

YOU LOOK ONLY ONCE (YOLO)

**A Hands-on Introduction to the State of the Art
Real Time Object Detection Deep Learning Model**

Pouria Rouzrokh

MD MPH MHPE

Bradley Erickson

MD PhD

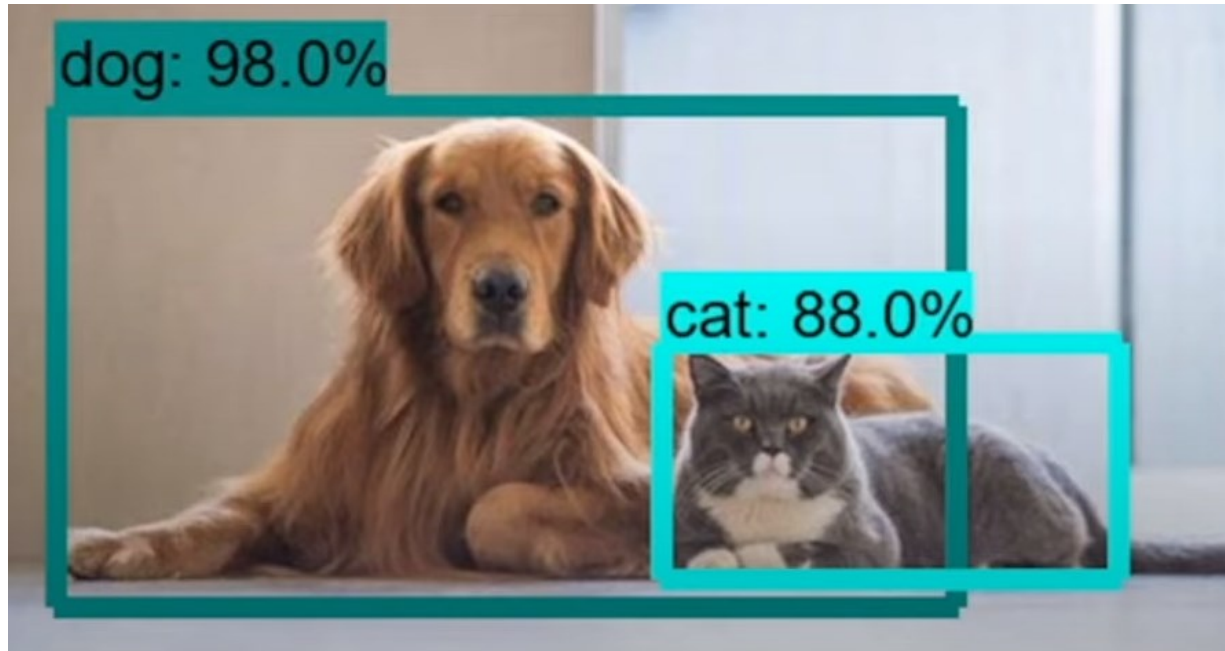
Radiology Informatics Laboratory | Mayo Clinic Artificial Intelligence Laboratory

Mayo Clinic, MN, USA

Part 1

INTRODUCTION

Defining Object Detection (1)



Object Detection:

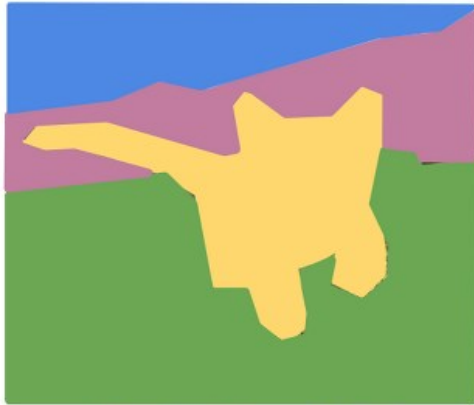
1. What kind of object is it?
2. Where is that object located?

Defining Object Detection (2)



Defining Object Detection (3)

