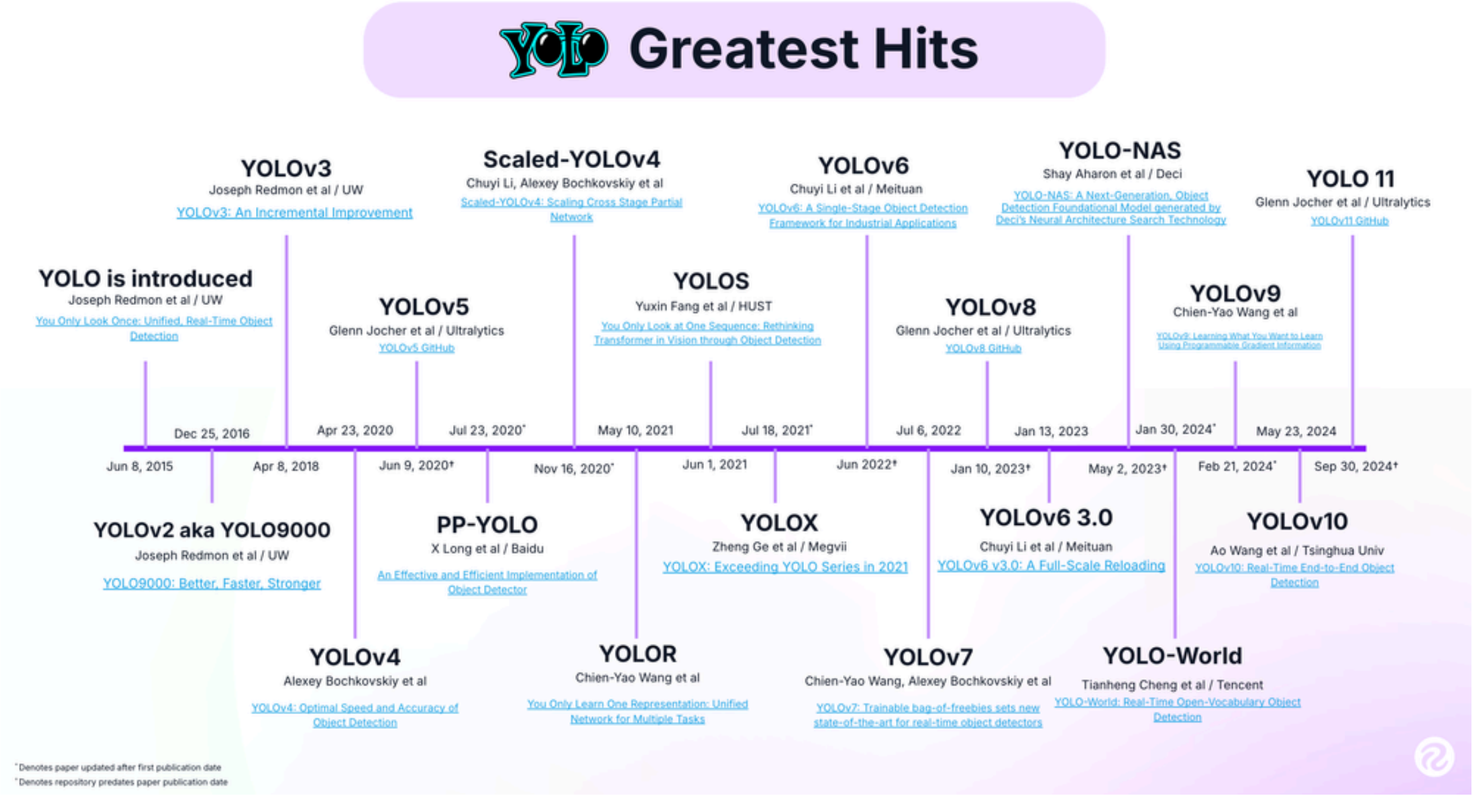
YOLO ALGORITHM

By Seitzhanova Altyn

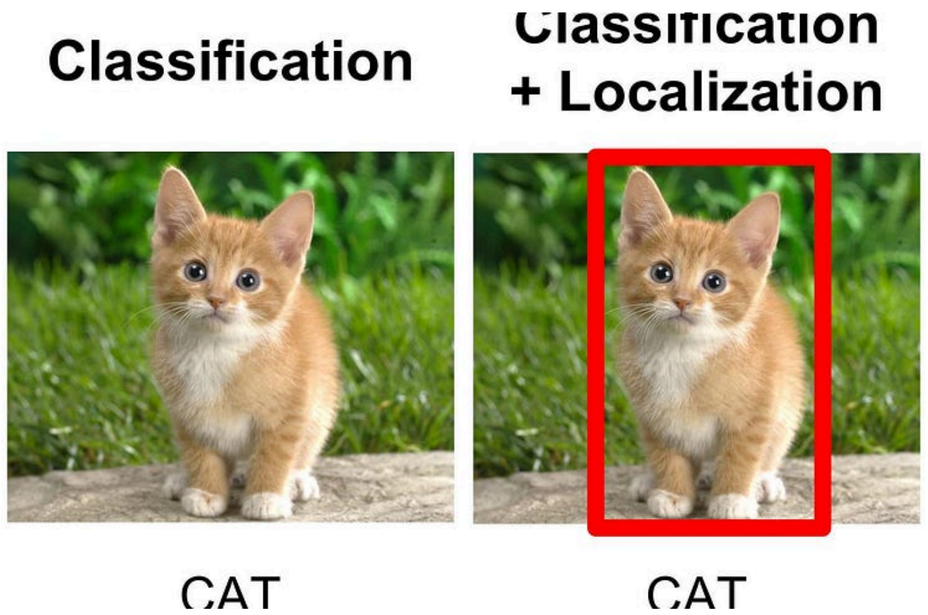
YOLO

YOLO (You Only Look Once) is a family of computer vision models that has gained significant fanfare since Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi introduced the novel architecture in 2016 at CVPR – even winning OpenCV's People Choice Awards.

The original YOLO (You Only Look Once) was written by Joseph Redmon in a custom framework called Darknet. Darknet is a very flexible research framework written in low level languages and has produced a series of the best realtime object detectors in computer vision: YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOV6, YOLOV7, YOLOv8, YOLO-NAS, YOLO-World, YOLOv9, YOLOv10, YOLOv11 and YOLOv12.



The original YOLO model was the first object detection network to combine the problem of drawing bounding boxes and identifying class labels in one end-to-end differentiable network.



Is this a cat?

Cat=1

Dog=0

Where is cat?

