RSNA® Radiological Society of North America

MAYO CLINIC

YOU LOOK ONLY ONCE (YOLO)

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

Bradley Erickson

MD PhD

Pouria Rouzrokh

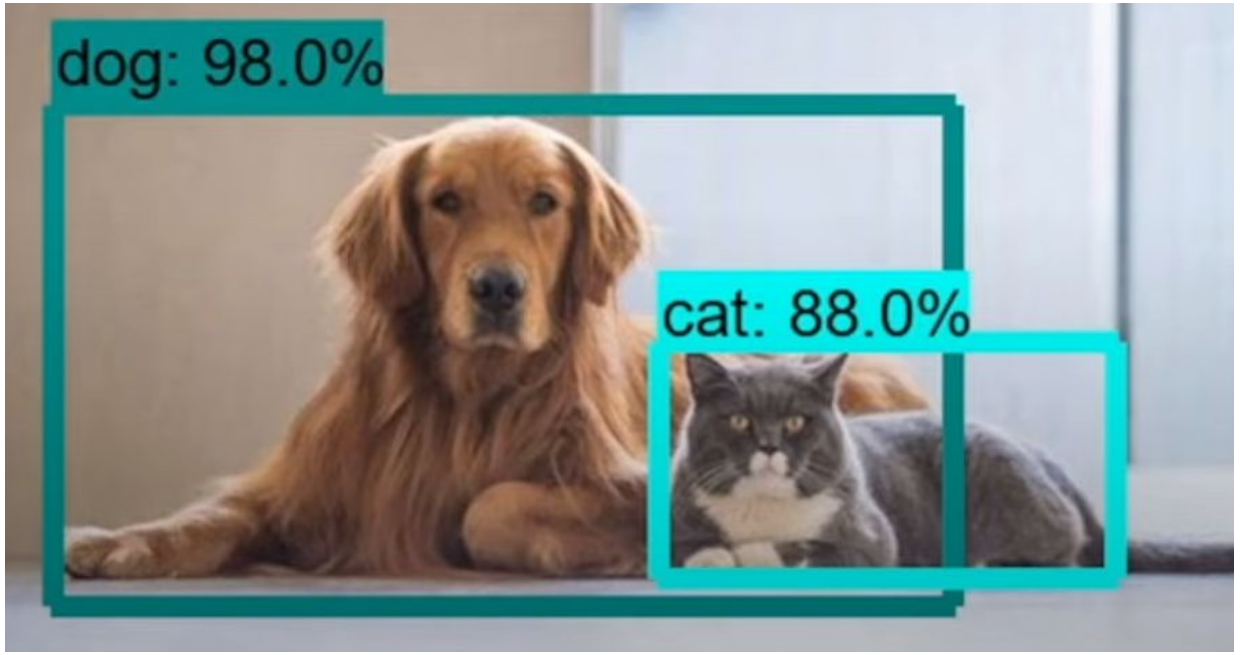
MD MPH MHPE

Radiology Informatics Laboratory
Mayo Clinic

Part 1

INTRODUCTION

Defining Object Detection (1)

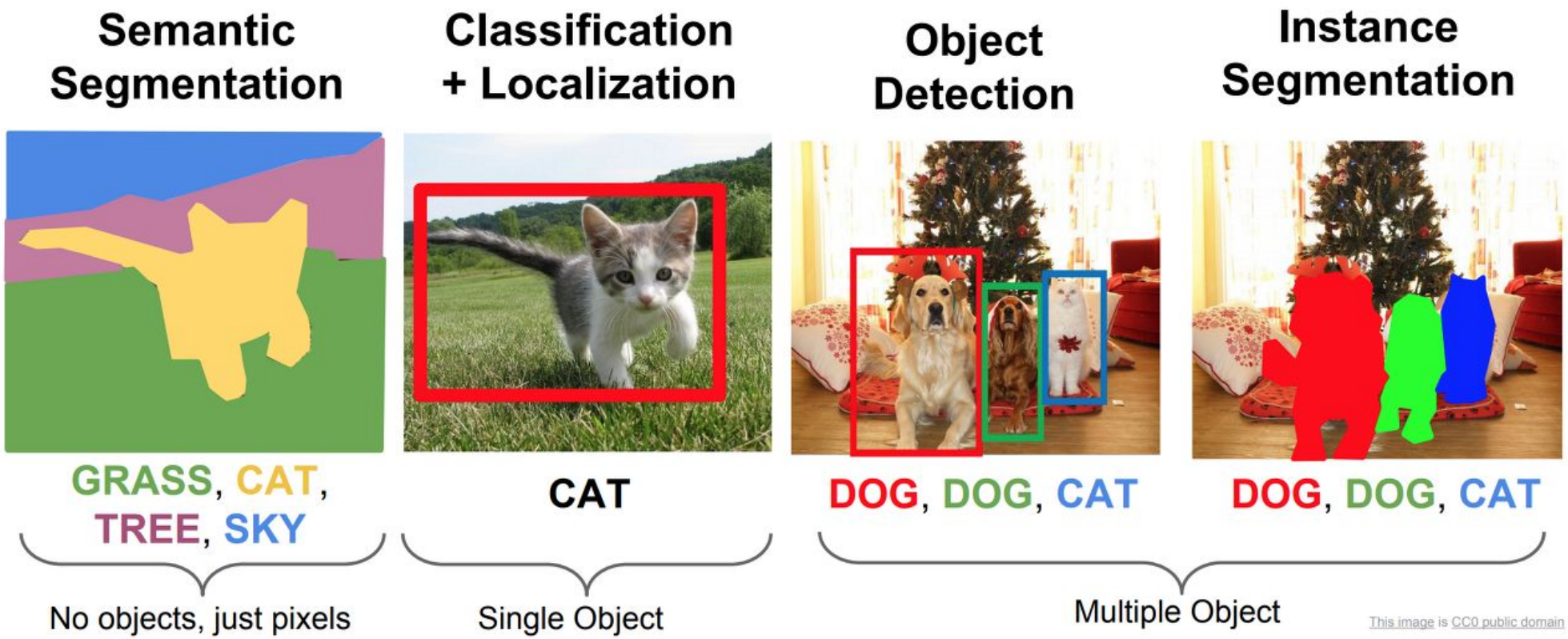


Object Detection:

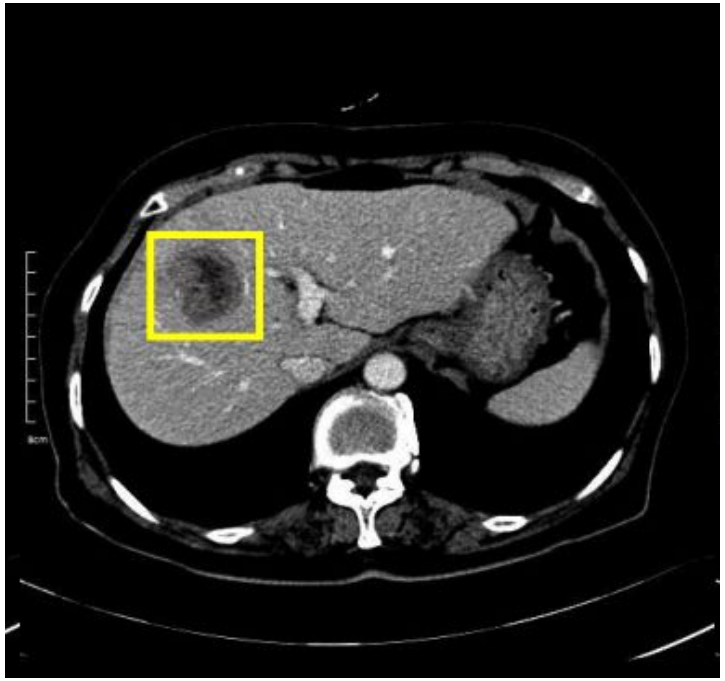
1. **What kind of object is it?**
2. **Where is that object located?**

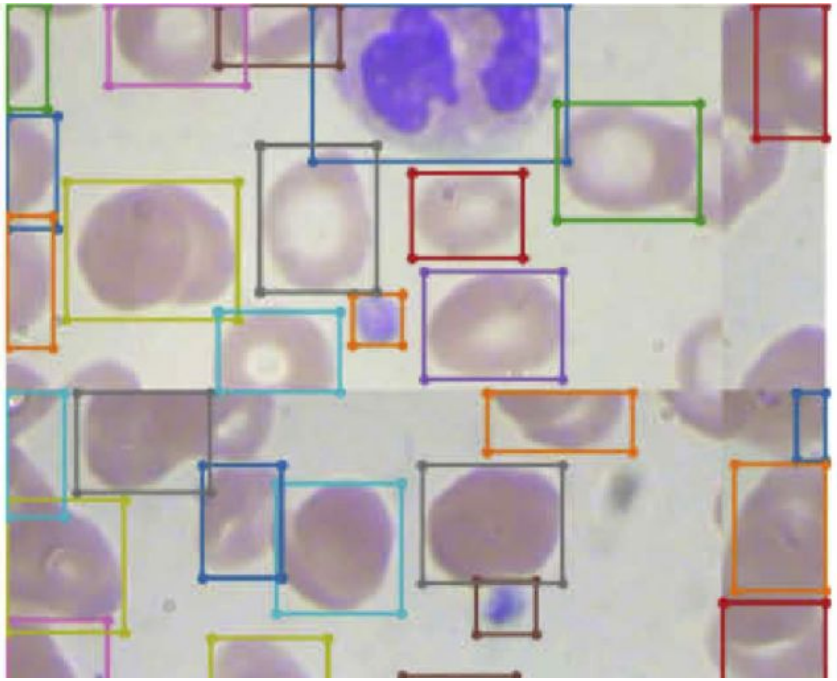
Defining Object Detection (2)