Semantic Segmentation



GRASS, CAT,
TREE, SKY

No objects, just pixels

Classification + Localization



CAT

Single Object

Object Detection



DOG, DOG, CAT

Multiple Object

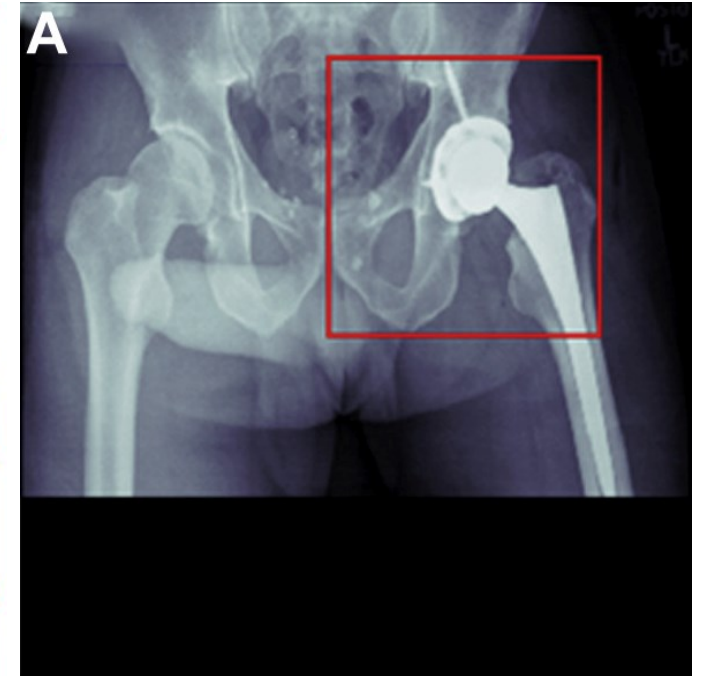
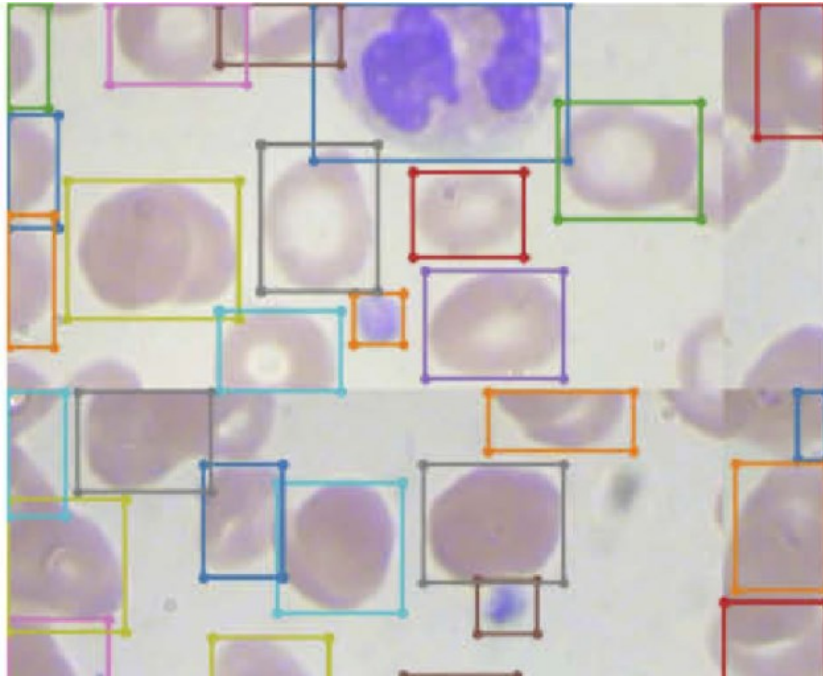
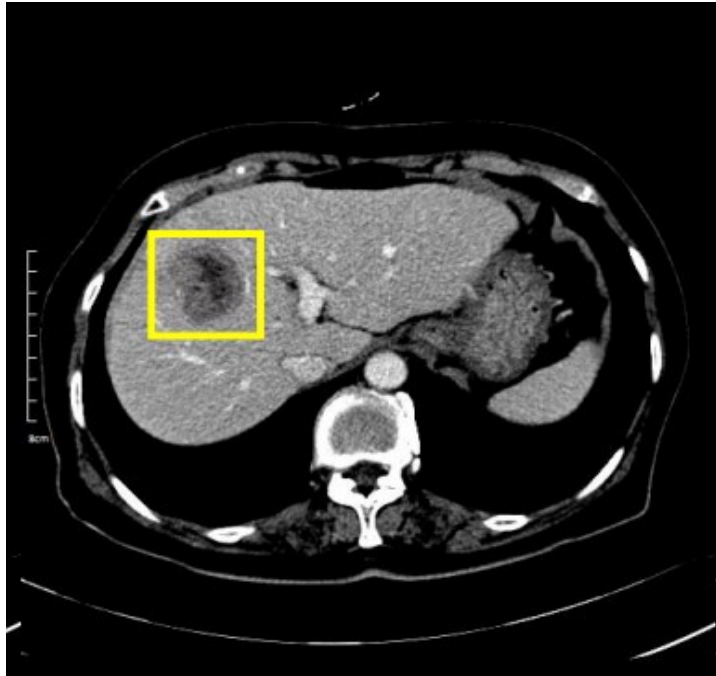
Instance Segmentation



DOG, DOG, CAT

[This image is CC0 public domain](#)

Medical Applications of Object Detection



Sources:

- Liver Lesion Detection from Weakly-labeled Multi-phase CT Volumes with a Grouped Single Shot MultiBox Detector
- Improved detection performance in blood cell count by an attention-guided deep learning method
- Deep Learning Artificial Intelligence Model for Assessment of Hip Dislocation Risk Following Primary Total Hip Arthroplasty From Postoperative Radiographs

Real Time Object Detection (1)

Real-time object detection is the task of doing object detection in real-time with fast inference while maintaining a base level of accuracy.

The model should be able to detect objects and make inferences within microseconds!

Examples of Real-time Object Detection Models

Faster-RCNN (as opposed to RCNN and Fast-RCNN)

EfficientDet

MM-Detection

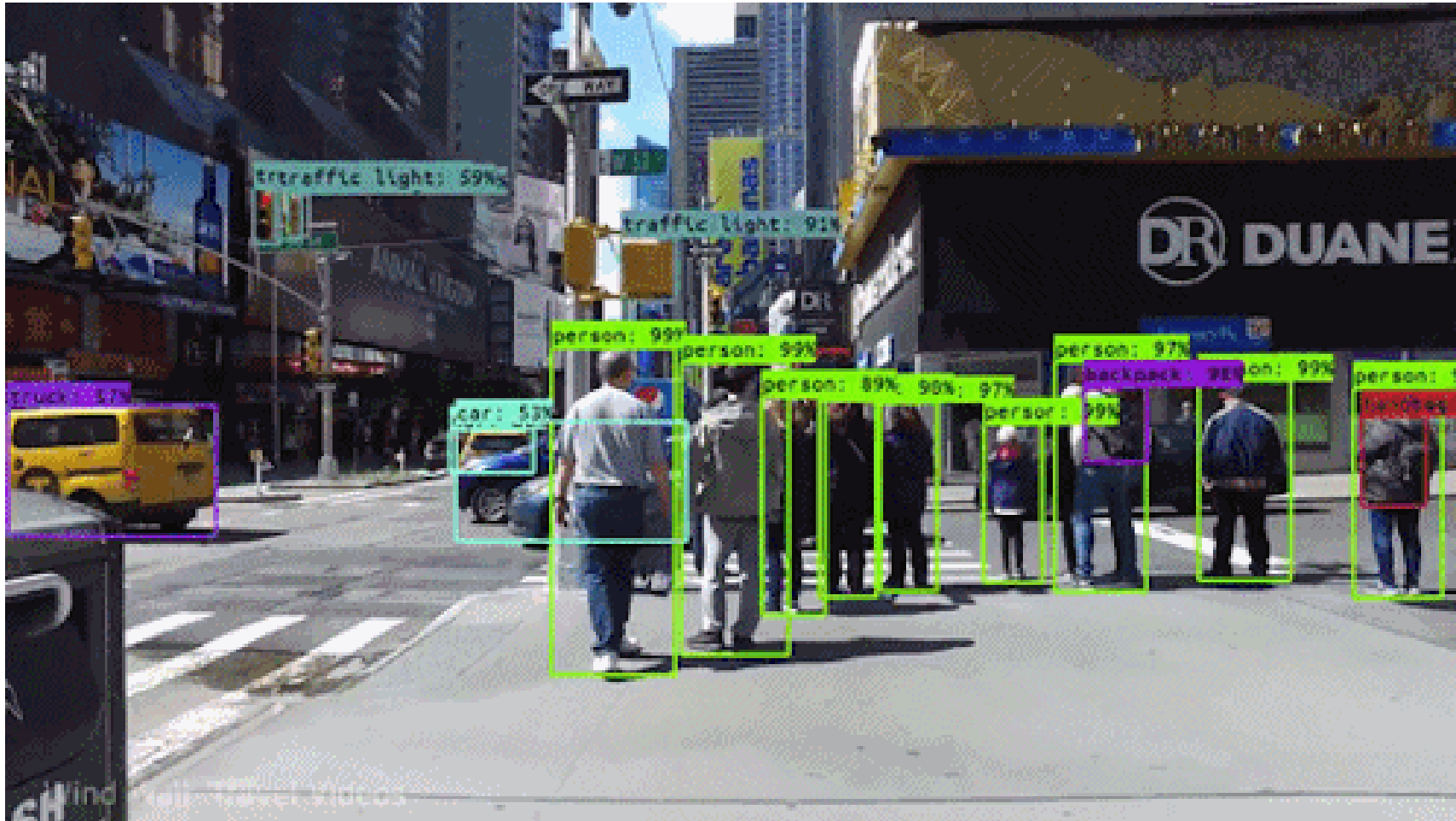
Single Shot Detection (SSD)

