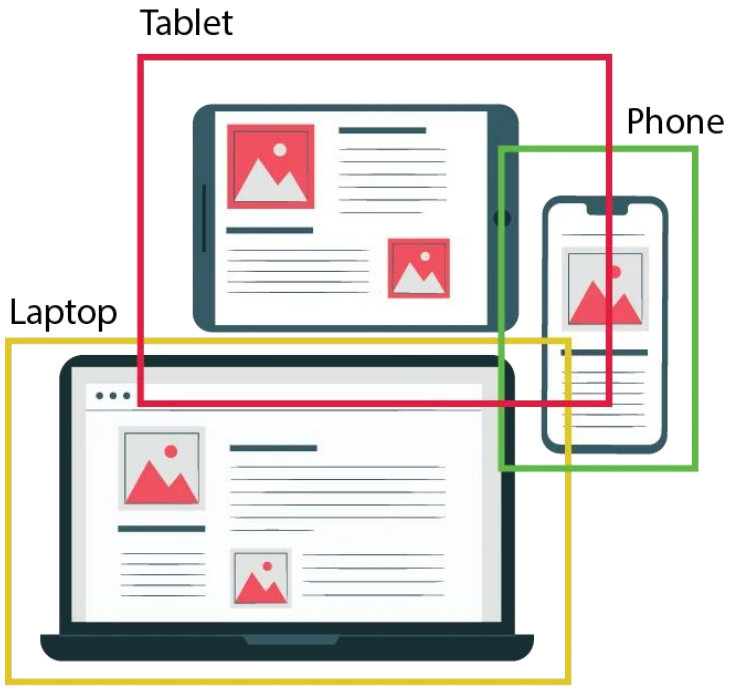
YOLO Network

Object Detection in real time



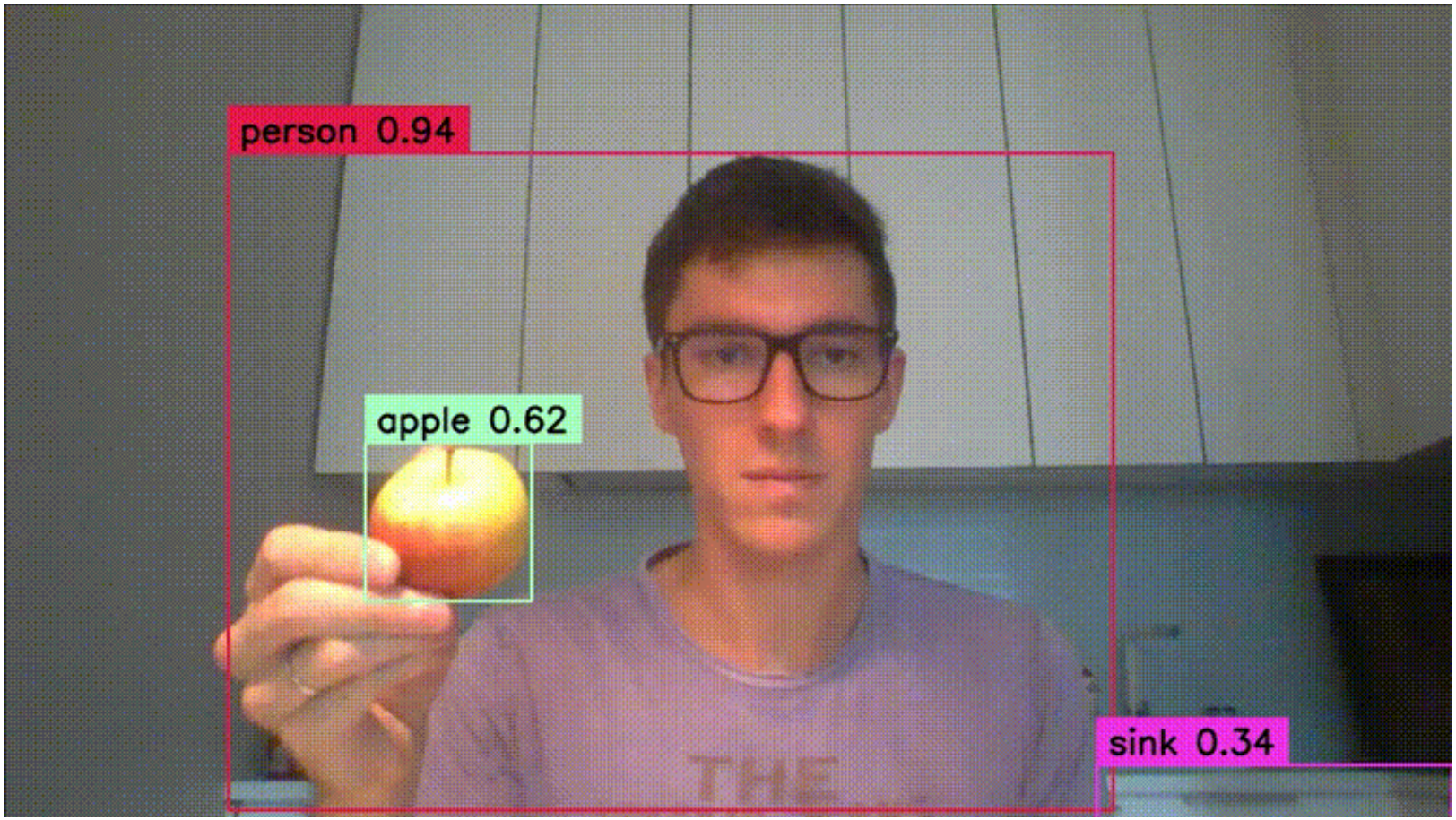
*Name: Akhmedov **Yusufjon** Muhammad ugli*

