Object Detection using the YOLOS model

Unified Real-Time object Detection

B-Tech Project

Under:- **Dr. Akash Yadav**

By:-**Abhishek Yadav**
21IT3003

Table of Contents

Objective

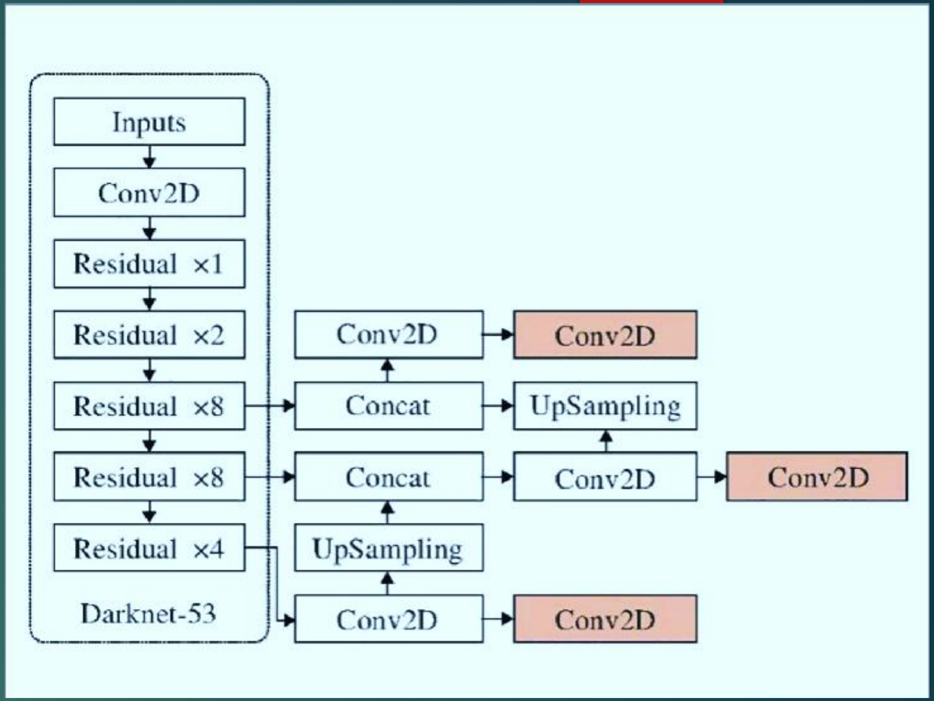
The objective of the provided code is to demonstrate object detection using the YOLOS (You Only Look Once) model. Object detection is a computer vision task that involves identifying and locating objects of interest within an image. The YOLOS model, implemented here through the Ultralytics library, is a popular algorithm for real-time object detection.

