

# GE 461 Introduction to Data Science

Spring 2025

#### Applications: Computer Vision Hamdi Dibeklioğlu

#### Computer Vision





According to Hollywood





#### Computer Vision





#### Computer Vision



#### Computer Vision

Computer vision tries to get computers to extract information from images

• Can you name some things that influence what we see?

- Can you name some things that influence what we see?
  - What objects are where in the world (and how they are deformed)
  - Lighting conditions of the environment
  - Position and orientation of the eyes (viewpoint)
  - Your own brain!!!







Ventral or "what" stream



• Count the red crosses

Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х
Х	Х	Х	Х	Х	Х	Х

Count the red crosses







• Computer vision is much more difficult than you might initially think:

- Your brain needs 25% of the cortex just to solve vision
- The general computer vision problem is still largely unsolved
- Main problem: vision needs to deal with enormous variations in the signal
  - Some of these variations are relevant and others should be ignored

#### Gap between "pixels" and "meaning"





What we see

#### Computer Vision: History

- An MIT undergraduate summer project\*, in 1966, aimed to solve background/foreground segmentation and object detection/classification.
- It has been 59 years and we still work on the same problems.

#### Goals - General

The primary goal of the project is to construct a system of programs which will divide a vidisector picture into regions such as likely objects likely background areas chaos. We shall call this part of its operation FIGURE-GROUND analysis. It will be impossible to do this without considerable analysis of

shape and surface properties, so FIGURE-GROUND analysis is really inseparable in practice from the second goal which is REGION DESCRIPTION. The final goal is OBJECT IDENTIFICATION which will actually name objects by matching them with a vocabulary of known objects.

\*Seymour A Papert. The summer vision project, MIT, 1966.

#### Computer Vision

- Making useful decisions about real physical objects and scenes based on images (Shapiro & Stockman, 2001)
- Extracting descriptions of the world from pictures or sequences of pictures (Forsyth & Ponce, 2003)
- Analyzing images and producing descriptions that can be used to interact with the environment (Horn, 1986)
- Designing representations and algorithms for relating images to models of the world (Ballard & Brown, 1982)

#### Computer Vision

• How do we describe the variations within the *class* "chair"?



- Invariance to some variations can be obtained using hand-crafted models
- We generally try to *learn* invariance to the remaining variations from <u>examples</u>

Slide Credit: L. van der Maaten

### **Object Recognition**

• Observation: chairs contain relatively lots of edges



horizontal edges  $\rightarrow$ 

Slide Credit: L. van der Maaten

## Basics: What is an image?

- Assume an image as a function,  $f: \mathbb{R}^2 \to \mathbb{R}$ :
  - f(x, y) gives the intensity at position (x, y)
  - Realistically, an image is defined over a rectangle:
    - $f:[a,b] \times [c,d] \rightarrow [0,255]$

• Color image = Three functions combined together:

• 
$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix}$$

### Basics: An image as a function



## **Basics: Digital Images**

- In computer vision we operate on digital (discrete) images:
  - Sample the 2D space on a regular grid
  - Quantize each sample (round to the nearest integer)
  - Each sample is a pixel (picture element)
  - If we assume each pixel as 1 byte, values range from 0 to 255





- Range transformation (pixel processing): g(x, y) = h(f(x, y))
- Example: ?

• Range transformation (pixel processing): g(x, y) = h(f(x, y))

h

• Example: Noise filtering



g



- Domain transformation (geometric transform):  $g(x, y) = f(h_x(x, y), h_y(x, y))$
- Example: ?

- Domain transformation (geometric transform):  $g(x, y) = f(h_x(x, y), h_y(x, y))$
- Example: Rotation / Translation



### **Basics: Analysis Pipeline**



Input Feature extraction - Classification Output

# Surveillance

- Face Recognition
- Object Detection
- Tracking



# Surveillance

- Anomaly Detection
- Action & Activity Recognition





- Detection & Segmentation
- Tracking



• Future Prediction



- Gaze Estimation
- Mood Recognition
- Fatigue Monitoring



# Photography

 Face Detection / Smile Shutter • Focus Tracking • Image Stabilization Color Enhancement 200 80 800 43 990 0101 SONY 1921



## **Special Effects**





## Special Effects (3D)



# Map & 3D Model Generation

- 3D reconstruction
- Structure from motion

## Weather Forecasting



- Spectral image analysis
- Image segmentation

## Robotics







## **Quality Control & Monitoring**



## Healthcare

- Facial expression analysis
- Pose estimation
- Medical image analysis
- Motion magnification
- Subtle motion tracking







# Neuromarketing

- Gaze estimation / tracking
- Facial emotion recognition
- Age & gender estimation
- Appreciation recognition





## **Visual Search**

#### https://lens.google.com/



# Gaming / Human Computer Interaction





- Pose and motion tracking
- Gesture recognition

## **Sports Analytics**



# Image Analytics: Identity Recognition / Tagging



# Image Analytics: Captioning



two birds sitting on top of a tree branch.

an elephant standing next to rock wall.

body of water.

## Image Synthesis/Modification







## Style Transfer





+





=

=





## Text Recognition / Document Digitization





AUGUST 29 2018 CHRISTOPHER KARR BERGLAND GABBY DEAREST READ THIS MORNING you I MISS STUDY SCIENCE PSYCHOLOGICAL A NEW WELL BEING SOCIALE THE PRO ABOUT You THANK OF HANDWRITTEN BENEFITS OF EXPRESSOR THE BOTH ECIATIONS NOTES RECIPIENT THE AND, POST WERE APPRECIATION AGAIN you OUT TURNS THANK ALONG ALL ITIS RIGHT STATIONERY THIS ENGRAVED GIVING FOR **KEEPS** THAT AGIFT TRULY THANKS OF LOVE WITH ETERNAL CHes



#### And many more applications...

