

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

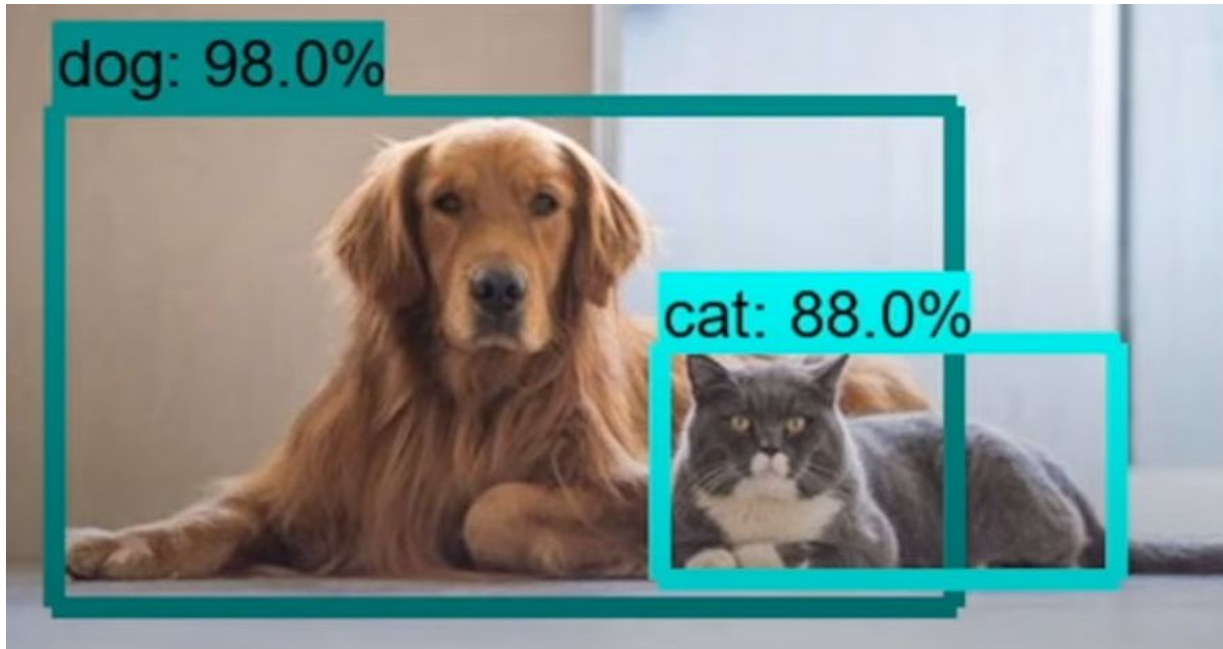
MD MPH MHPE

Radiology Informatics Laboratory
Mayo Clinic

Part 1

INTRODUCTION

Defining Object Detection (1)



Object Detection:

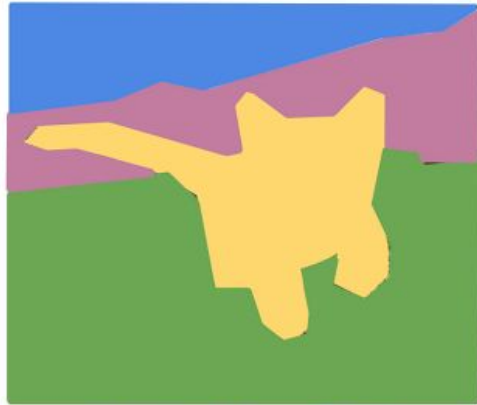
1. What kind of object is it?
2. Where is that object located?

Defining Object Detection (2)



Defining Object Detection (3)