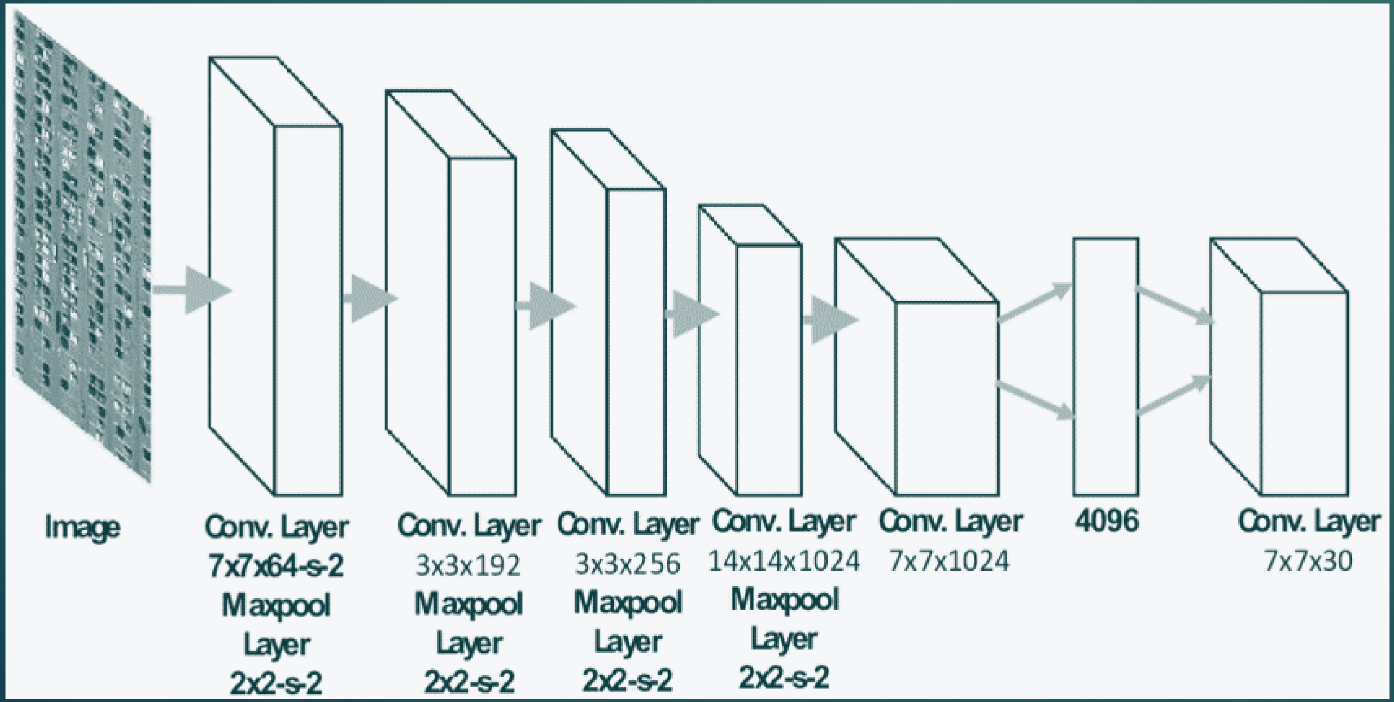
We'll cover image preprocessing, the selecion and fine-tuning of the YOLOS model, as well as post-processing steps for visualizaton and result interpretion.

Network Architecture:

- YOLOv3, one of the popular versions, uses a deep neural network with a Darknet-53 backbone.

- The Darknet-53 architecture is a convolutional neural network (CNN) that extracts features from the input image.





How Does It Work?

- S*S*(C+5)+AB

1. Grid Division:

- The input image is divided into a grid, typically, a fixed-size grid like 19x19 or 45x45.

