YOLOV8

Real time detection and tracking of moving objects

- BAH Thahir
- KARMOUT Ismail

AGENDA

1. INTRODUCTION
2. What is YOLO?
3. How Yolo works?
4. YOLO Architecture
5. YOLO Versions
6. Real-World Applications
7. Advantages and Limitations
8. CONCLUSION

INTRODUCTION

Since the beginning of this 21st century, the mandkind enter in the era of intelligent system. Real time detection and object tracking are one of it applications using in fields like security and surveillance, autonomous vehicules, robotics and many other industries. And today we will talk about one of the best deep learning algorithm used for this. Called YOLO.

What is YOLO?

which stands for

You Only LooK Once

- **Real-time object detection algorithm**
- **Introduced in 2015**
- **Research paper “You Only Look Once: Unified, Real-Time Object Detection”**

Joseph Redmon

Ali Farhadi

ross girshick

Santosh Divvala

What Makes YOLO Popular for Object Detection?

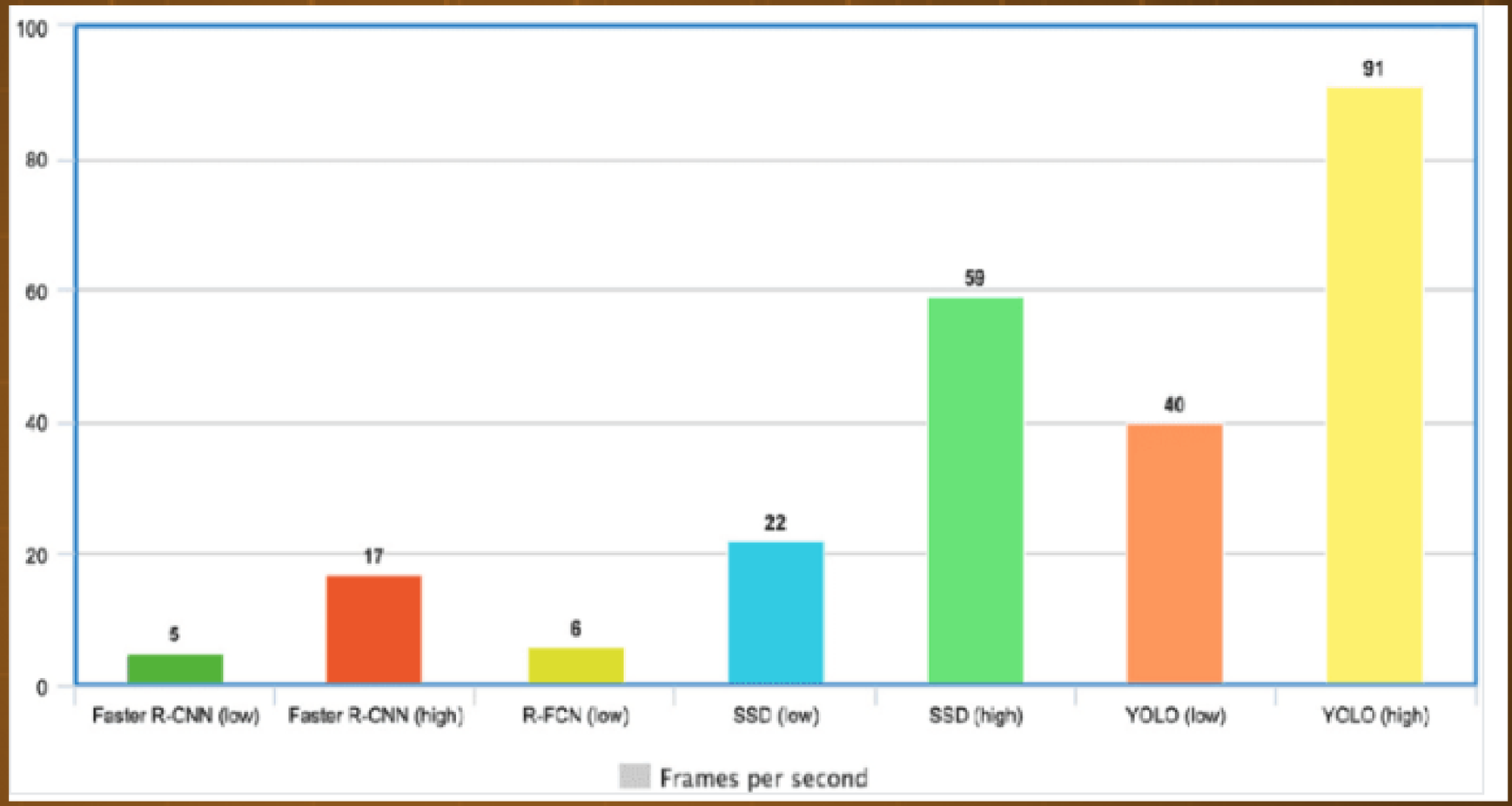
- **Speed**
- **Detection accuracy**
- **Good generalization**
- **Open-source**

Let’s have a closer look at each one of them.

1- Speed

YOLO is extremely fast because it does not deal with complex pipelines. It can process images at 45 Frames Per Second (FPS).

From the graphic below, we observe that YOLO is far beyond the other object detectors with 91 FPS.



2- High detection accuracy

YOLO is far beyond other state-of-the-art models in accuracy with very few background errors.

3- Better generalization

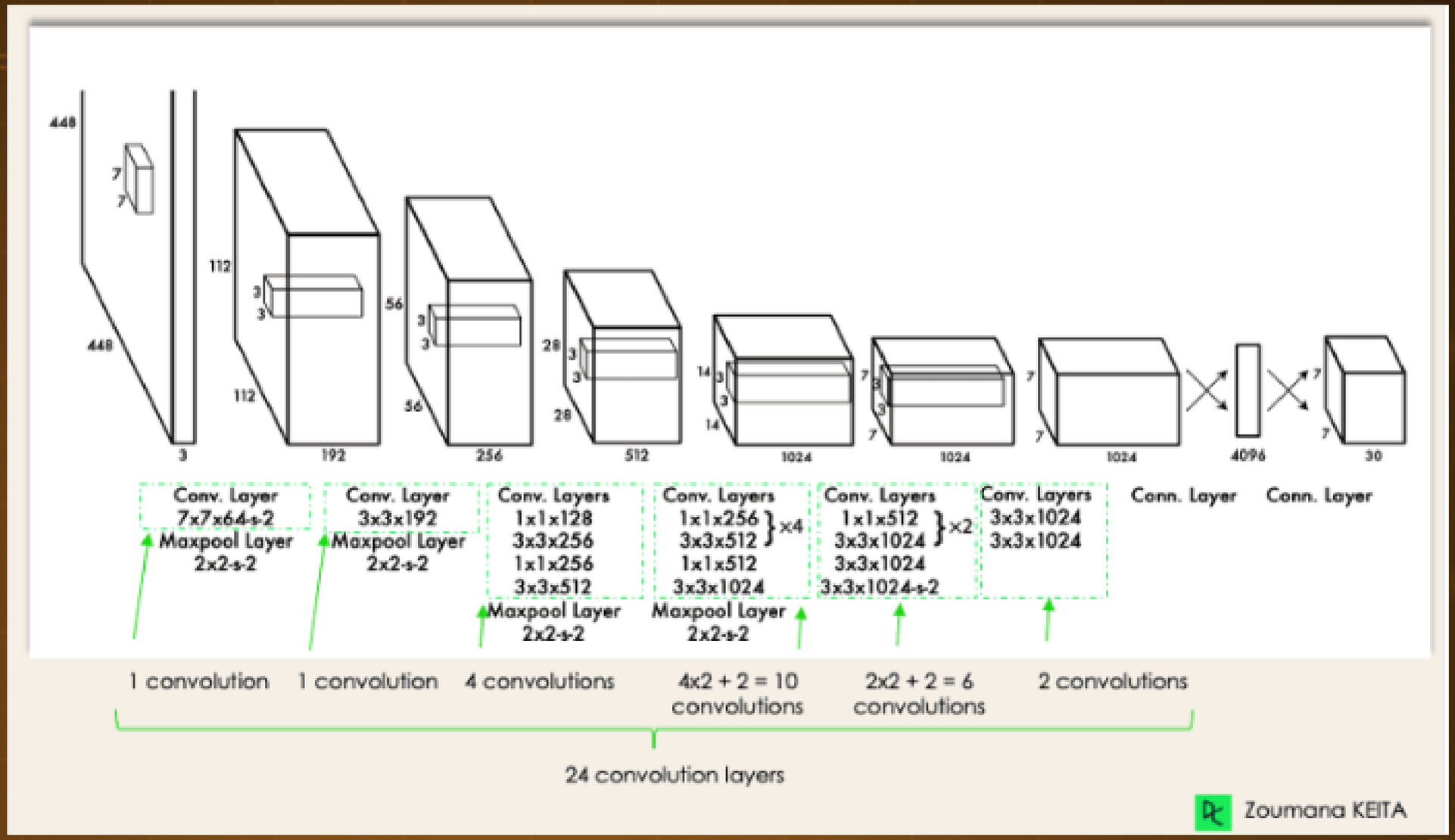
YOLO pushed a little further by providing a better generalization for new domains, which makes it great for applications relying on fast and robust object detection.