You Look Only Once (YOLO)



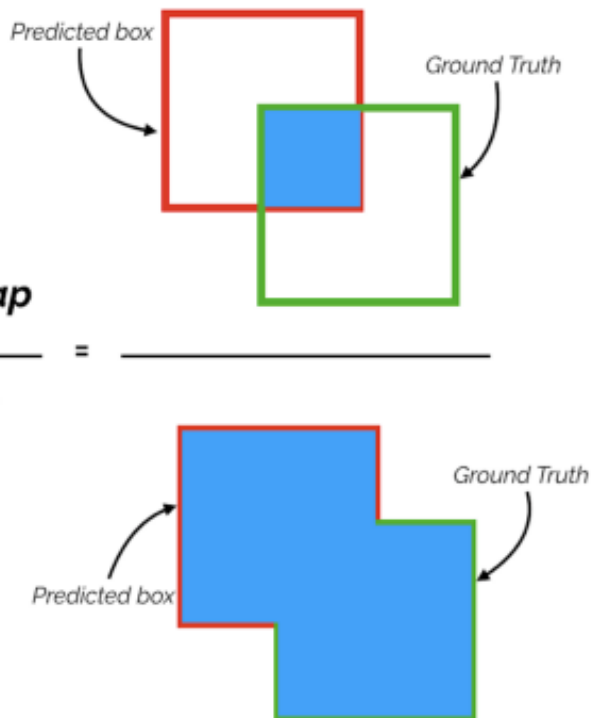
Source: <https://theconversation.com/whos-to-blame-when-a-self-driving-car-has-an-accident-150941>

Real Time Object Detection (2)



Object Detection Metrics (1)

Intersection over Union (IoU) = $\frac{\text{Area of Overlap}}{\text{Area of Union}}$ = $\frac{\text{Area of Intersection}}{\text{Area of Union}}$



The diagram illustrates the Intersection over Union (IoU) metric. It consists of two parts. The top part shows two overlapping rectangles: a red rectangle labeled 'Predicted box' and a green rectangle labeled 'Ground Truth'. The intersection of these two rectangles is shaded blue. The bottom part shows the union of the two rectangles, which is a single blue shape, with the red 'Predicted box' and green 'Ground Truth' outlines still visible. The formula for IoU is given as:
$$\text{Intersection over Union (IoU)} = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$

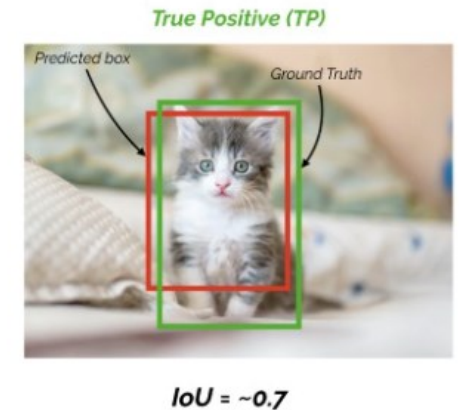
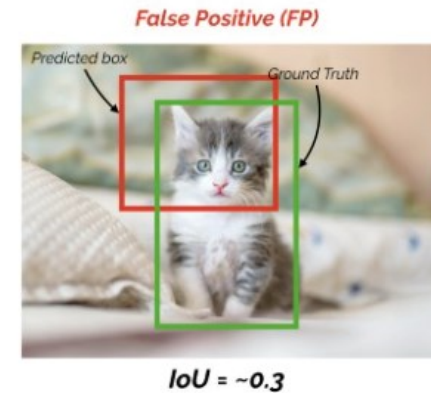
Source: <https://towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2>

Object Detection Metrics (2)

Mean Average Precision (mAP):

- For object detection tasks, we calculate **Precision** and **Recall** using IoU value for a given IoU threshold.
- The general definition for the Average Precision (AP) is finding the area under the precision-recall curve above.
- The mean Average Precision or mAP score is calculated by taking the mean AP over all classes and/or overall IoU thresholds, depending on different detection challenges that exist.
- mAP is usually used as the standard metric for evaluating the performance of object detection models.

If IoU threshold = 0.5



$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$


Source: <https://towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2>

Part 2

DATA

What Kind of Data Do We Need?

images	Folder	--
train	Folder	--
im1.jpg	JPEG image	63 KB
im2.jpg	JPEG image	90 KB
im3.jpg	JPEG image	30 KB
val	Folder	--
im4.jpg	JPEG image	105 KB
im5.jpg	JPEG image	57 KB
im6.jpg	JPEG image	67 KB
labels	Folder	--
train	Folder	--
im1.txt	Plain Text	308 bytes
im2.txt	Plain Text	78 bytes
im3.txt	Plain Text	72 bytes
val	Folder	--
im4.txt	Plain Text	39 bytes
im5.txt	Plain Text	35 bytes
im6.txt	Plain Text	77 bytes
README.txt	Plain Text	190 bytes
coco6.yaml	YAML Document	2 KB






im2.jpg
JPEG image - 90 KB

Information [Show Less](#)

Created **Saturday, 10 July 2021 at 16:29**

Modified **Saturday, 10 July 2021 at 16:29**

 Rotate Left  Markup  More...

