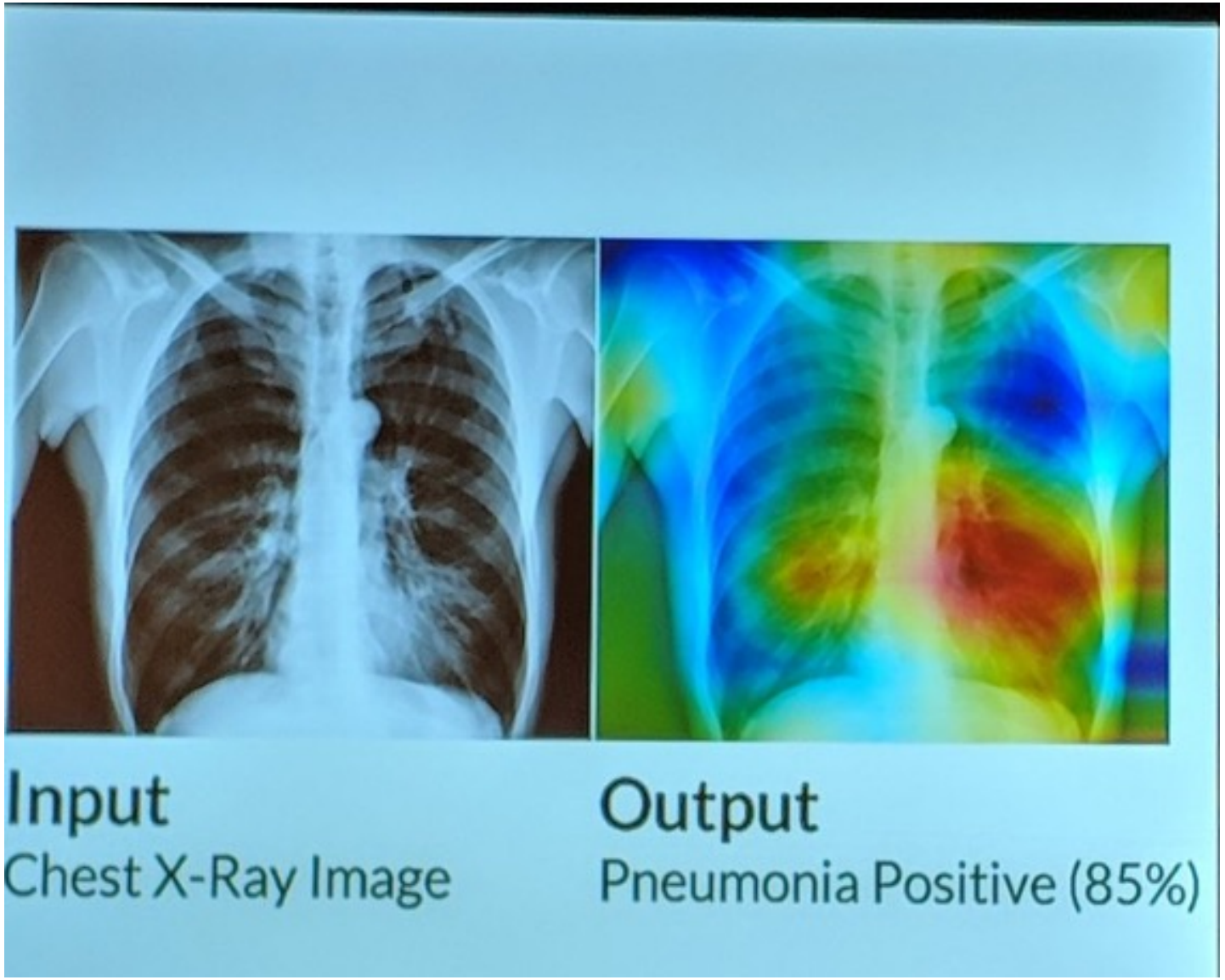
Why computer vision matters?