Semantic Segmentation



GRASS, CAT,
TREE, SKY

No objects, just pixels

Classification + Localization



CAT

Single Object

Object Detection



DOG, DOG, CAT

Multiple Object

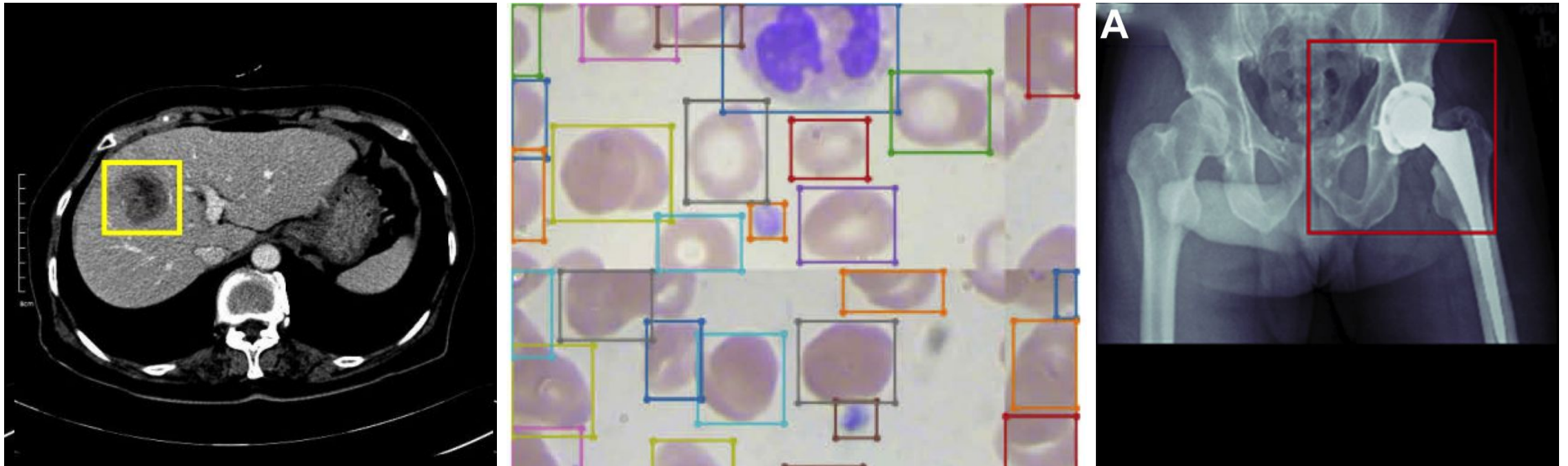
Instance Segmentation



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

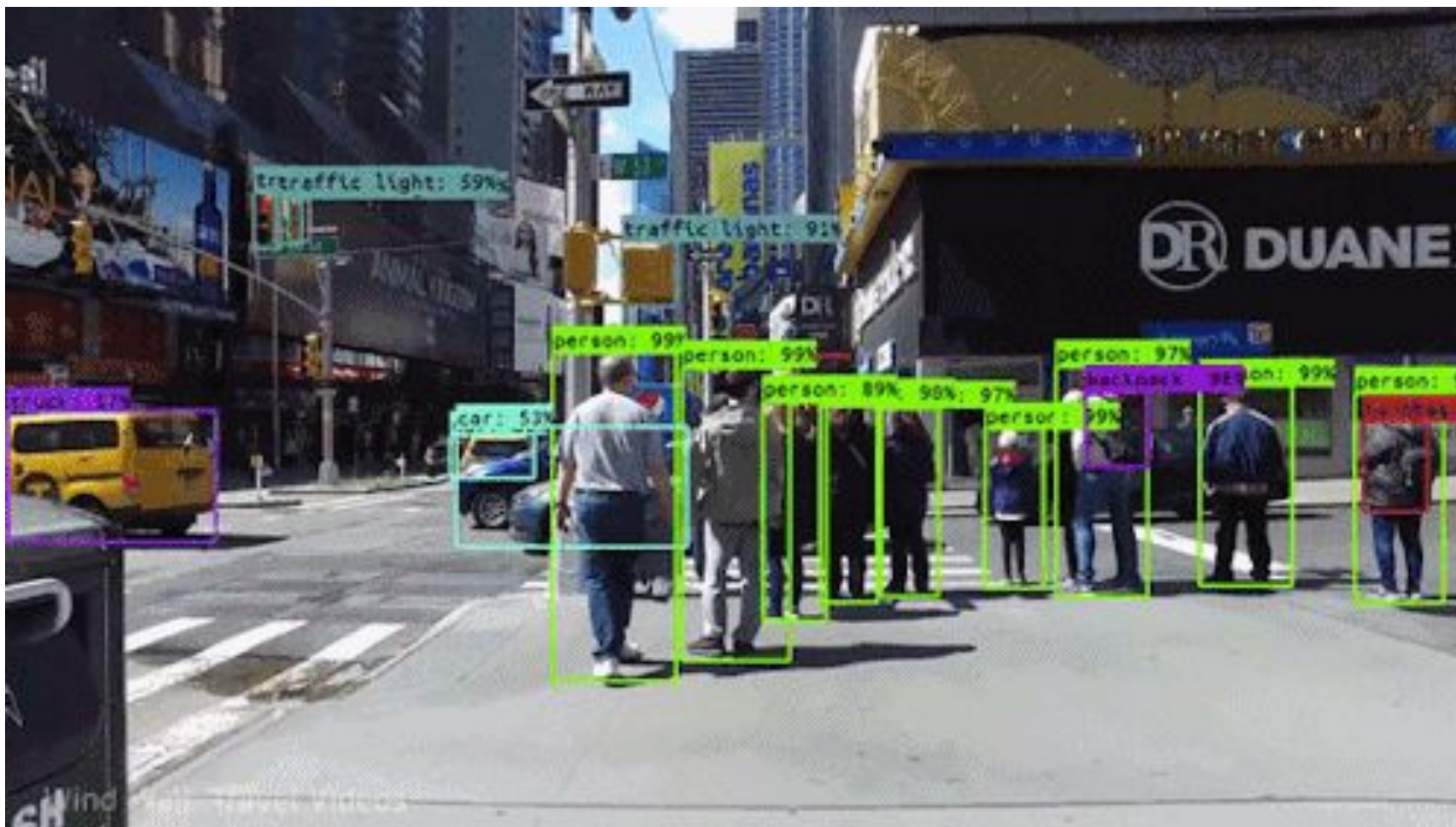
Single Shot Detection (SSD)

You Look Only Once (YOLO)



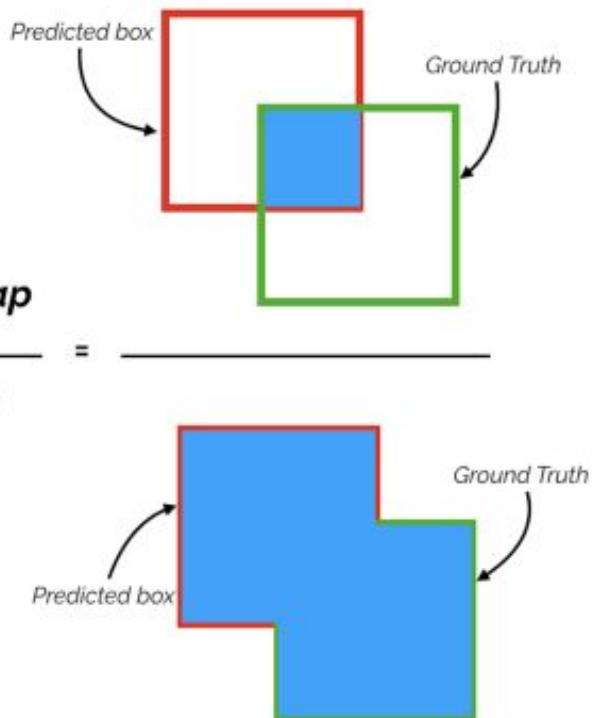
Source: <https://theconversation.com/whos-to-blame-when-a-self-driving-car-has-an-accident-150941>

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. The 'Predicted box' and 'Ground Truth' labels are placed around this union shape, with arrows pointing to its boundaries.

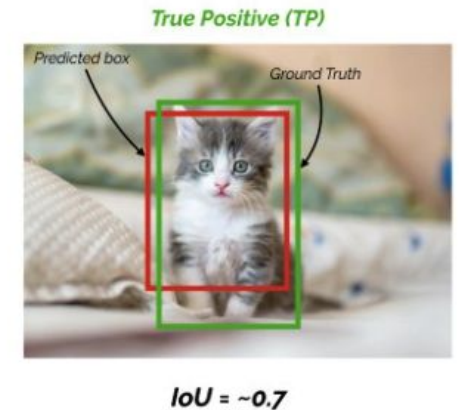
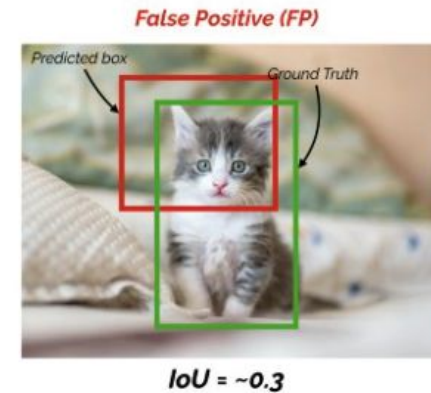
Source: <https://towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2>

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...