*Major: **Computer Engineering***

*Student Number: **202438404***

WHAT IS OBJECT DETECTION ACTUALLY?

Object detection is a computer vision technique for identifying and localizing objects within an image or a video.



WHAT IS YOLO Network

WHY IS YOLO popular

HOW DOES YOLO work

QUESTION MARK

- YOLO – (You Only Look Once) real-time object detection algorithm

- 2015 – Start of history by Joseph Redmon, Santosh Divalla, Ross Girshick, Ali Farhadi – YOLO v1




Extremely fast – the mean Average Precision (mAP) is 45 FPS ~ 91 FPS
Also compared to other detectors (SSD, R-CNN, etc.)

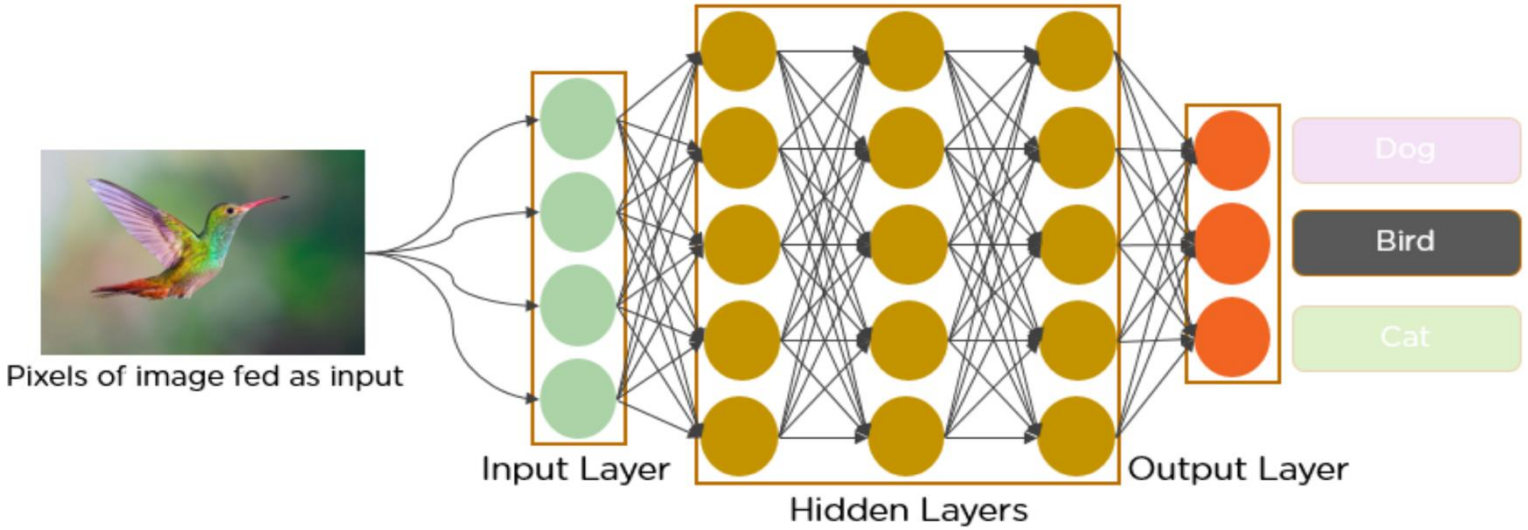
High detection accuracy
with few background errors

