Real-Time Object (Face Mask) Detection Method Based on YOLOv3

Reza Adinepour

Shahrood University of Technology

Dec. 2022, Shahrood

Content

- Image Understanding
- Object Detection
- What is YOLO?
- Architecture of YOLOv3
- Loss Function
- YOLO Network Train Setting

Image Understanding

What objects are There?

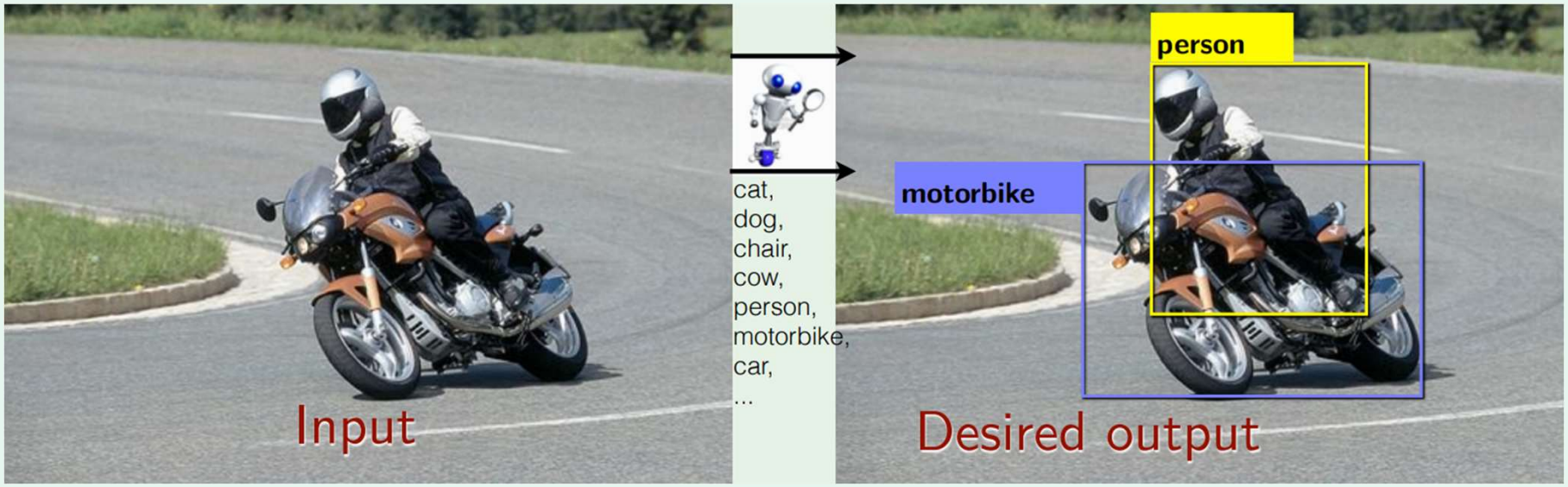
Human VS Robot:

I see a table with pencil, notebook, paper, and pen holder

I don't see anything except number(0, 1)