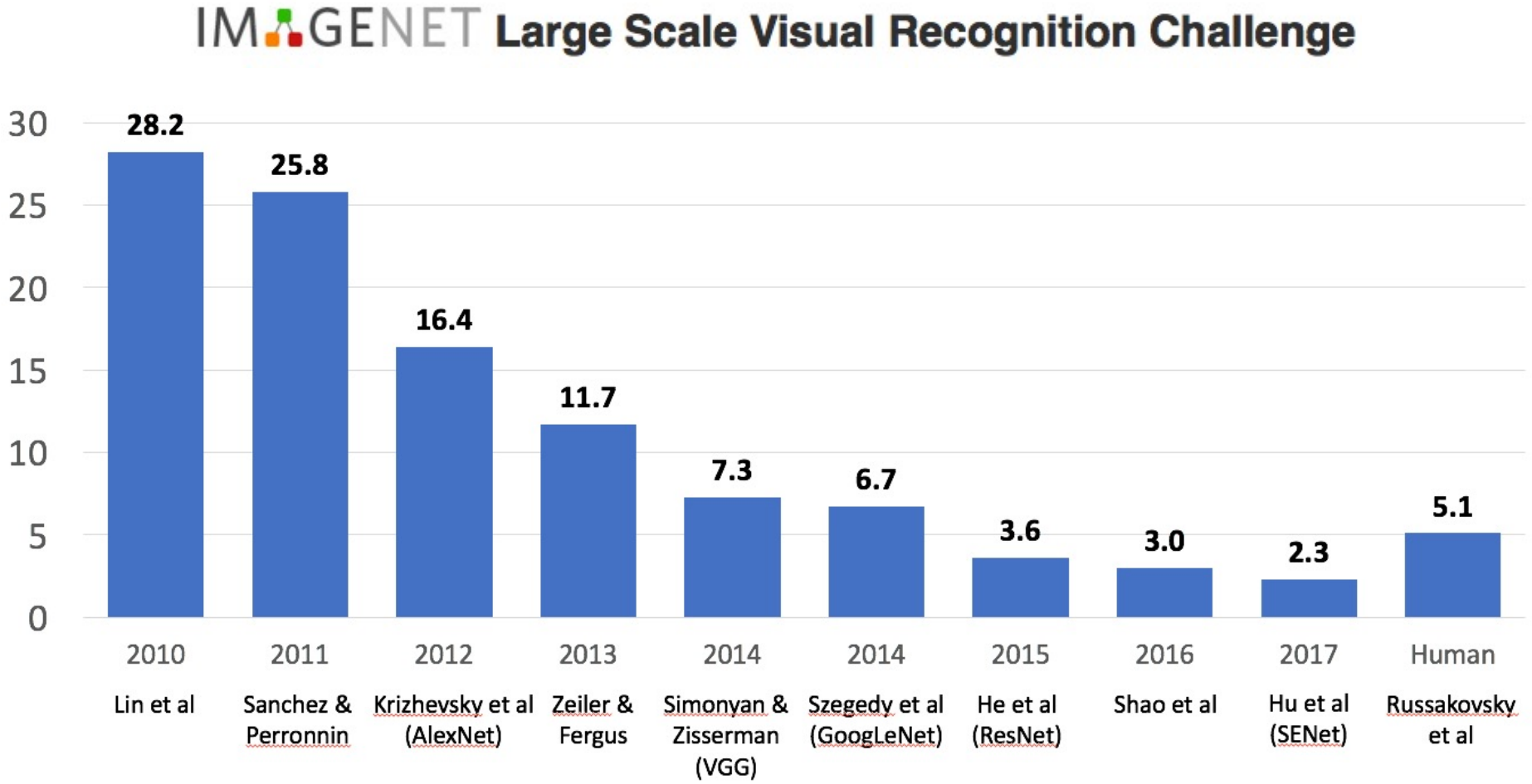


Medical Imaging

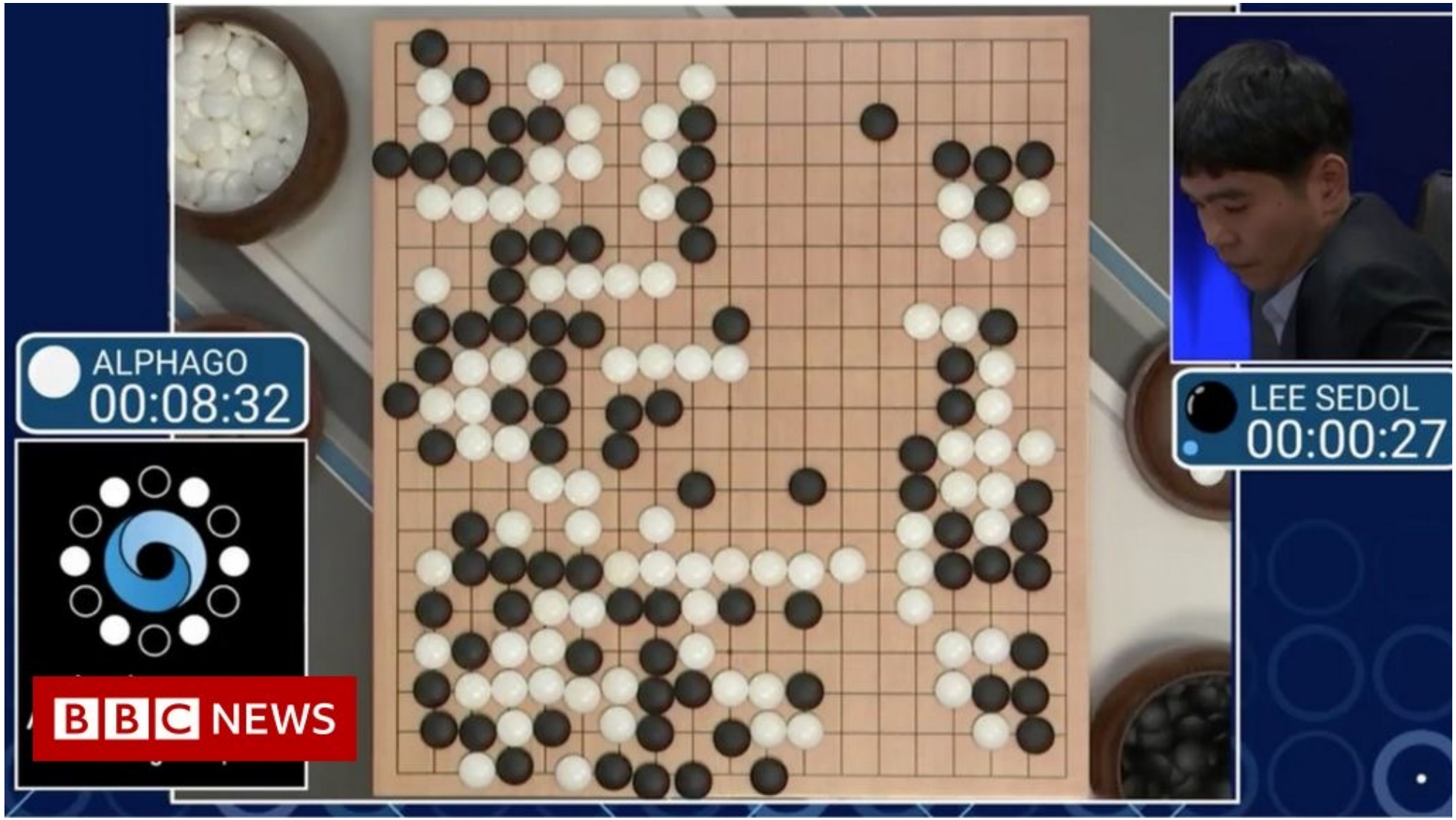




We have made significant progress

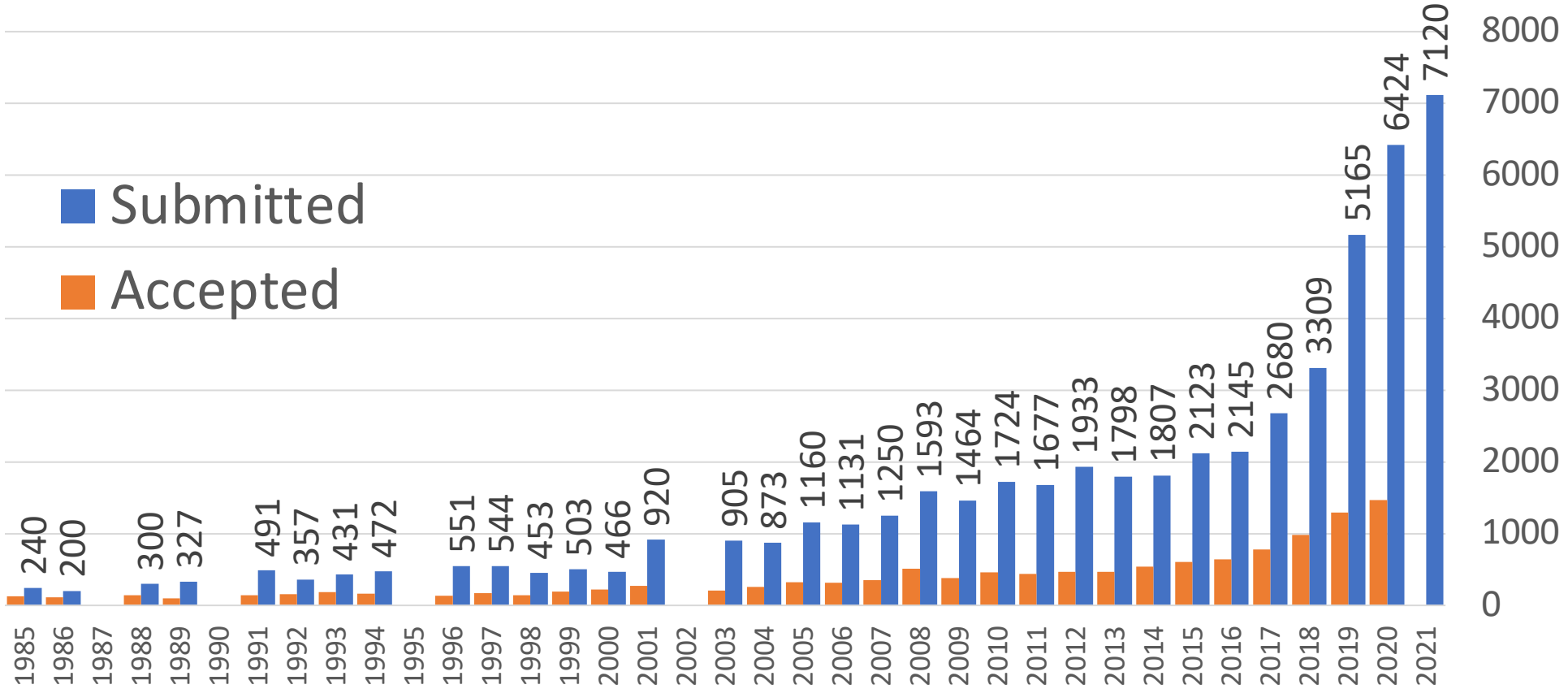


We have made significant progress



Explosion of Computer Vision

Publications at top Computer Vision conference:
Number of submitted and accepted papers at CVPR



Slide from Justin Johnson

Yet, computer vision is far from solved…

1. Many of the applications started less than 10 years ago. More new applications will appear in the next decade.

2. This is a rapidly growing area: new technologies, new applications, and new problems.

3. Computer vision is far from solved.

Course Logistics

Lecture and Office Hours

- Lecture
  - SL 031, in-person, Monday and Thursday, 11:45am-1:25pm
  - Recordings will be available
- Office hours
  - Huaizu Jiang: via Zoom
    - Monday: 1:45pm-2:45pm
    - Wednesday: 1-2pm
    - Email to schedule 1-1
  - TAs: TBD

Waitlist

- Sorry, I can do nothing about it…

Prerequisites

- Python, Numpy
  - Python Numpy tutorial: https://compsci682-fa21.github.io/notes/python-numpy-tutorial/
  - Jupyter Notebook tutorial: https://compsci682-fa21.github.io/notes/jupyter-tutorial/
- Data structures: students will be writing code that builds representations of images, features, and geometric constructions.
- Math: matrix analysis and probability.