2. Bounding Box Prediction

3. Object Confidence

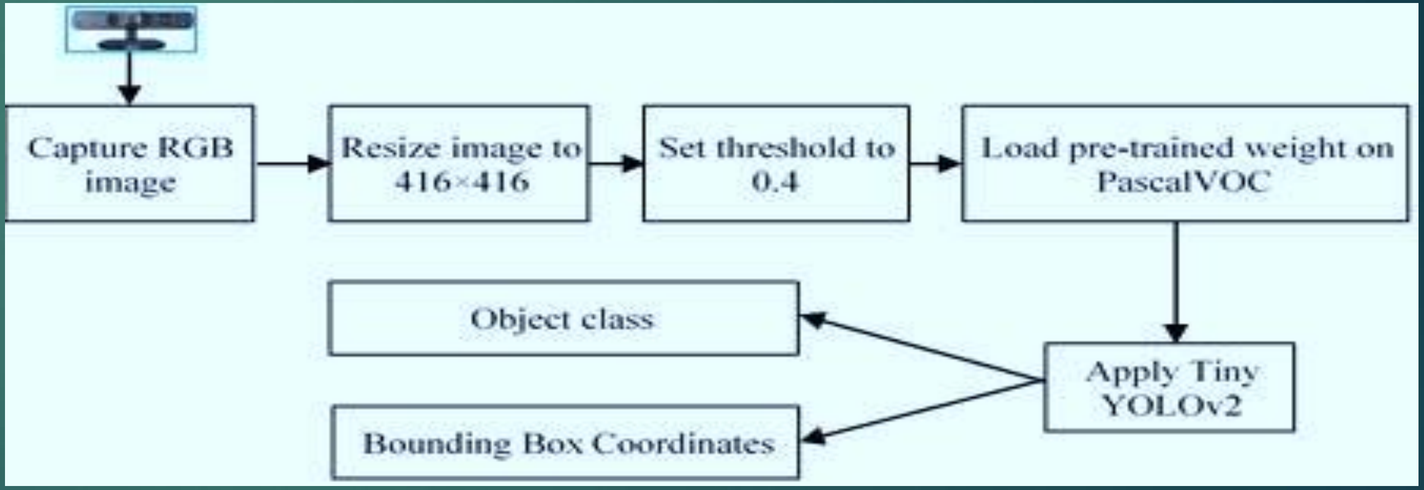
4. Class Prediction

5. Anchor Boxes

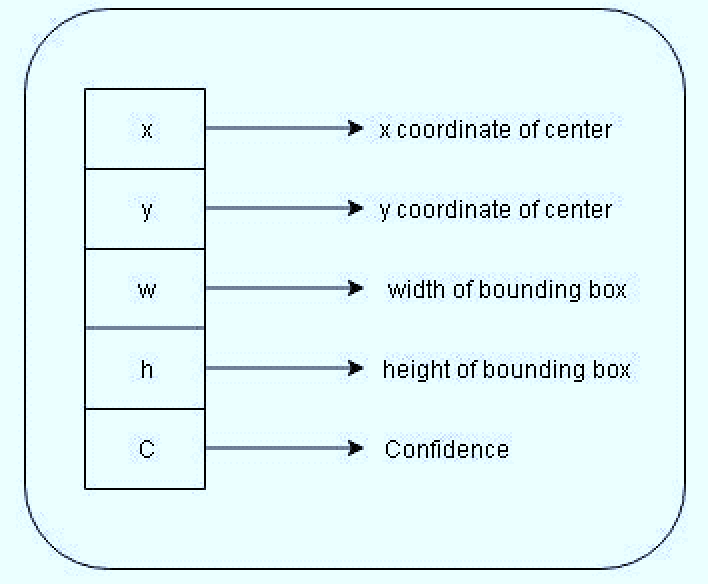
6. Single Forward Pass

7. Non-Maximum Suppression (NMS)

8. Output:



- *C = Pr(object) * IoU*
- IoU: Intersection over Union between the predicted box and the ground truth.
- If no object exists in a cell, its confidence score should be zero.



The loss function defined in YOLO as follows

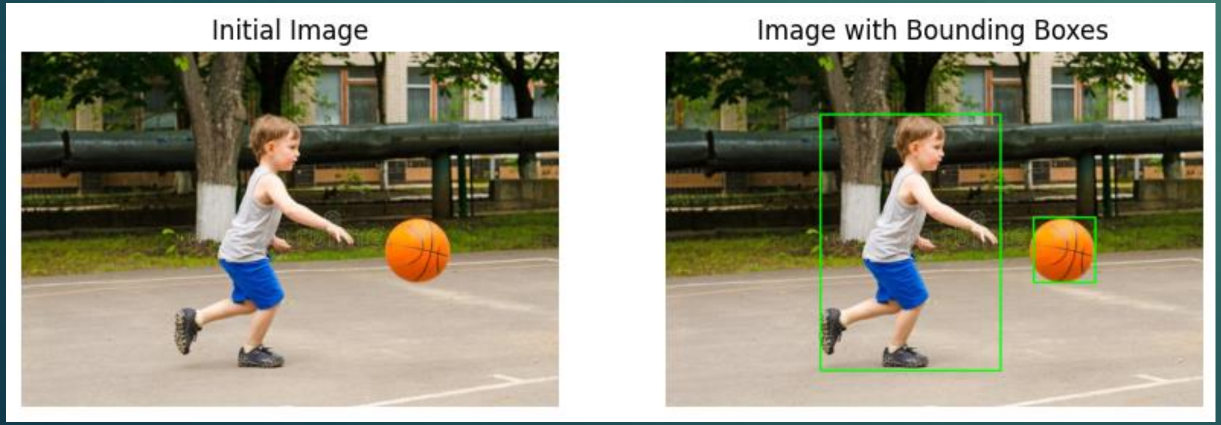
$$
\begin{aligned}
&\lambda_{\textbf{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\
&+ \lambda_{\textbf{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right] \\
&+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(C_i - \hat{C}_i \right)^2 \\
&+ \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i \right)^2 \\
&+ \sum_{i=0}^{S^2} \mathbb{1}_{i}^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2
\end{aligned}
$$

