CSE 152A: Computer Vision
Manmohan Chandraker

Lecture 0: Introduction
In-person Virtual classrooms

• In-person Virtual lectures
  – Only host shares the screen
  – Keep microphone muted masks on
  – But please do speak up (remember to unmute keep masks on!)
  – Slides uploaded on webpage just before class

• In-person Virtual interactions
  – Ask and answer plenty of questions
  – “Raise hand” feature on Zoom when you wish to speak
  – Post questions on chat window
  – Happy to try other suggestions!

• Lectures recorded and uploaded on Kaltura
  – Available under “My Media” on Canvas
Enrollment logistics

• Waitlist
  – You are welcome to attend lectures even if on waitlist
  – To limit TA workload, we can grade only enrolled students
  – Suggestion: wait as long as you can for waitlist to clear

• Canvas
  – All enrolled and waitlisted students should have access

• Announcements will be posted on Piazza
Important Announcement

• The university requires us to determine which students commence academic activity.

• Not completing this may result in students being billed for unearned financial aid.

• To do this automatically, we are using a survey in Canvas that every student must fill out by the end of Friday of Week 2 to ensure that they are certified.

• Link: https://canvas.ucsd.edu/courses/39510/quizzes/117413
Computer Vision
Defining computer vision

Wall-E: Fact and Fiction (Minh Do, Princeton University)

CSE 152A, FA22: Manmohan Chandraker
Defining computer vision

• Old: Computer programs that can
  • Process image information
  • Recognize instances of objects
  • Find distances of objects

• Modern: Understanding the world based on visual cues
  • Determining factors that govern image formation
  • Recognition across variations
  • Estimate semantic properties of a scene
  • Recognize complex actions
  • Predict long-term behaviors
Studying computer vision

• Images are everywhere around us
Studying computer vision

- Images are everywhere around us
- We are all users of computer vision

![Advance Driver Assistance Systems (ADAS)](image1)

Transportation

![Design](image2)

Design

![Communication](image3)

Communication

![E-Commerce](image4)

E-Commerce
Studying computer vision

- Images are everywhere around us
- We are all users of computer vision
- Deep and attractive scientific problems
  - How do we recognize objects?
  - Why do newborn babies respond to face-like shapes?

[Farroni et al., 2005]
We Use Computer Vision
Computer vision in living rooms

Microsoft Kinect Xbox

Sportvision first down line

CSE 152A, FA22: Manmohan Chandraker
Vision to explore the world

Image from Microsoft Virtual Earth

CSE 152A, FA22: Manmohan Chandraker
Vision to explore other worlds

- Panorama stitching
- Stereo imaging
- Navigation
- ....

Image from NASA’s Mars Exploration Rover Spirit
Vision to explore all worlds

Including virtual ones!

*The Matrix* movies, ESC Entertainment, XYZRGB, NRC

CSE 152A, FA22: Manmohan Chandraker
Vision to recognize humans
Vision to recognize objects
Vision for shopping
Vision for sports
Vision for health
Vision on phones
New devices

- Time-of-flight sensors
- Structured light systems
- Light field cameras
- Coded apertures
Organizing Computer Vision
Broad classes of vision applications

- Sense
- Understand
- Interact
- Reconstruct
- Recognize
- Reorganize
Broad classes of vision applications

- Sense
- Understand
- Interact

Scenes

People
Broad classes of vision applications

- Sense
- Understand
- Interact

- Human-Human
- Human-Machine
- Machine-Machine

CSE 152A, FA22: Manmohan Chandraker
Significant progress in recent years

Sense
Understand
Interact

Autonomous Driving Systems
Deep learning is revolutionizing AI

- Tic-tac-toe (1952)
- Checkers (1994)
- Chess (1997)
- Atari (2015)
- Go (2016)
Computer vision is also riding the wave

- Autonomous driving (Google, Tesla, Mobileye, ....)
- Augmented reality (HoloLens, Oculus, MagicLeap)
- Social networks (Google, Facebook, ....)
- Mobile applications
Autonomous Driving
Autonomous navigation
The hardness of the problem

- Finding locations
- Localize objects
- Estimate distances
- Understand relations
- Be aware of traffic rules
- Predict future behaviors
- Understand intentions
- Interdependent decisions
Object detection
Semantic segmentation
Real-time navigation

Song and Chandraker, CVPR 2014
Future behavior prediction
Augmented Reality
Vision in augmented reality devices

Gaze tracking

Head pose estimation

Object detection

Material and lighting estimation

Depth estimation

Semantic segmentation
Gaining perspective on computer vision

Important for Autopilot:
Do not hit a *police* car!
Gaining perspective on computer vision

Important for Autopilot: Do not hit a police car!