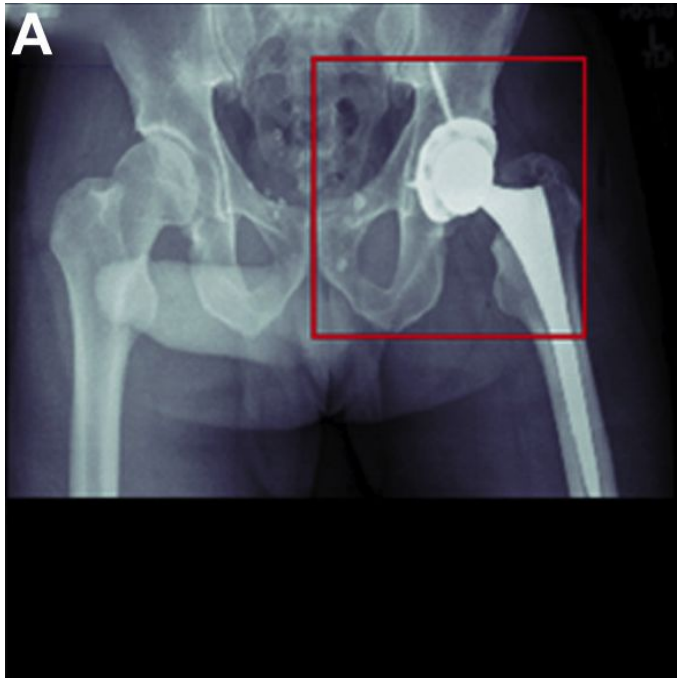
Defining Object Detection (3)



Medical Applications of Object Detection







Sources:

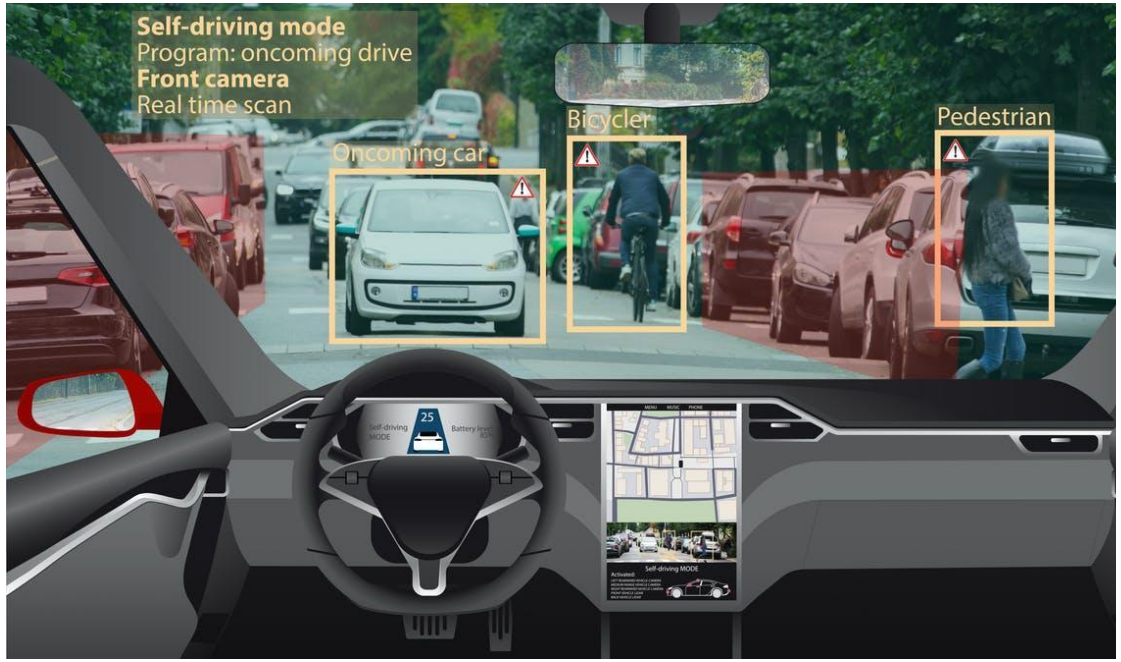
- Liver Lesion Detection from Weakly-labeled Multi-phase CT Volumes with a Grouped Single Shot MultiBox Detector
- Improved detection performance in blood cell count by an attention-guided deep learning method
- Deep Learning Artificial Intelligence Model for Assessment of Hip Dislocation Risk Following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

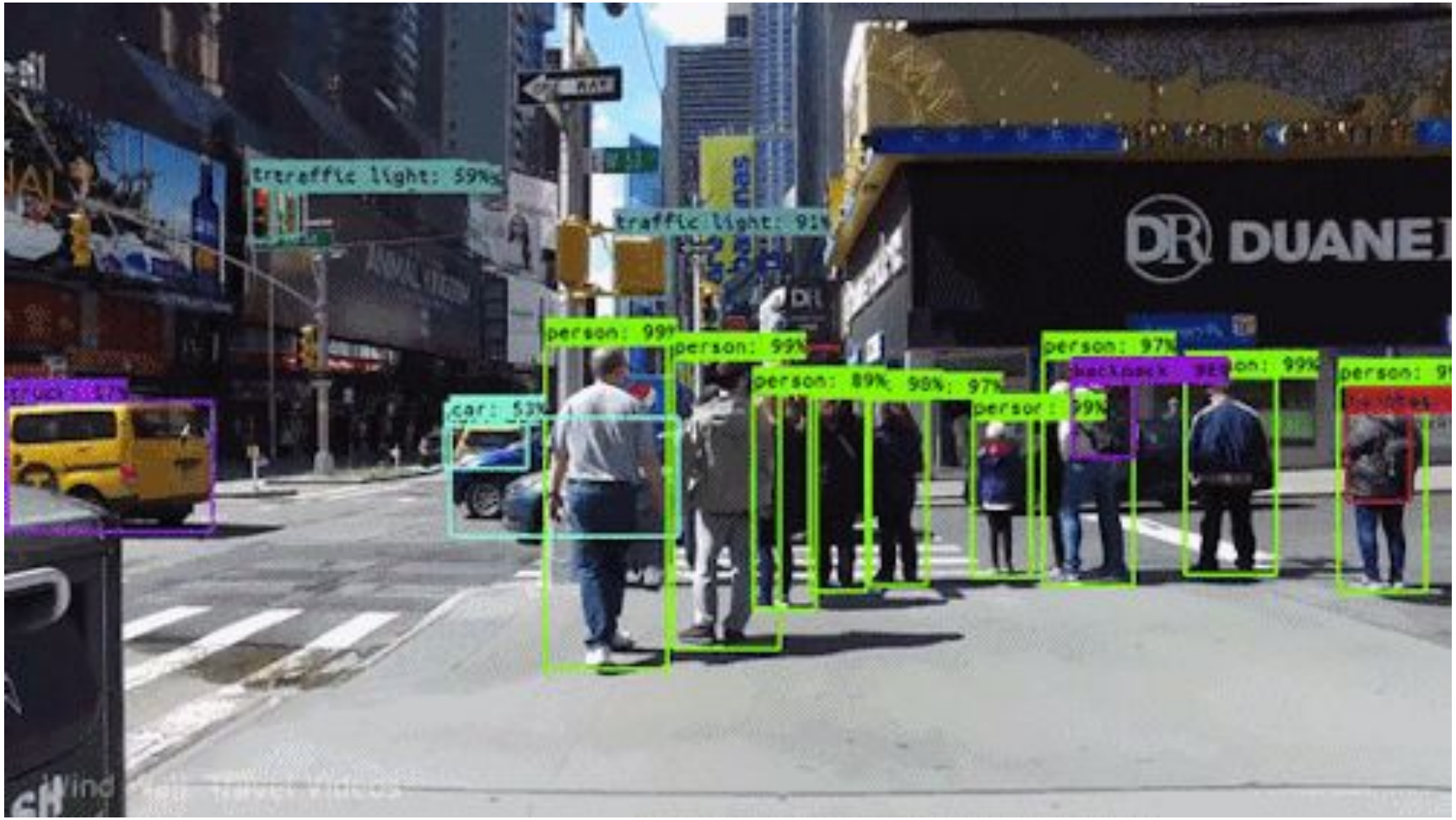
The model should be able to detect objects and make inferences within microseconds!