In this model, we taken 5 lambda coord. and 0.5 lambda noobj

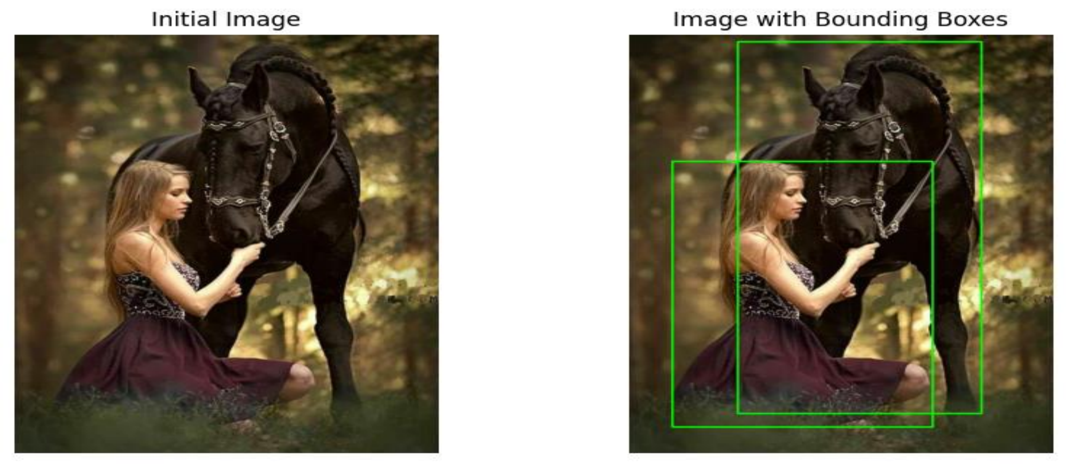
Results and discussion

Detected dog with confidence 0.989 at location [164.05, 1.45, 646.14, 662.74]
Detected cat with confidence 0.996 at location [569.34, 107.56, 975.35, 647.13]

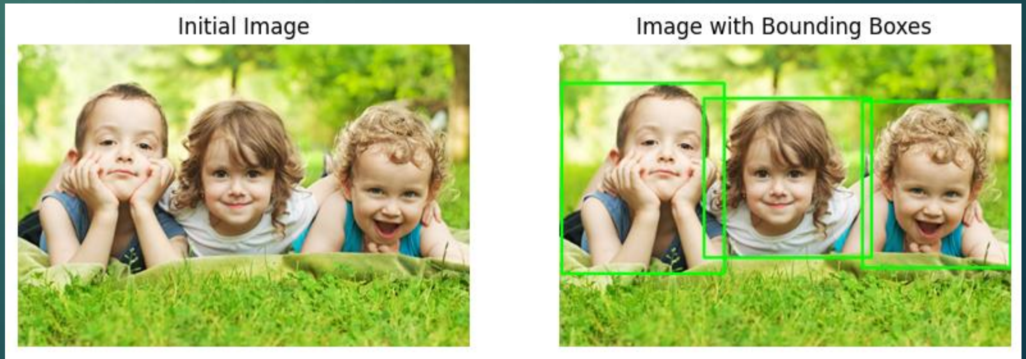
Detected person with confidence 0.993 at location [230.81, 95.43, 498.0, 476.48] Detected sports ball with confidence 0.989 at location [547.14, 248.66, 639.59, 345.28]



Detected horse with confidence 0.965 at location [144.58, 12.68, 468.57, 674.67]
Detected person with confidence 0.995 at location [57.55, 225.47, 403.14, 698.4]