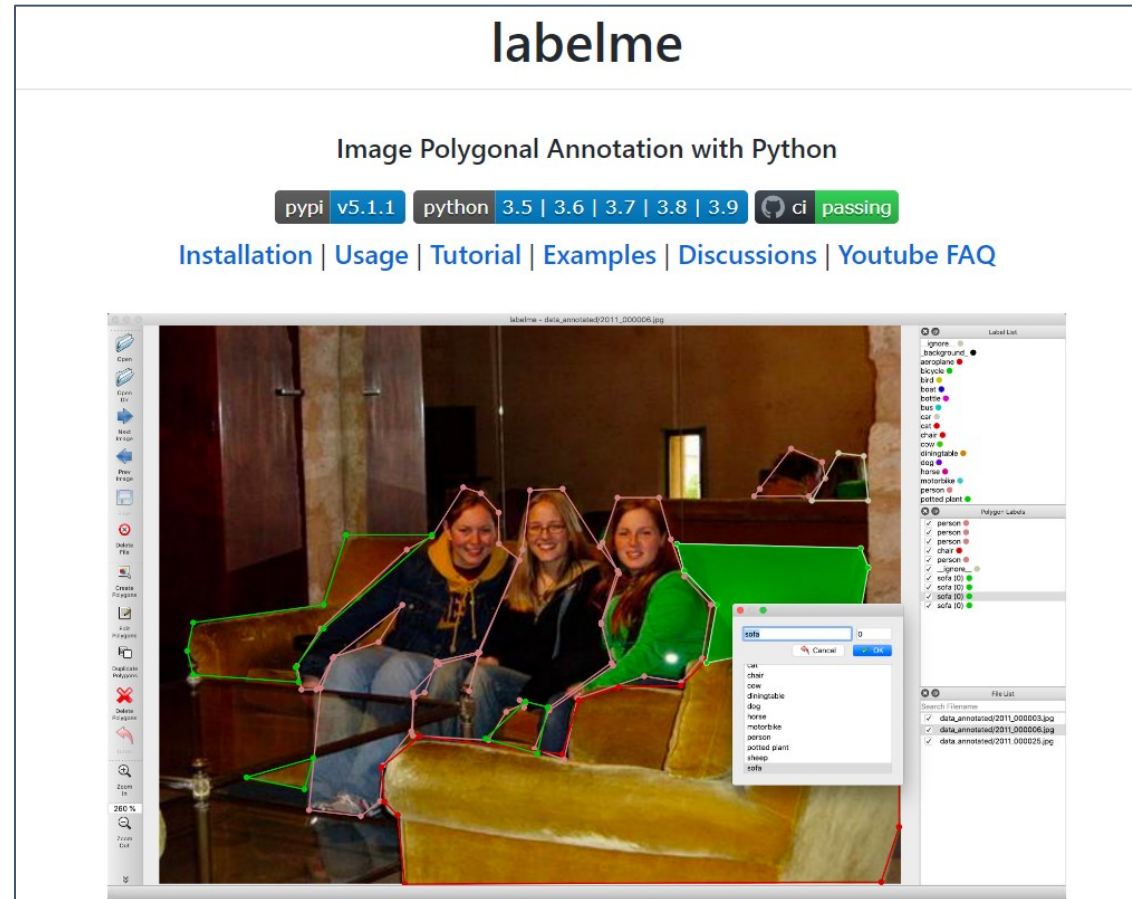
Im1.txt

```
1 120 156 456 519
2 89 120 311 421
```

```
....
....
```

C Xmin Ymin Xmax Ymax
C Xcenter Ycenter Width Height

How to Label Custom Datasets? (1)



<https://github.com/wkentaro/labelme>

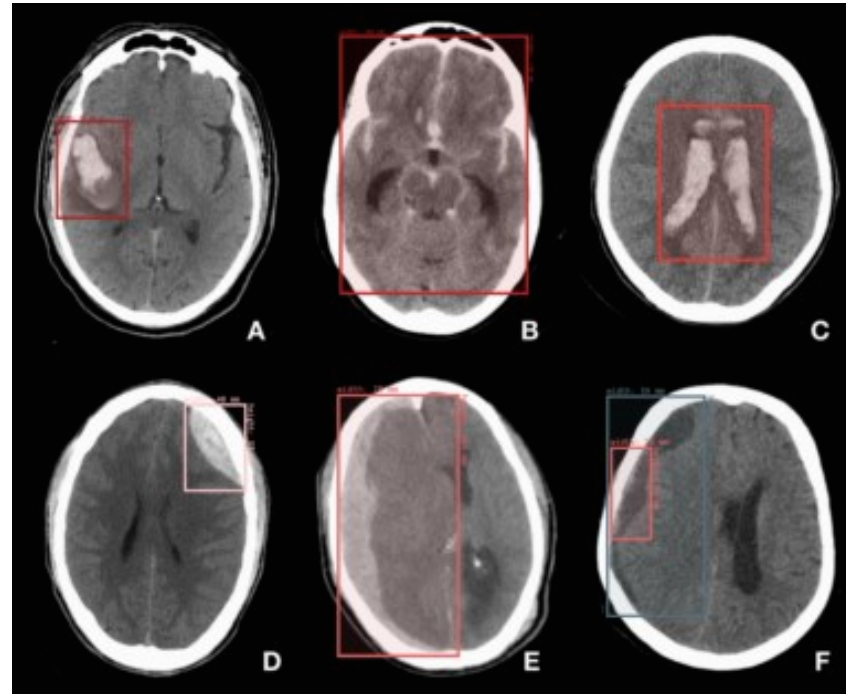
How to Label Custom Datasets? (2)



A Nice Tutorial on How to Use labelme:

<https://www.youtube.com/watch?v=ydHI8SUE58Y>

Let's Prepare the Data for Our Workshop!