Popular?

Open source – publicly available for everyone

Forking - Many improvements in a limited time
GitHub Forks

- A DL model - identifies and locates multiple objects in an image using a **single forward pass** through a convolutional neural network (CNN).



Example of simple CNN network processing of image of bird!

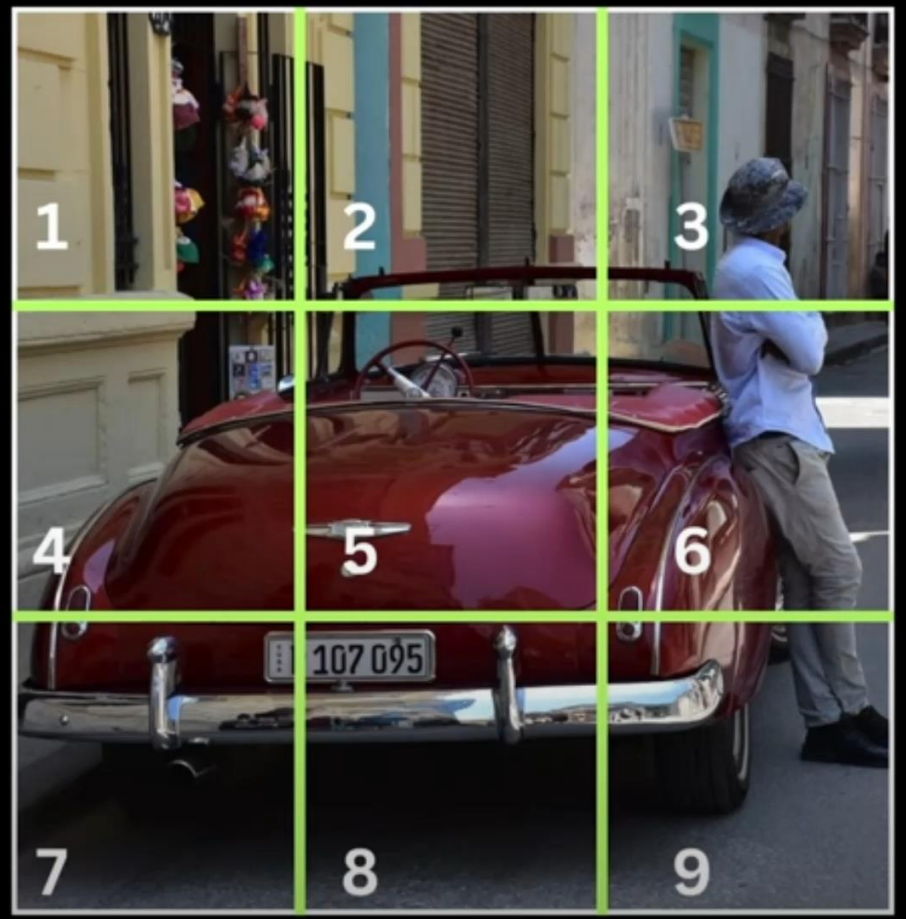
YOLOv1 was a groundbreaking start — simple but limited.