4- Open source

Making YOLO open-source led the community to constantly improve the model. This is one of the reasons why YOLO has made so many improvements in such a limited time.

YOLO Architecture

YOLO architecture has overall 24 convolutional layers, four max-pooling layers, and two fully connected layers.

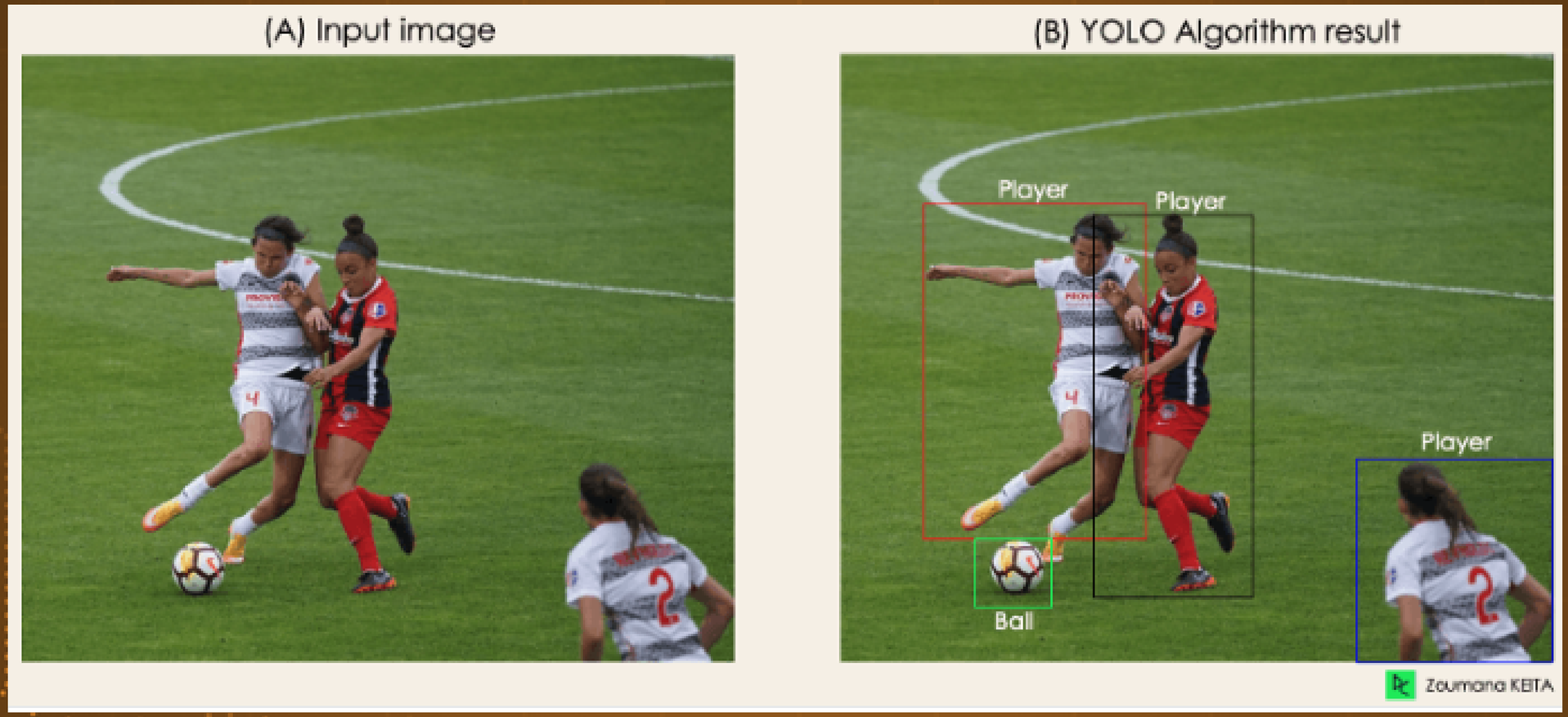


How Does YOLO Object Detection Work?

let ' s have a high-level overview of how the YOLO algorithm performs object detection using a simple use case.

We will understand the whole process of how YOLO performs object detection;

- " **How to get image (B) from image (A)** "



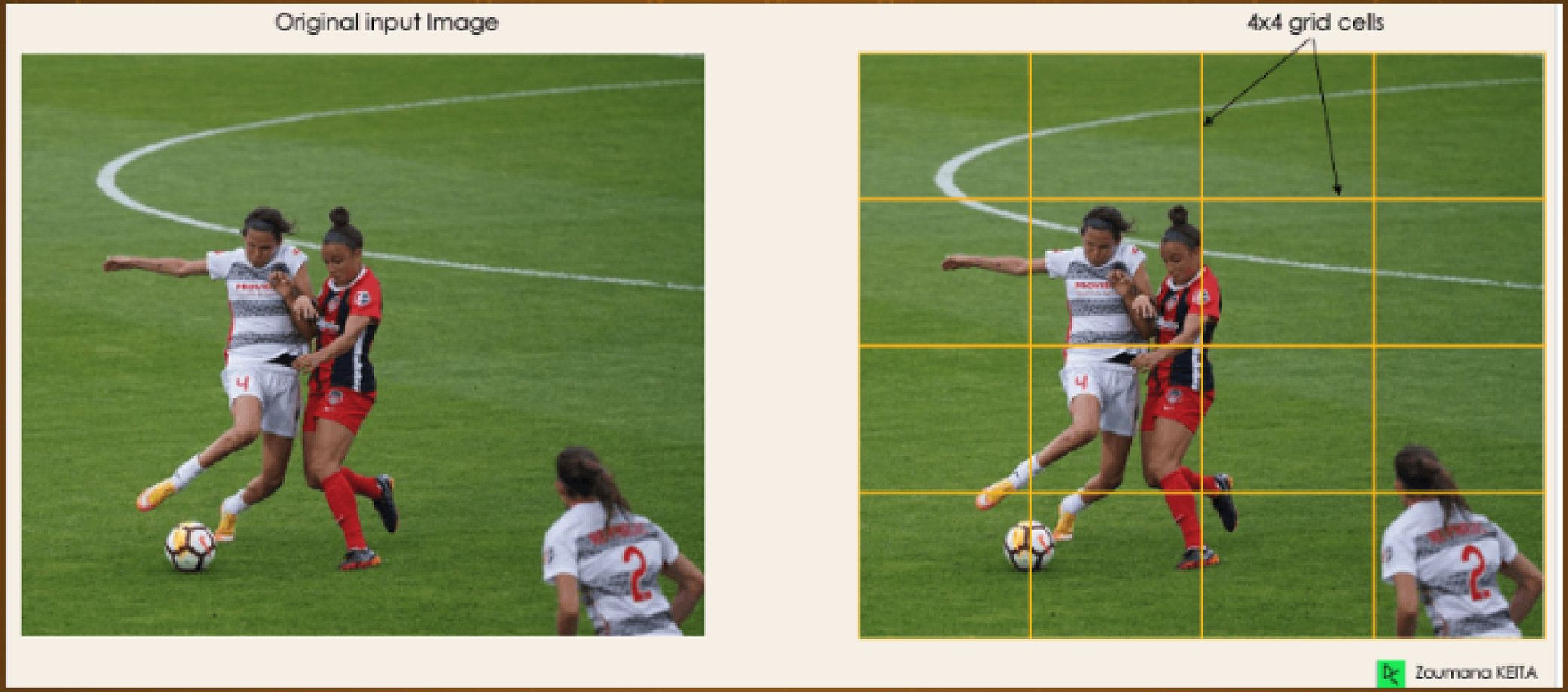
The algorithm works based on the following four approaches

