Computer Vision

Image Segmentation

Presented by: Besma Guesmi

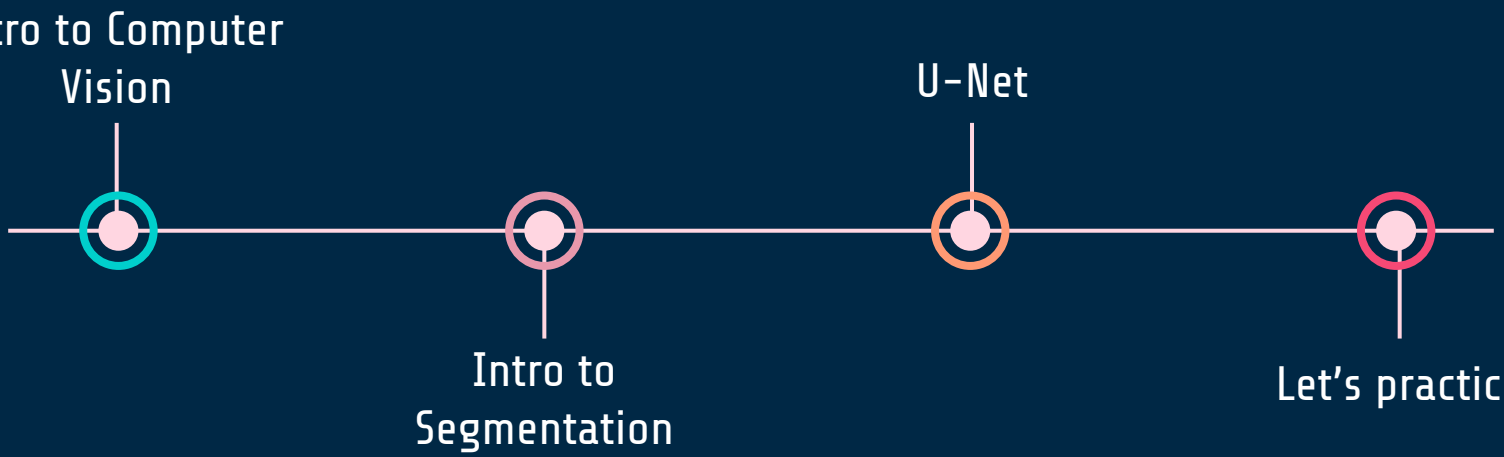
Software Engineer Specializing in
Computer Vision

"AI is the new electricity. Just as electricity transformed numerous industries starting 100 years ago, AI is now poised to do the same."

—Andrew Ng

Co-Founder @Coursera

OUR Agenda


Intro to Computer Vision
Intro to Segmentation
U-Net
Let’s practice

What is Computer Vision?

01

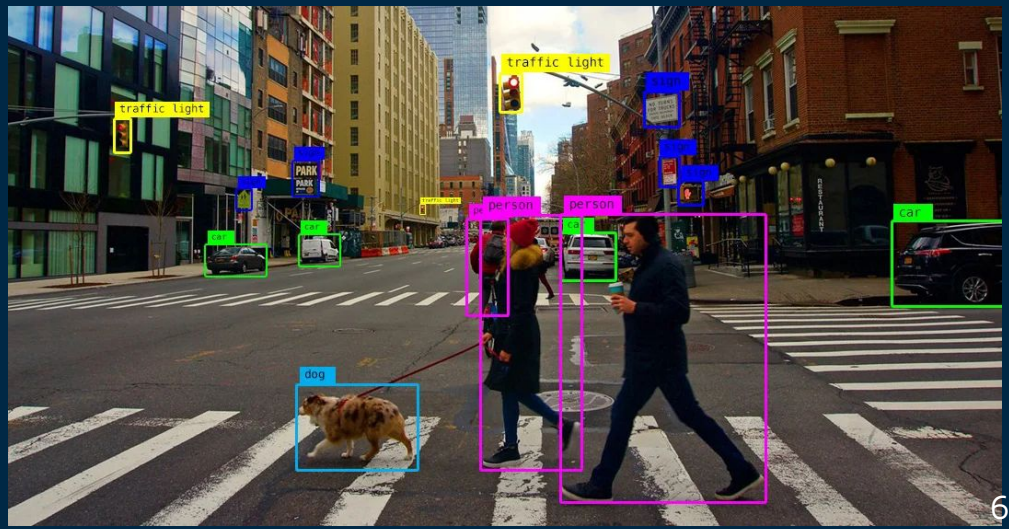
What is Computer Vision?