Websites

- Course website
  - Canvas
- Piazza https://piazza.com/northeastern/fall2021/cs5330
  - For discussions and QA
  - You can make anonymous posts (visible to the instructor and TAs only)
- Assignments and project submission
  - Canvas

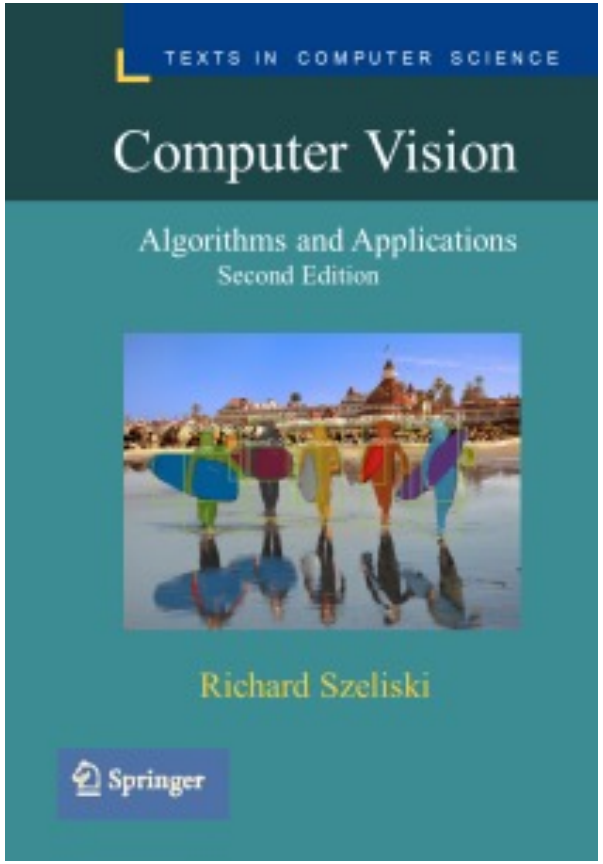
Topics

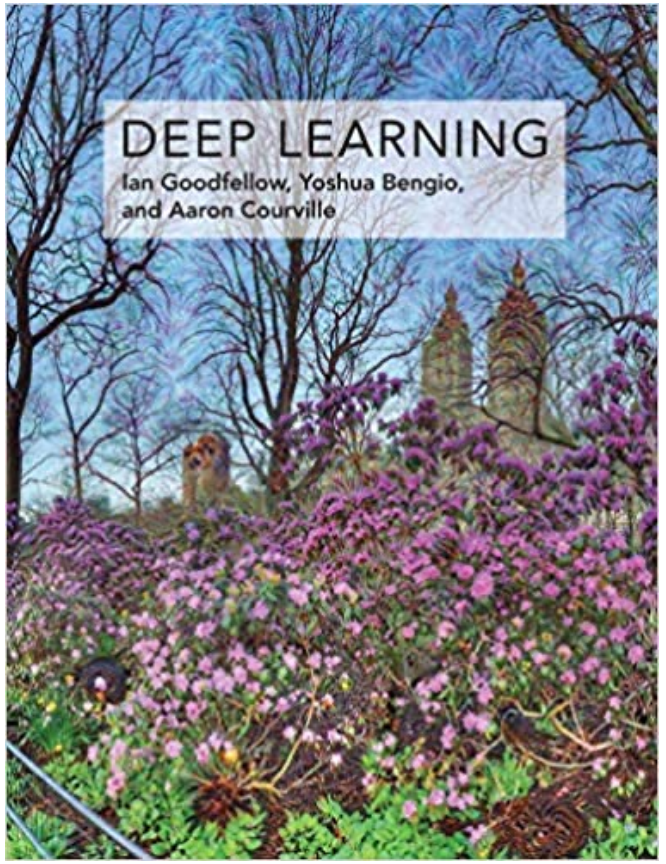
	Class Date	Topic	Reading
Week 0	Thur, Sep 09	Introduction to computer vision and course logistics	
	Image Formation and Filtering (Szeliski chapters 2 and 3)		
	Mon, Sep 13	Camera Models	Szeliski 2.1
Week 1	Thur, Sep 16	Light and Shading	Szeliski 2.2, 2.3
	Mon, Sep 20	Image Filtering	Szeliski 3.2, 3.3
	Correspondence and Alignment		
Week 2	Thur, Sep 23	Interest points and corners	Szeliski 7.1.1, 7.1.2
	Mon, Sep 27	Local image features	Szeliski 7.1.3
Week 3	Thur, Sep 30	Model fitting, Hough Transform	Szeliski 7.4.2 and 2.1
	Mon, Oct 4	RANSAC and transformations	Szeliski 8.1 and 2.1
	Perspective and 3D Geometry		
Week 4	Thur, Oct 7	Camera Calibration, Epipolar Geometry	Szeliski 11
	Mon, Oct 11	Stereo Vision	
Week 5	Thur, Oct 14	Optical Flow	
	Mon, Oct 18	Structure from Motion	