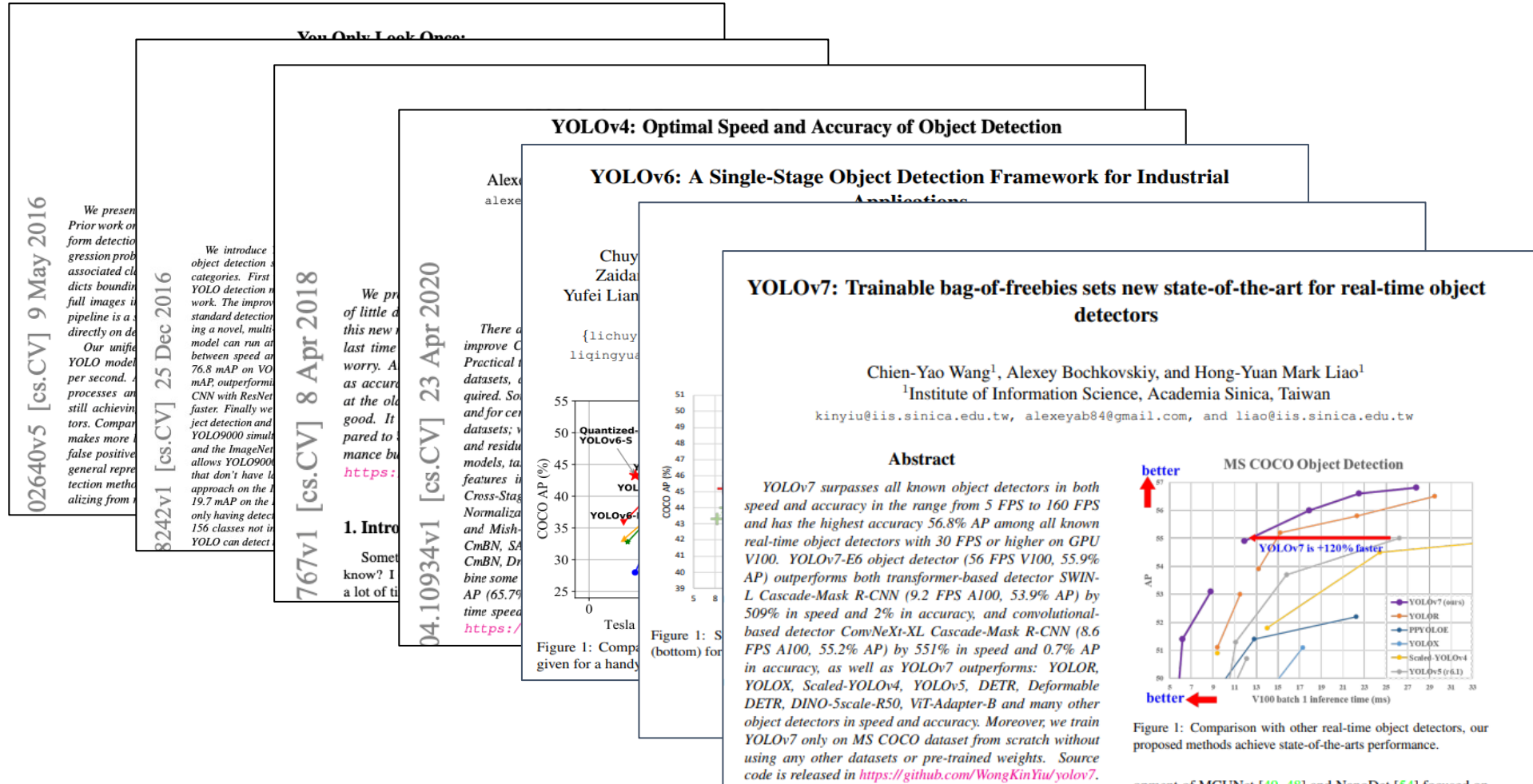


In this workshop, we will train an object detection
deep learning model to detect
brain hemorrhage lesion on Head CT scans!