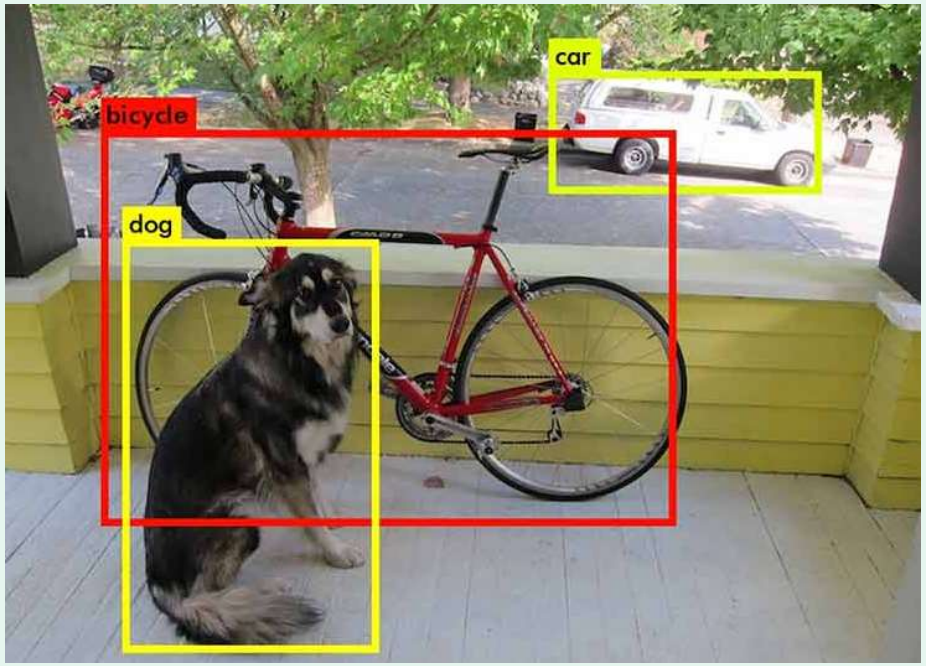
Formalizing the object detection task

We want robot to see like a human.



Object Detection Algorithms

- Sliding Window
- Regions Proposal
- YOLO

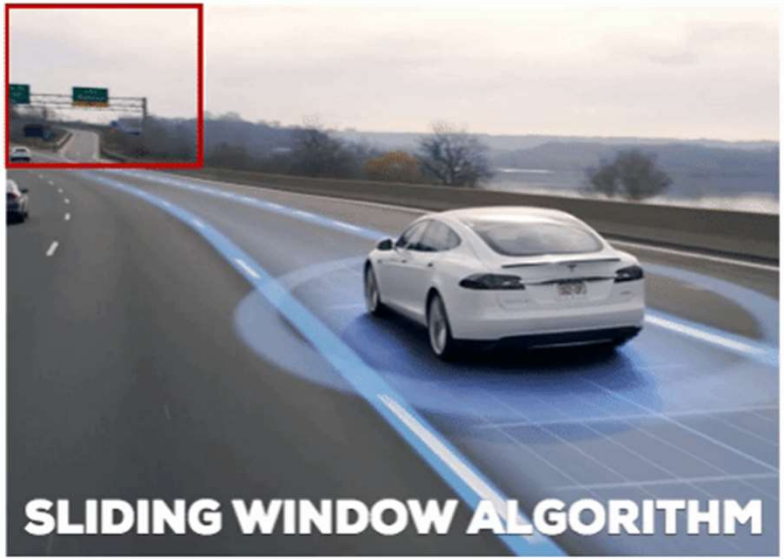


Sliding Window

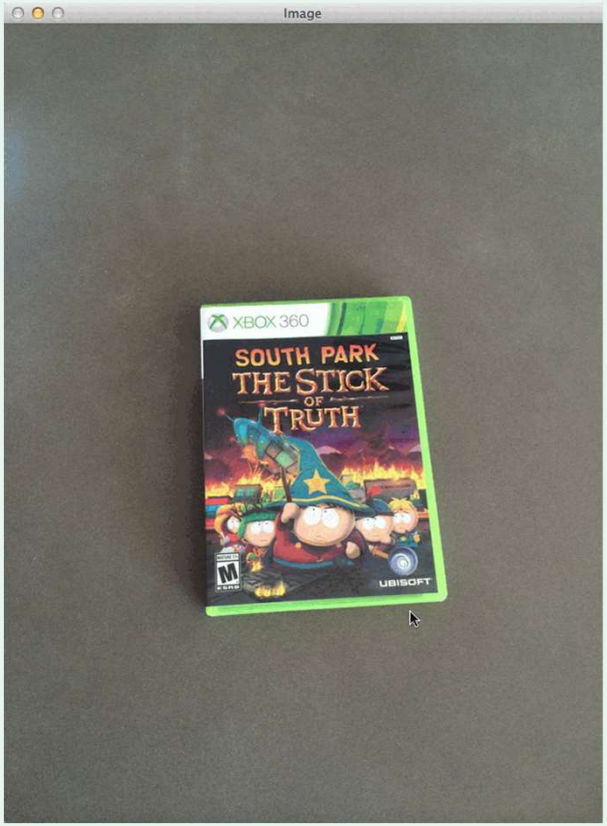
in this algorithm make n × n window and moving this slide a window from left to right and top to bottom.