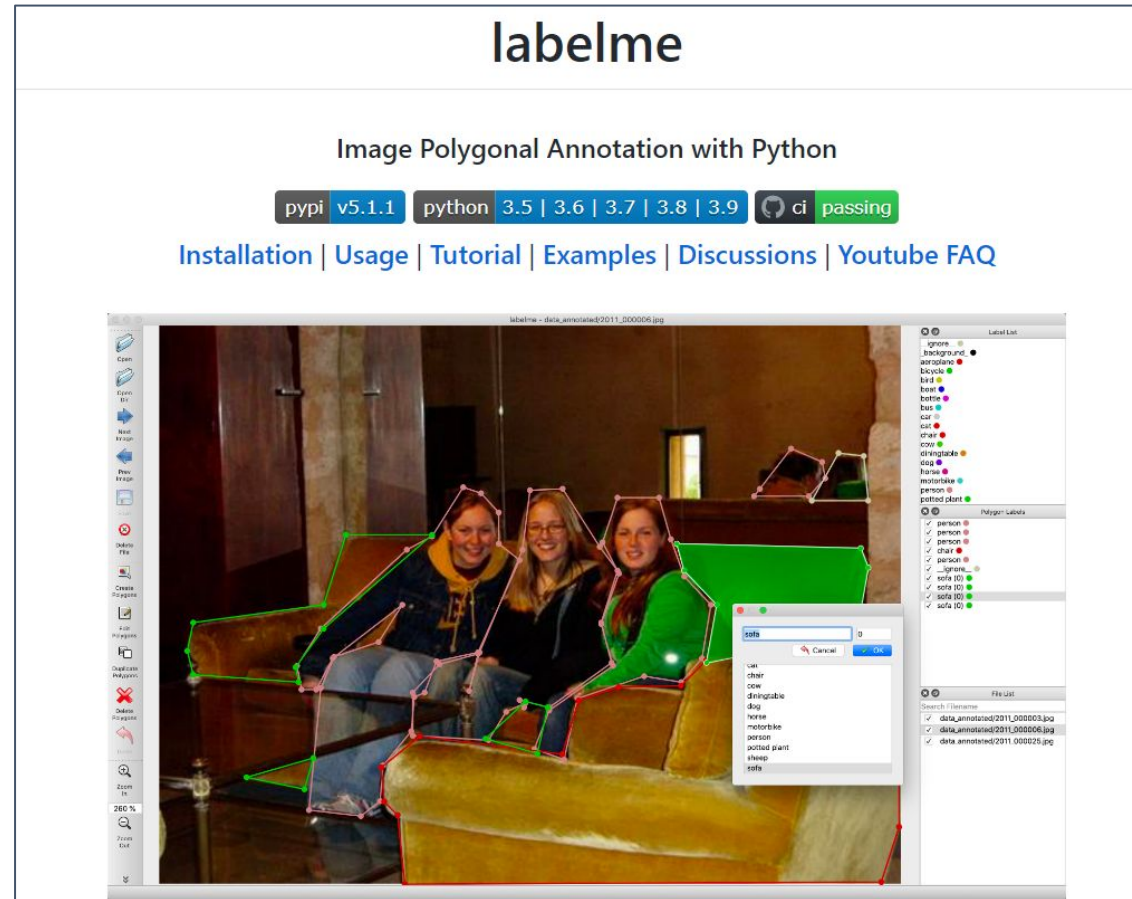
Im1.txt

```
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>

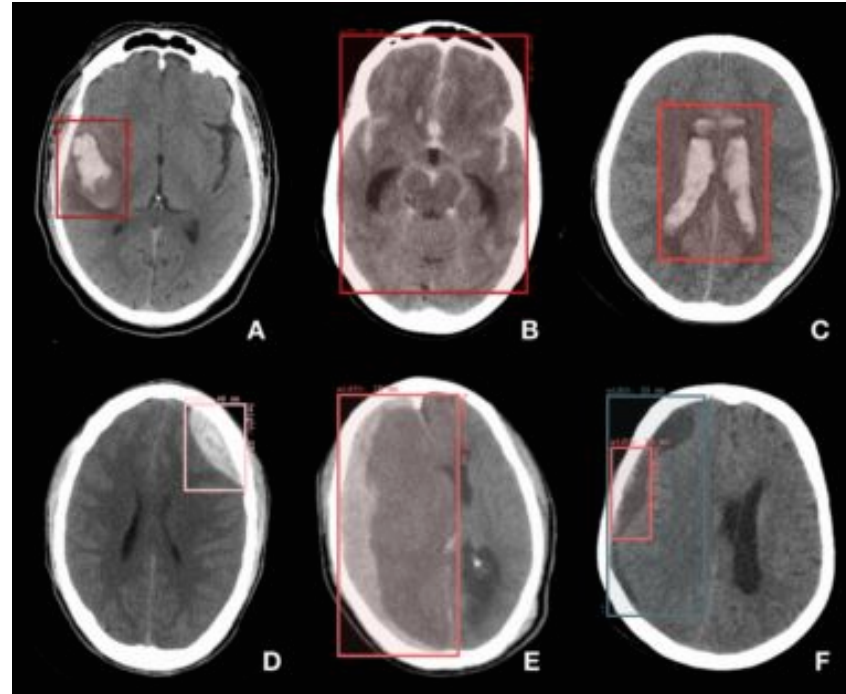
How to Label Custom Datasets? (2)



A Nice Tutorial on How to Use labelme:

<https://www.youtube.com/watch?v=ydHI8SUe58Y>

Let's Prepare the Data for Our Workshop!

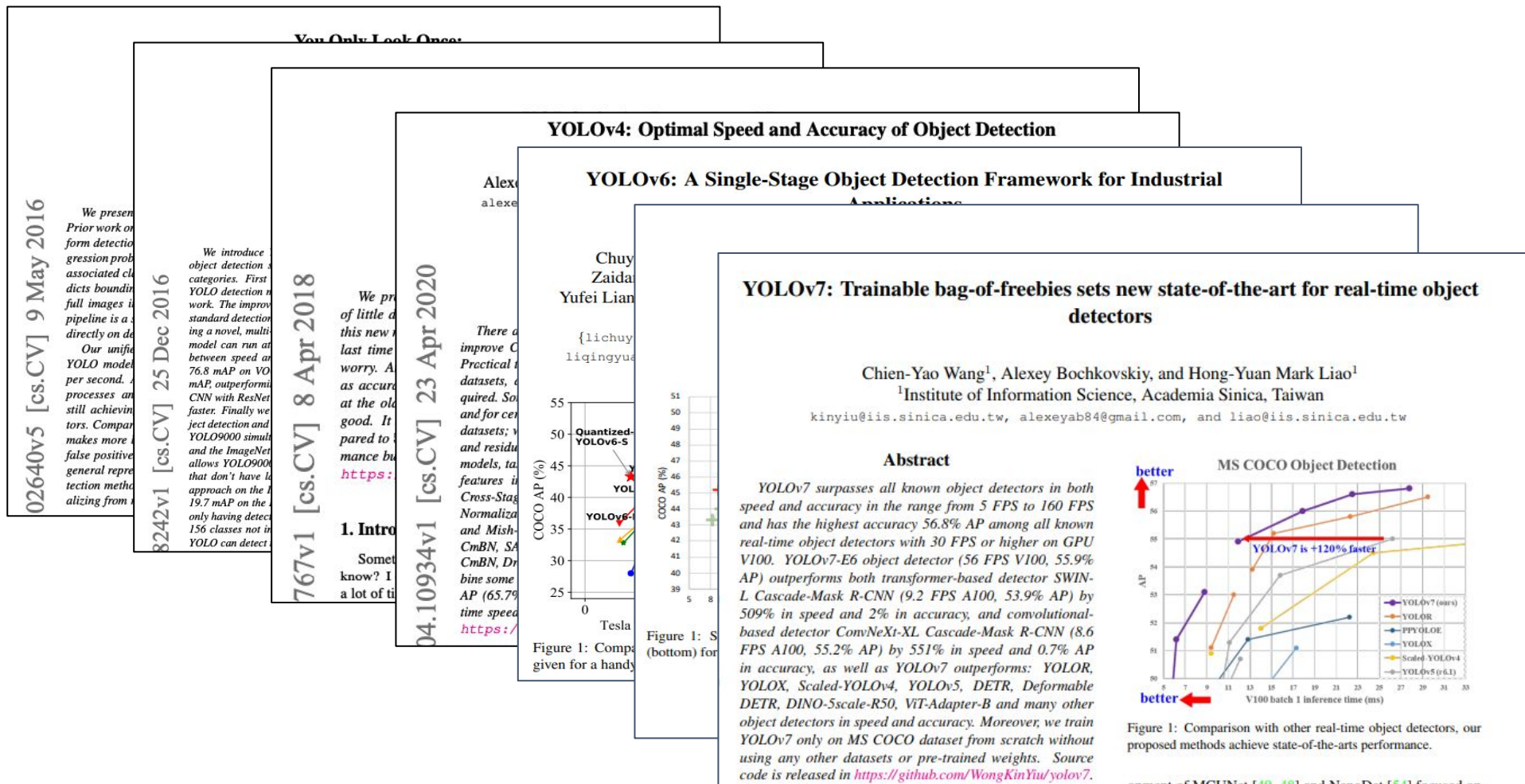


In this workshop, we will train an object detection
deep learning model to detect
brain hemorrhage lesion on Head CT scans!

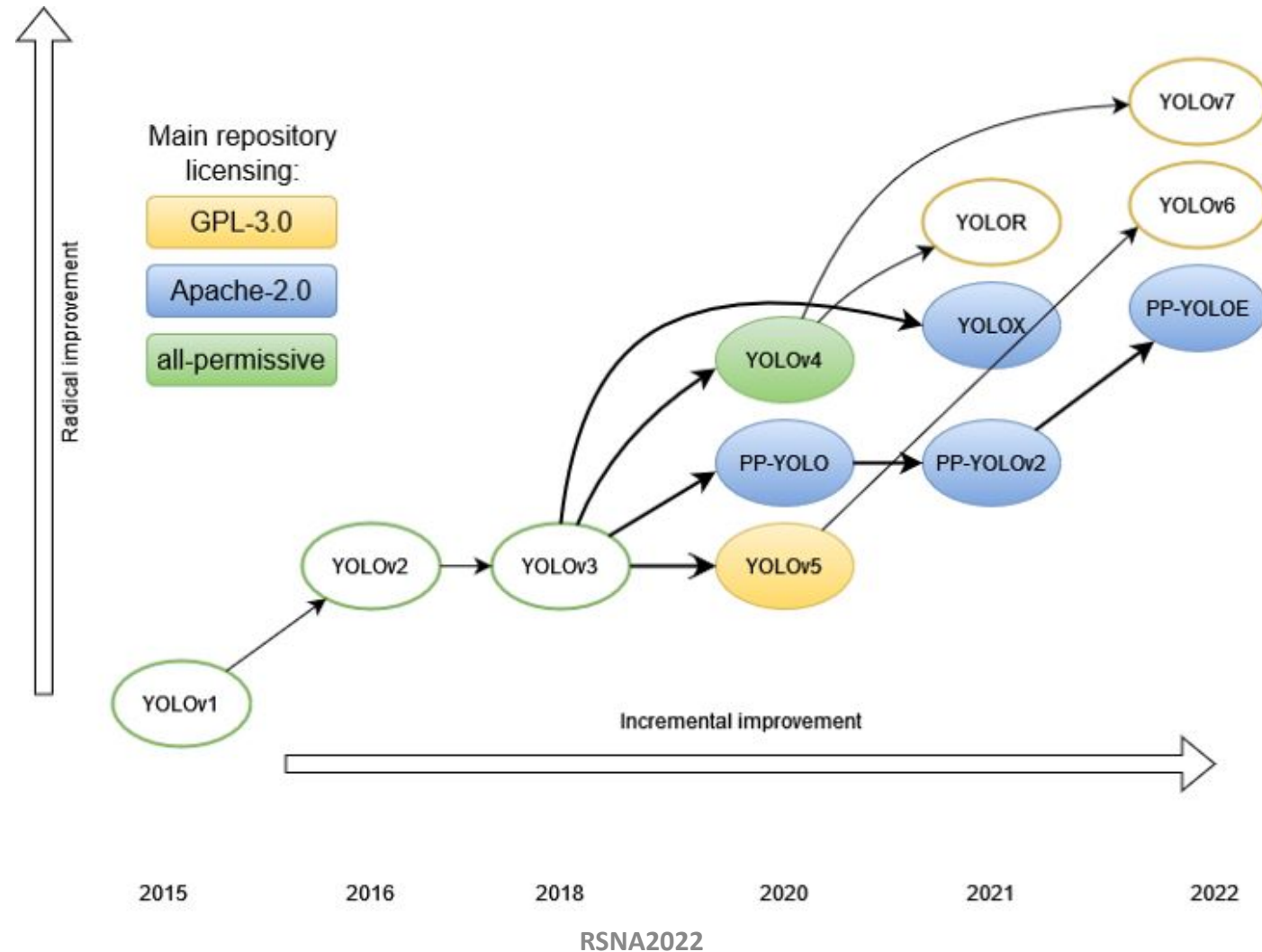
Part 3

