Computer Vision

Main problems and mostly used architectures

AUA, DS330 Deep Learning 2021, Hovhannes Tamoyan

Main Problems

Classification

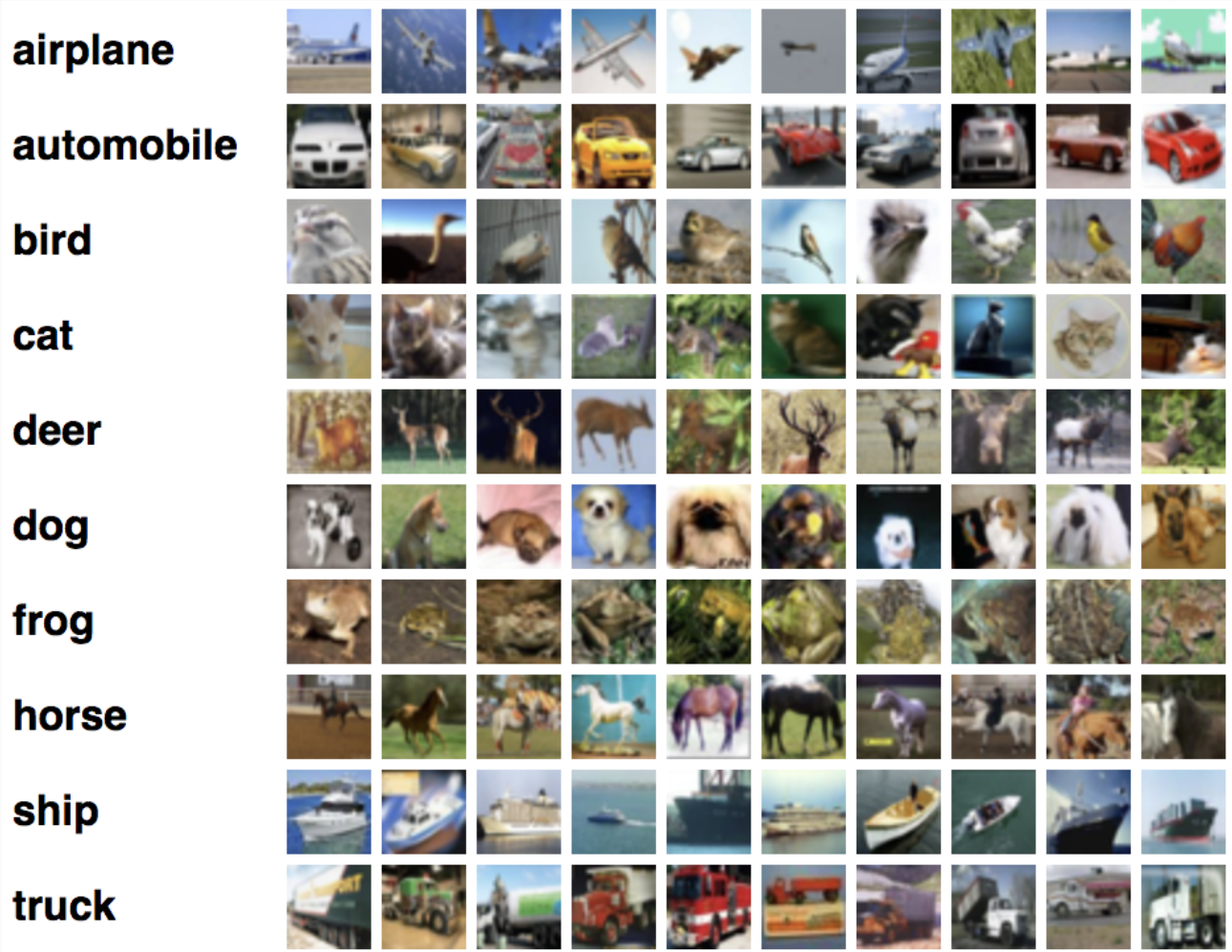
Semantic Segmentation

Object Detection, Instance Segmentation

Classification

Given an image x, the goal is to classify the image by assigning it to a specific label from the given labels set: y_i

$$P(x|\theta) = y_c \text{ where } c \in \{1...n\}, n \text{ is the number of classes}$$


airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck

Common Architectures

AlexNet

VGG

ResNet

Inception v3

GoogLeNet

Semantic Segmentation

Given an image x, the goal is to label each pixel of the x with a corresponding class of what is being represented.