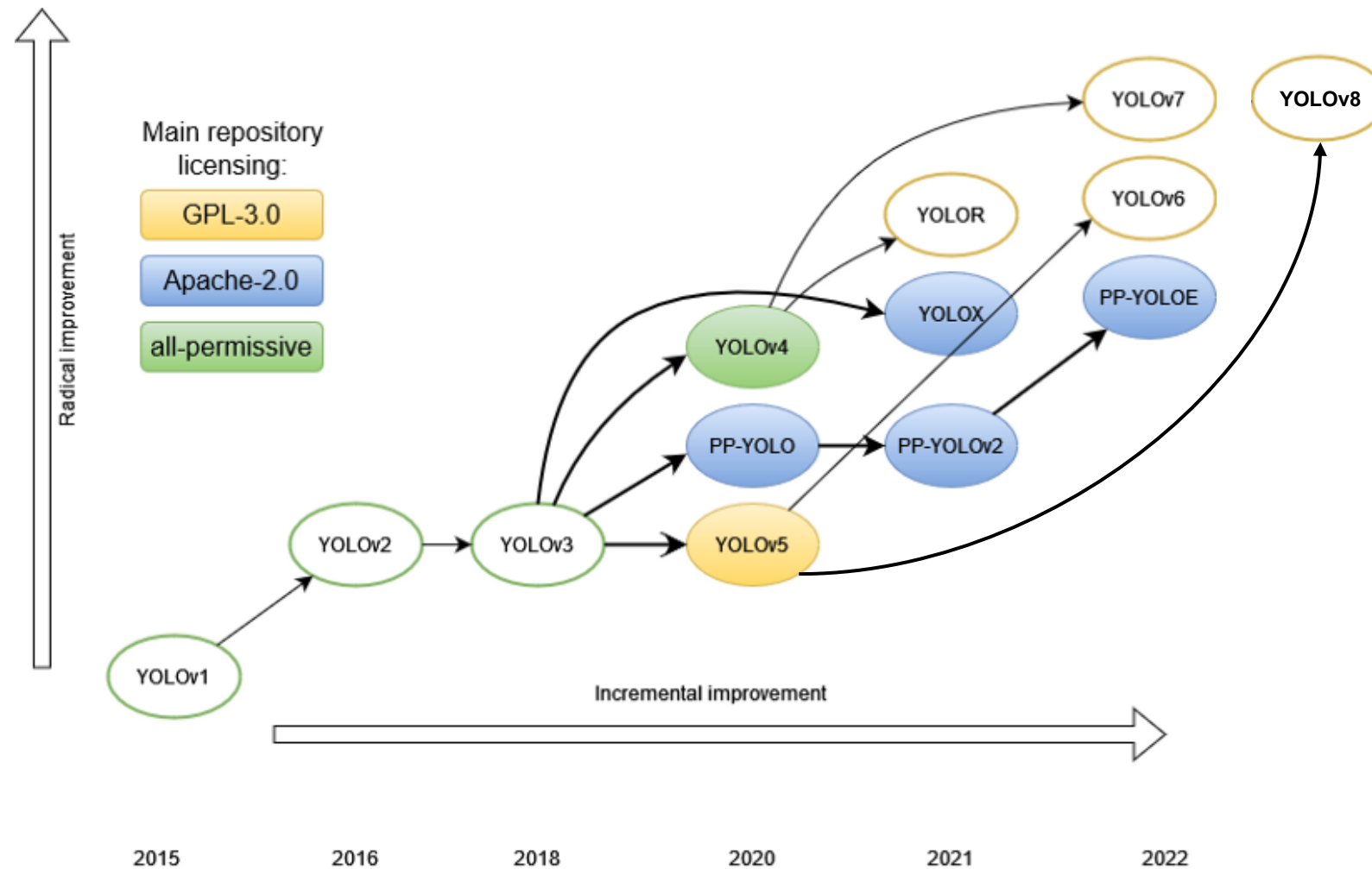
Part 3

YOLO

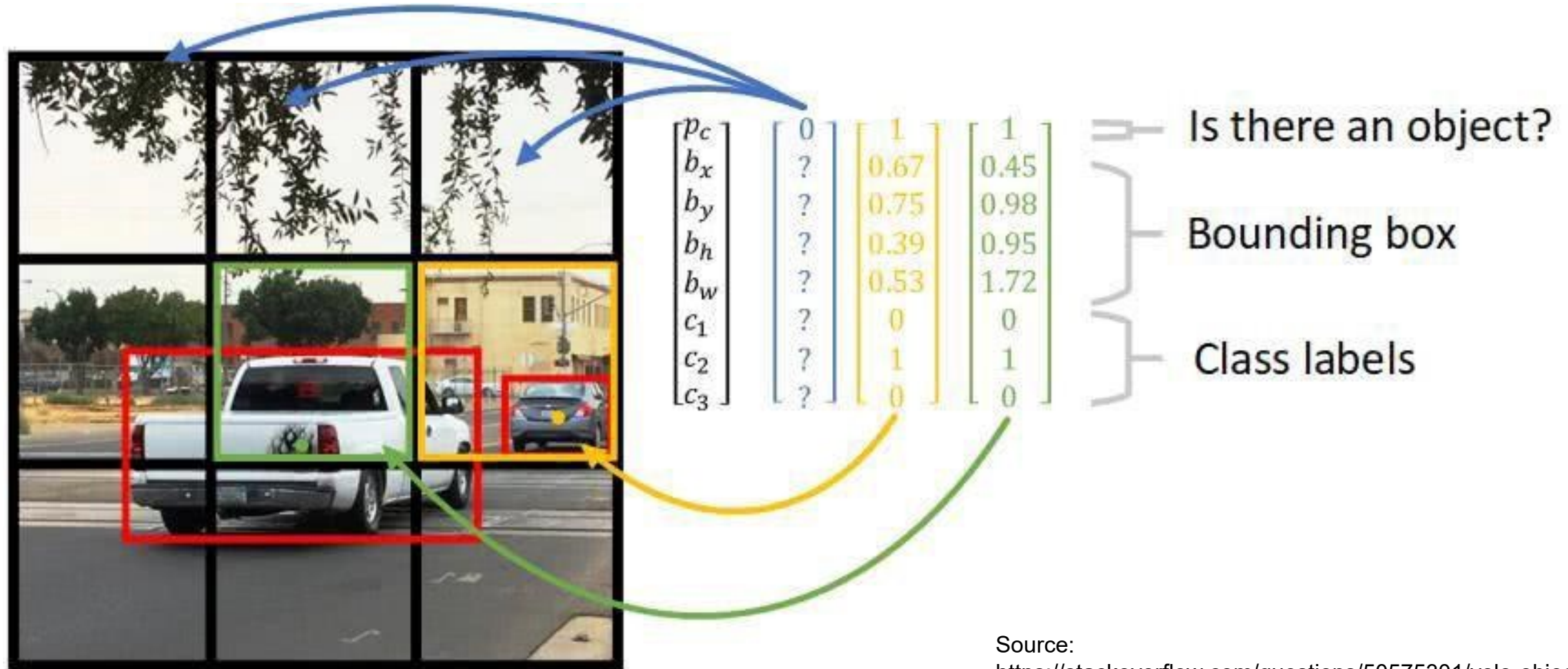
You Look Only Once (YOLO)



You Look Only Once (YOLO)