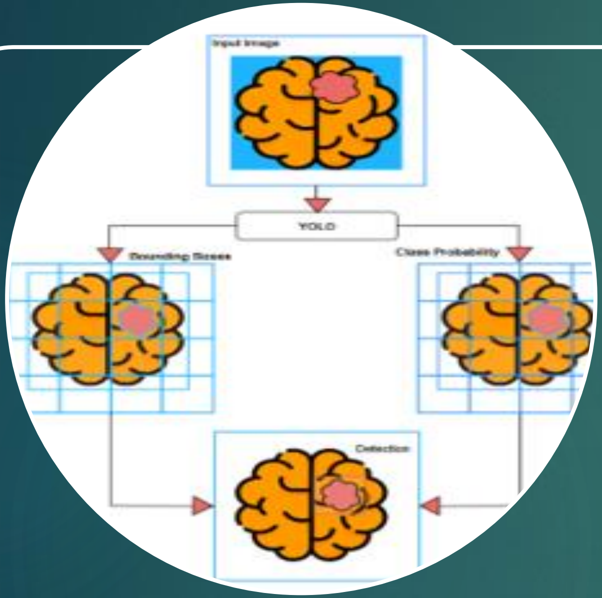


Detected person with confidence 0.987 at location [2.47, 38.71, 163.55, 227.68]
Detected person with confidence 0.994 at location [143.87, 53.64, 309.17, 211.93]
Detected person with confidence 0.972 at location [300.14, 56.8, 447.88, 221.98]



YOLO Applications

- Object detection in images and videos
- Industrial automation



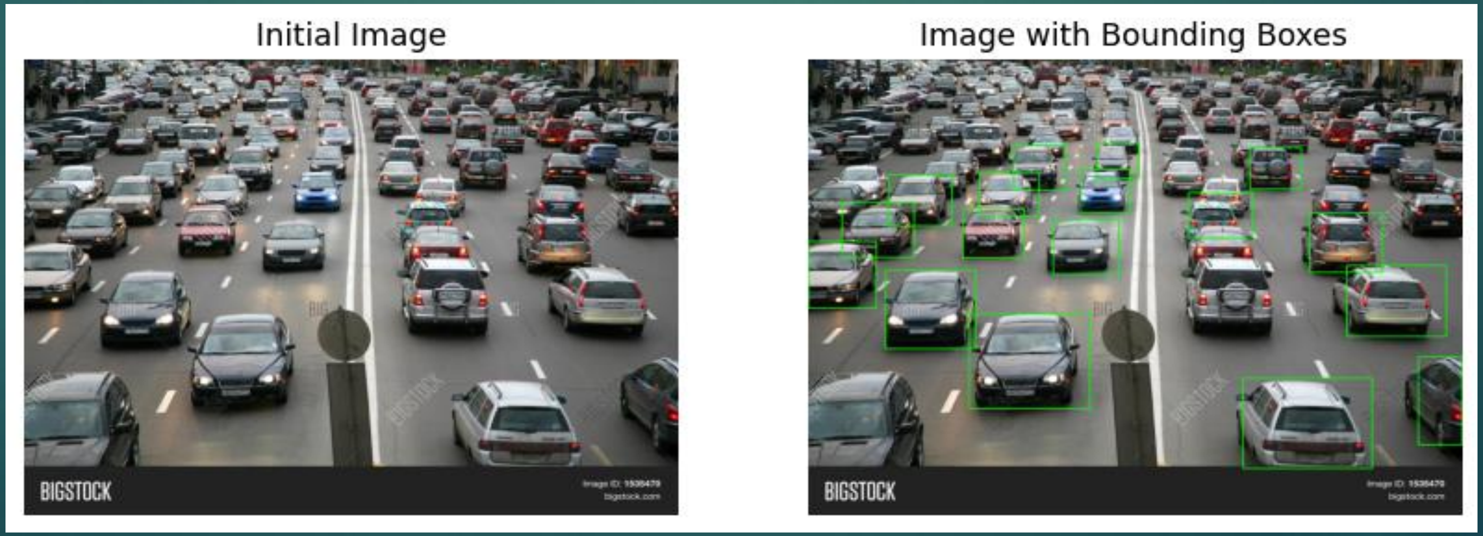
Medical image analysis

Autonomous vehicles

Surveillance systems

Limitations Of YOLO

- Spatial constraints on bounding box predictions as each grid cell only predicts two boxes and can have only one class.
- It is difficult to detect small objects that appear in groups.
- It struggles to generalize objects in new or unusual aspect ratios as the model learns to predict bounding boxes from data itself.



Conclusion

- In conclusion, this project has allowed us to explore the fascinating world of object detection using the YOLOS model. We've seen how preprocessing, model selection, fine-tuning, and post-processing come together to create an effective object detection pipeline.

Future direction

Opportunities for improvement in YOLO models

Integration with other technologies (e.g., deep reinforcement learning)

Emerging trends in object detection

Thank You