- A field of Computer Science focuses on analysing and processing visual images or videos intelligently like human
- Enable machines to learn and understand the images at pixel level through training and validation
- Machines retrieve visual information, handle it and interpret the result by getting training from special ML Software algorithms.

What is Computer Vision?

Make computers understand images and video.

- **What kind of scene?**
- **Where are the cars?**
- **How far is the building?**



How Computer Vision Works?

02

How Computer Vision works?

Computer Vision works in three basics steps:

- **Acquiring an image (image capture)**
- **Preprocessing the image (noise reduction, image enhancement…)**
- **Understanding the image (Classification, detection, segmentation)**

How Computer Vision works?

- CV algorithms that we use today are based on pattern recognition
- We train computers on massive amount of visual data - computer process the image, label objects on them, and find patterns on those images.

Computer Vision Applications

03

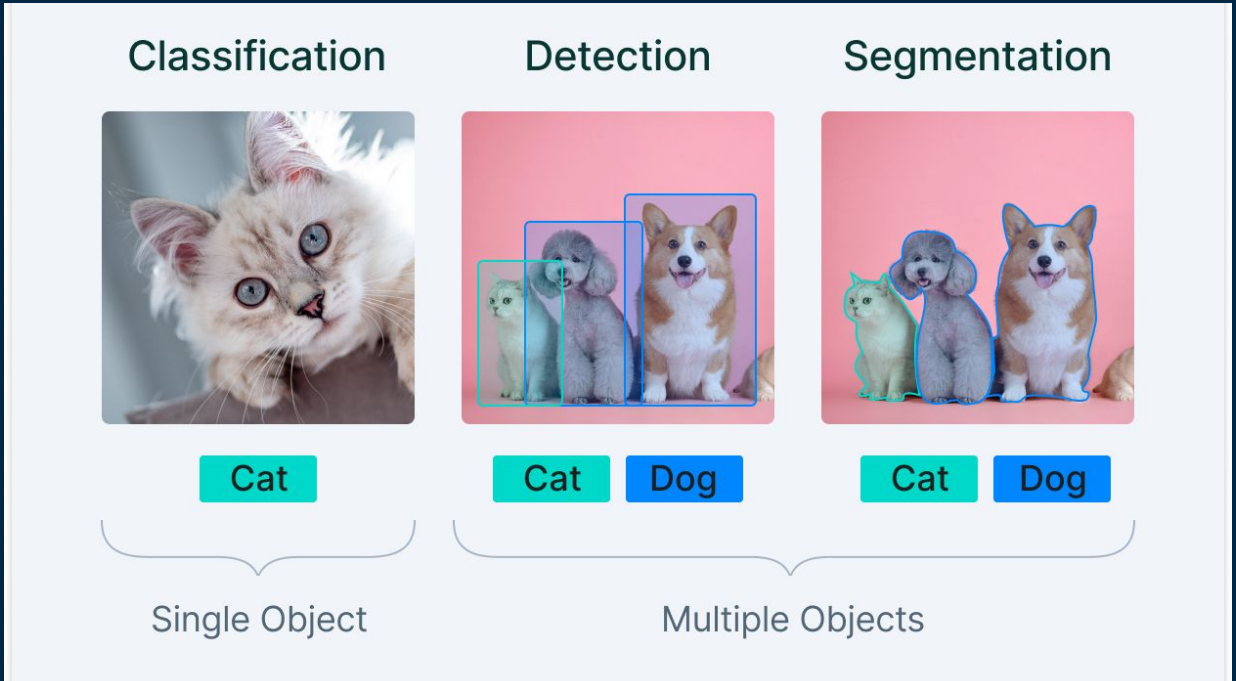
Computer Vision applications

- CV is one of the ML (Machine Learning) area where core concepts are already being integrated into major products that we use every day, like:
  - Self driving cars
  - Facial recognition
  - Healthcare
  - etc