1. It uses a single CNN to divide the image into a 7×7 grid and predicts bounding boxes and classes in one pass.

2. It’s fast but struggles with small or overlapping objects and has limited accuracy.

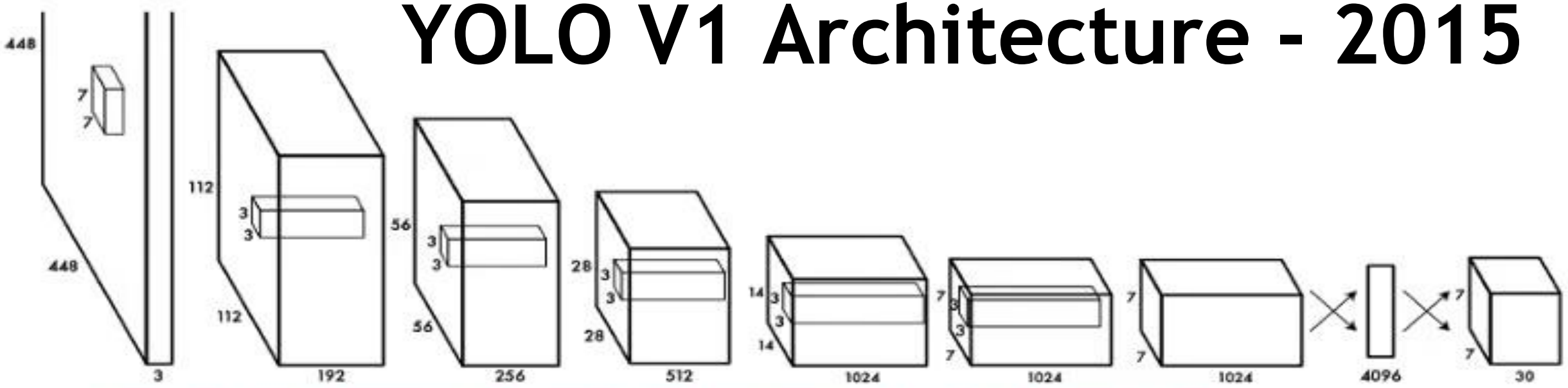
YOLO

- Divide image into SxS grid cells



Example	Implementation
S = 3	S = 7

YOLO V1 Architecture - 2015



Conv. Layer 7x7x64-s-2	Conv. Layer 3x3x192	Conv. Layers 1x1x128 3x3x256 1x1x256 3x3x512	Conv. Layers 1x1x256 3x3x512 ×4 1x1x512 3x3x1024	Conv. Layers 1x1x512 3x3x1024 ×2 3x3x1024 3x3x1024-s-2	Conv. Layers 3x3x1024 3x3x1024	Conn. Layer	Conn. Layer
Maxpool Layer 2x2-s-2	Maxpool Layer 2x2-s-2	Maxpool Layer 2x2-s-2	Maxpool Layer 2x2-s-2				
1 convolution	1 convolution	4 convolutions	4x2 + 2 = 10 convolutions	2x2 + 2 = 6 convolutions	2 convolutions		