YOLO

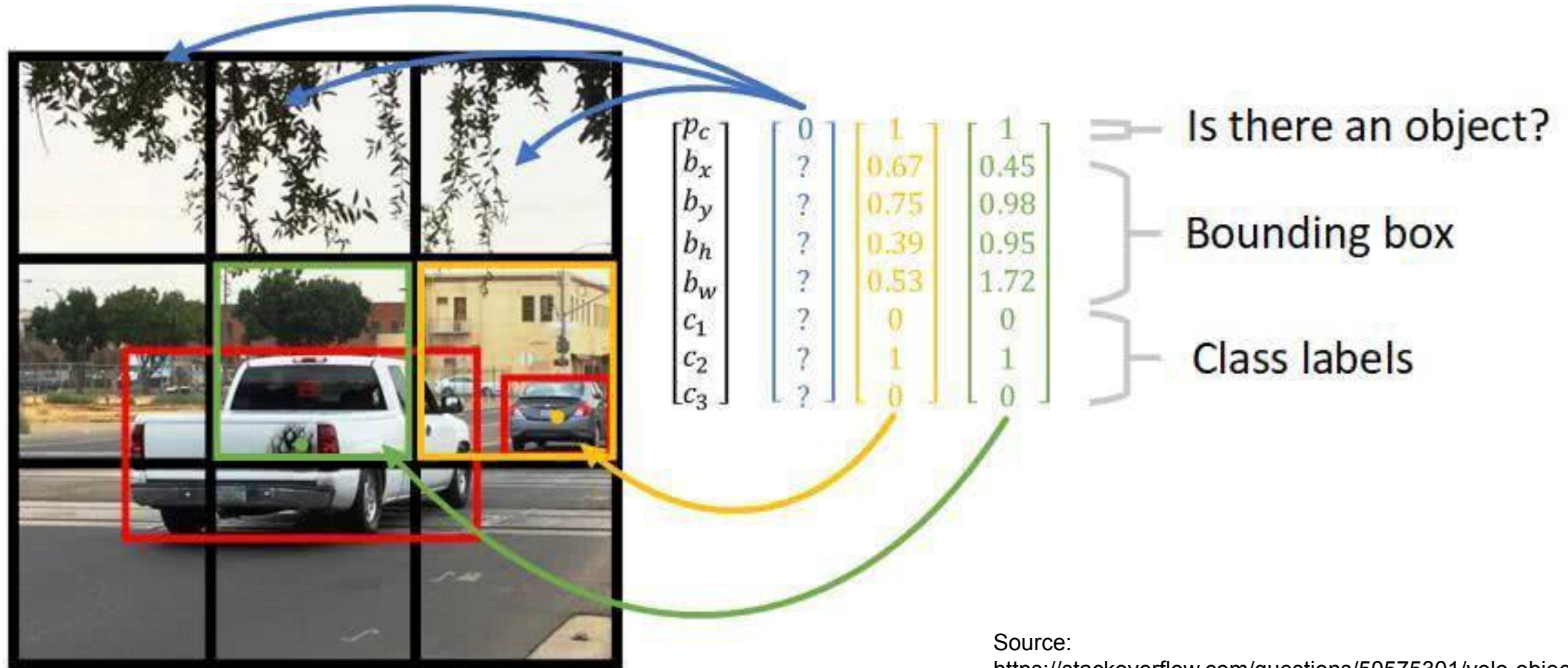
You Look Only Once (YOLO)



You Look Only Once (YOLO)



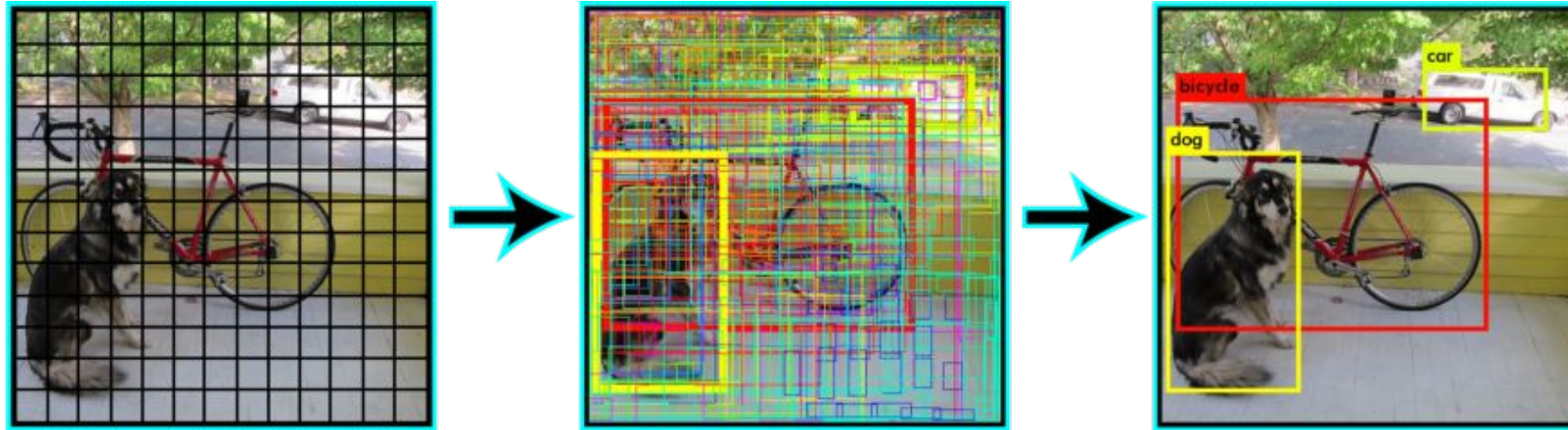
How YOLO works (1)



Source:

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

How YOLO works (2)

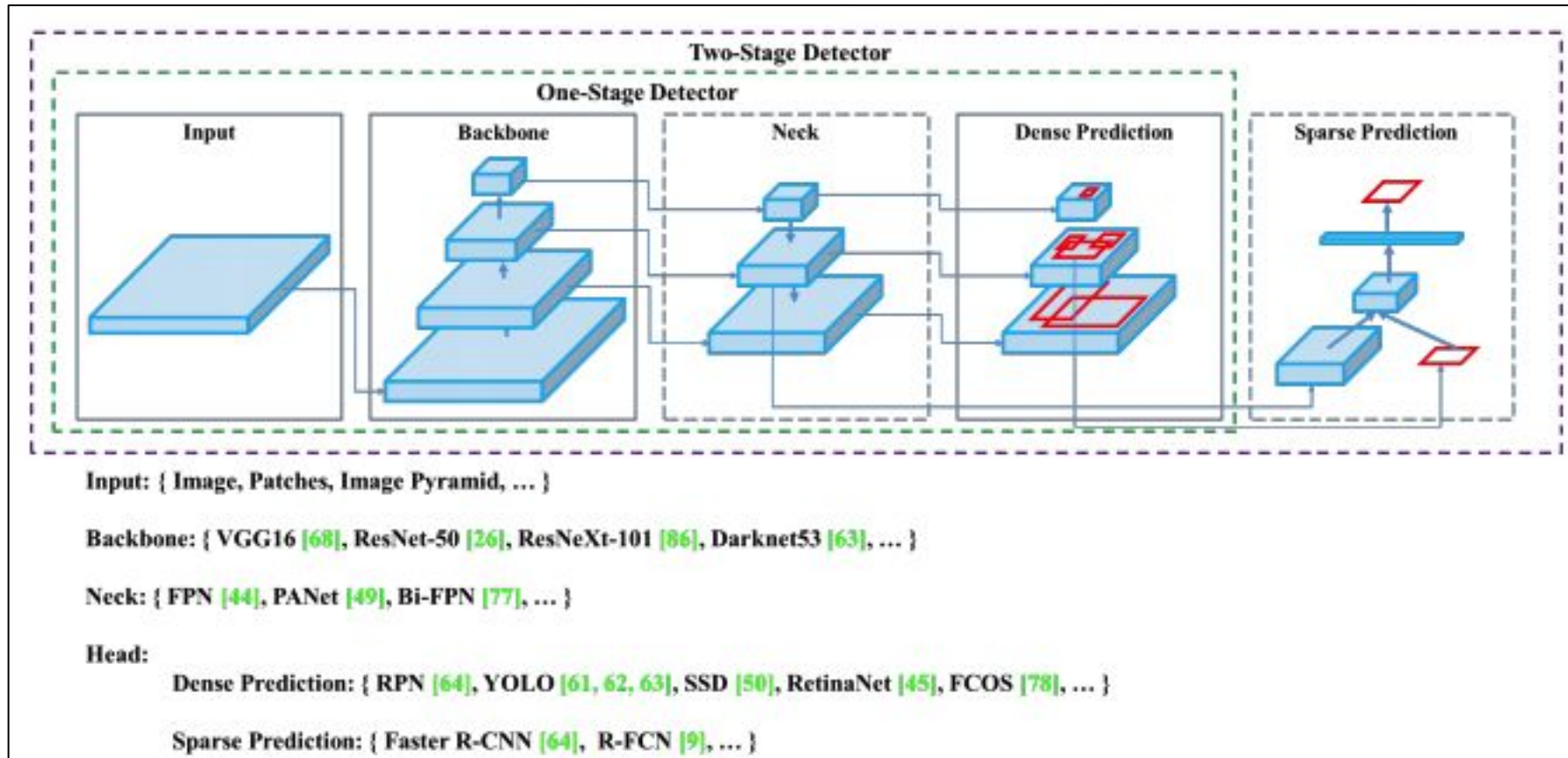


Non-Max Suppression:

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/>

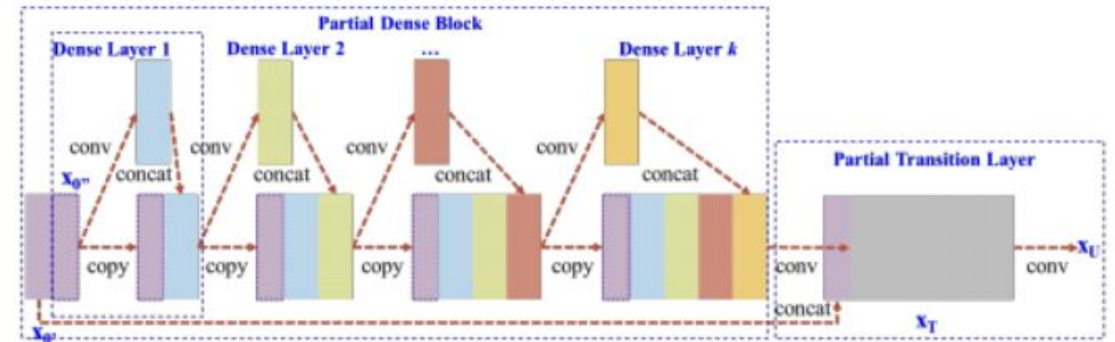
Features of Modern YOLO Models (1)



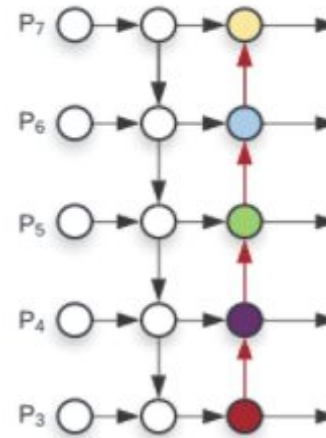
Source: YOLOv4: Optimal Speed and Accuracy of Object Detection