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)

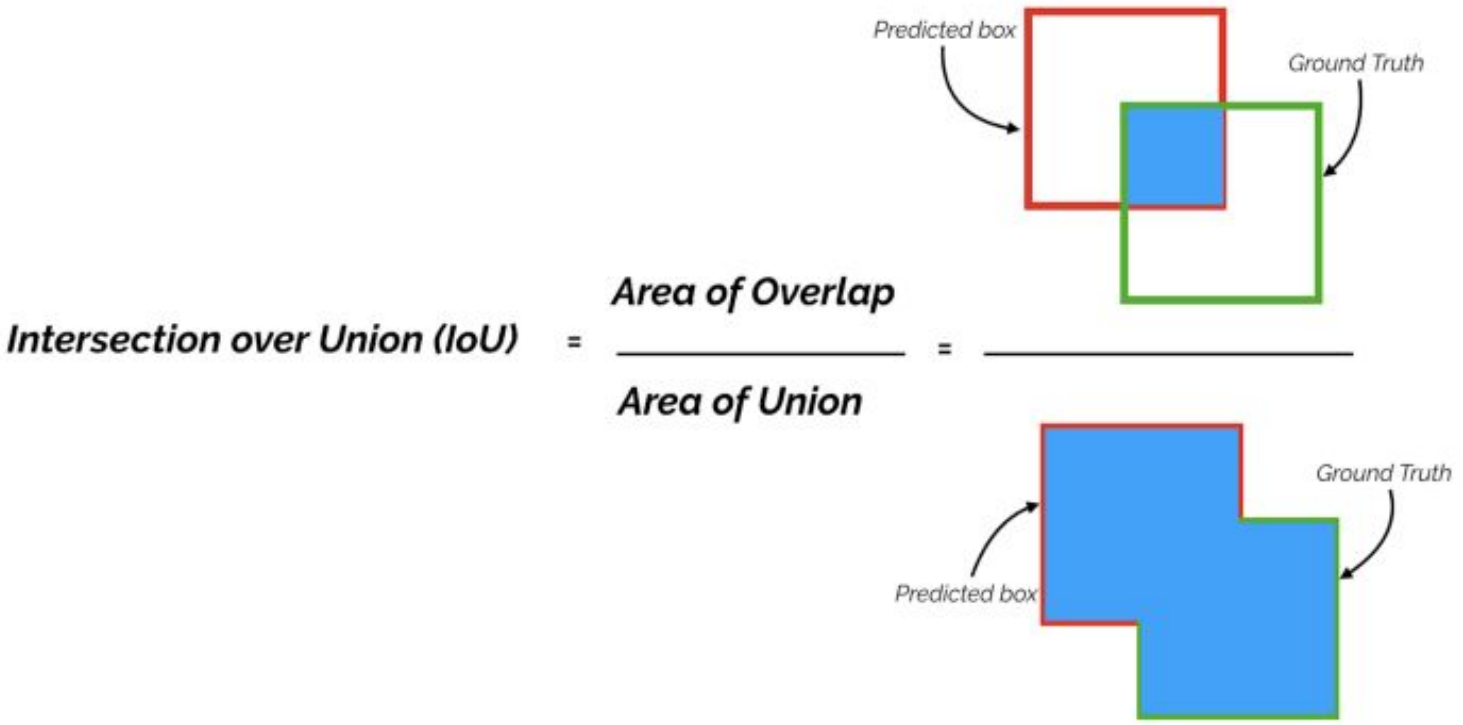


Source: https://theconversation.com/whos-to-blame-when-a-self-driving-car-has-an-accident-150941

Real Time Object Detection (2)



Object Detection Metrics (1)

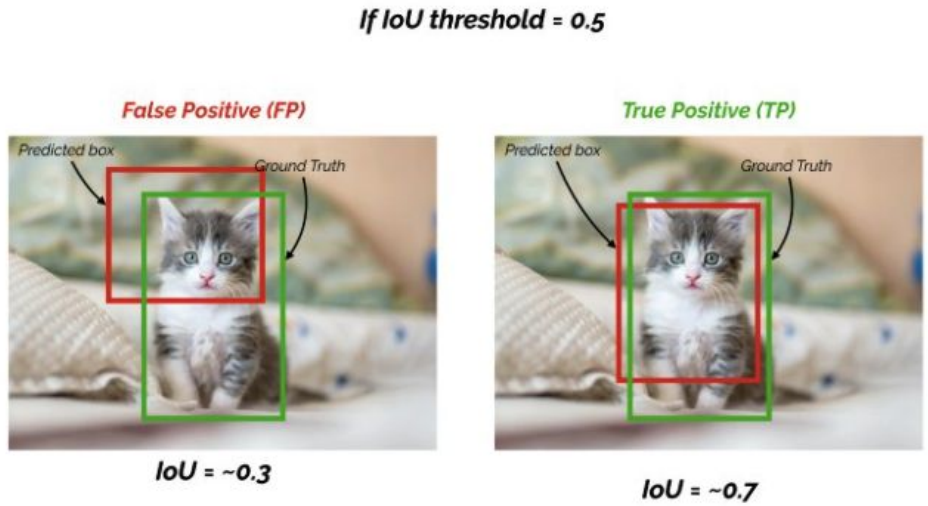


Source: https://towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2

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.



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

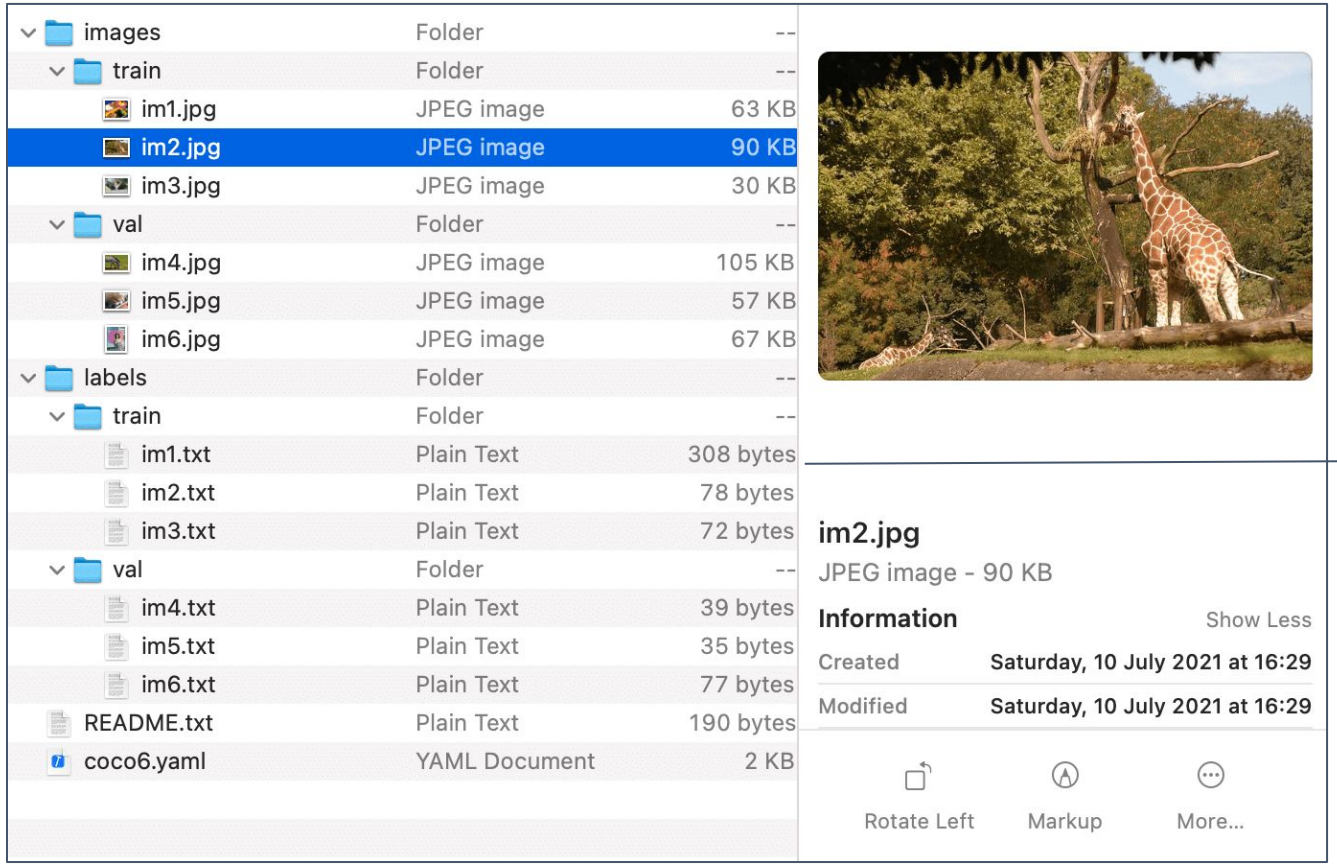
$$Precision = \frac{TP}{TP + FP}$$

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

Part 2

DATA

What Kind of Data Do We Need?