Computer Vision Tasks

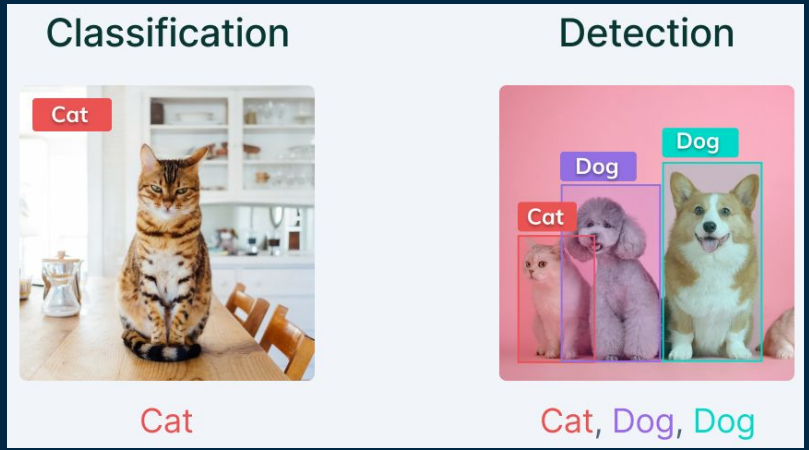
04

Computer Vision Tasks



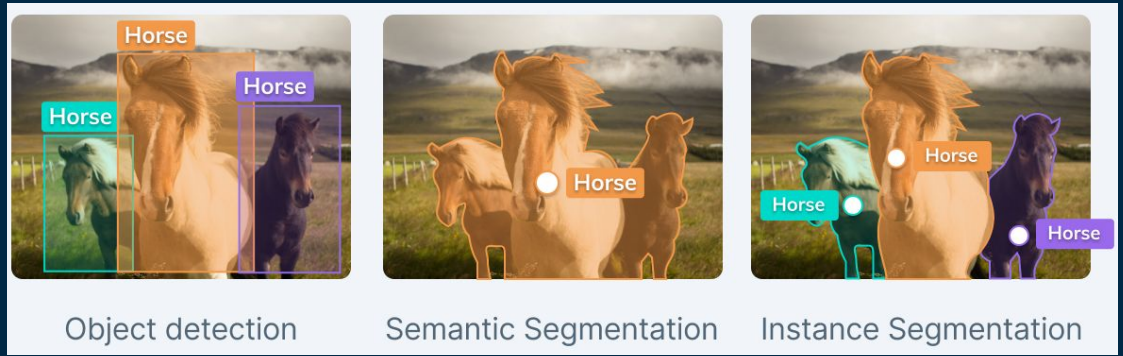
Object detection vs. image classification

- **Image classification** assigns a single label to an entire image using a classifier, but it doesn't pinpoint where the labelled object is.
- **Object detection** goes further by not only classifying objects, but also drawing bounding boxes around them to precisely locate them in the image.



Object detection vs image segmentation

- **Semantic segmentation** marks all pixels of that class, but doesn't outline object boundaries.
- **Object detection** doesn't segment objects, but precisely locates them with bounding boxes.
- **Instance segmentation** combines both by first detecting objects and then segmenting them within the detected boxes, yielding distinct regions for each instance.



What is Image Segmentation?

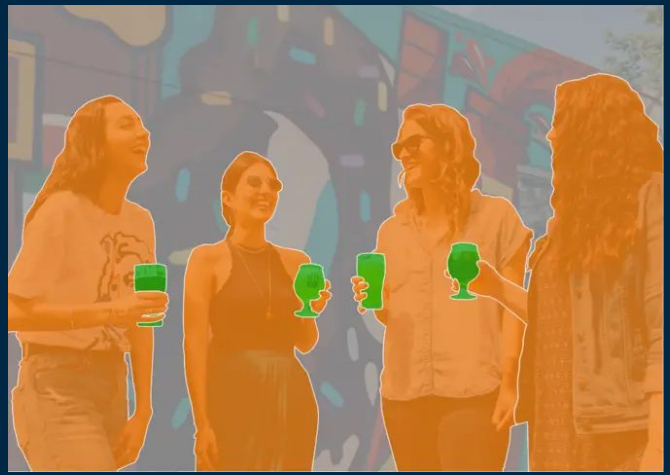
05

Image Segmentation

- Image segmentation is a computer vision technique that partitions a digital image into discrete groups of pixels—image segments—to inform object detection and related tasks.