$$P(x_{i,j}|\theta) = y_c \text{ where } i, j \in \{H, W\}, c \in \{1...n\},$$

H and W are the x image height and width respectively, n is the number of classes

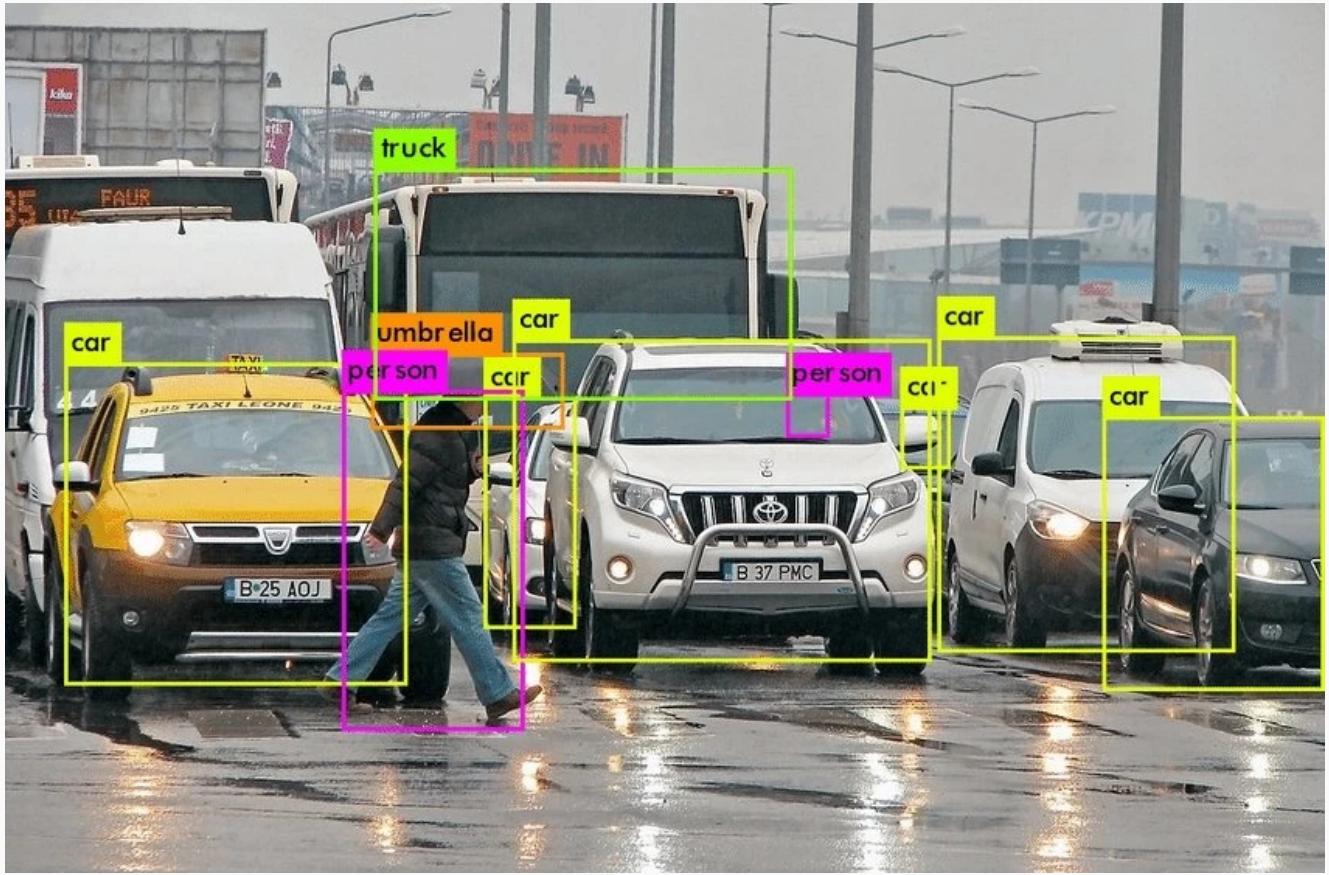
Common Architectures

ResNet

Object Detection, Instance Segmentation

Given an image x, the goal is to identify and locate (draw boundaries) objects in it. Instance segmentation receives the window of an object and applies segmentation.