Disadvantages:

1. Very low
2. We don't know size of objects for make size of window

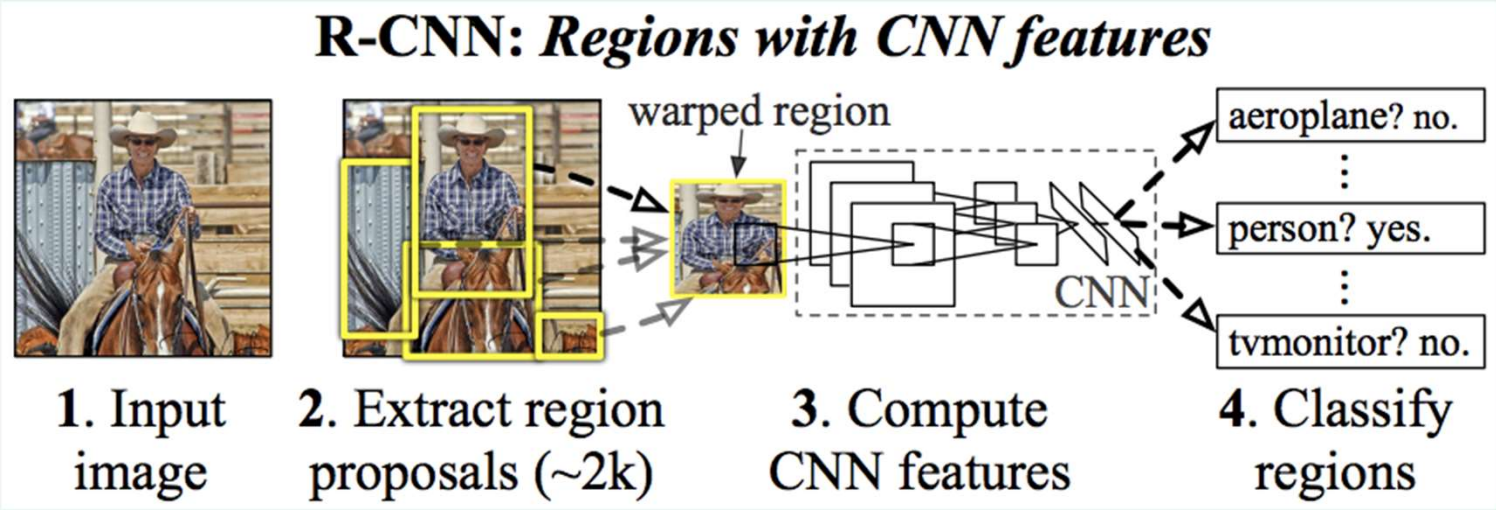


SW Disadvantages Object Size



Regions Proposal

Like R-CNN. In this algorithm, first: extract region proposals, second: compute CNN feature third: classify region



YOLO

- YOLO is a shortened from of "You Only Look Once" like human. And it uses convolutional neural networks for objects detection.
- YOLO can detect multiple objects on a single image.
- YOLO network solved problem with regression approach. (direct mapping from input image pixel to coordinates of bounding box and confidence of classes)

YOLO Versions

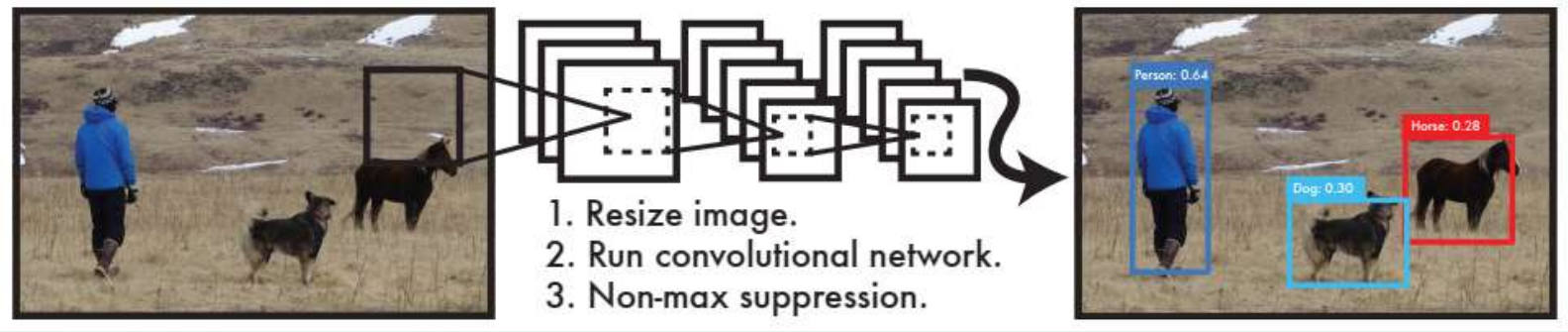
$$\text{YOLO Net} \begin{cases} Original\ YOLO \rightarrow 45fps \begin{cases} V1 \\ V2 \\ V3 \\ V4 \end{cases} \\ Fast\ YOLO \rightarrow 150\ fps \end{cases}$$