Image Segmentation Type

- **Binary image segmentation — Pixels classified as object or background.**
- **Multi-class Segmentation: Pixels categorized into multiple classes.**
- **Instance Segmentation: — Identifying and distinguishing individual instances of objects.**

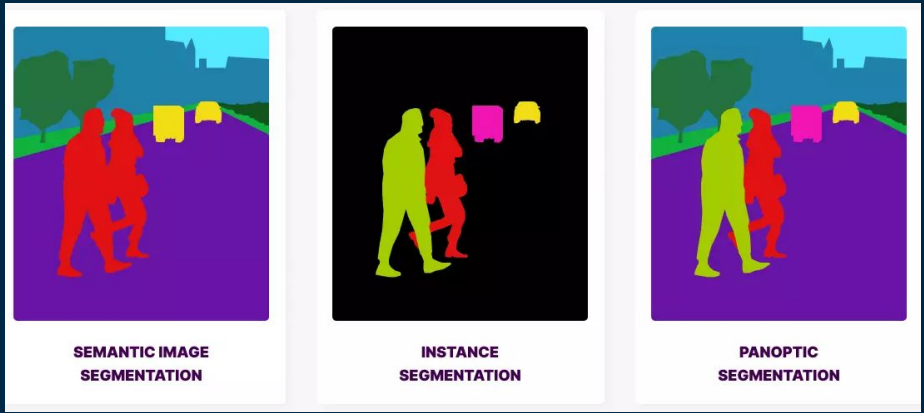


Semantic Segmentation

Instance Segmentation

Type of Segmentation

- **Semantic Segmentation** — this involves arranging the pixels in an image based on semantic class.
- **Instance segmentation** — this technique involves classifying pixels based on the instances of an object instead of classes.
- **Panoptic segmentation** — a combination of semantic and instance segmentation. It predicts the identity of each object, separating every instance of each object in the image.

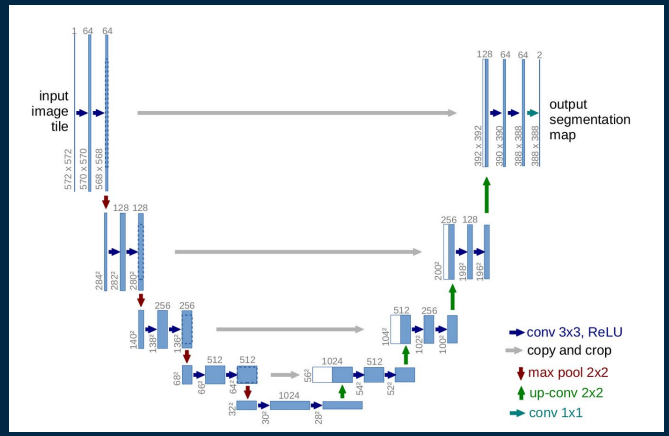


Intro to U-Net

07

U-Net architecture

- **Contracting Path (Encoder)** — is a series of layers that extract image features using progressively deeper and narrower filters. The encoder can be pre-trained on a similar task, such as image recognition, which allows it to leverage its existing knowledge to perform segmentation tasks.
- **Bottleneck** — Compressed representation that capturing essential features.
- **Expansive Path (Decoder)** — is a series of layers that gradually convert the encoder's output into a segmentation mask corresponding with the input image's pixel resolution.
- **Skip connections** — multiple long-range neural network connections allow the model to identify features at different scales to enhance model accuracy.

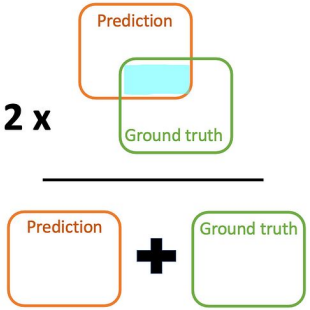


Segmentation Metrics

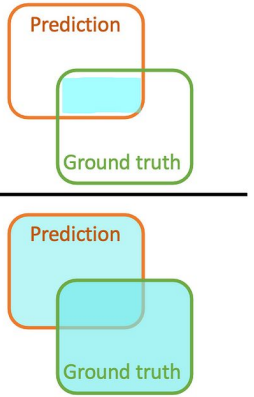
08

Metrics

$$\text{Dice} = \frac{2\,X\,Area\ of\ overlap}{Total\ area} =$$



$$\text{IoU} = \frac{Area\ of\ overlap}{Area\ of\ union} =$$



Any questions?

bessmagsm@gmail.com

THANK YOU!

Besma Guesmi