24 convolution layers

Zoumana KEITA

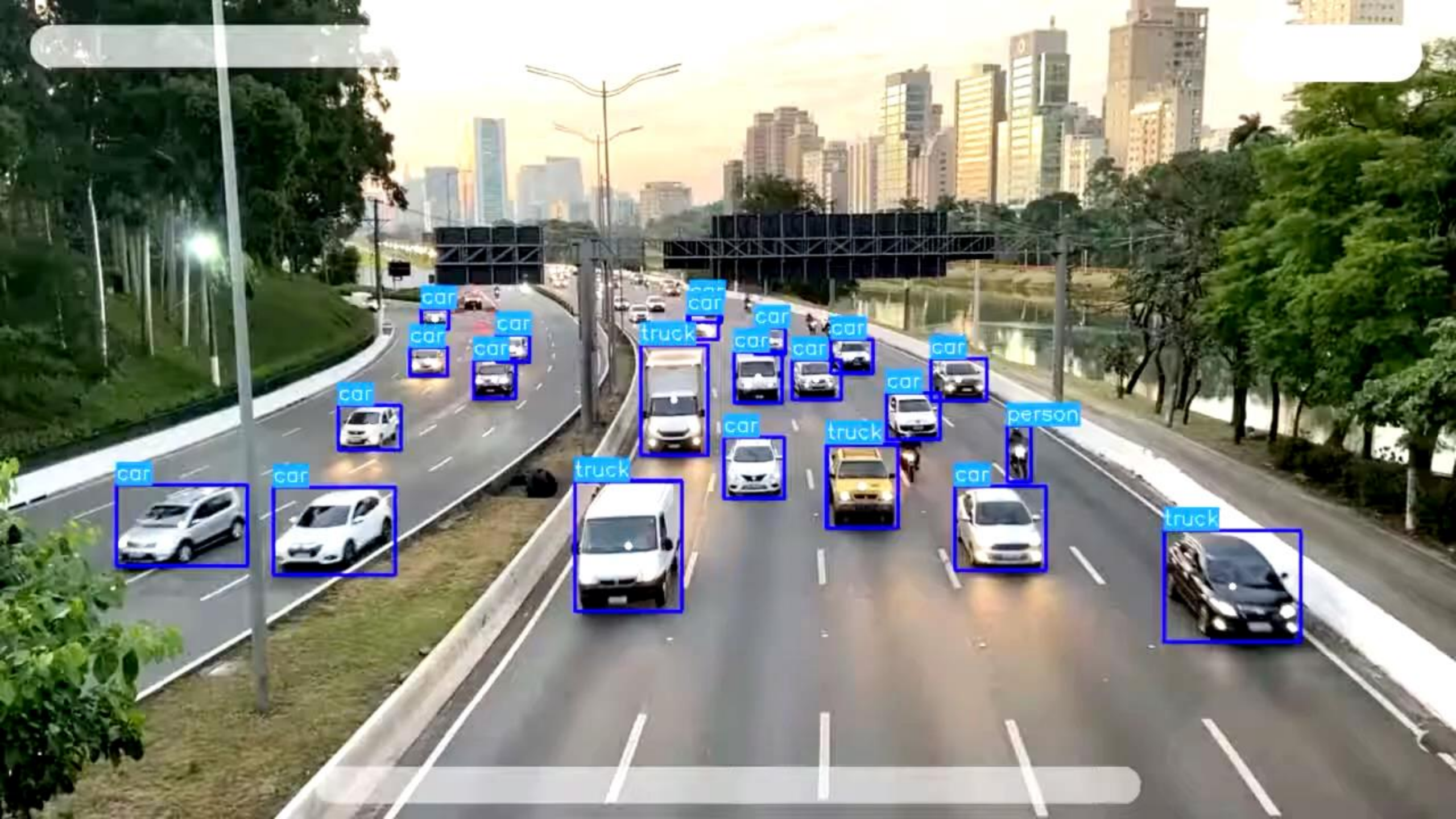
https://paperswithcode.com/method/yolov1

