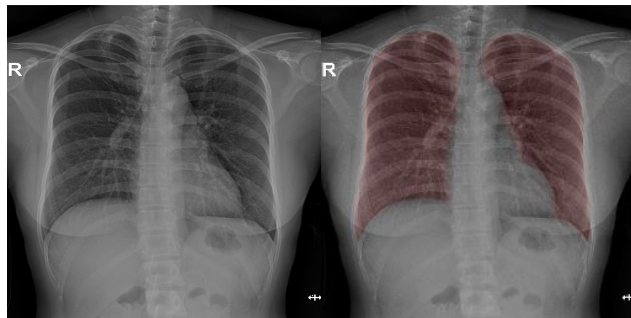
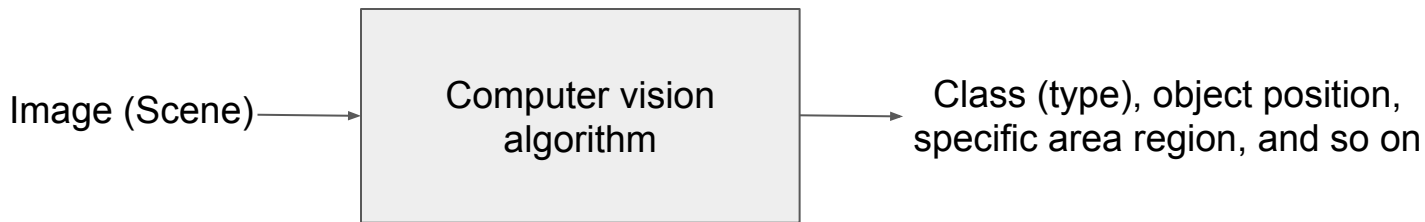


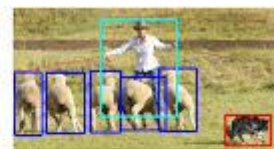
Introduce computer vision  
using deep neural network

# What is the computer vision?

- image or a video is taken as input, and the goal is to understand (including being able to infer something about it) the image and its contents.



(a) Image classification



(b) Object localization



(c) Semantic segmentation



(d) This work

# Computer vision task

**Classification**



**Semantic Segmentation**



GRASS, CAT,  
TREE, SKY

No objects, just pixels

**Classification + Localization**



CAT

Single Object

**Object Detection**



DOG, DOG, CAT

Multiple Object

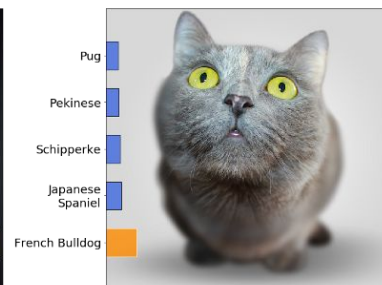
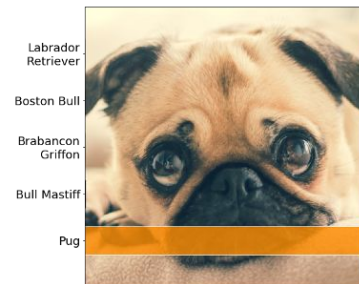
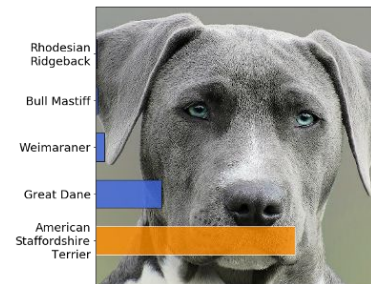
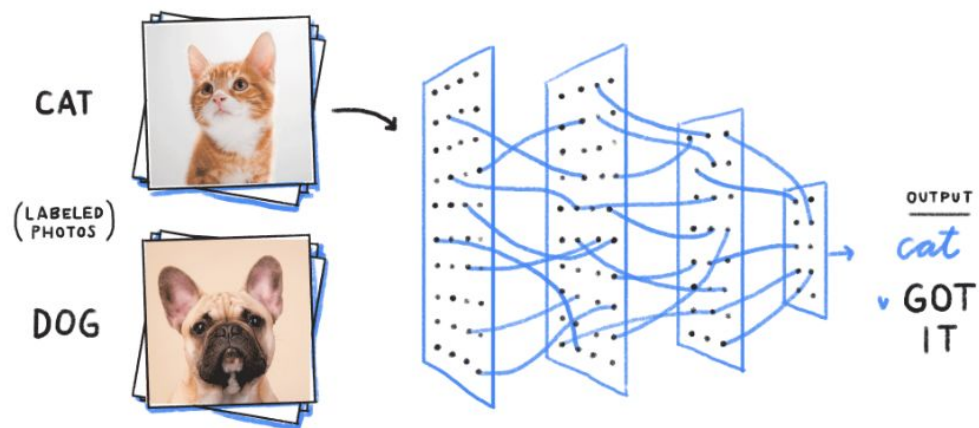
**Instance Segmentation**