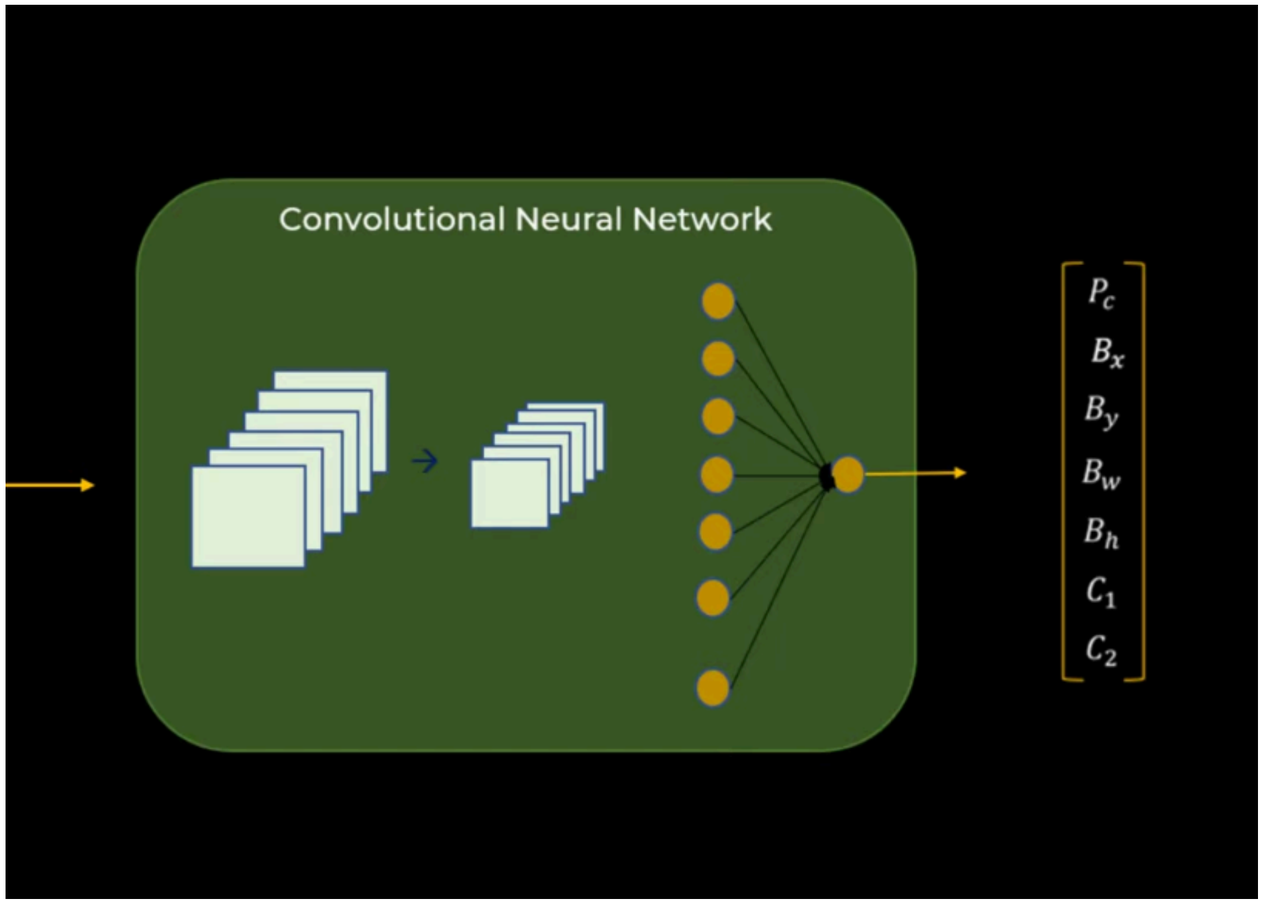
Cat=1

Dog=0

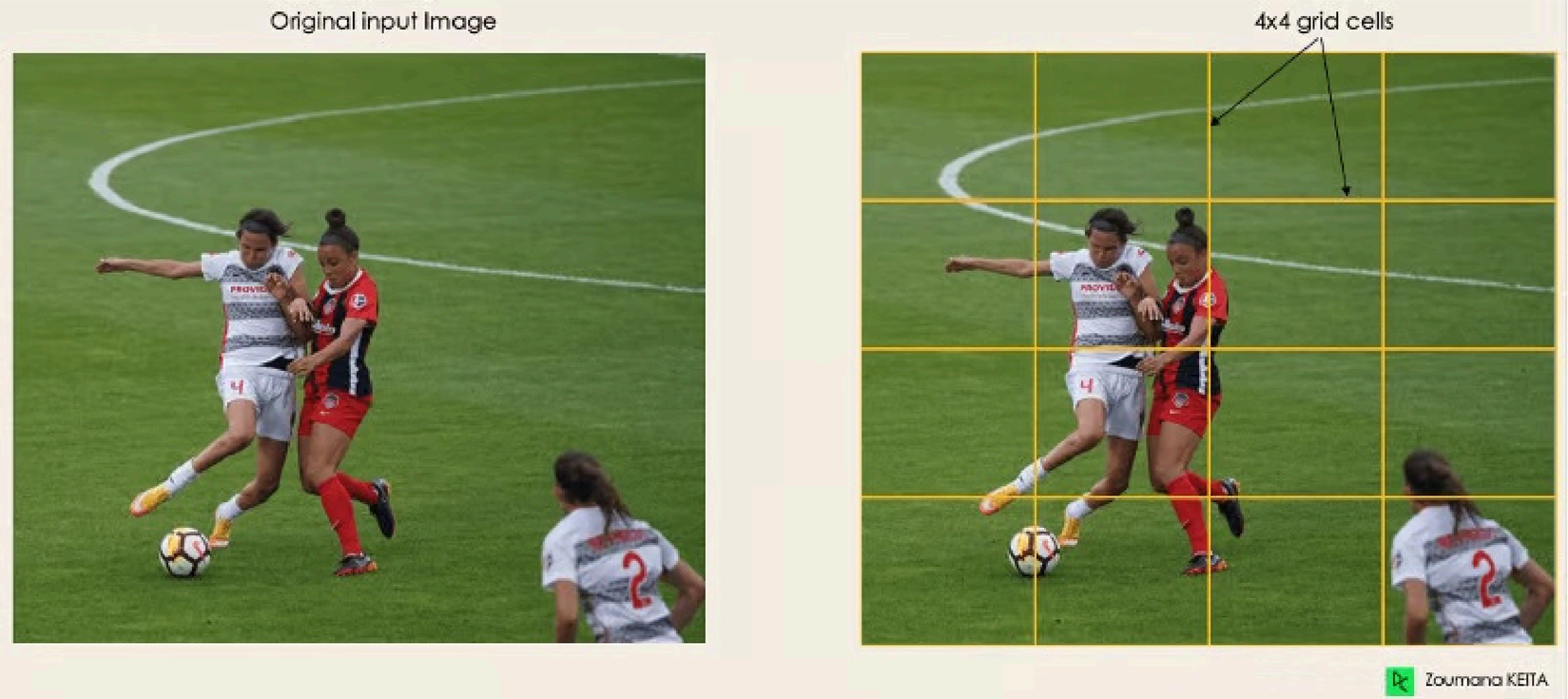
+

Bounding box

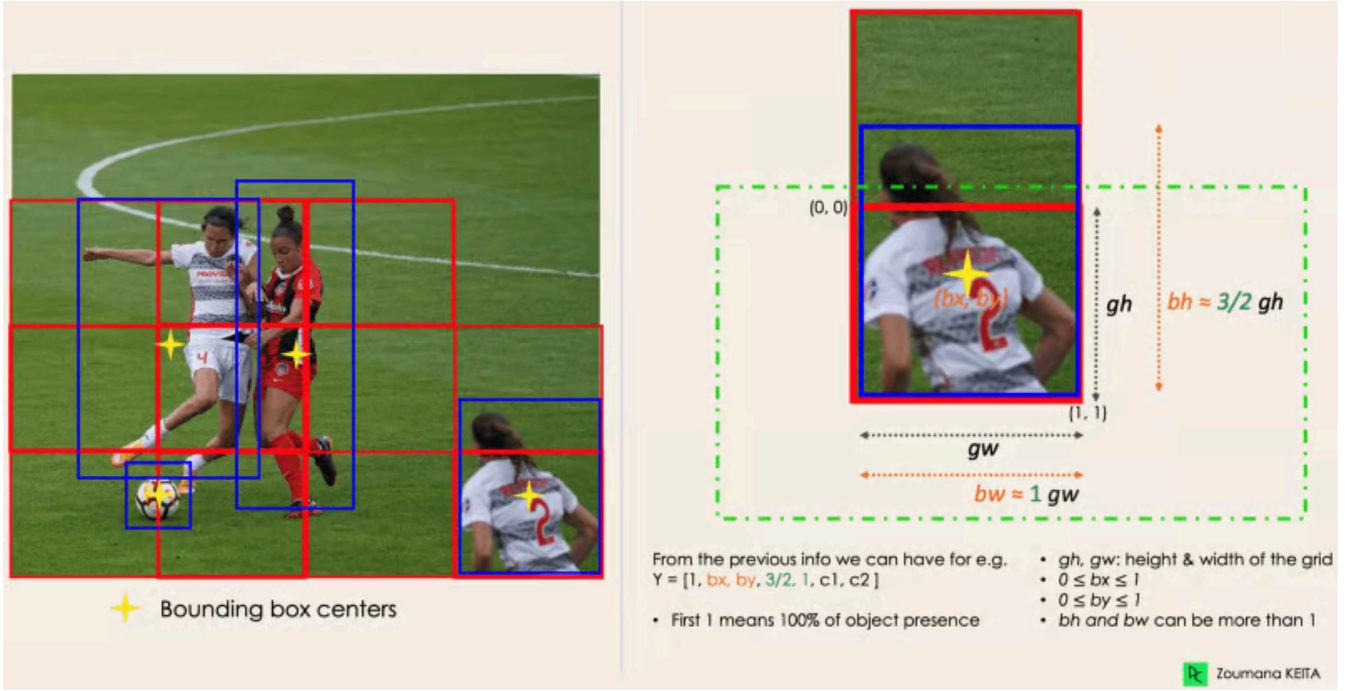
OBJECT LOCALIZATION



$$\begin{bmatrix} P_c \\ B_x \\ B_y \\ B_w \\ B_h \\ C_1 \\ C_2 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 50 \\ 70 \\ 60 \\ 70 \\ 1 \\ 0 \end{bmatrix}$$



The first step by dividing the original image (A) into NxN grid cells of equal shape, where N, in our case, is 4, as shown in the image on the right. Each cell in the grid is responsible for localizing and predicting the class of the object that it covers, along with the probability/confidence value.



The next step is to determine the bounding boxes corresponding to rectangles, highlighting all the objects in the image. We can have as many bounding boxes as there are objects within a given image.