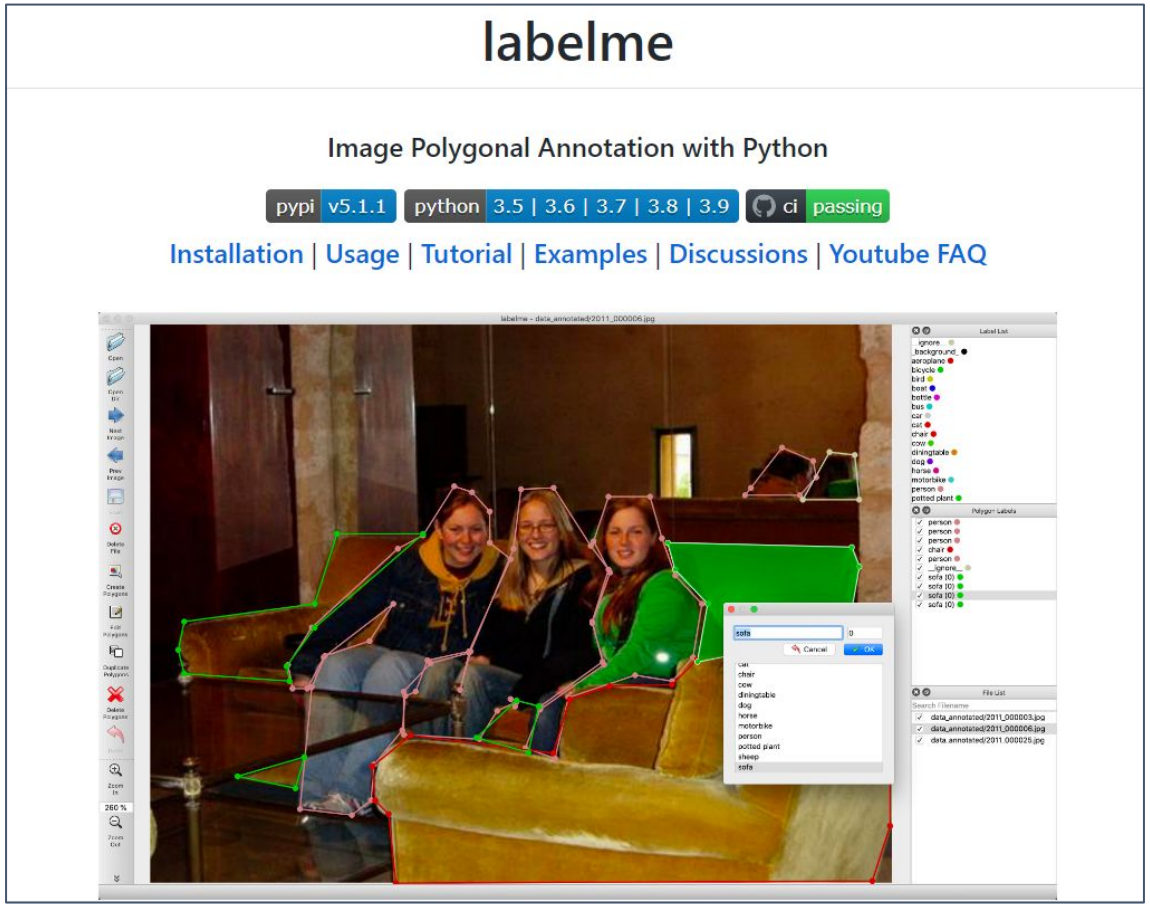
Im1.txt

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)



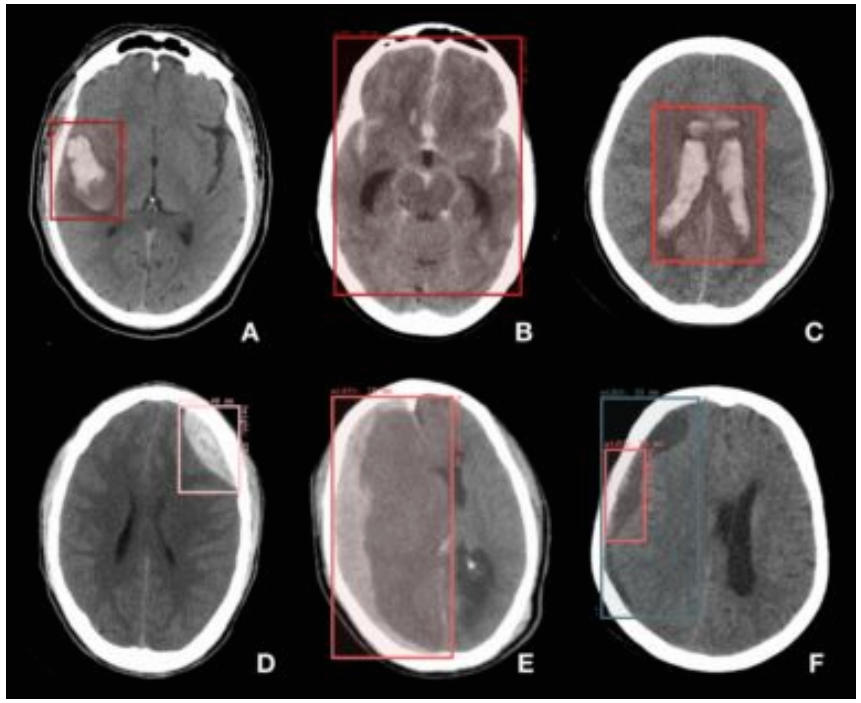
https://github.com/wkentaro/labelme

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!

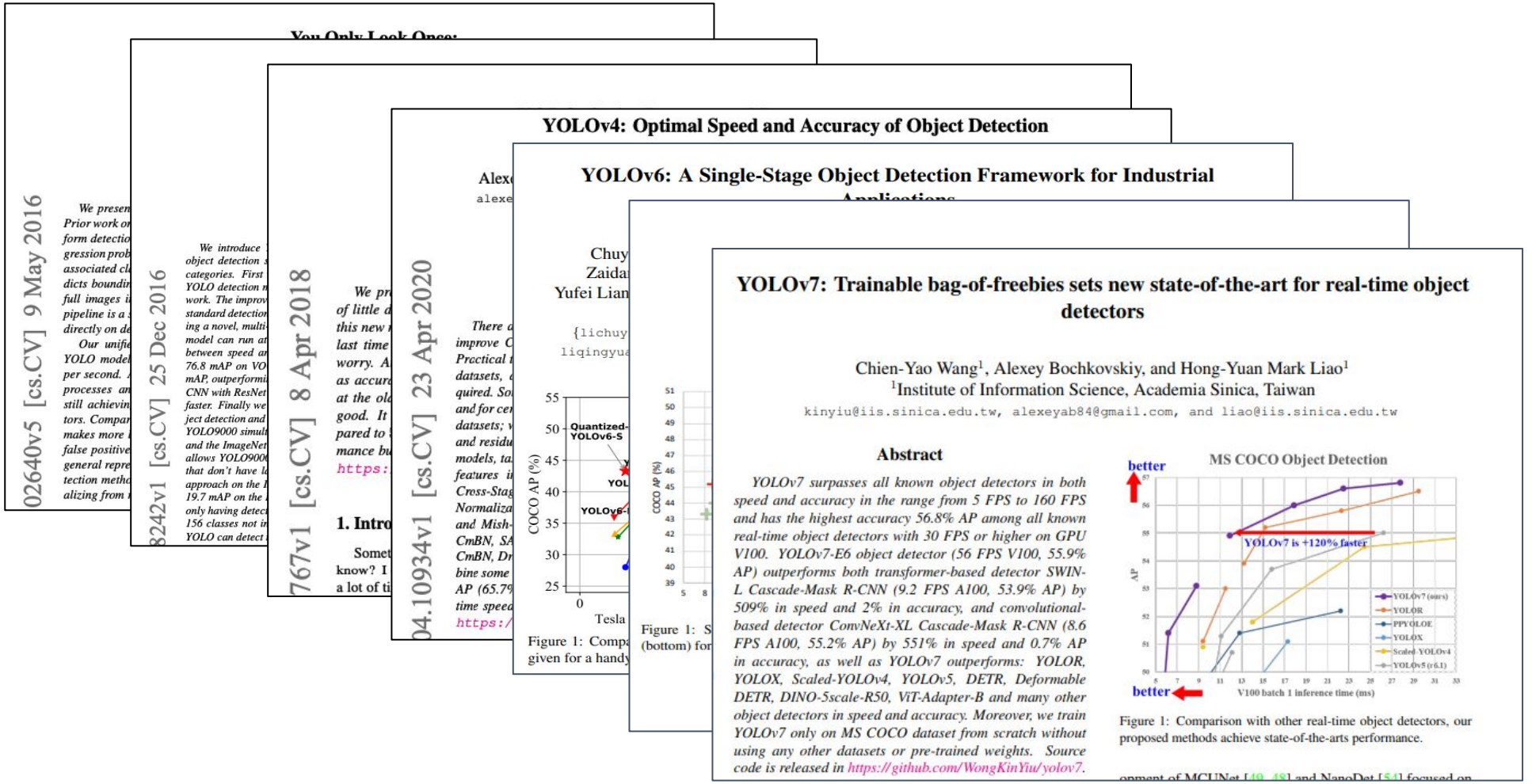


In this workshop, we will train an object detection deep learning model to detect
brain hemorrhage lesion on Head CT scans!