Topics

	Fundamentals of Deep Neural Networks		
Week 6	Thur, Oct 21	Basic Concepts of Machine Learning	
	Mon, Oct 25	Optimization: Stochastic Gradient Descent and Backpropagation	
Week 7	Thur, Oct 28	Neural Networks 1	
	Mon, Nov 1	Neural Networks 2	
Week 8	Thur, Nov 4	Convolutional Neural Networks 1	
	Mon, Nov 8	Convolutional Neural Networks 1	
	Recognition, Detection and Segmentation via DNNs		
Week 9	Thur, Nov 11	Object Detection 1: R-CNN	
	Mon, Nov 15	Object Detection 2: Fast and Faster R-CNN	
Week 10	Thur, Nov 18	Segmentation 1: Mask R-CNN, FCN	
	Mon, Nov 22	Segmentation 2: Panoptic Segmentation	
Week 11	Thur, Nov 25	No class, Thanksgiving recess	
	Advanced Topics		
	Mon, No 29	Novel view synthesis and video frame interpolation	
Week 12	Thur, Dec 2	Vision and language (VQA, visual grounding, etc)	
	Mon, Dec 6	Learning with unlabeled data	
Week 13	Thur, Dec 9	Transformers in computer vision	
	Mon, Dec 13	backup class	
Week 14	Thur, Dec 16	backup class	

Textbooks (not required)





https://szeliski.org/Book/

https://www.deeplearningbook.org

Grading

Assignment	Percentage
Programming Assignment 1	14%
Programming Assignment 2	16%
Programming Assignment 3	17%
Programming Assignment 4	16%
Programming Assignment 5	17%
Final Project (group-based)	20%

A (93%),
A- (90%),
B+ (87%),
B (83%),
B- (80%),
C- (65%).

First programming assignment will be out next Monday after the class, due in two weeks.

Late Policy

- Every student has five late days in total.
- After using all of the late days, homework will be penalized 25% every 24 hours.
- Late days can not be used for the final project (we will set up milestones throughout the class to help you on track).
- If stuff happens, email the faculty.

Collaborations and Academic Integrity

- 1. Discussions, with natural language, with your peers are allowed.
- 2. No discussions should be made in Python – no copy of code (not from your peers, not from the Internet).
- 3. We actively detect code copying.
- 4. I very much hate to deal with such cases. But I have no many options.

Asking for help

- 30-minute rule: no progress at all if you work on your own for more than 30 minutes -> ask for help
- Piazza!
- Office hours
- Use university resources during pandemic (for physical and mental issues)

Questions?