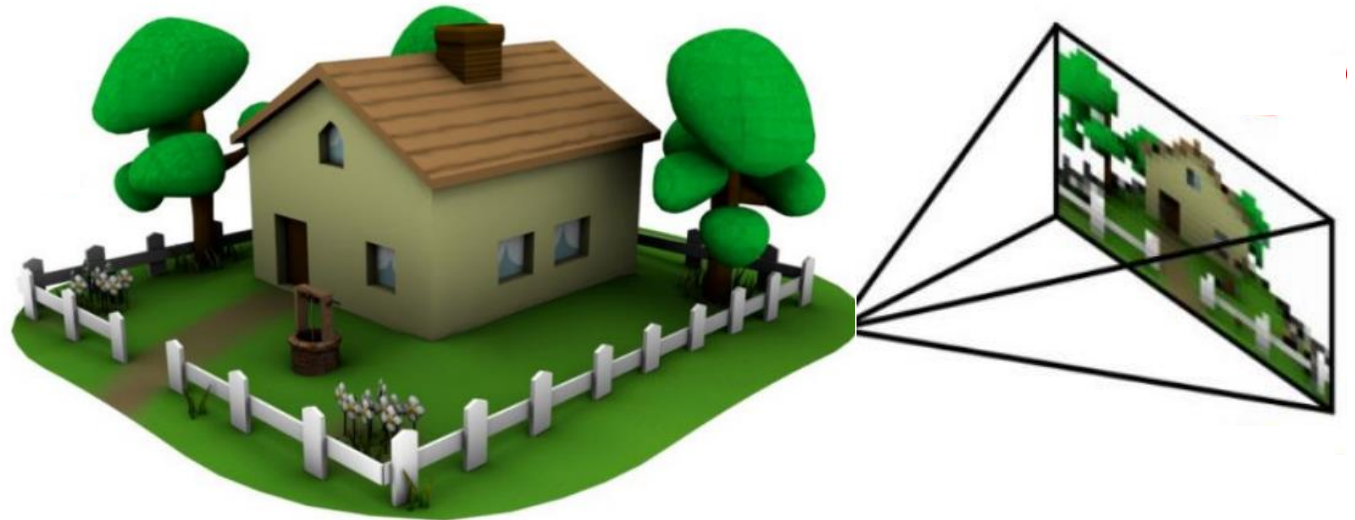# Introduction

Computer Vision

Goal of Computer Vision

extract semantic information from digital image data

to be used for decision making support or automated systems

challenging problem:
images are only 2D projections
of the 3D world

nowadays heavily powered by artificial intelligence (AI), especially machine learning (ML)

Applications of Computer Vision

autonomous driving

facial recognition

automated inspection

medical imaging
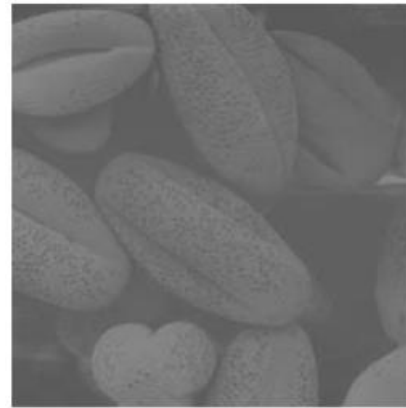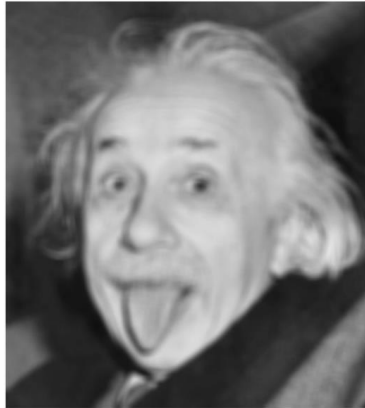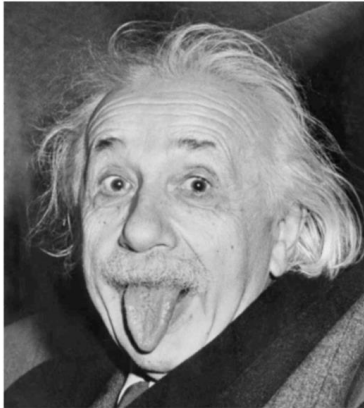
optical character recognition

augmented reality

and many more ...