- **Residual blocks**
- **Bounding box regression**
- **Intersection Over Unions or IOU**
- **Non-Maximum Suppression**

Let’s have a closer look at each one of them.

1- Residual blocks

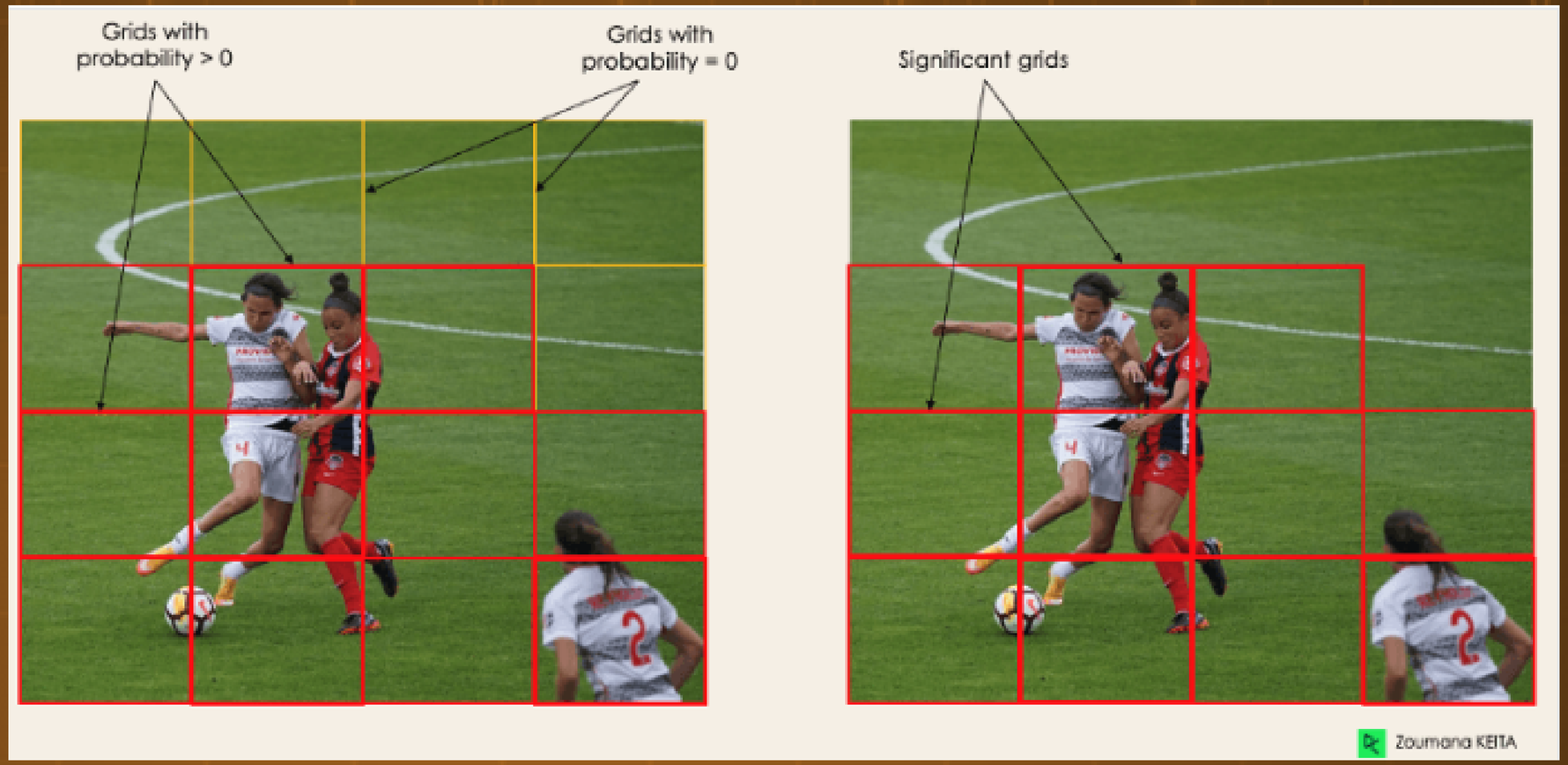
This first step starts by dividing the original image (A) into NxN grid cells of equal shape, where N in our case is 4 shown on the image on the right. Each cell in the grid is responsible for localizing and predicting the class of the object that it covers, along with the probability/confidence value.