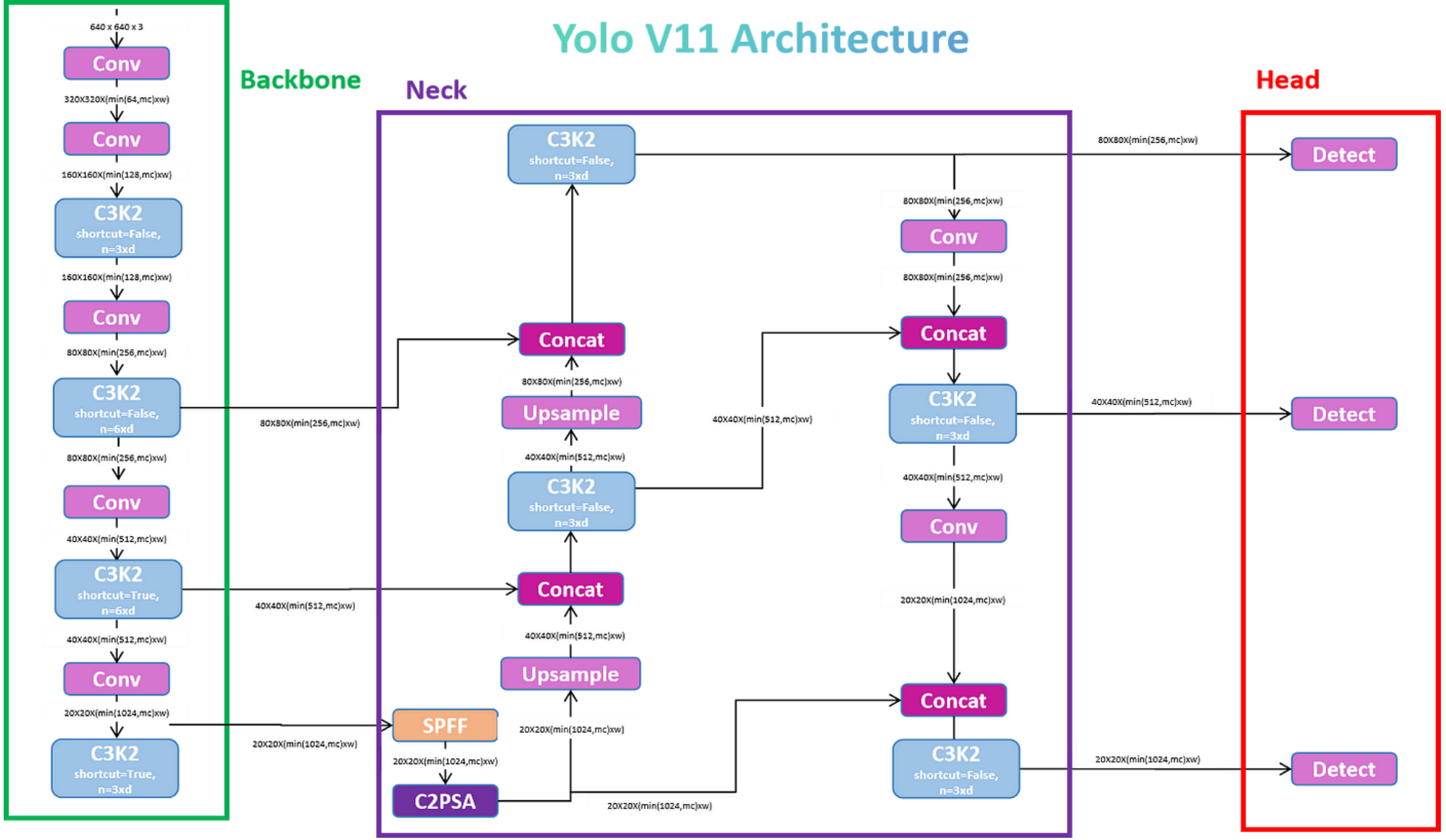
YOLOv8 (your "YOLOv11") is a powerful, modern, and production-ready system capable of high-precision detection across tasks and devices.

1. YOLOv8 is a modern, high-performance object detection model built by Ultralytics.

2. It uses advanced CNN architecture with multi-scale detection and anchor-free methods for better accuracy.

3. YOLOv8 supports detection, segmentation, and classification, and is optimized for real-time use on various devices.


car
car
car
car
car
car
truck
car
car
car
car
car
car
car
truck
car
truck
person
car
truck