Image Processing

transformations from image to image

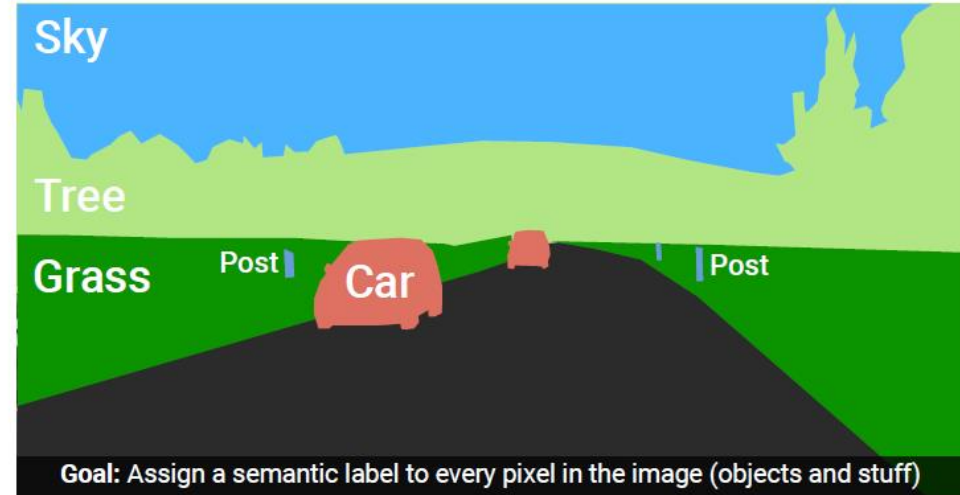(such as scaling, smoothing, sharpening, or contrast stretching)

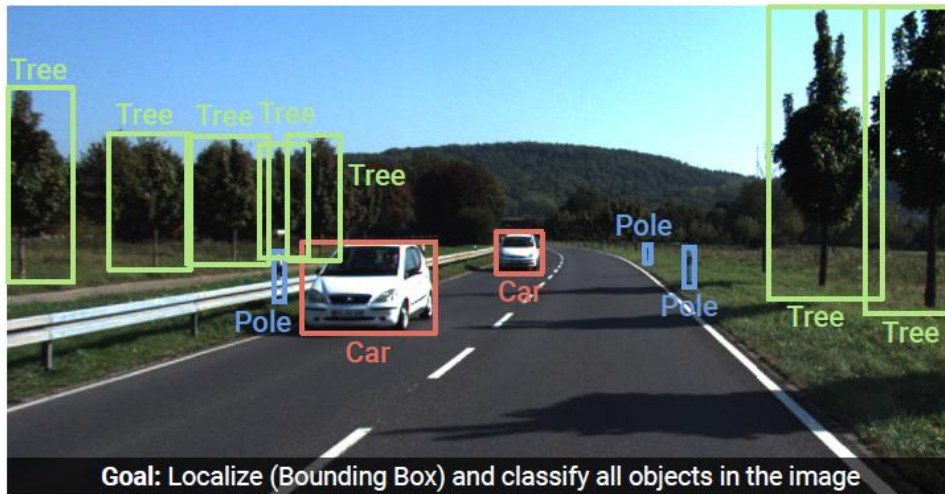to facilitate either machine perception or just human interpretation
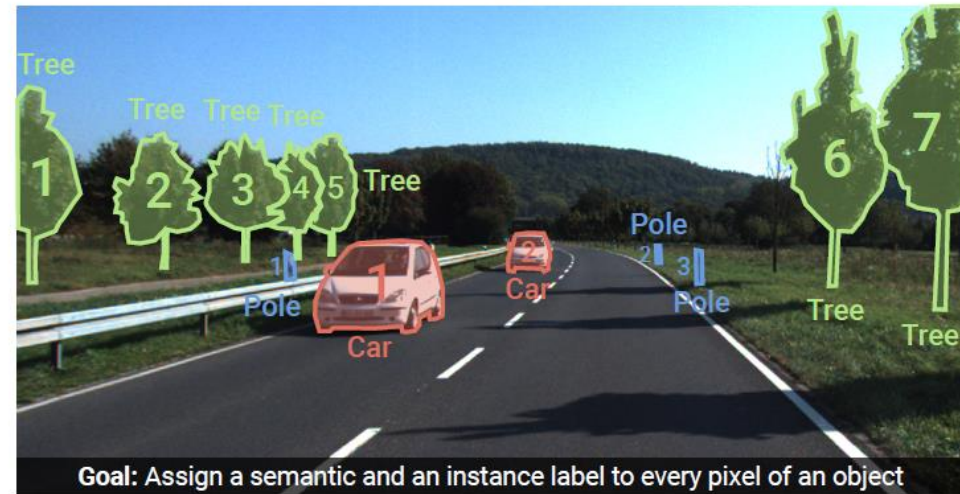
Image Understanding (Recognition)