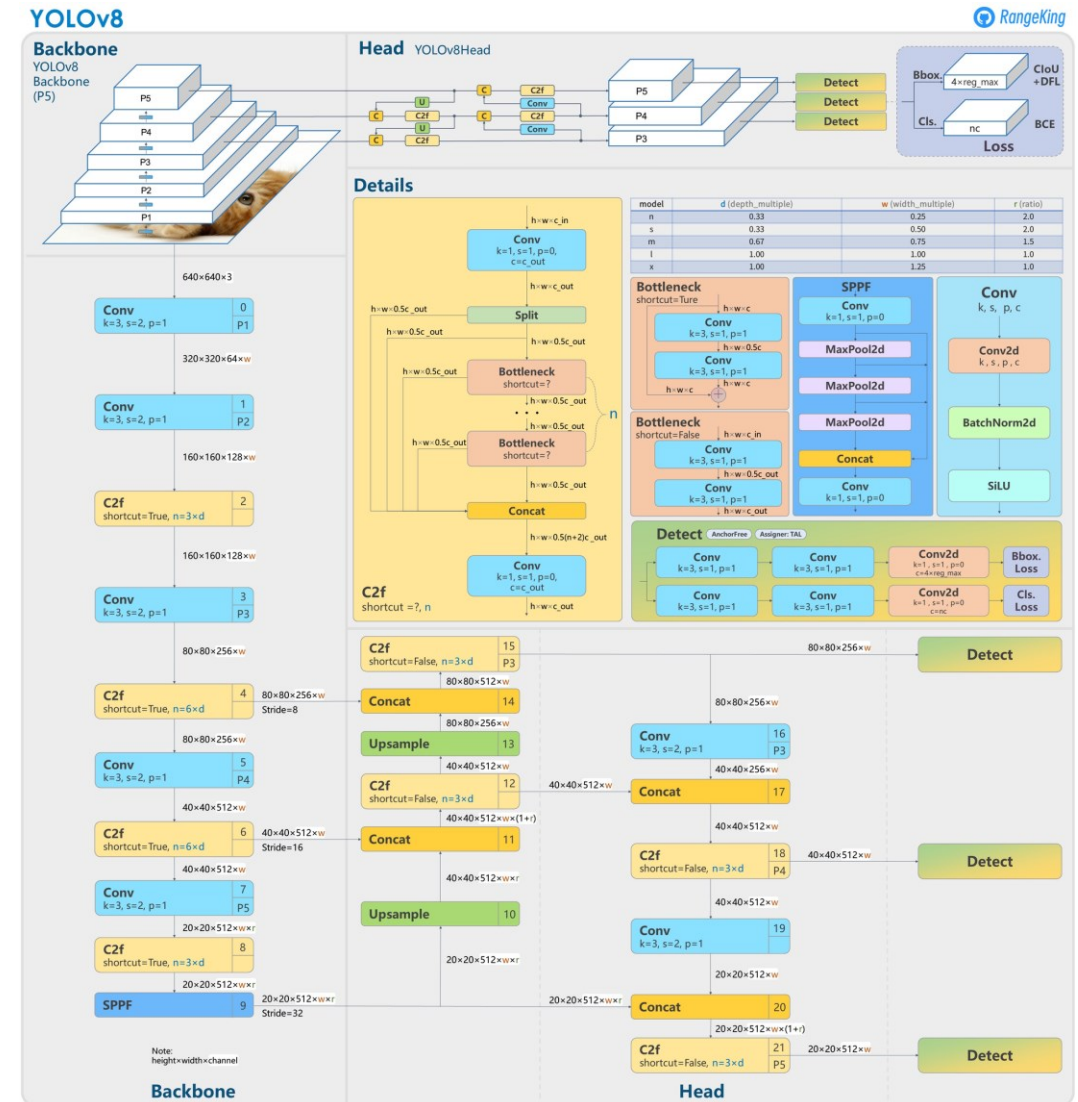
What Does YOLO Output?



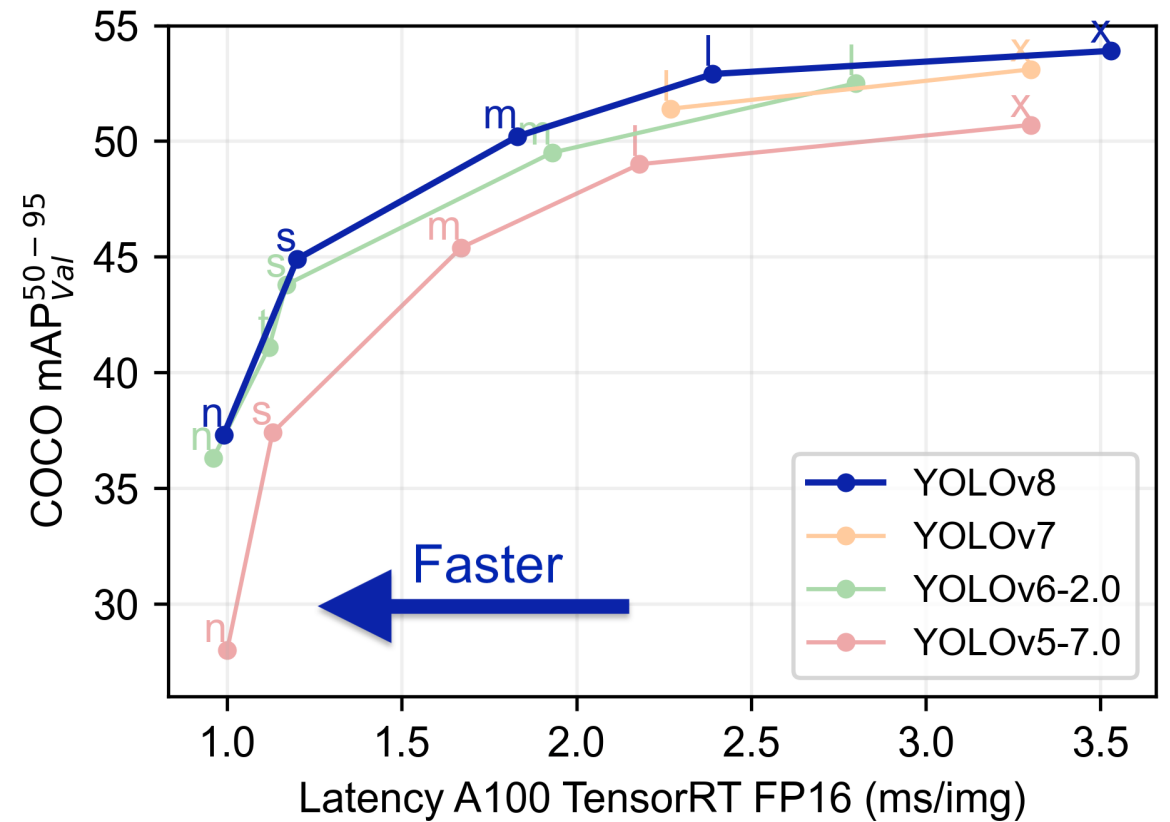
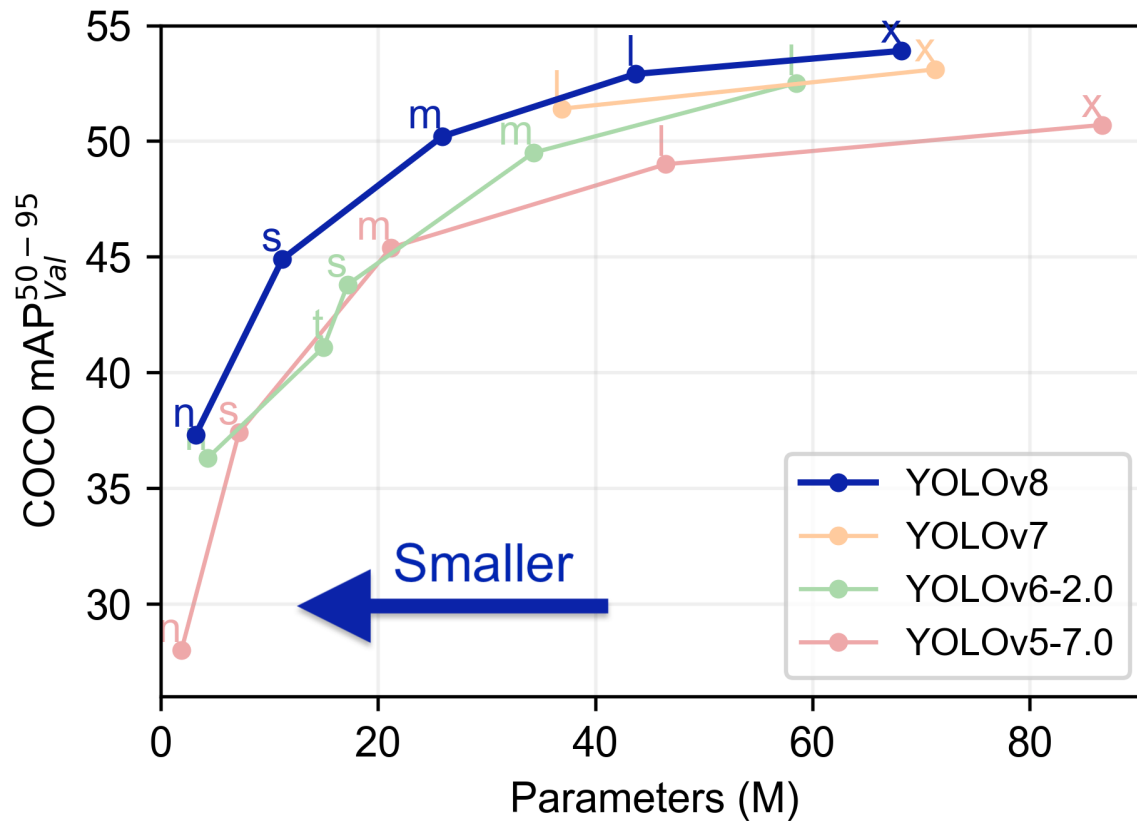
Source:
<https://stackoverflow.com/questions/50575301/yolo-object-detection-how-does-the-algorithm-predict-bounding-boxes-larger-than>