Features of Modern YOLO Models (2)

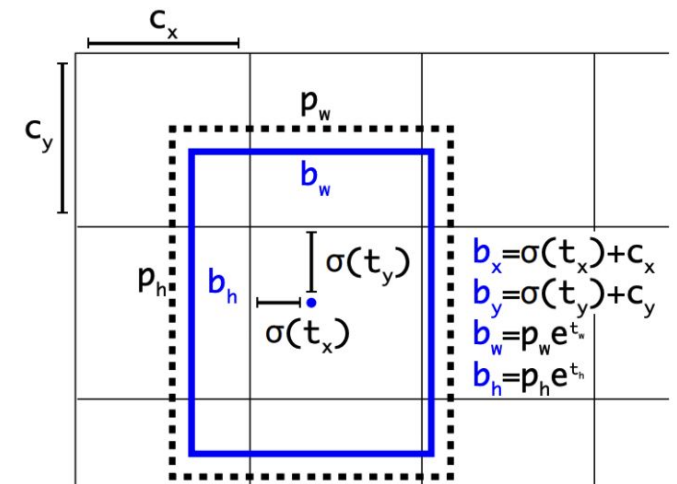
- Use **CSPDarknet53** as their backbone.
- Use **PANet** as the neck.
- Use **Anchor-based** detection on Head.
- Use **Bag-of-Freebies** techniques to improve data augmentation.
 - Mosaic augmentation
 - Self-adversarial augmentation
- Use **Bag-of-Specials** techniques to increase the performance.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



(b) Cross Stage Partial DenseNet



(b) PANet

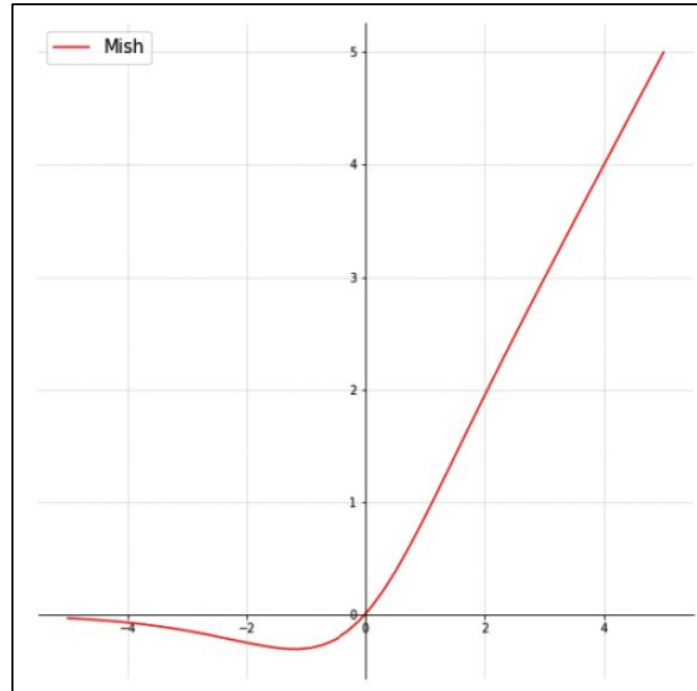