The YOLO Detection System

Processing images with YOLO is simple and straightforward.

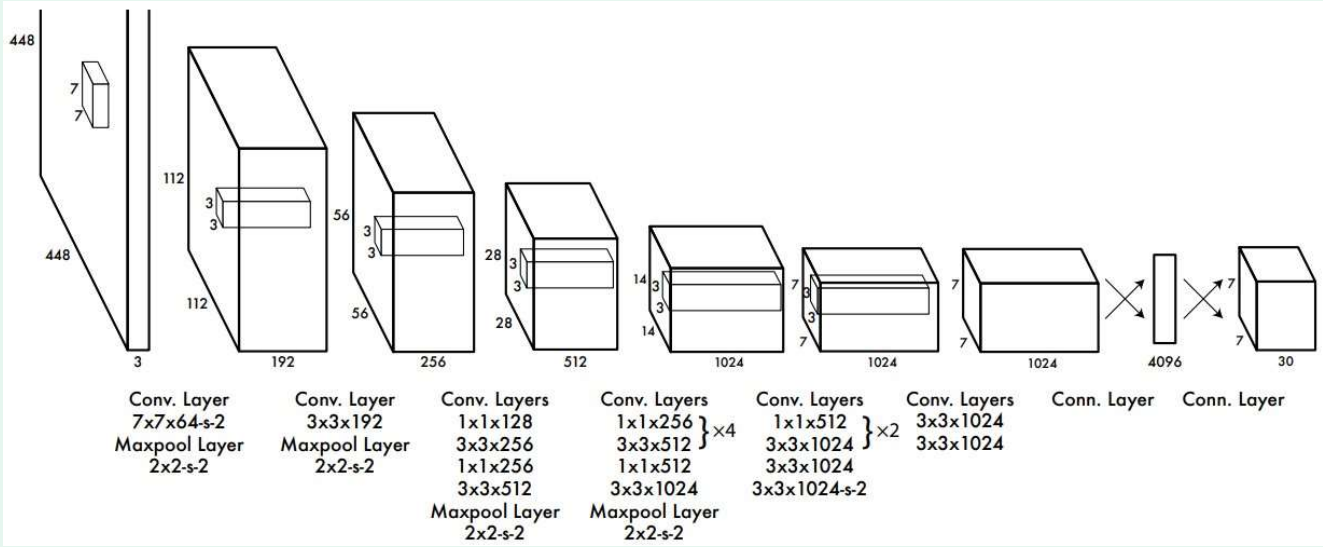
Our system:

1. resizes the input image to 448 × 448 × 3
2. runs a single convolutional network on the image, and
3. thresholds the resulting detections by the model's confidence.