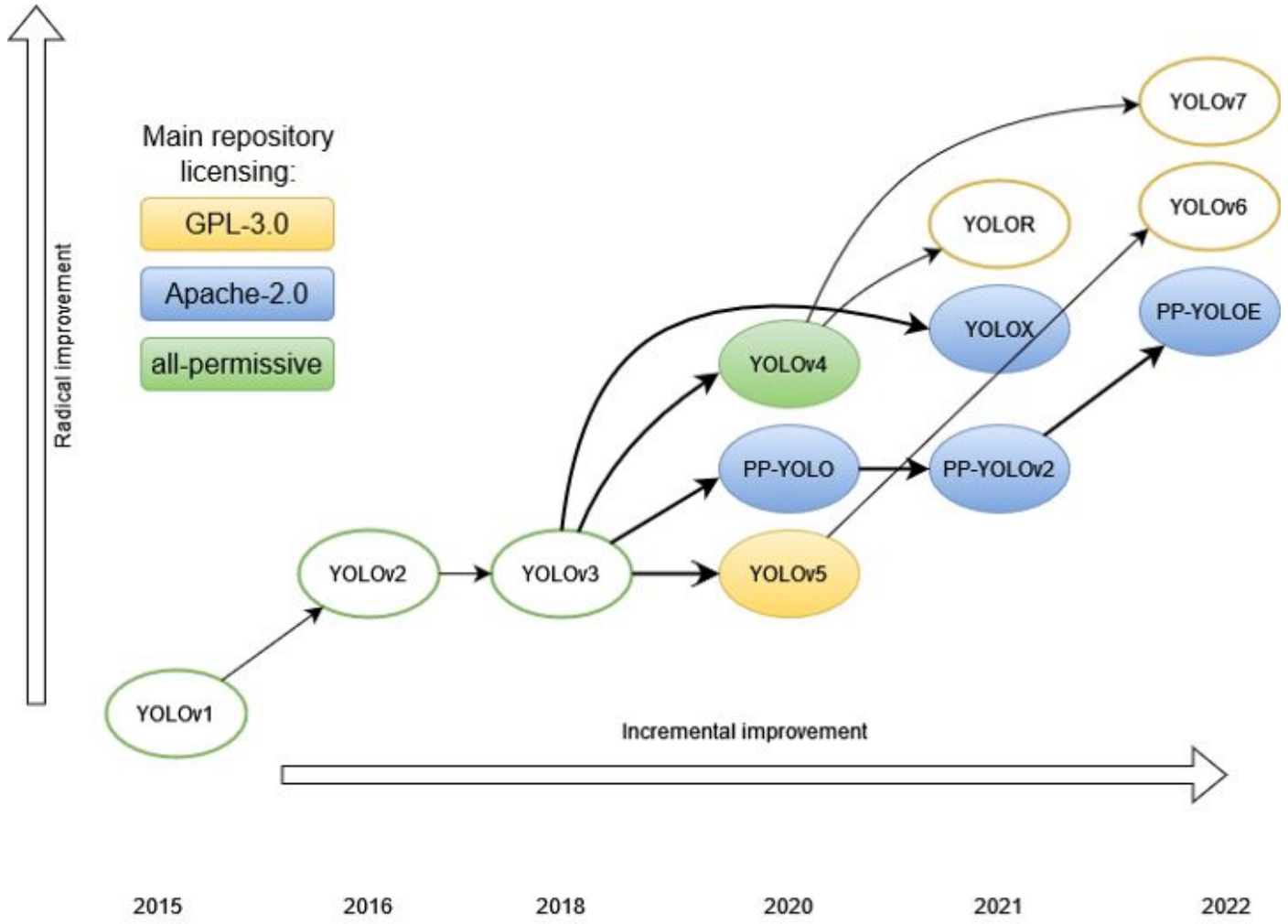
Part 3

YOLO

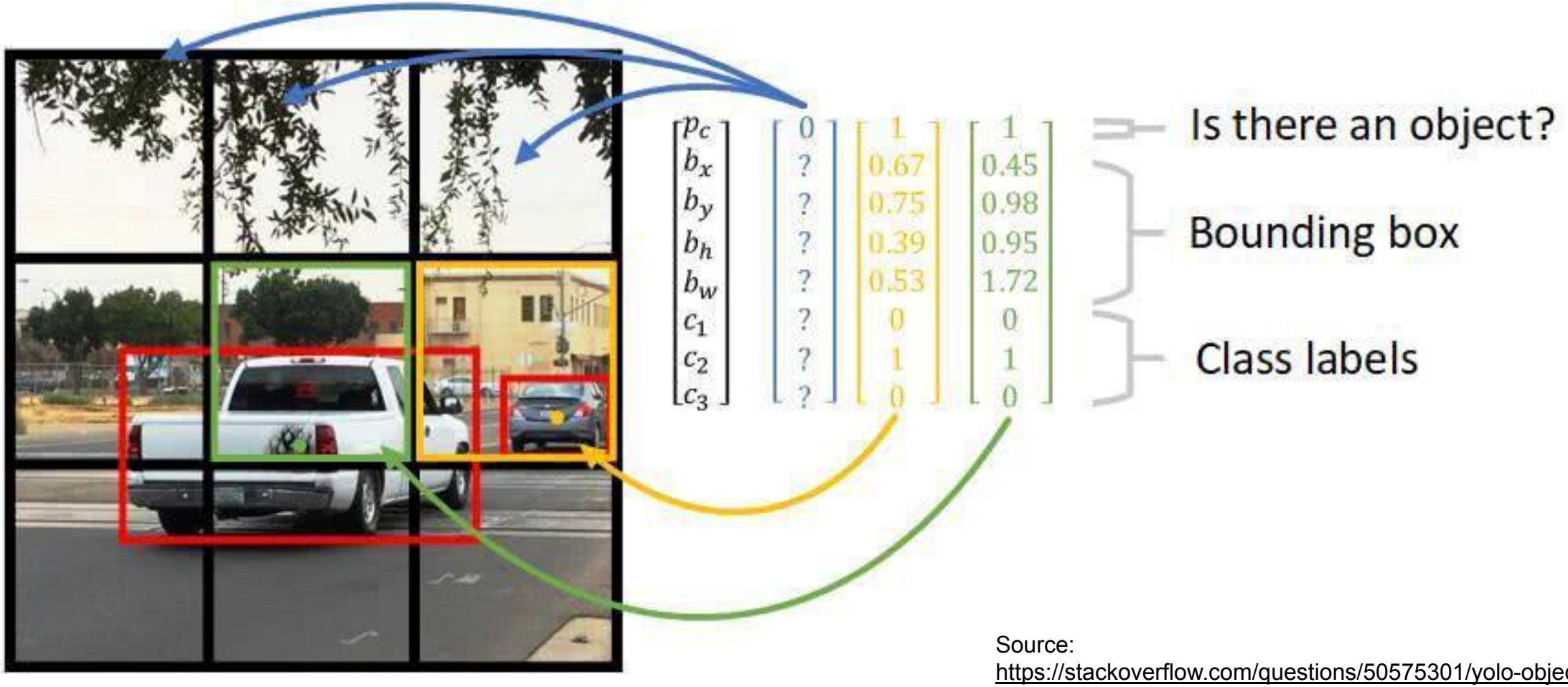
You Look Only Once (YOLO)



You Look Only Once (YOLO)

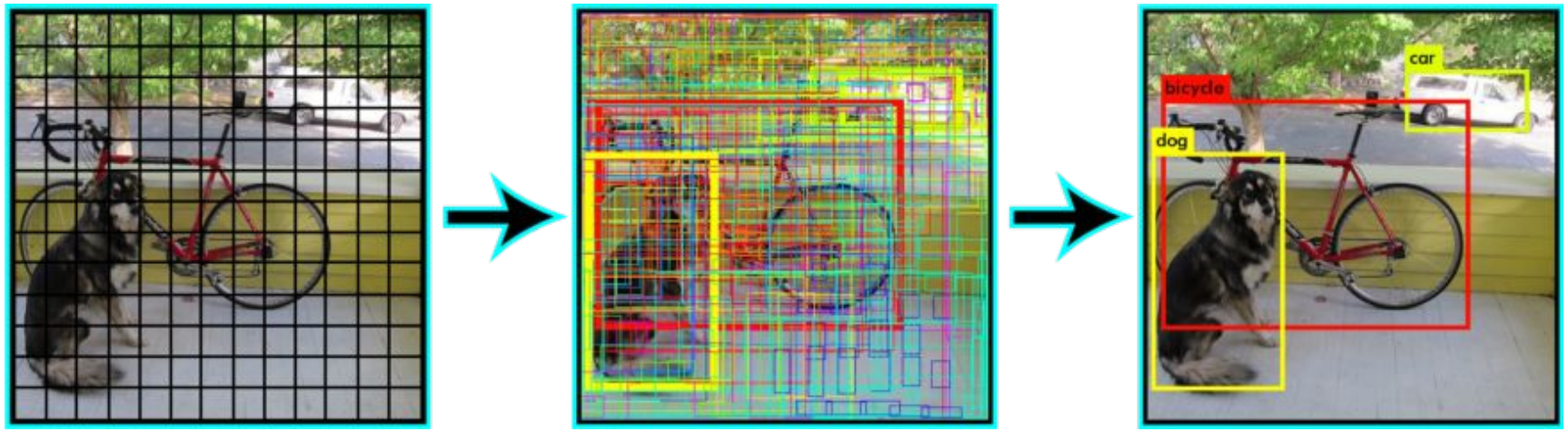


How YOLO works (1)



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

How YOLO works (2)