Source: YOLOv4: Optimal Speed and Accuracy of Object Detection

Features of Modern YOLO Models (3)



Mosaic augmentation



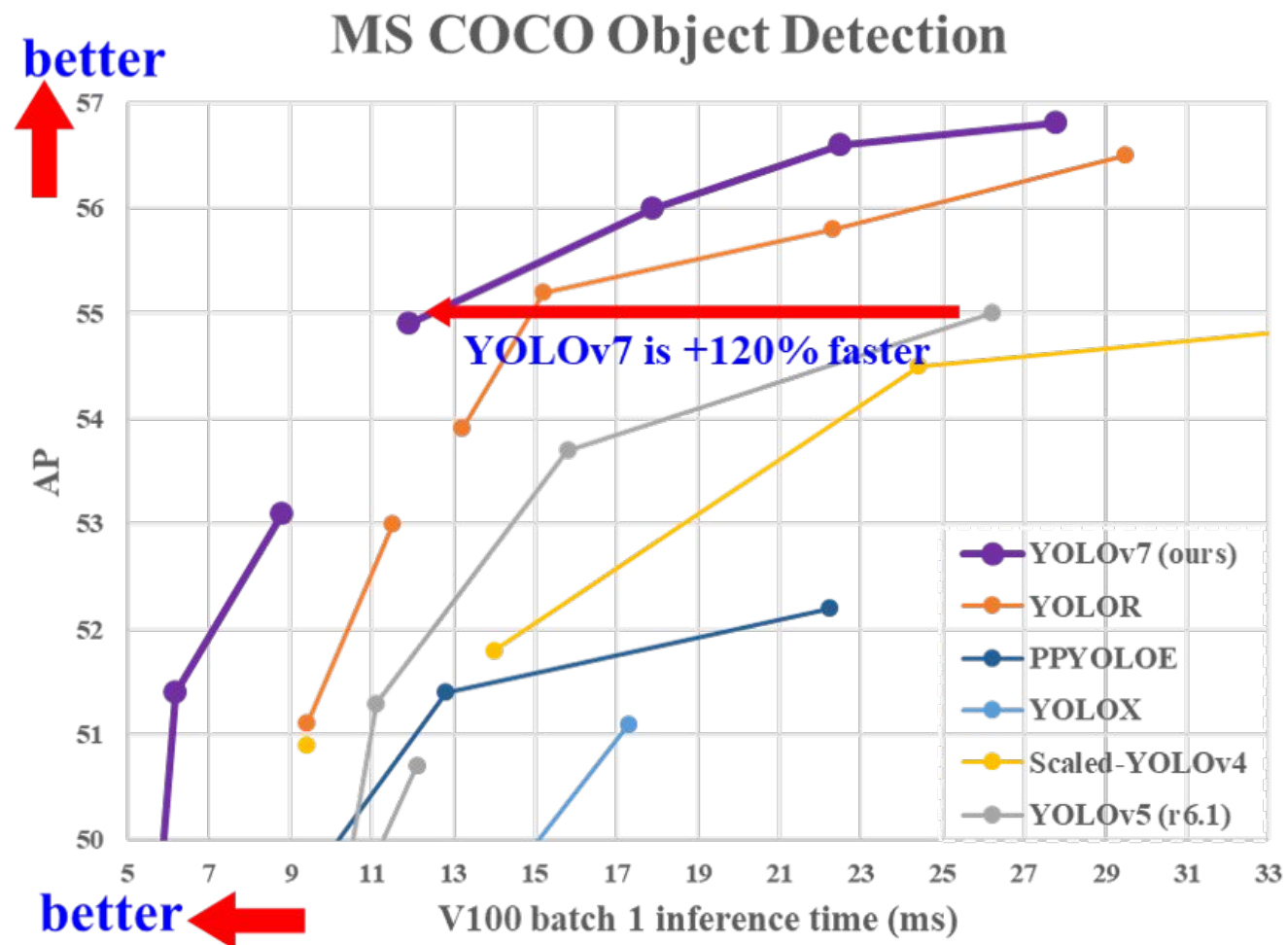
Mish activation function



DropBlock regularization

Source: [YOLOv4: Optimal Speed and Accuracy of Object Detection](#)

YOLOv7: State-of-the-art YOLO Model



Part 4

Ultralytics YOLOv5

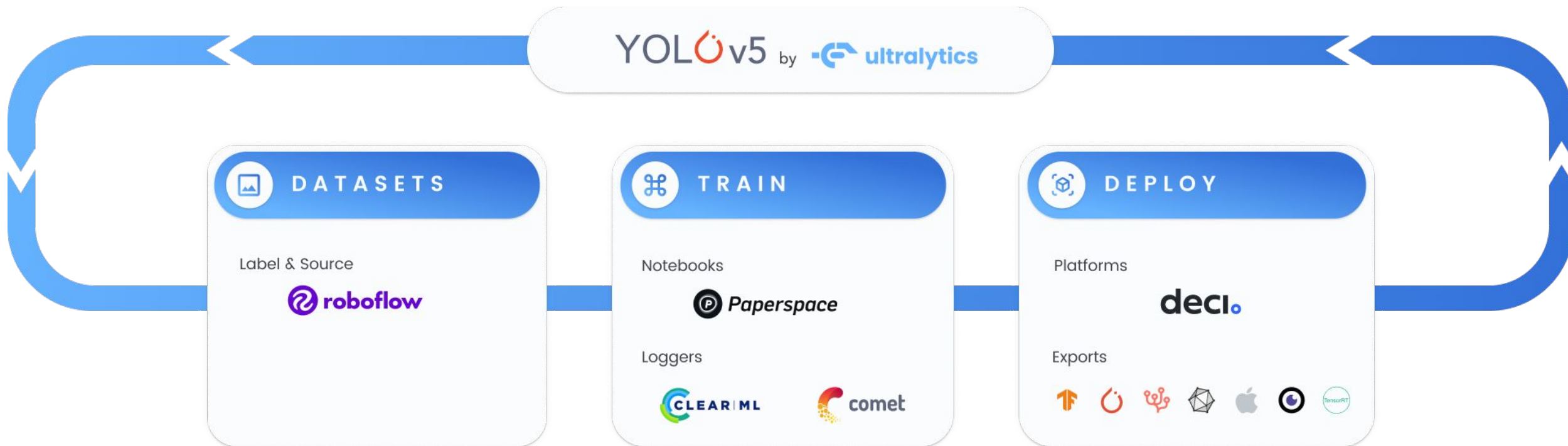
Why to Still Talk About YOLO v5? (1)

- 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.