$$P(x|\theta) = \{z_i\} \text{ where } z_i \in \text{set } Z,\ z_i = \{(\text{bounding-box, class})...\}$$


truck
car
umbrella
car
person
car
person
car
car
B 25 AOJ
B 37 PMC

Common Architectures

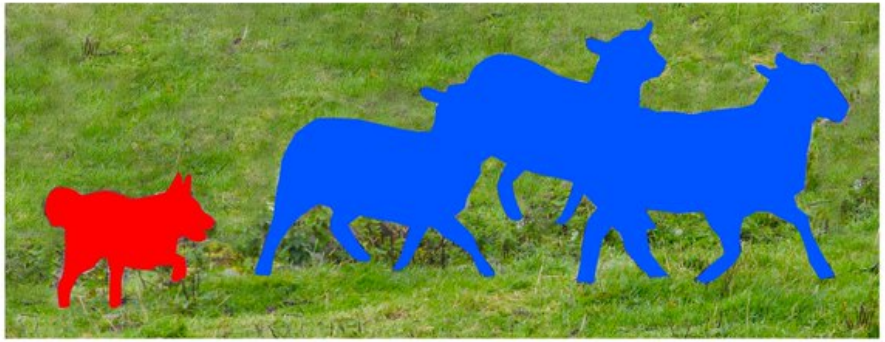
R-CNN

RetinaNet

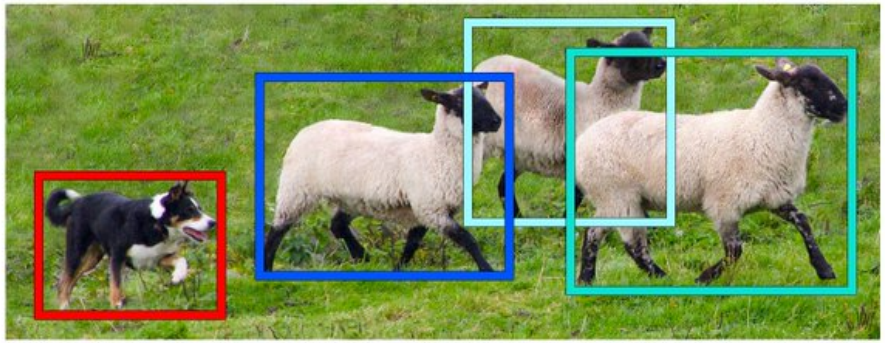
SSD

Yolo

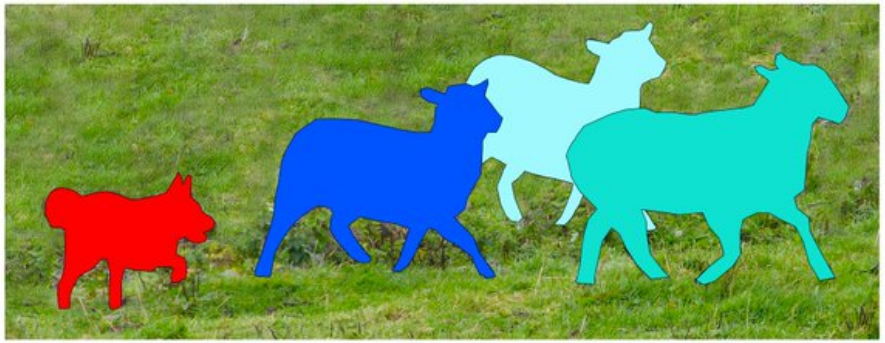
Image Recognition



Semantic Segmentation



Object Detection



Instance Segmentation

Thank You