YOLO convolutional Layer

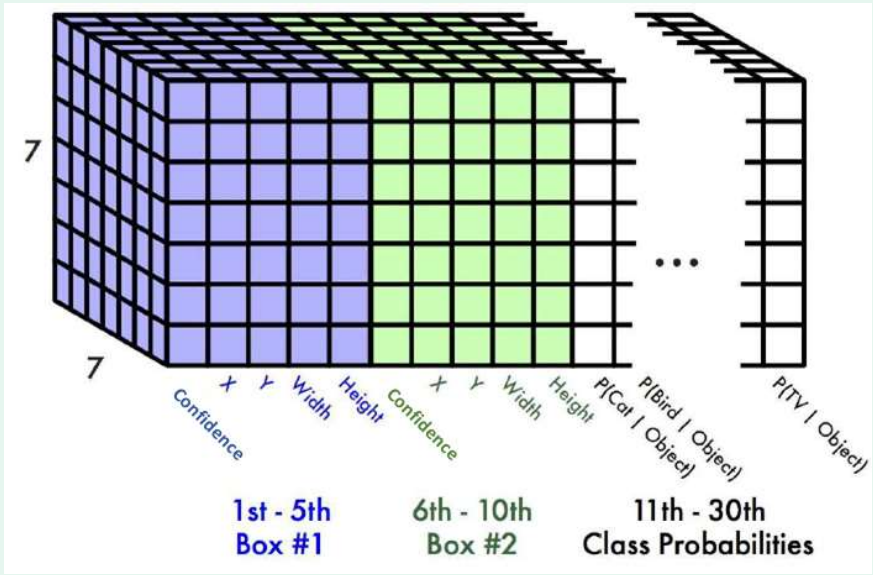
Our detection network has 24 convolutional layers followed by 2 fully connected layers. (YOLOv3 have 53 convolutional layers)



Output of YOLO Network

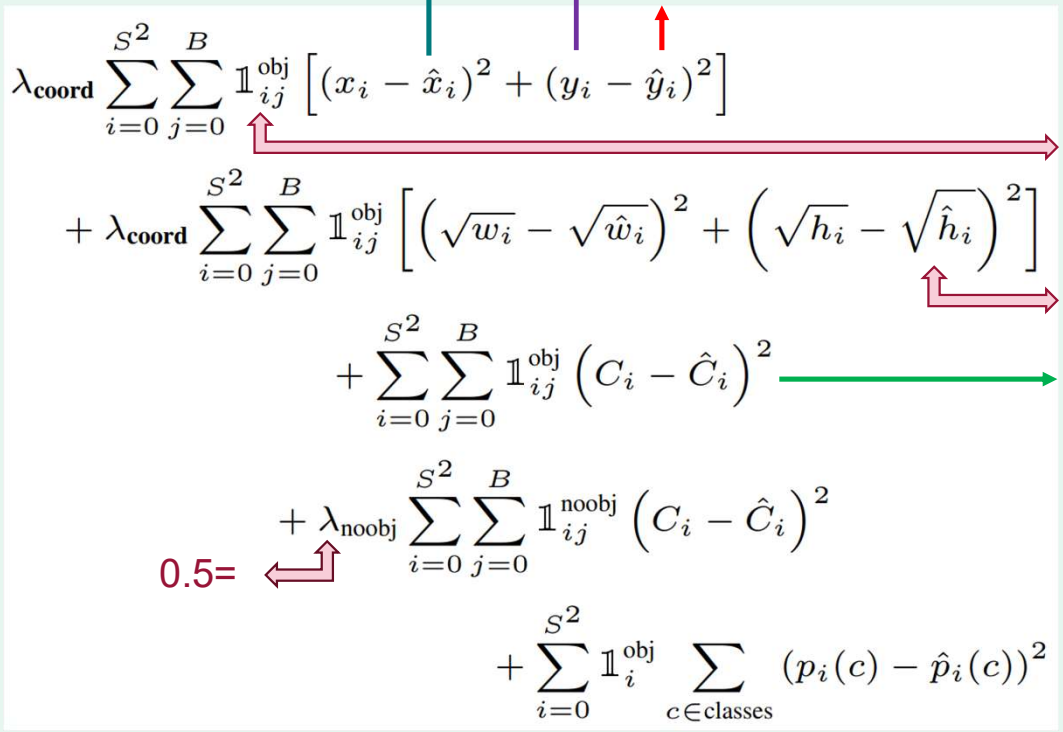
Output of network is a $7 \times 7 \times 30$

Each 7×7 entry contains a vector with length of 30 (contains x, y, h, w, confidence)



LOSS Function in YOLO

Mean Squared Error:



YOLO Advantages and Disadvantages:

Advantages:

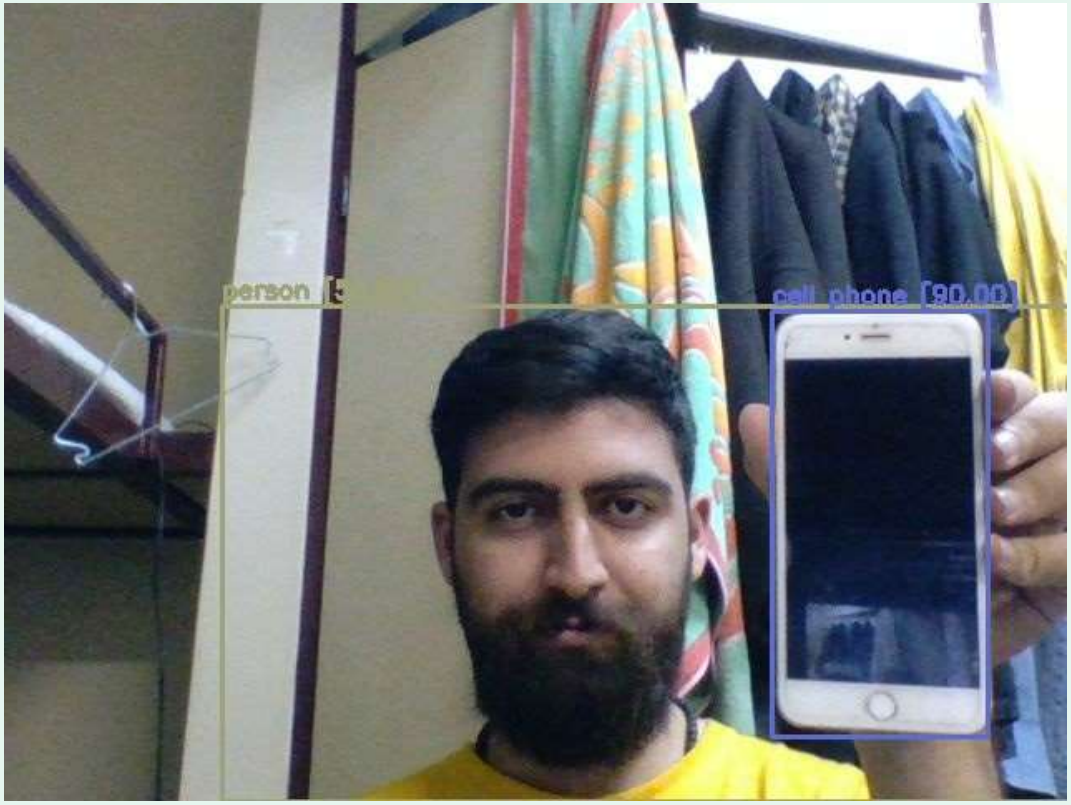
1. Very Fast
2. High generalizability for test data

Disadvantages:

1. Challenge in detect small objects
2. Challenge in detect objects with abnormal size

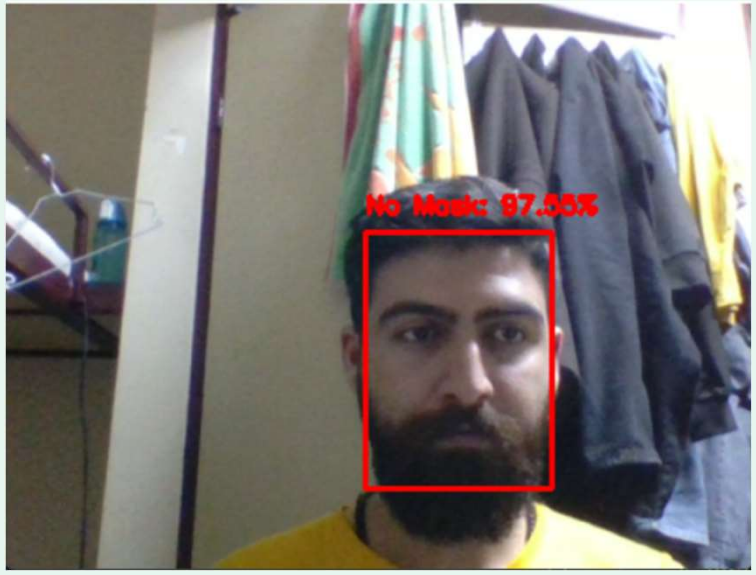
Real Time Object Detection Using YOLO

run code!