How Does YOLO Work?

Modern YOLO architectures leverage complicated data handling pipelines, model architectures, and training schemas.



YOLOv8: State-of-the-art YOLO Model



How to use a YOLO model without coding?

You can use **Ultralytics hub** to train (fine-tune) a custom YOLOv8 model on your own data:

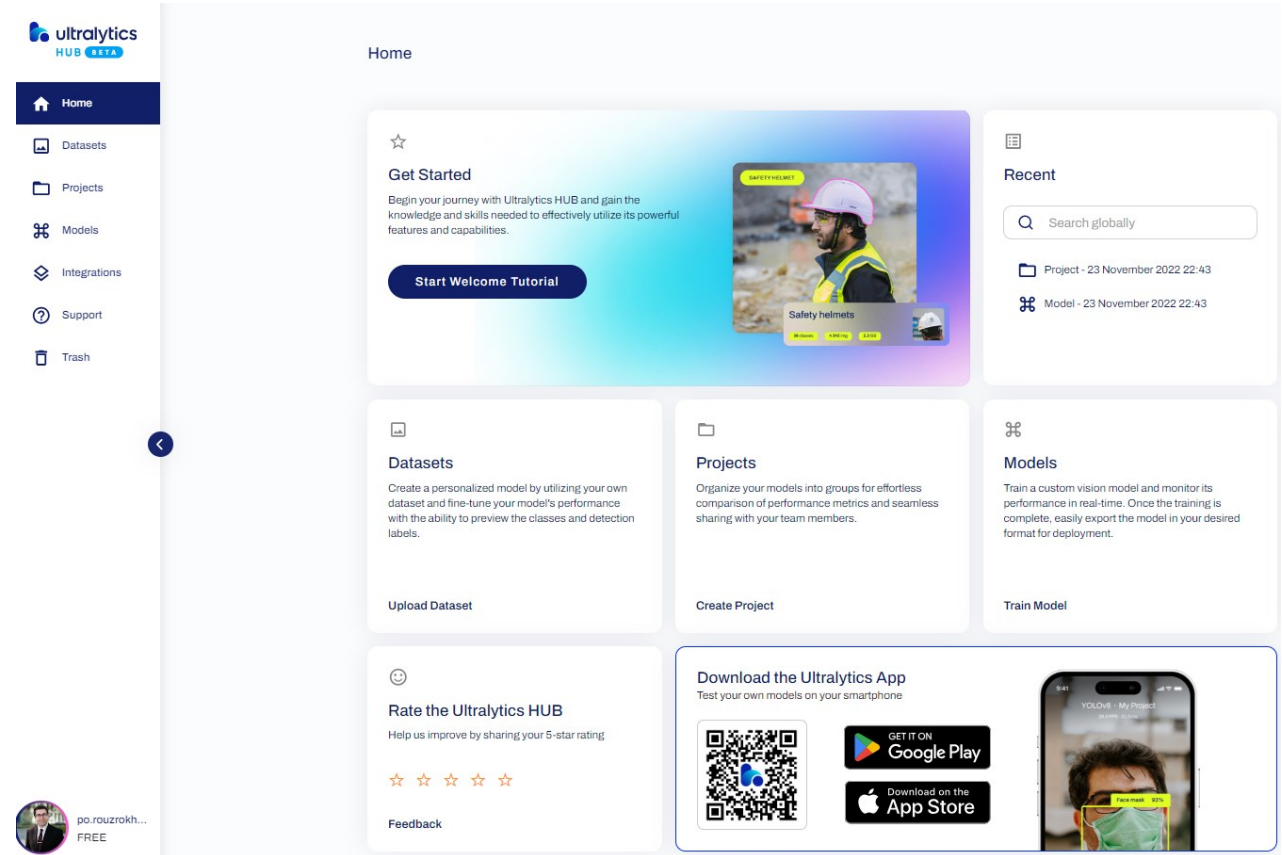
<https://hub.ultralytics.com/home>

Advantages:

- No coding required.
- Easy use cases.

Disadvantages:

- Less flexibility in deployment
- Data privacy issues



How to use a YOLO model with coding?

This is what we will learn today! But before that...

Do you have any questions?

Thank you for your attention!
Please open the Google Colab notebook Prepared for this workshop!