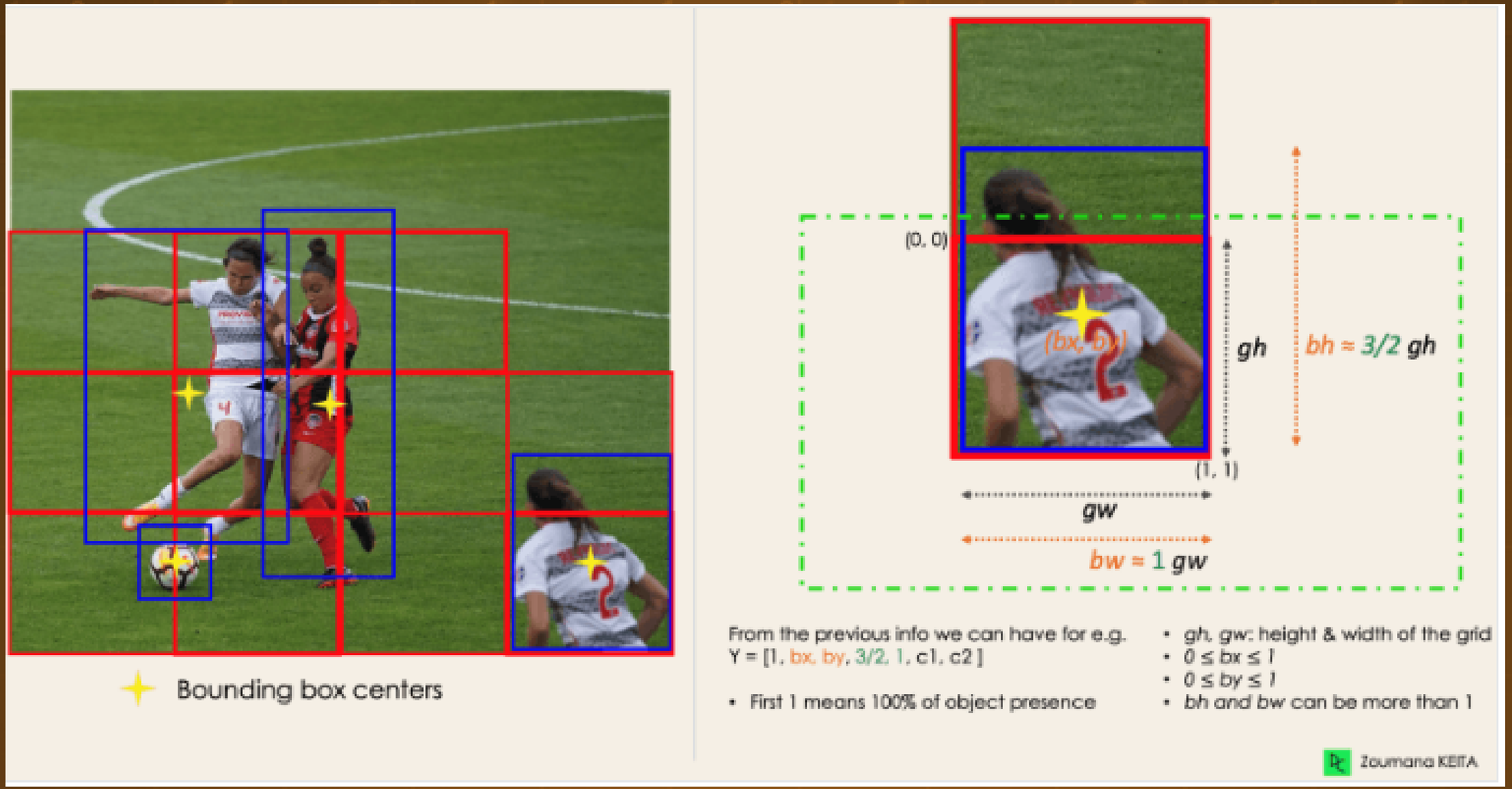
2- Bounding box regression

The next step is to determine the is to determine the bounding boxes which correspond to rectangles highlighting all the objects in the image

$$Y = \begin{pmatrix} pc, \\ bx, \\ by, \\ bh, \\ bw, \\ c1, \\ c2 \end{pmatrix}$$

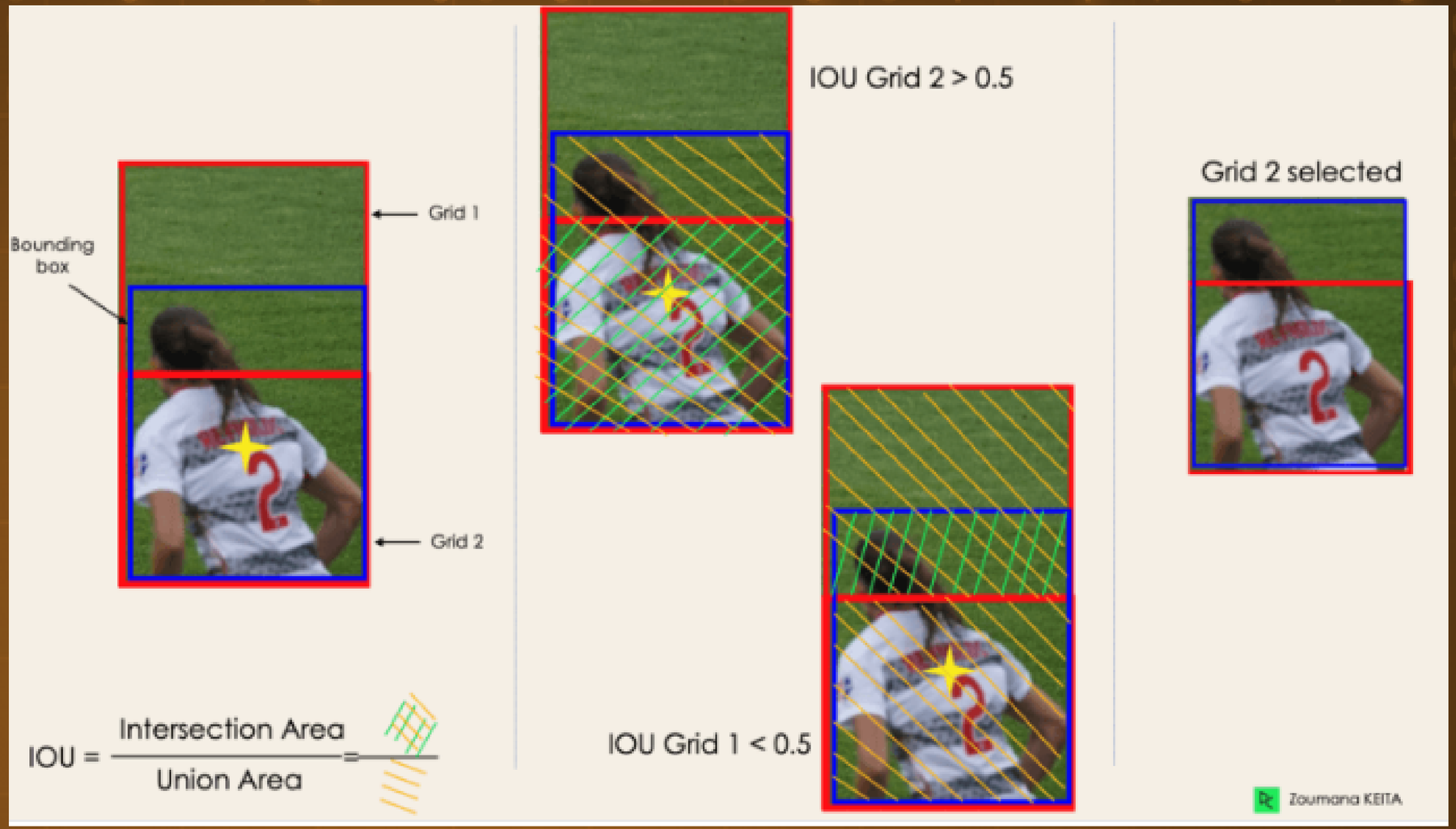


We determine the center of each bounder box, their width and height and the class of each image



3- Intersection Over Unions or IOU

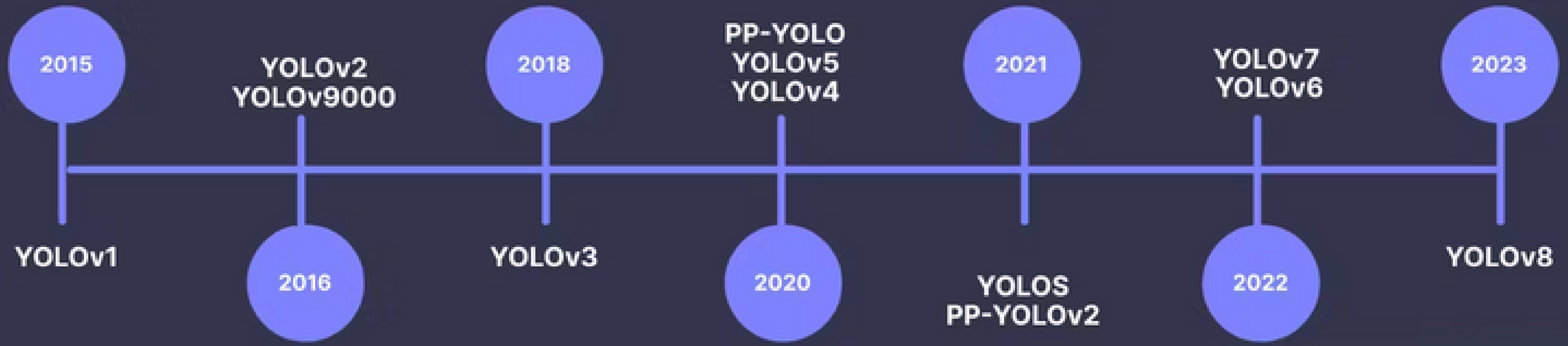
Below is an illustration of applying the grid selection process to the bottom left object. We can observe that the object originally had two grid candidates, then only “Grid 2” was selected at the end.