Important for Robots: Stay clear of humans!

CSE 152A, FA22: Manmohan Chandraker
Gaining perspective on computer vision

Important for Autopilot: Do not hit a police car!

Important for Robots: Stay clear of humans!

[CSE 152A, FA22: Manmohan Chandraker]
Fairness to Society
A Few Challenges in Computer Vision
The challenge of computer vision

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".
Why is computer vision difficult?

- Viewpoint
- Lighting
- Scale
- Deformation
Why is computer vision difficult?

- Intra-class variation
- Background clutter
- Motion (Source: S. Lazebnik)
- Occlusion

CSE 152A, FA22: Manmohan Chandraker
Solving Challenges: Building Models
Representation of images

Diagram showing the camera's COP (Center of Projection) and the projection plane (PP) with coordinates (x', y', -d) and (x, y, z). The diagram illustrates the mapping between the 3D world and the 2D image plane.
Representation of shape
Representation of motion
Representation of lighting and material
Solving Challenges: Machine Learning
Define a cat
Machine learning

• Typically in CS: write a program to execute a set of rules
• Computer vision: sometimes very hard to define rules
• Machine learning: develop own program based on examples
• Training data: input-output pairs
Where can machine learning help?
Verification: is that a bus?
Detection: are there cars?
Identification: who is that person?
Scene categorization

- outdoor
- city
- traffic
- ...

CSE 152A, FA22: Manmohan Chandraker
Machine learning is a key player

• What is it?
  • Object and scene recognition

• Who is it?
  • Identity recognition

• Where is it?
  • Object detection

• What are they doing?
  • Activities

• All of these are **classification** problems
  • Choose one class from a list of possible candidates
Rapid progress in data-driven vision

Big labeled datasets
Deep learning
GPU technology

Error rates on ImageNet Visual Recognition Challenge, %

Sources: ImageNet; Stanford Vision Lab
Take-home message

• Computer vision is a key branch of AI
• Enables several modern applications around us
• A lot of highly visible and high-impact activity
• Huge industry interest
• This is a great time to study computer vision!
Course Details
Course details

• Class webpage:
  – http://cseweb.ucsd.edu/~mkchandraker/classes/CSE152A/Fall2022/

• Instructor email:
  – mkchandraker@eng.ucsd.edu

• Grading
  – 30% final exam
  – 45% homework assignments
  – 20% mid-term
  – 5% self-study exercise
  – Several ungraded quizzes

• Aim is to learn together, discuss and have fun!
Course details

• TAs
  – Meng Song: mes050@eng.ucsd.edu
  – Mallikarjun Swamy: mswamy@ucsd.edu
  – Rishikanth Chandrasekaran: r3chandr@ucsd.edu
  – Vishal Vinod: vvinod@ucsd.edu

• Tutors
  – Nick Chua, Navya Sharma, Ang Li: {nchua, n1sharma, a3li}@ucsd.edu

• Discussion section: M 3-3:50pm

• TA office hours and tutor hours to be posted on webpage

• Piazza for questions and discussions:
  – https://piazza.com/ucsd/fall2022/cse152a/
Course details

• Homework assignments
  – 4 graded assignments and 1 ungraded
  – Programming in Python might be required
  – Submit PDF to Gradescope before deadline

• Late policy
  – 15% penalty for every 12 hours of delay
  – Submission allowed only up to 72 hours after deadline
Course details

• Collaboration policy
  – It is expected that you complete your academic assignments on your own and in your own words and code. The assignments have been developed by the instructor to facilitate your learning and to provide a method for fairly evaluating your knowledge and abilities. So, to facilitate learning, you are authorized to discuss assignments with others. However, to ensure fair evaluations, you are not authorized to use the answers developed by another, copy the work completed by others, or write your academic assignments in collaboration with another person.

• Academic integrity policy
  – Integrity of scholarship is essential for an academic community. The University expects that both faculty and students will honor this principle and in so doing protect the validity of University intellectual work. For students, this means that all academic work will be done by the individual to whom it is assigned, without unauthorized aid of any kind.
Course details

• No textbook required

• Optional:
  – Computer Vision: Algorithms and Applications
Overall goals for the course

• Introduce fundamental concepts in computer vision

• Enable one or all of several such outcomes
  – Pursue higher studies in computer vision
  – Join industry to do cutting-edge work in computer vision
  – Gain appreciation of modern computer vision technologies

• This is a great time to study computer vision!