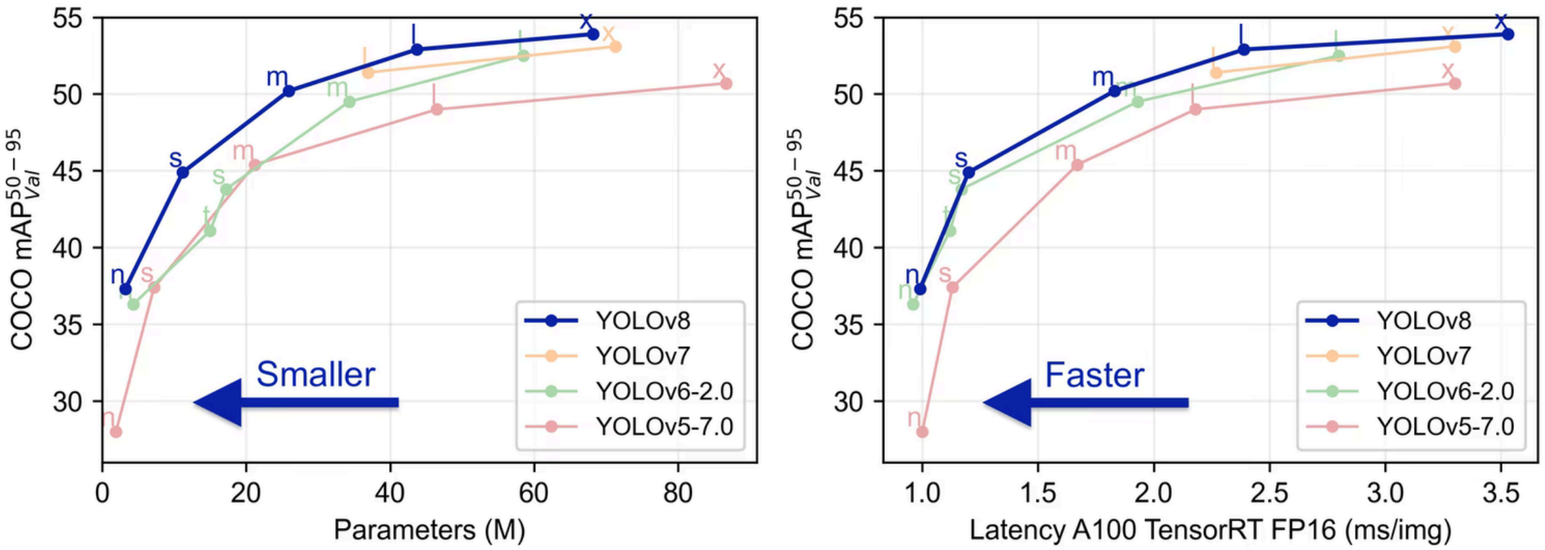
YOLO determines the attributes of these bounding boxes using a single regression module in the following format, where Y is the final vector representation for each bounding box.

Y = [pc, bx, by, bh, bw, c1, c2]

This is especially important during the training phase of the model.

- Most of the time, a single object in an image can have multiple grid box candidates for prediction, even though not all are relevant. The goal of the IOU (a value between 0 and 1) is to discard such grid boxes to only keep those that are relevant. Here is the logic behind it:
- The user defines its IOU selection threshold, which can be, for instance, 0.5.
- Then, YOLO computes the IOU of each grid cell, which is the Intersection area divided by the Union Area.
- Finally, it ignores the prediction of the grid cells having an IOU ≤ threshold and considers those with an IOU > threshold.

- YOLOv8 was released by Ultralytics on January 10th, 2023, offering cutting-edge performance in terms of accuracy and speed. Building upon the advancements of previous YOLO versions, YOLOv8 introduced new features and optimizations that make it an ideal choice for various object detection tasks in a wide range of applications.



KEY FEATURES

- Advanced Backbone and Neck Architectures:YOLOv8 employs state-of-the-art backbone and neck architectures, resulting in improved feature extraction and object detection performance.
- Anchor-free Split Ultralytics Head: YOLOv8 adopts an anchor-free split Ultralytics head, which contributes to better accuracy and a more efficient detection process compared to anchor-based approaches.
- Optimized Accuracy-Speed Tradeoff: With a focus on maintaining an optimal balance between accuracy and speed, YOLOv8 is suitable for real-time object detection tasks in diverse application areas.
- Variety of Pre-trained Models: YOLOv8 offers a range of pre-trained models to cater to various tasks and performance requirements, making it easier to find the right model for your specific use case.

CONCLUSION

- Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vel dolor ante. Nullam feugiat egestas elit et vehicula.
- Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis vel dolor ante. Nullam feugiat egestas elit et vehicula.

THANK YOU