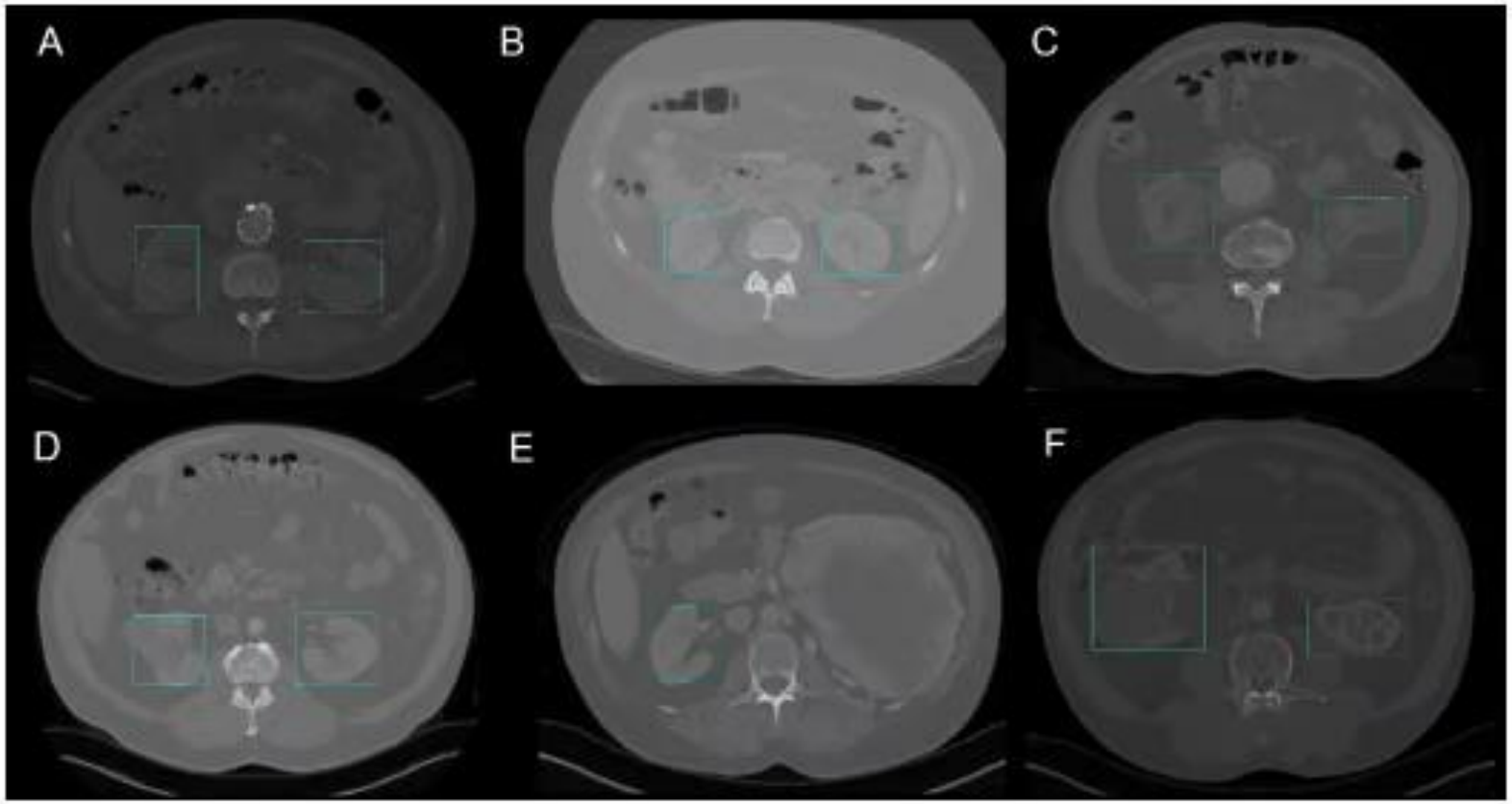


Yolov8 -> Ultraliytcs 2024 -> “Yolo 11”