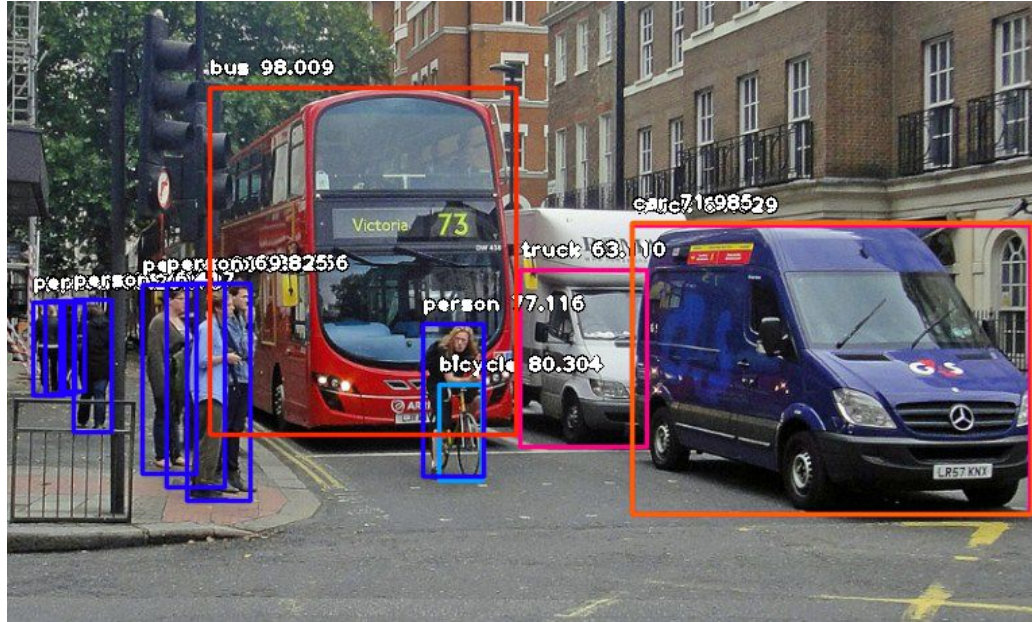
DOG, DOG, CAT

This image is CC0 public domain

# Image classification

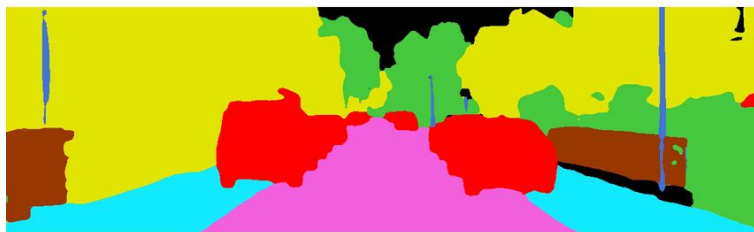










# Object detection (object localization + classification)



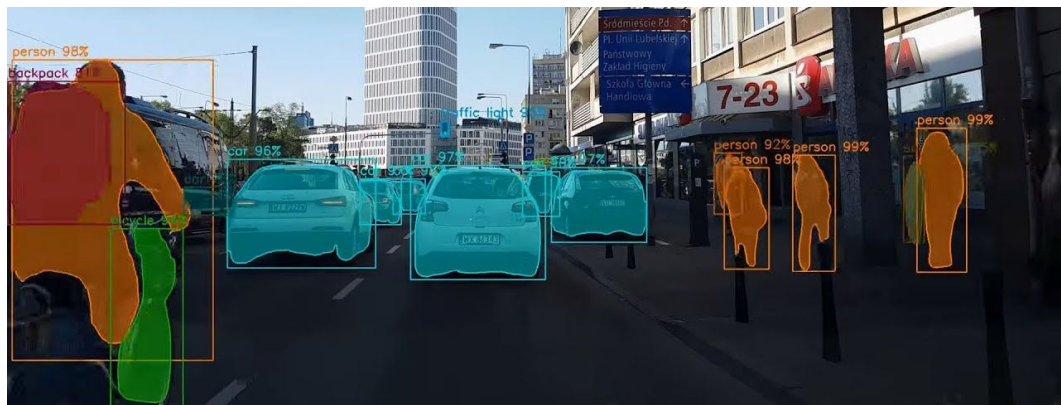


# Image segmentation



 Road	 Sidewalk	 Building	 Fence
 Pole	 Vegetation	 Vehicle	 Unlabel

**Semantic segmentation (pixel-wise classification)**

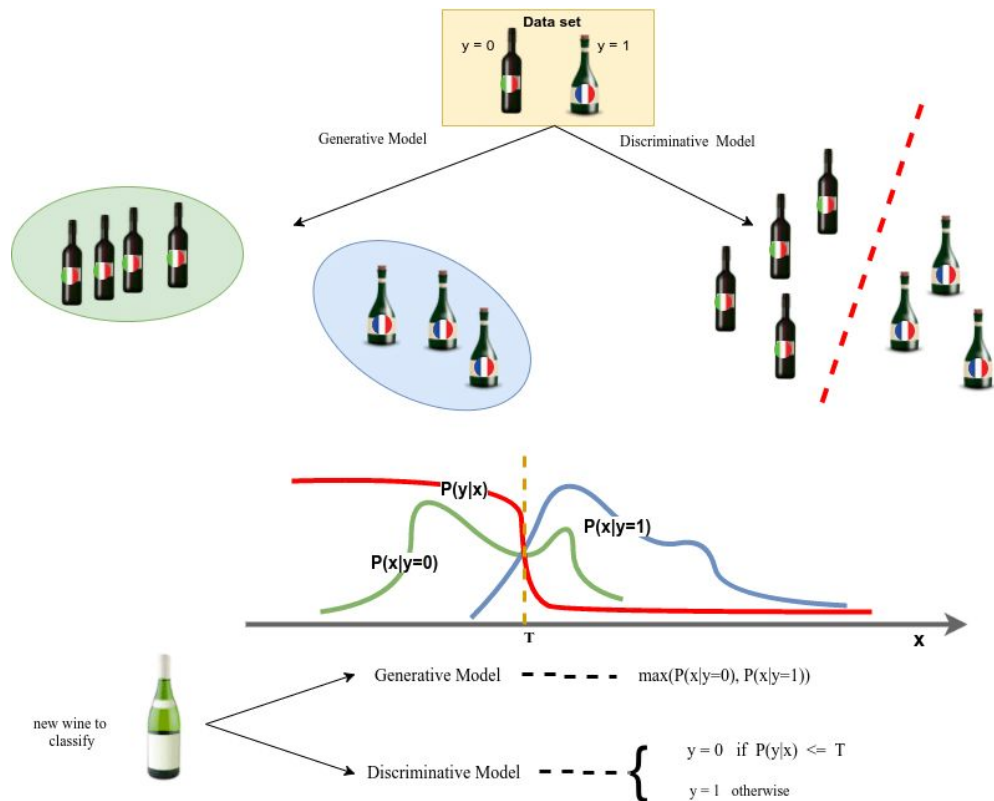


**Instance segmentation**

# Image generation

[https://www.youtube.com/watch?v=xq\\_wlxeax8](https://www.youtube.com/watch?v=xq_wlxeax8)

# Generative model vs discriminative model





# Perception (autonomous driving car field)

<https://www.youtube.com/watch?v=O8kdeaZoMbM> (Perception)

<https://www.youtube.com/watch?v=gfWjsKsEry0> (autonomous driving text using perception)