Assignment No 7

Name : Samiksha More

PRN : 22420127

Roll : 381070

Problem Statement: Object Detection using YOLO and Pretrained Model

Objective

* To implement object detection using the YOLO (You Only Look Once) algorithm.
* To leverage pretrained models (e.g., YOLOv3, YOLOv5, YOLOv8) for high accuracy and reduced training time.
* To analyze the efficiency of YOLO in detecting multiple objects in real-time scenarios.

Theory

YOLO is a deep learning-based object detection framework that treats detection as a regression problem. Unlike traditional approaches (R-CNN, Fast R-CNN, etc.) that perform region proposals and classification separately, YOLO divides the image into grids and simultaneously predicts:

* Bounding boxes (location of objects).
* Confidence scores (probability that the box contains an object).
* Class probabilities (the category of the object).

Key theoretical aspects:

* Single forward pass → faster detection.
* Grid-based approach → better for real-time performance.
* Pretrained models trained on large datasets like COCO or ImageNet improve accuracy and generalization.

Methodology

1. Dataset Selection
   * Use pretrained YOLO weights trained on COCO dataset (80 object categories).
2. Model Loading
   * Load pretrained YOLO model and configuration files.
3. Preprocessing
   * Input images resized to the required resolution (e.g., 416×416).
   * Normalization applied for consistent feature extraction.
4. Inference
   * Pass the image through YOLO.
   * Detect multiple bounding boxes with class labels and confidence scores.
5. Postprocessing
   * Apply Non-Maximum Suppression (NMS) to remove redundant overlapping boxes.
   * Display results with bounding boxes and labels.

Advantages

* Real-time performance due to single forward pass.
* High accuracy with pretrained models.
* Can detect multiple objects simultaneously.
* Easily adaptable for different use cases (traffic monitoring, healthcare, security).

Limitations

* Struggles with very small objects compared to region-based methods.
* Accuracy decreases in crowded or overlapping object scenarios.
* Requires high computational power (GPU) for real-time inference.
* Pretrained models may not perform well on domain-specific datasets without fine-tuning.

Applications

* Autonomous vehicles: Pedestrian and traffic sign detection.
* Surveillance systems: Detecting suspicious activities in real time.
* Healthcare: Identifying medical instruments or anomalies in scans.
* Retail and inventory: Automated checkout systems and stock monitoring.
* Robotics: Object localization for navigation and manipulation.

Working / Algorithm

1. Input image is divided into S × S grid.
2. Each grid predicts B bounding boxes with confidence scores.
3. Each bounding box consists of:
   * Coordinates (x, y, w, h).
   * Confidence (probability of object presence × IoU score).
4. For each bounding box, class probabilities are predicted.
5. Final prediction = bounding box with highest confidence after Non-Maximum Suppression.

Conclusion

YOLO and its pretrained models (YOLOv3, YOLOv5, YOLOv8) provide an efficient and real-time solution for object detection. By framing detection as a regression problem, YOLO avoids complex pipelines and achieves faster inference while maintaining competitive accuracy. Its ability to detect multiple objects simultaneously makes it a widely adopted framework across industries. However, limitations with small object detection and high computational requirements must be considered in practical applications.