4- Non-Max Suppression or NMS

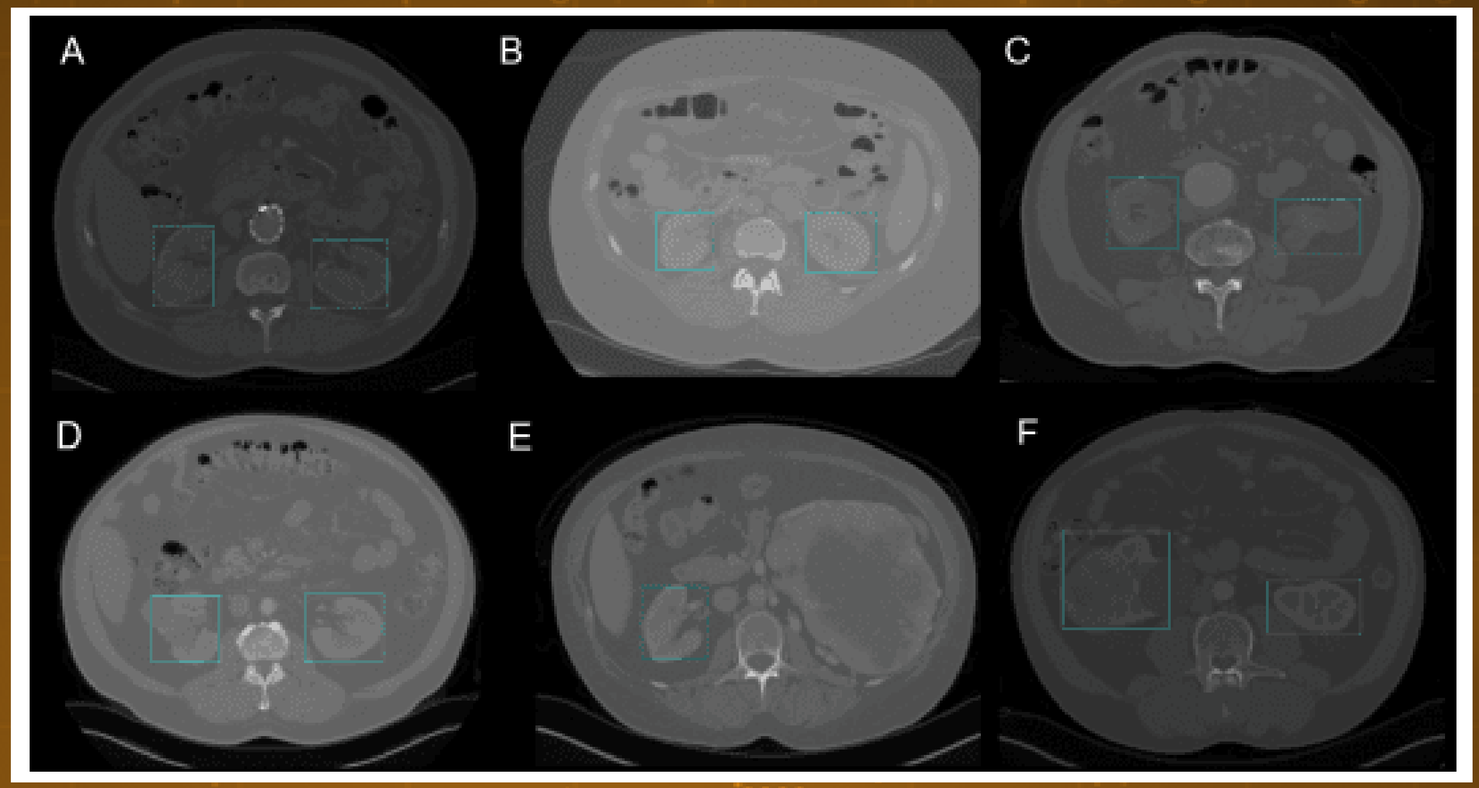
Setting a threshold for the IOU is not always enough because an object can have multiple boxes with IOU beyond the threshold, and leaving all those boxes might include noise. Here is where we can use NMS to keep only the boxes with the highest probability score of detection.

YOLO: Release Dates Timeline



Real world Applicatons

Healthcare



2D Kidney detection by YOLOv3 (Image from Kidney Recognition in CT using YOLOv3)