Image Classification

Semantic Segmentation

Object Detection

Instance Segmentation

Course Schedule

Part 1: Old-School Computer Vision	**Part 2: Machine Learning**
1. **Digital Image Processing** image formation, intensity transformations, compression, spatial filtering, Fourier transform, aliasing, image pyramids 2. **Edges and Features** Canny & Marr-Hildreth edge detectors, Hough transform, feature detection (Harris corner detector) and description (SIFT), feature matching, eigenfaces 3. **Projective Geometry** geometric transformations, image alignment, stereo vision, camera calibration	4. **Image Classification: From Classic ML to Deep Learning** supervised learning, linear regression, bias-variance tradeoff, neural networks, convolutional neural networks, transfer learning, transformers (language models, ViT, DINO, CLIP) 5. **Segmentation and Detection** semantic segmentation, object detection (R-CNN, YOLO), object tracking, instance segmentation, promptable segmentation 6. **Image Synthesis** generative models, GAN, VAE, diffusion

Main Areas of AI/ML

empowered by one key component: learning from data (ML)

tabular data

language models

computer vision

control
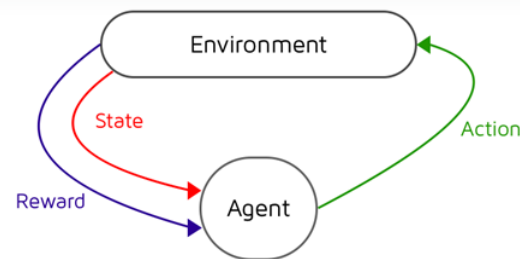
When to Use ML (Learning from Data)
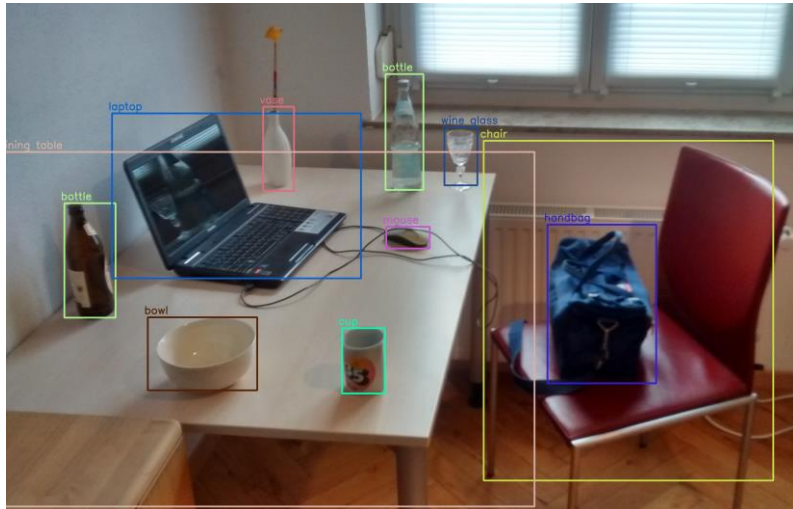
automation