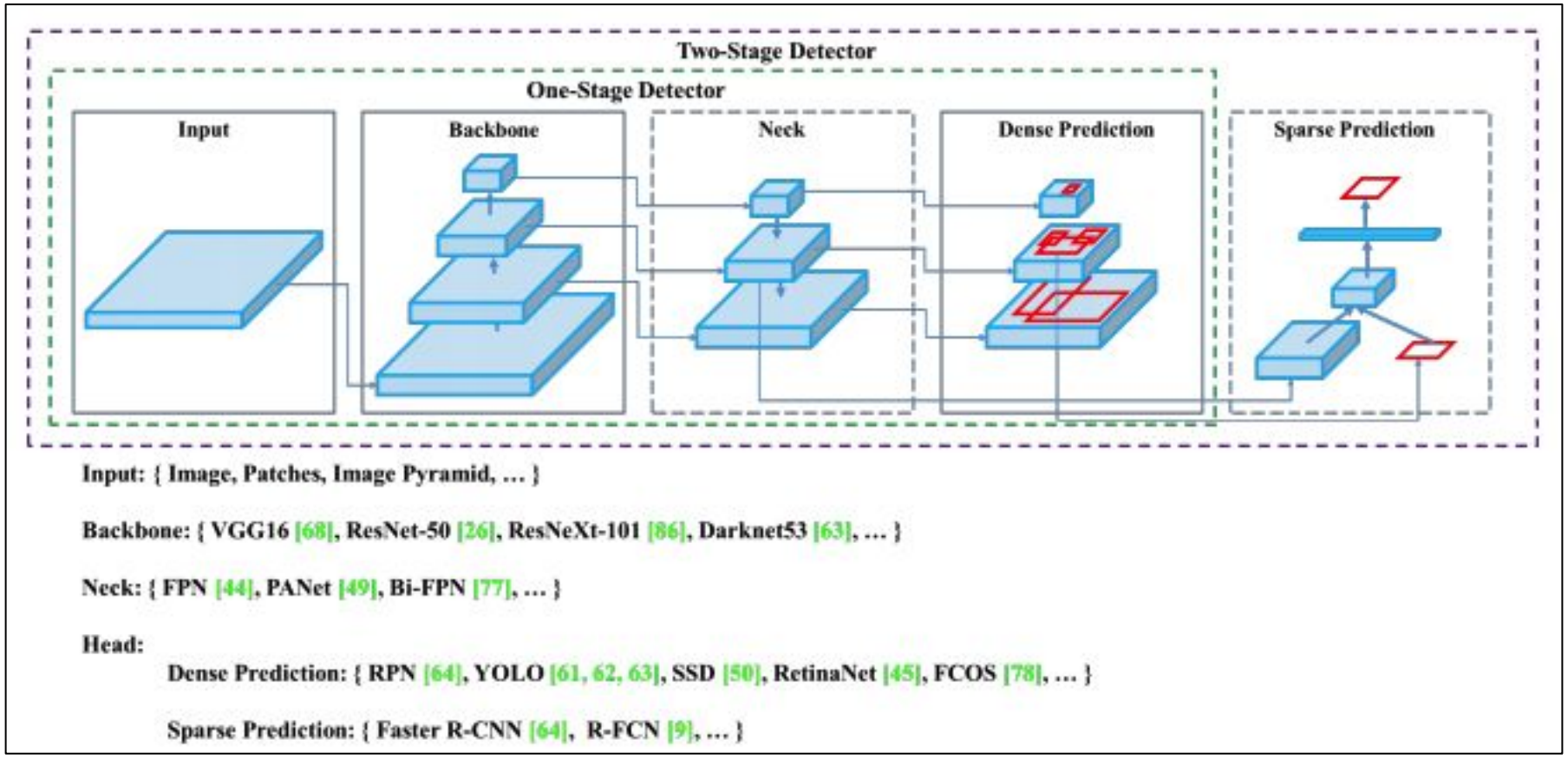


Non-Max Suppression:

To select the best bounding box, from the multiple predicted bounding boxes, these object detection algorithms use non-max suppression. This technique is used to “suppress” the less likely bounding boxes and keep only the best one.

Source: https://www.pyimagesearch.com/2018/11/12/yolo-object-detection-with-opencv/

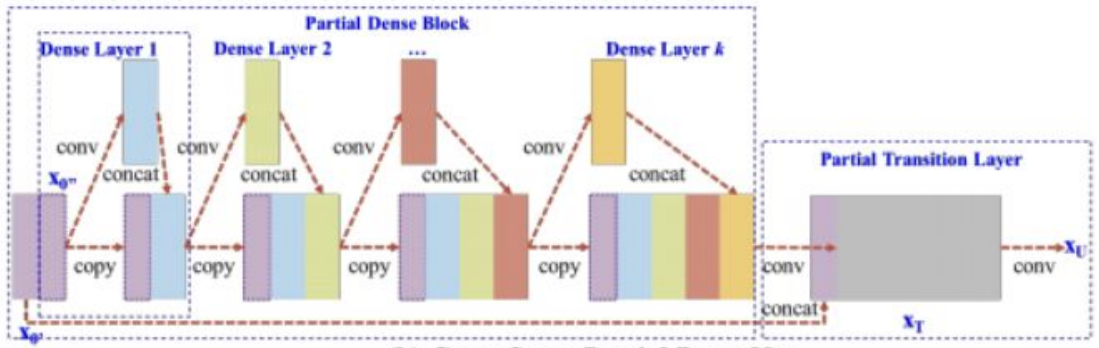
Features of Modern YOLO Models (1)

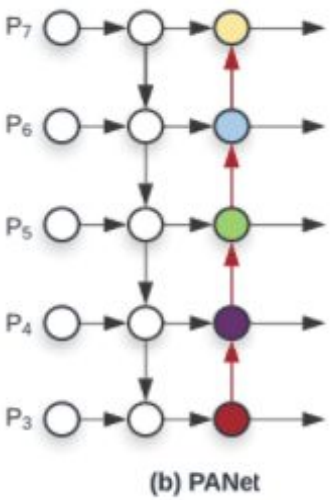


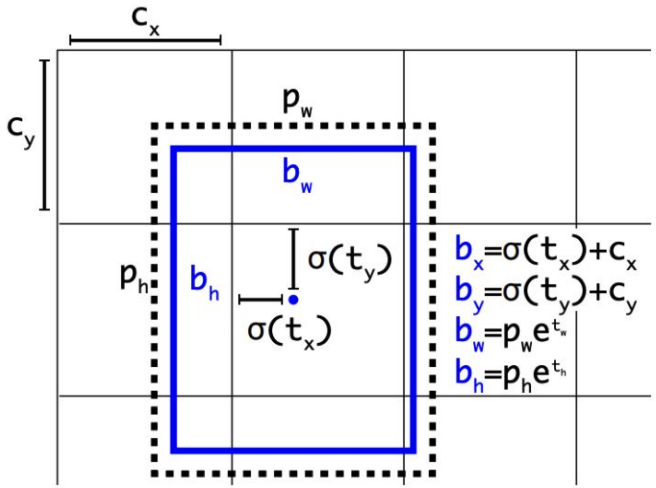
Source: YOLOv4: Optimal Speed and Accuracy of Object Detection

Features of Modern YOLO Models (2)

- Use **CSPDarknet53** as their backbone.
- Use **PANet** as the neck.
- Use **Anchor-based** detection on Head.
- Use **Bag-of-Freebies** techniques to improve data augmentation.
  - Mosaic augmentation
  - Self-adversarial augmentation
- Use **Bag-of-Specials** techniques to increase the performance.
  - Mish activation function
  - DropBlock regularization



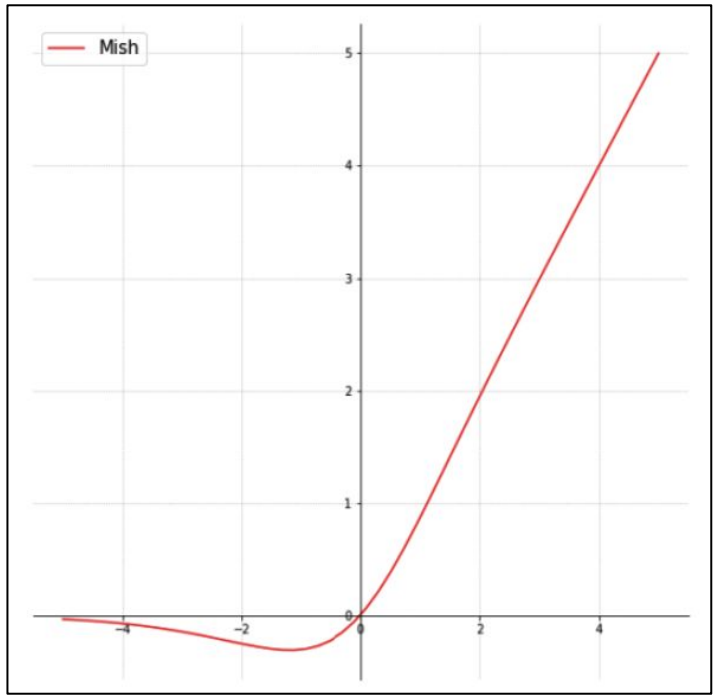




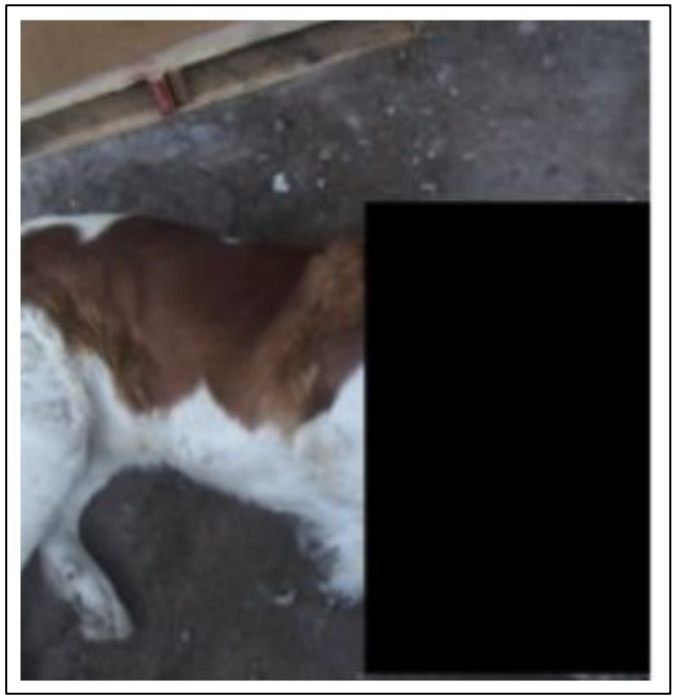
Source: YOLOv4: Optimal Speed and Accuracy of Object Detection