Agriculture

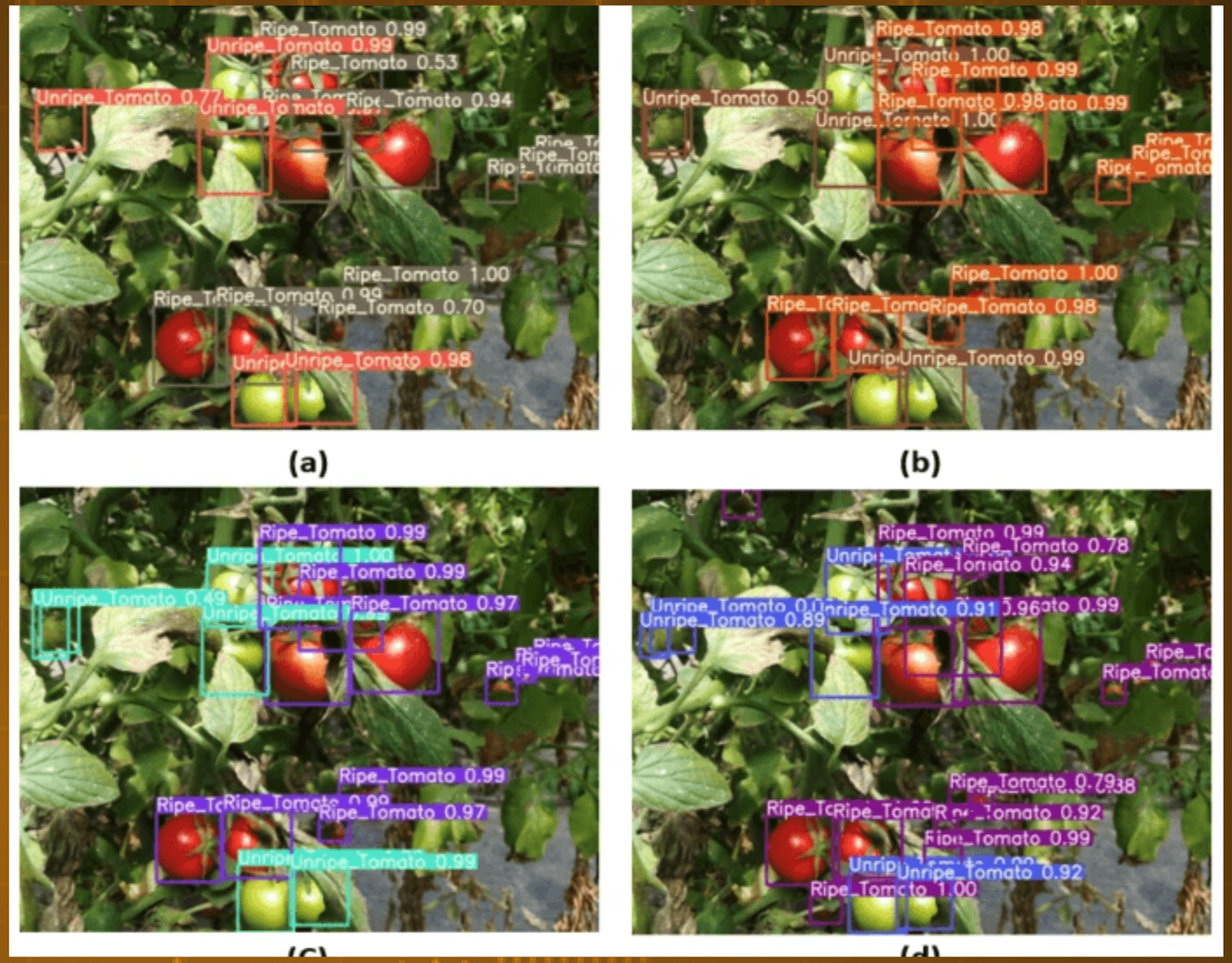
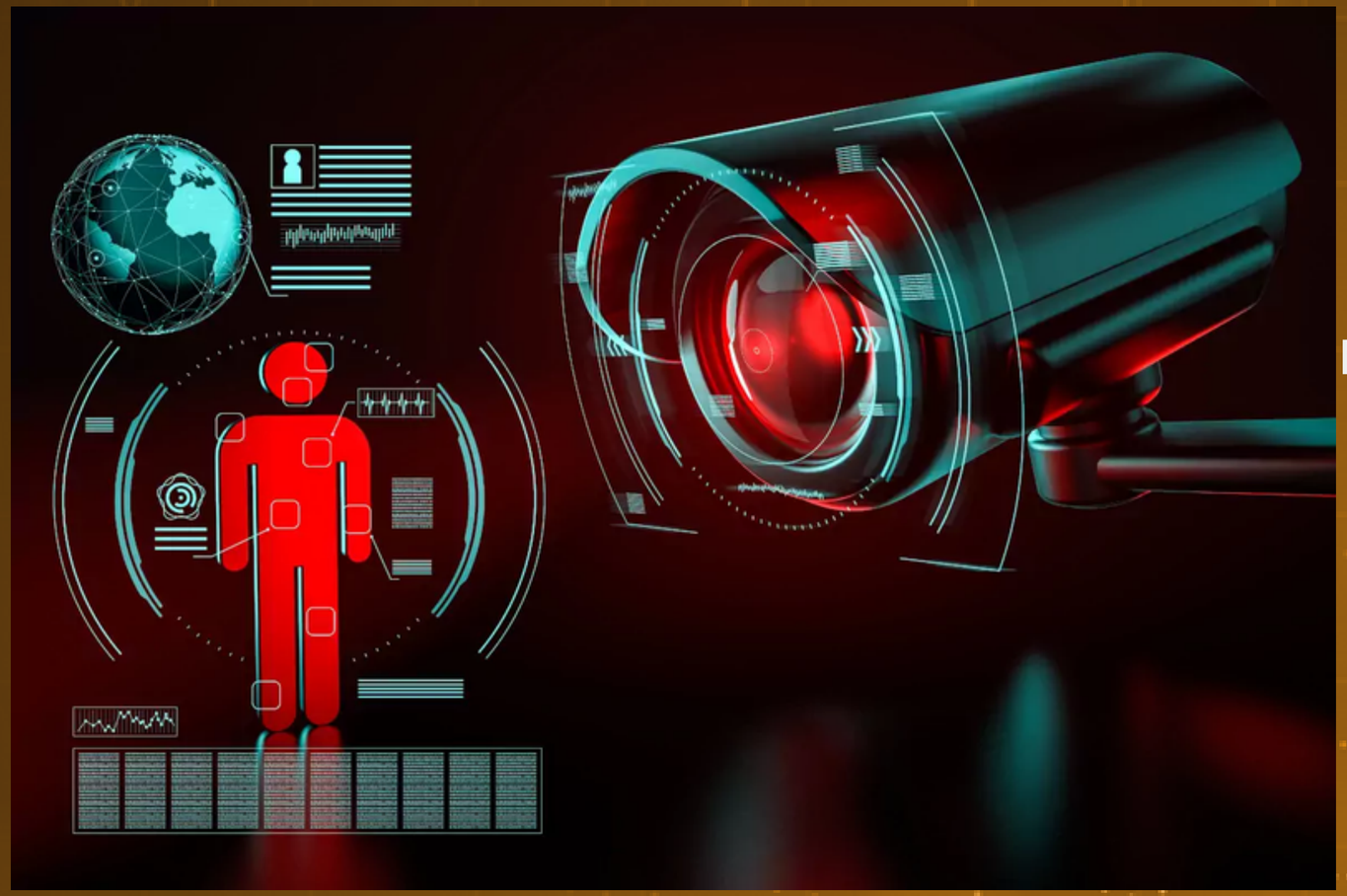


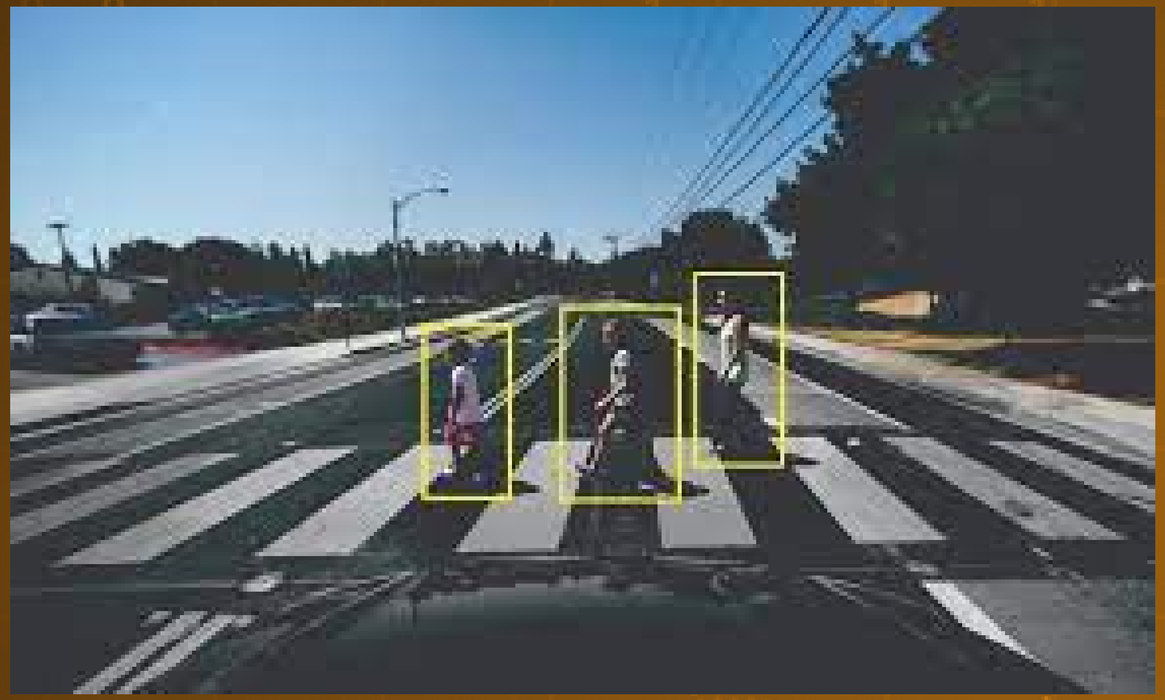
Image from Tomato detection based on YOLOv3 framework