Face Mask Detection

run code!

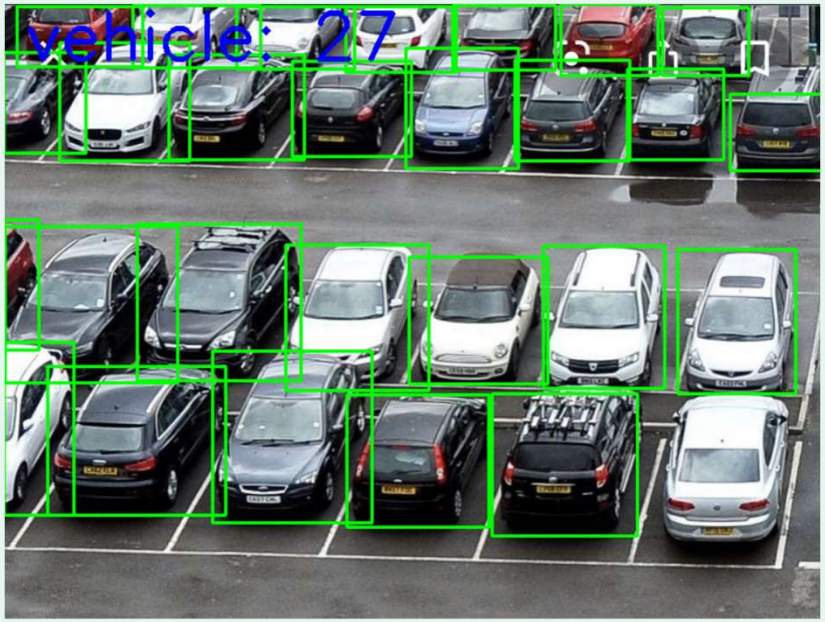


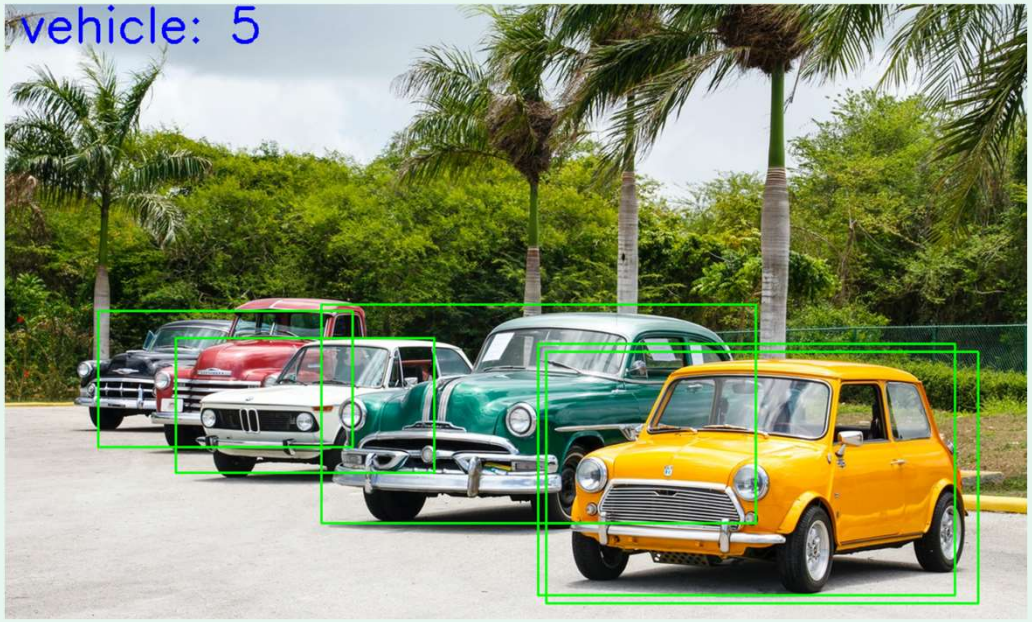
Object Detection on Image Using YOLO

run code!

Count Vehicles on Image Using YOLO

run code!





[Thank You]

-

Any Question?

Reza_Adinepour@shahroodut.ac.ir