Features of Modern YOLO Models (3)

Mosaic augmentation



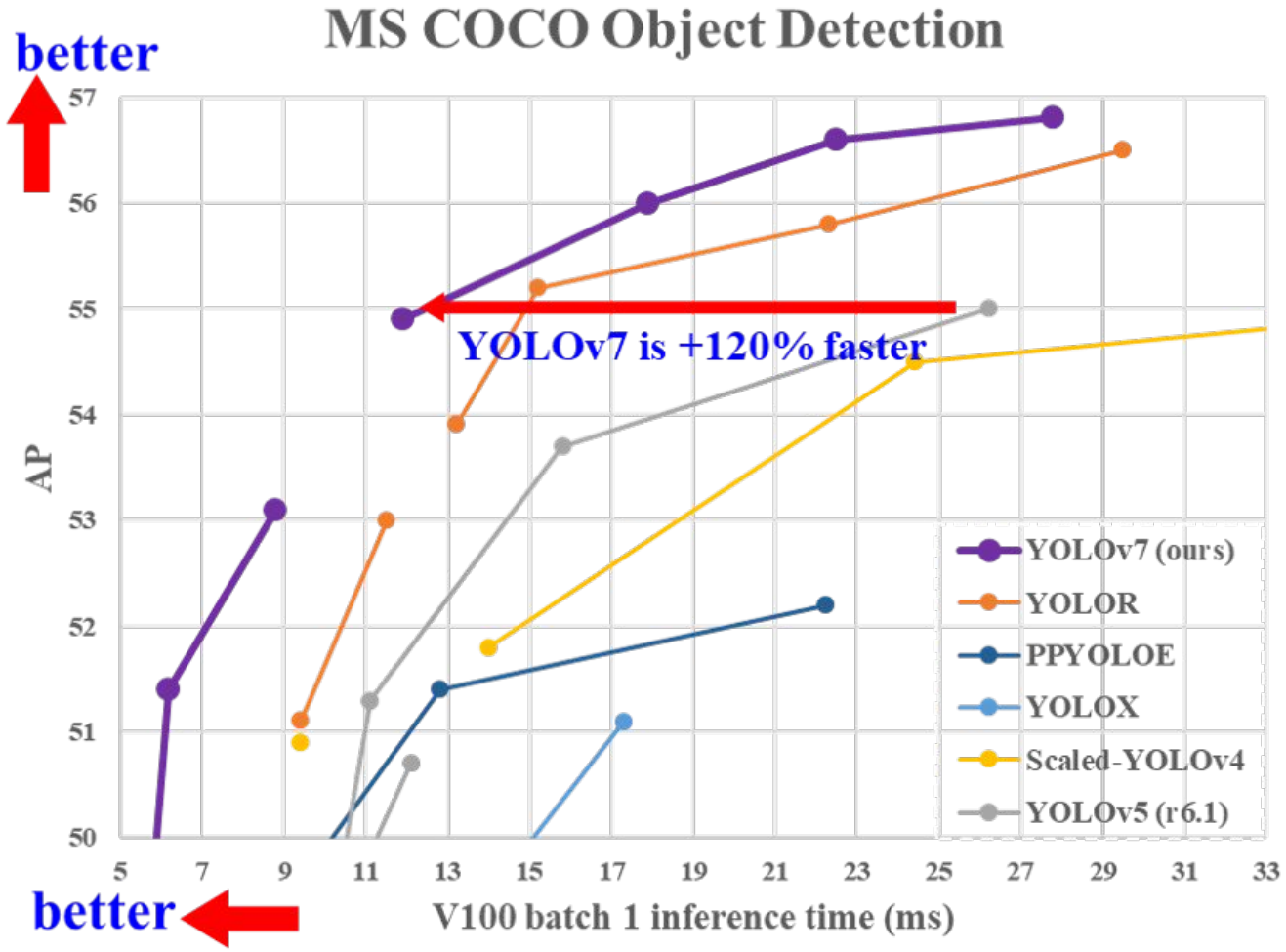
Mish activation function



DropBlock regularization

Source: YOLOv4: Optimal Speed and Accuracy of Object Detection

YOLOv7: State-of-the-art YOLO Model



Part 4

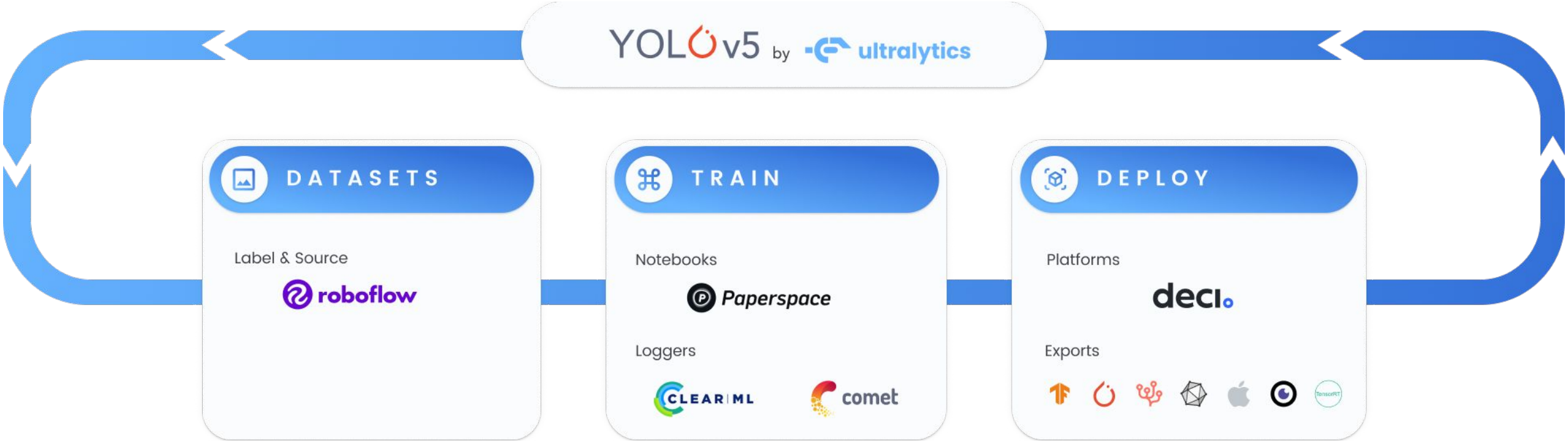
Ultralytics YOLOv5

Why to Still Talk About YOLO v5? (1)