Security surveillance



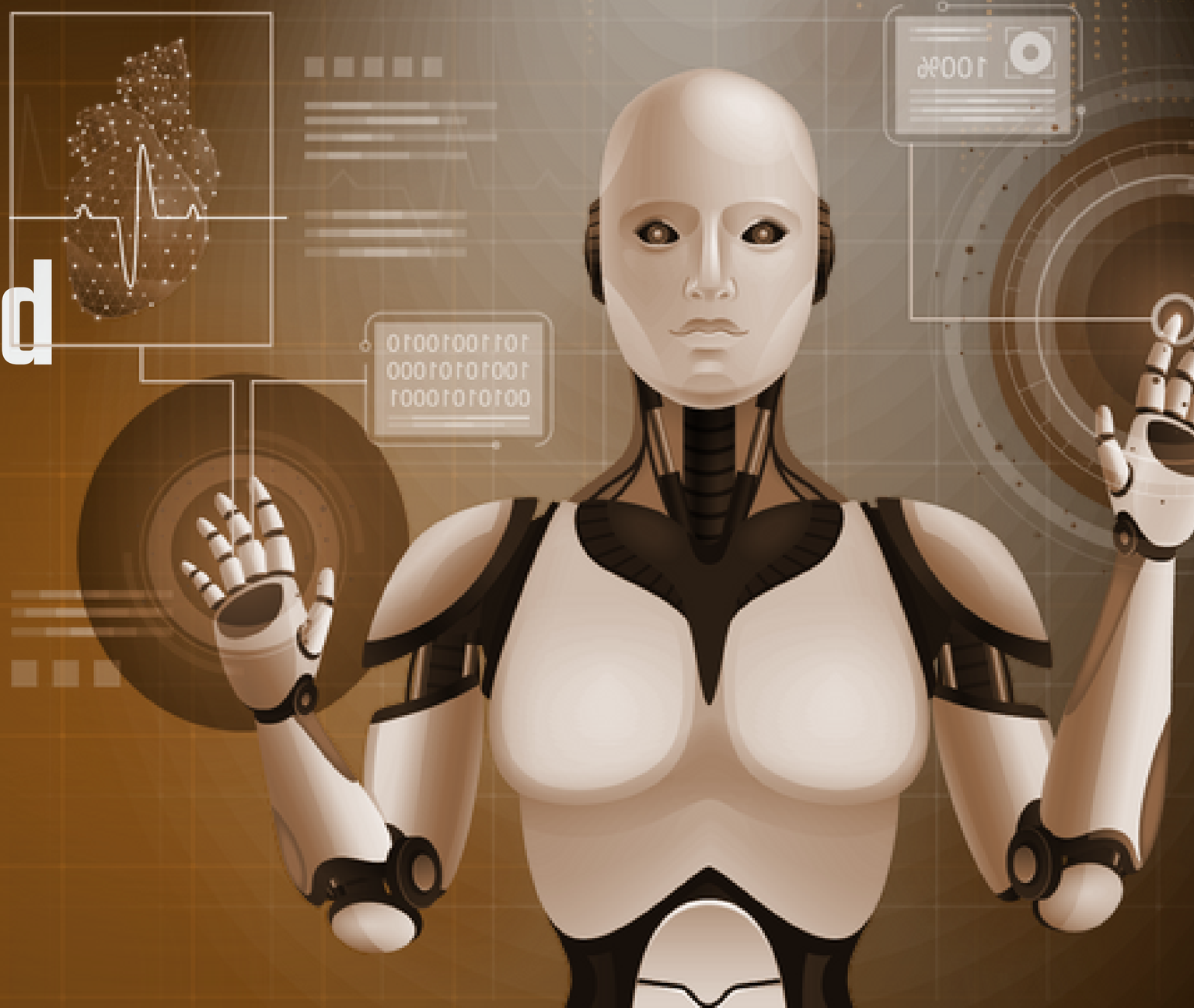
Even though object detection is mostly used in security surveillance, this is not the only application. YOLOv3 has been used during covid19 pandemic to estimate social distance violations between people.

Self car driving



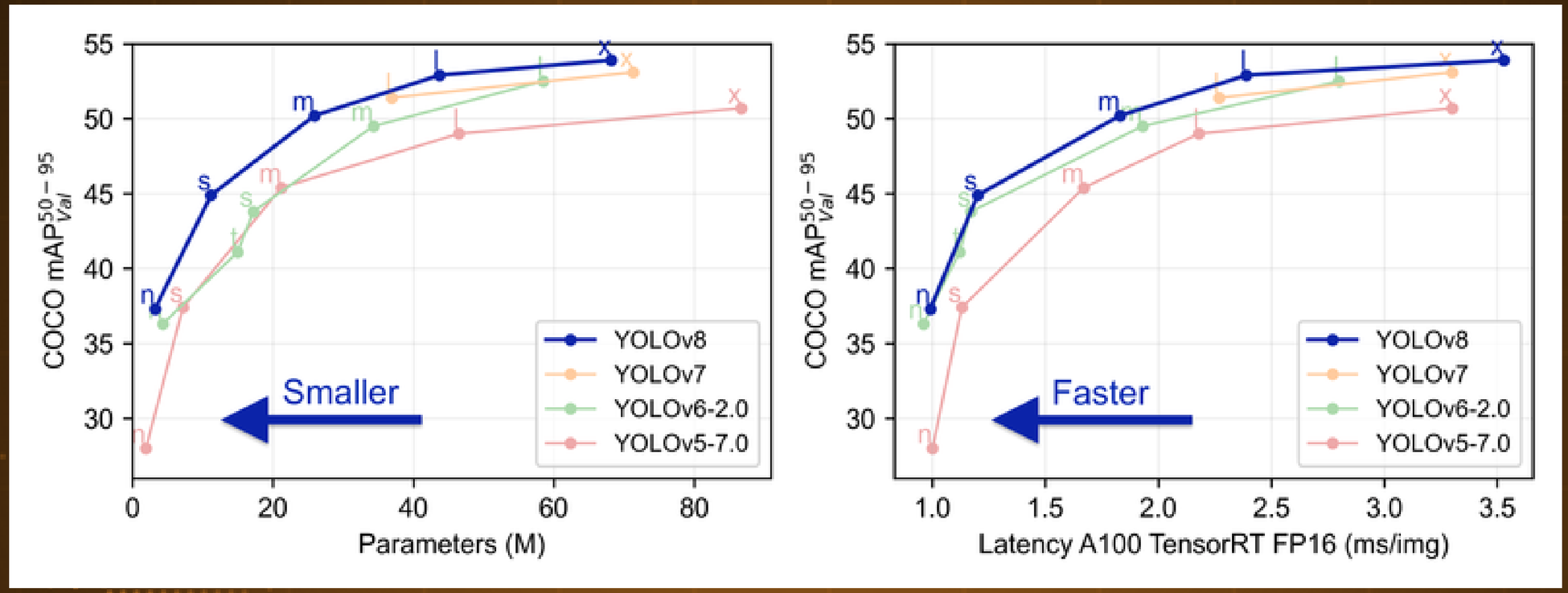
The real-time aspect of YOLO makes it a better candidate compared to simple image segmentation approaches.

Advantages and Limitations



Advantages

- **Improved Accuracy**



mAP: High Mean Average Precision

Low Latency

Advantages

- **Faster Speeds**

One of the key advantages of YOLOv8 is its ability to achieve fast inference speeds without compromising accuracy. By implementing advanced optimization techniques, the model can rapidly process input data, allowing for real-time object detection in dynamic environments. This accelerated performance opens up possibilities for applications that require quick response times, such as autonomous vehicles and live video analysis.