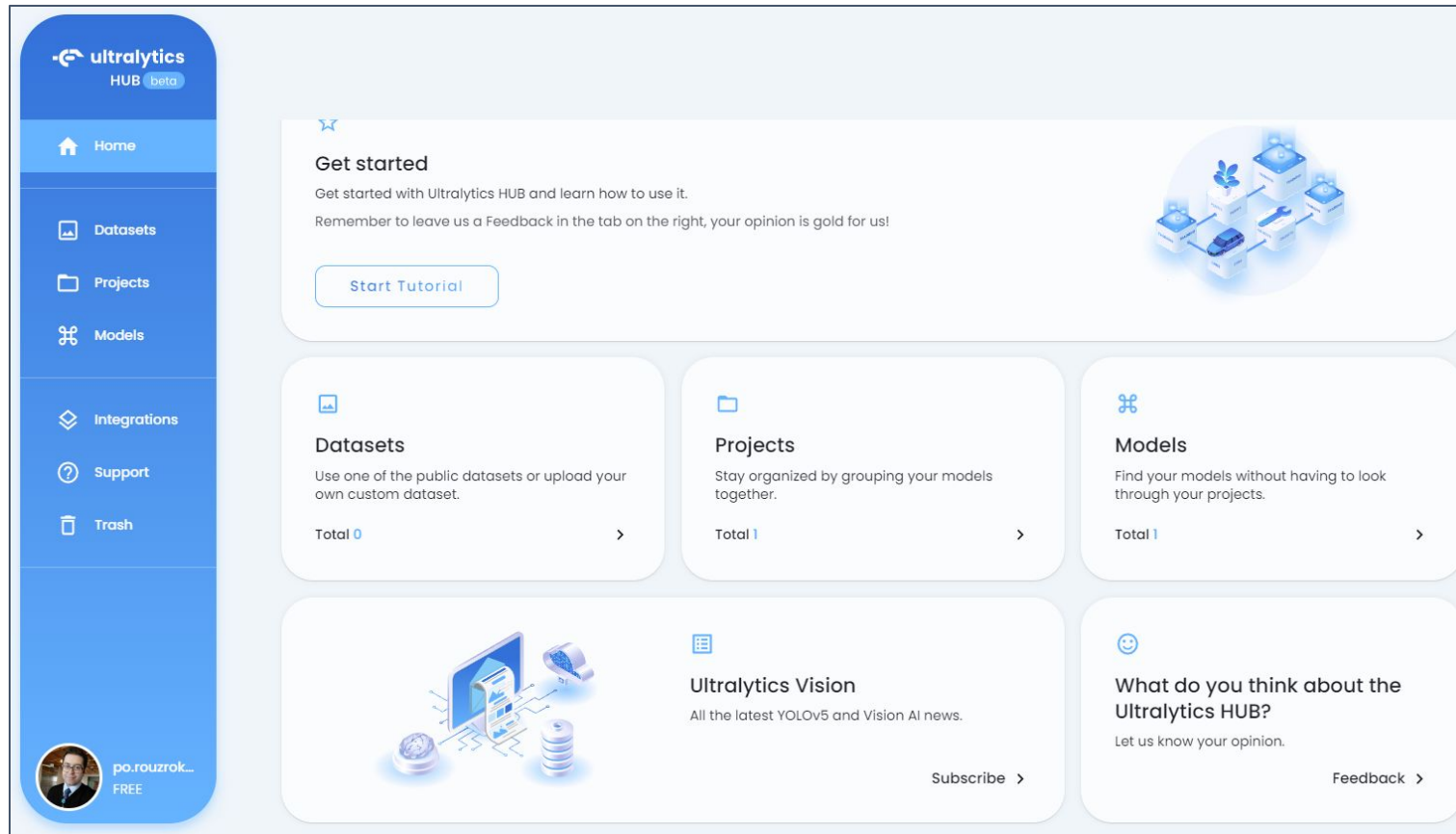


CI CPU testing **failing** DOI [10.5281/zenodo.4679653](https://doi.org/10.5281/zenodo.4679653)
[Open in Colab](#) [Open in Kaggle](#) [docker pulls 37k](#)

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)

Upload Dataset

Upload your custom dataset formatted for Ultralytics HUB

Dataset name

Dataset - 23 November 2022 22:42

Description

Dataset .zip file

Drag and drop your dataset here
or [browse](#) your computer

I need help preparing my dataset

[Cancel](#) [Upload Dataset](#)

Train a Model

Follow these 3 simple steps

Dataset
COCO128

Model
YOLOv5s

Train 3

Step 3 of 3 Train

[Ultralytics Cloud](#) [Google Colab](#)

Step 1
Copy the API key

857302fca72c484d69beecd093d1037bddcef4ef5d_ZofQcMzcbFm2Zig727b7

Use the API key to connect Google Colab to the model

Step 2
Follow the steps on the Google Colab

[Open Google Colab](#)

Waiting for connection

Advanced options

Epochs 100

Image size 640

Patience 100

Cache Strategy None RAM Disk

Device GPU CPU

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!