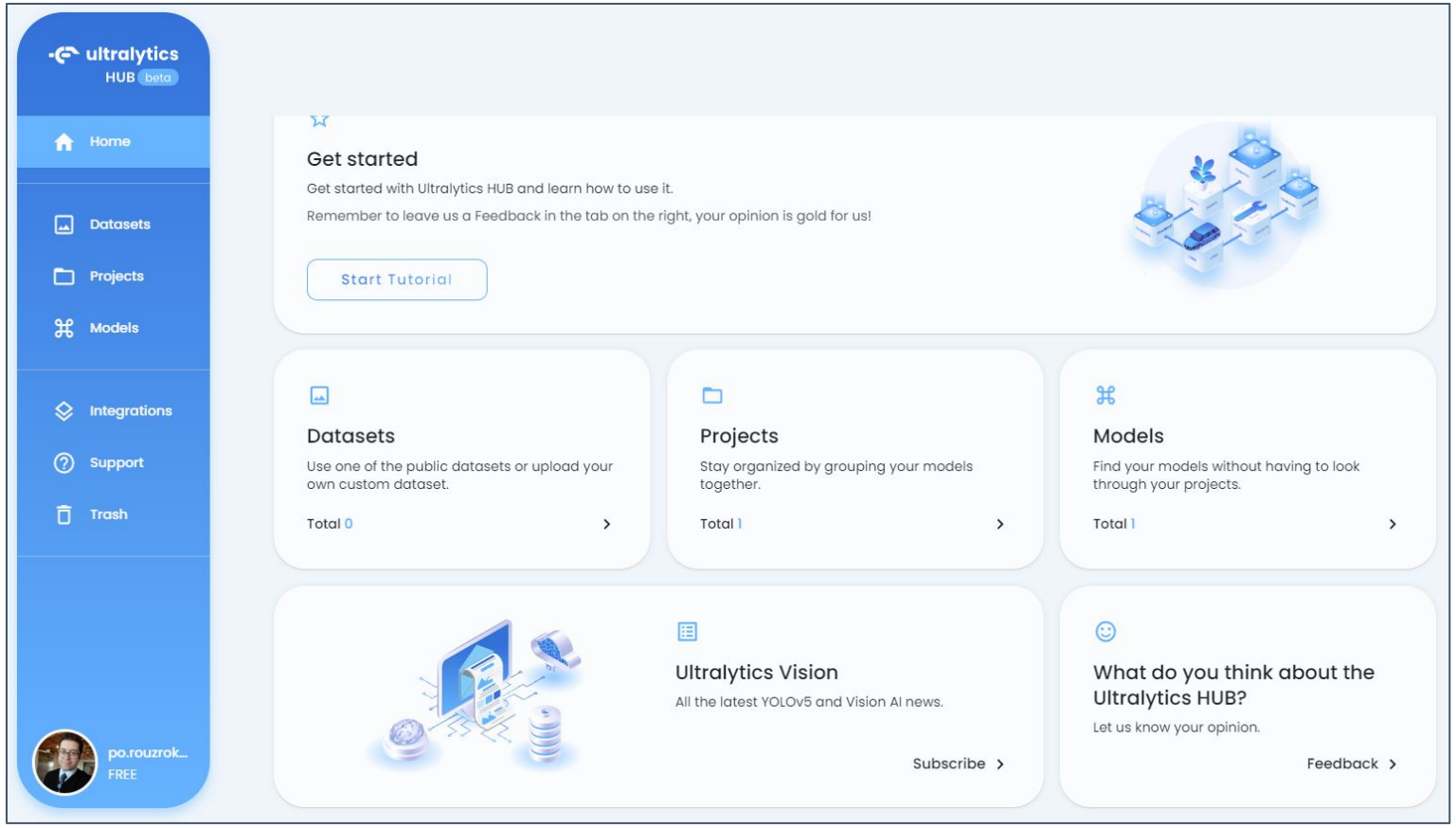
- **Not more the latest versions of YOLO.**
- **No scientific manuscript released yet.**

- Widely used in object detection tasks.
- Maintained and supported by Ultralytics.
- Easy Installation.
- Easy code-based training and inference on custom datasets.
- Available free no-code interface for training and inference on custom datasets!
- Easy deployment and integration.
- Still one of the best performing models.

Why to Still Talk About YOLO v5? (2)

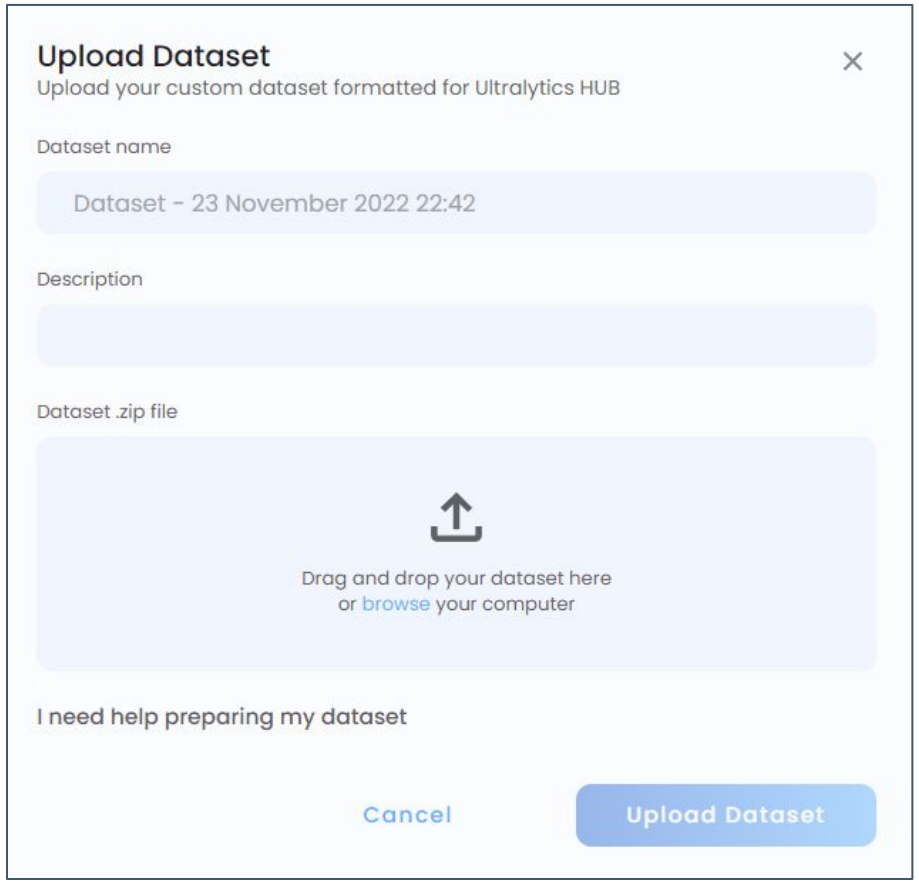


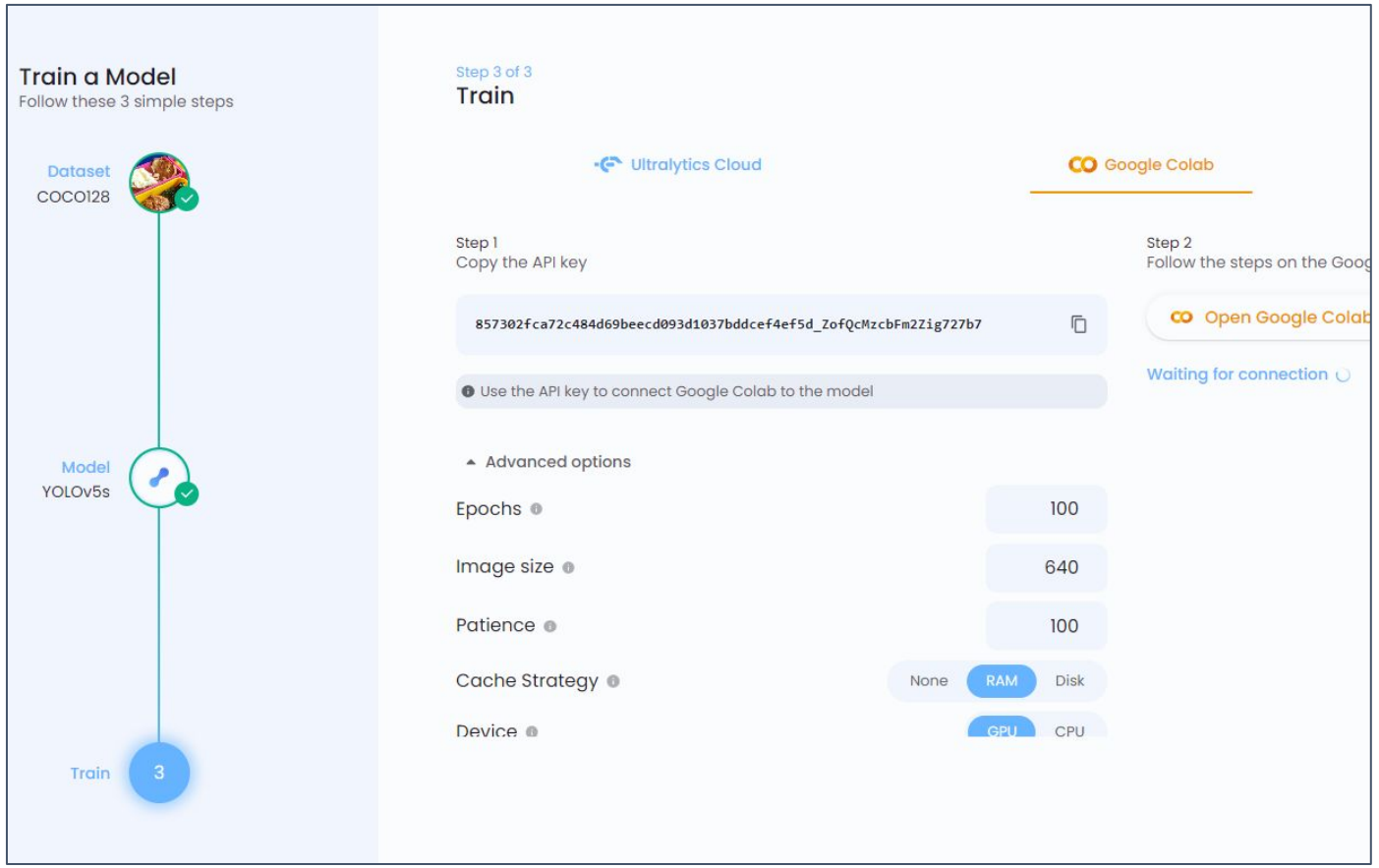
Ultralytics YOLOv5 HUB (1)



- **Web-based framework for training and applying YOLOv5 model.**
- **No coding needed (almost!)**
- **No installation needed!**
- **Free if you bring your own GPU.**

Ultralytics YOLOv5 HUB (2)





YOLOv5 Coding

But what if we want to train and apply our YOLOv5 model with Python coding?

Let’s open our Google Colab notebook!

Thank you for your attention!