Advantages

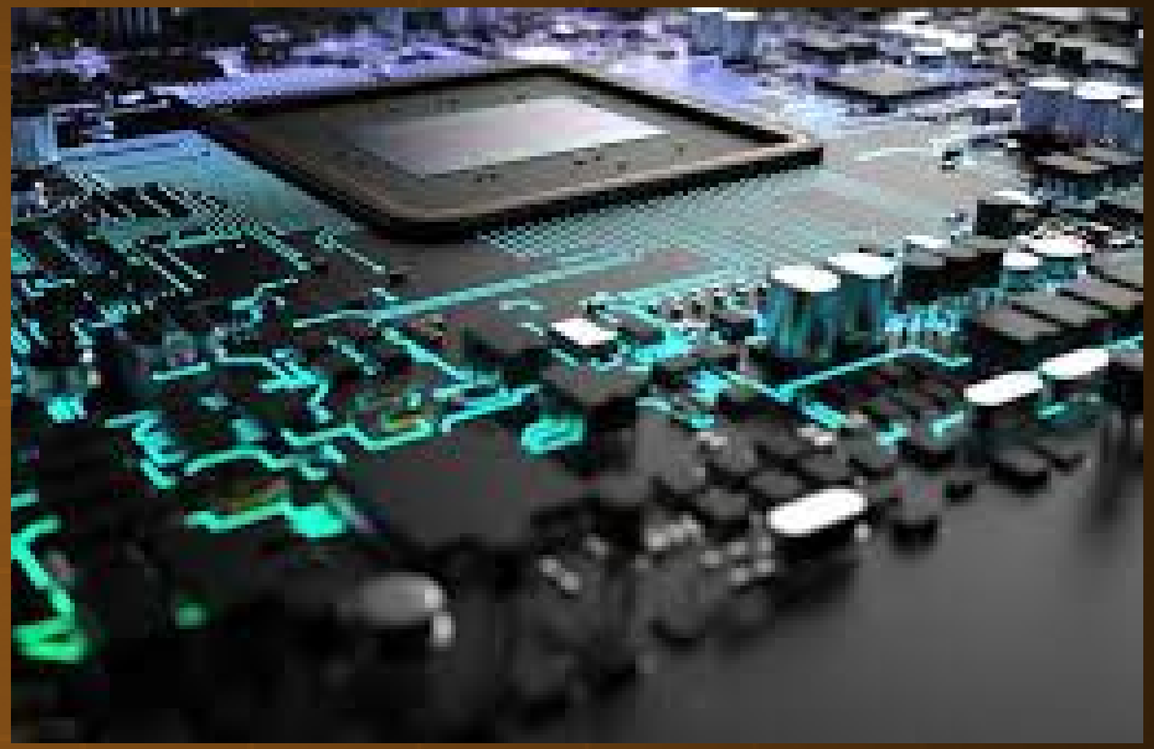
- **Fewer Parameters**

YOLOv8 is designed to be more efficient and streamlined, requiring fewer parameters compared to previous versions. This reduction in parameters results in a more lightweight model, reducing computational complexity and memory requirements.

Limitations

- **Resource Intensive**

YOLOv8 requires substantial computational resources, including GPUs, Users with limited hardware capabilities may face difficulties in using YOLOv8 efficiently.

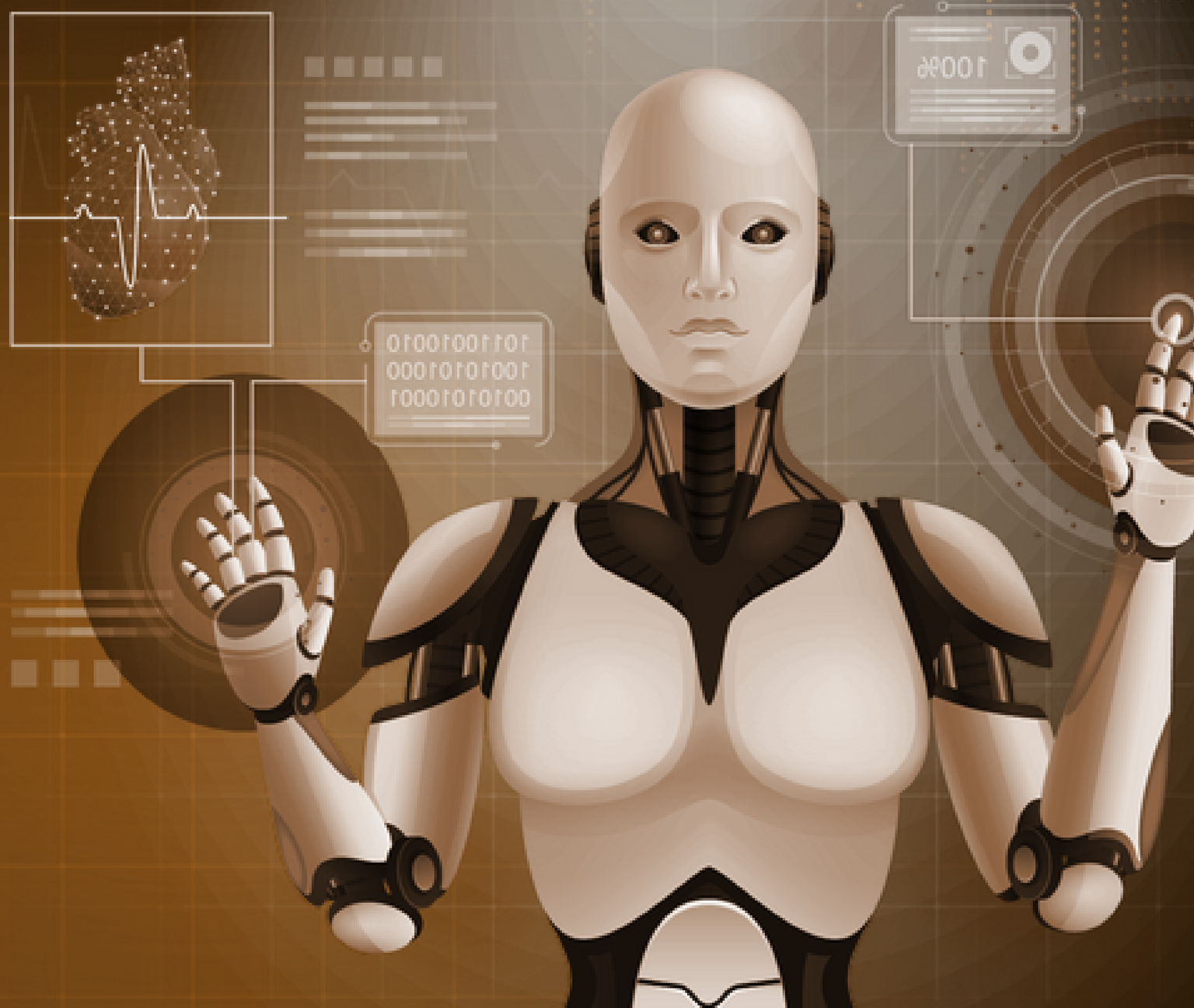


Limitations

- **Advanced Customization**

While YOLOv8 offers various pre-defined configurations and customization options, advanced customization and fine-tuning may require additional coding skills and computer vision expertise.

Application



CONCLUSION

YOLOv8 revolutionizes the field of object detection with its cutting-edge advancements. Its speed, accuracy, and flexibility make it a go-to choice for professionals in the computer vision industry. With ongoing advancements and emerging applications, YOLOv8 is set to shape the future of object detection and drive the evolution of computer vision technology.

THANKS!