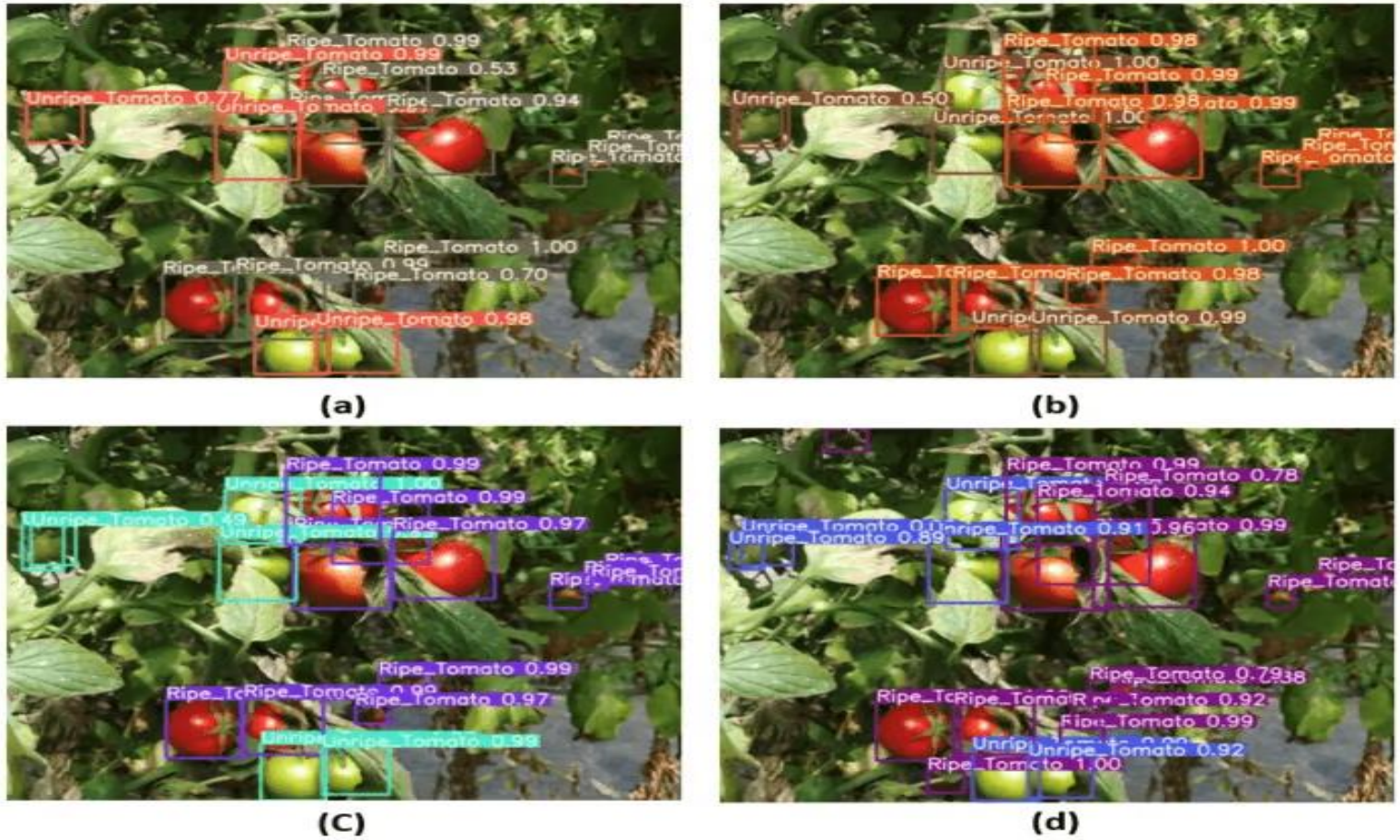
https://www.ultralytics.com/blog/ultralytics-yolo11-has-arrived-redefine-whats-possible-in-ai

YOLO Applications
Real life examples

Healthcare - Kidney Detection in CT using YOLO v3e



Agriculture - Harvesting robots are vision-based robots that were introduced to replace manual picking of fruits



Military - YOLO is used in drones to detect and track people, vehicles, weapons, and infrastructure in real time.