too complex for rules

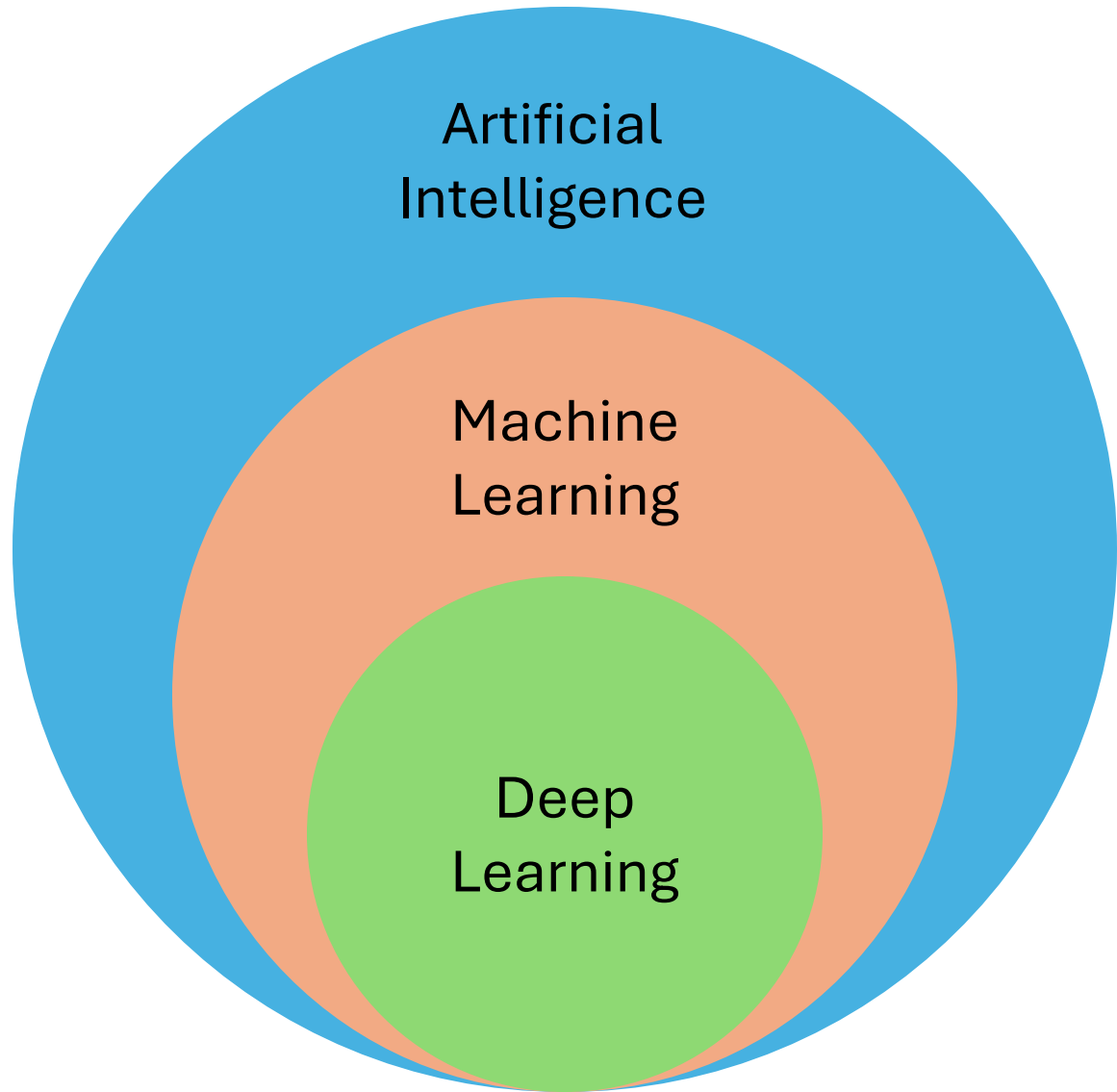object recognition, chat bot, ...

uncertainty

too complex for humans

AlphaFold

protein structure predictions, demand forecasting, ...

blend of diverse components from different domains (statistics, optimization, computer science, ...)

Deep Learning: special kind of ML methods using *deep* neural networks (e.g., CNNs, transformers)
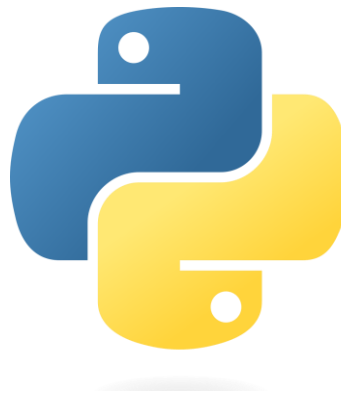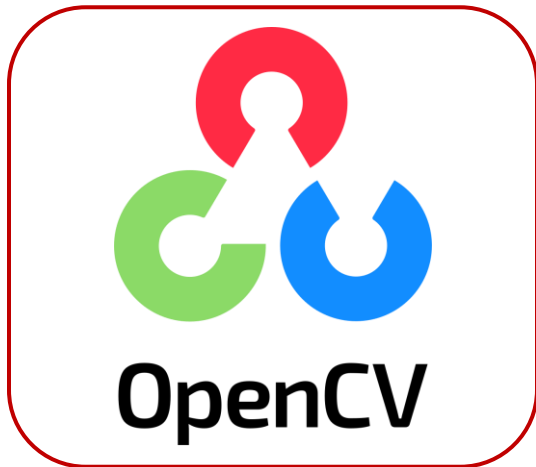
Used Programming Language: Python

- good compromise between rapid prototyping and production

- vast ecosystem

- very popular for data processing and ML: scientific Python stack

Used Python Libraries

computer vision:

deep learning:

Programming Environments

locally, best use

- a virtual environment to flexibly install packages (e.g., venv)

- an IDE of your choice (e.g., VS Code)

- both plain Python files or Jupyter notebooks are fine

but cloud-based environments also fine (e.g., Google Colab)

Literature

- [Computer Vision: Algorithms and Applications](#)
- Digital Image Processing, Gonzales & Woods

- [Deep Learning](#)
- [The Little Book of Deep Learning](#)
- [Understanding Deep Learning](#)