**Report: Object Detection and Comparison on MOT20 Video**

**1. Introduction**

In this study, we applied the YOLOv8 model for human detection using a video file from the MOT20 dataset. MOT20 is a challenging dataset containing videos of crowded human environments. Object detection was performed on the video, and the model’s accuracy was evaluated by comparing its results with ground truth annotations using the Intersection over Union (IoU) metric.

**2. Methodology**

2.1 Dataset Selection

The MOT20 dataset was selected due to its complexity and real-world human motion scenarios. The dataset consists of videos with labeled bounding boxes for human objects, making it ideal for evaluating object detection models.

2.2 Model Selection

The YOLOv8 (You Only Look Once) model was chosen for object detection. YOLOv8 is a state-of-the-art pre-trained model known for its high accuracy and real-time performance in detecting objects, especially human classes.

2.3 Processing Steps

1. Data Loading: The MOT20 video file was loaded using OpenCV.

2. Object Detection: The YOLOv8 model detected objects frame by frame, drawing bounding boxes around detected humans and assigning confidence scores.

3. Ground Truth Comparison: The detected results were compared with the provided ground truth labels using the IoU metric, which measures the overlap between predicted and actual bounding boxes.

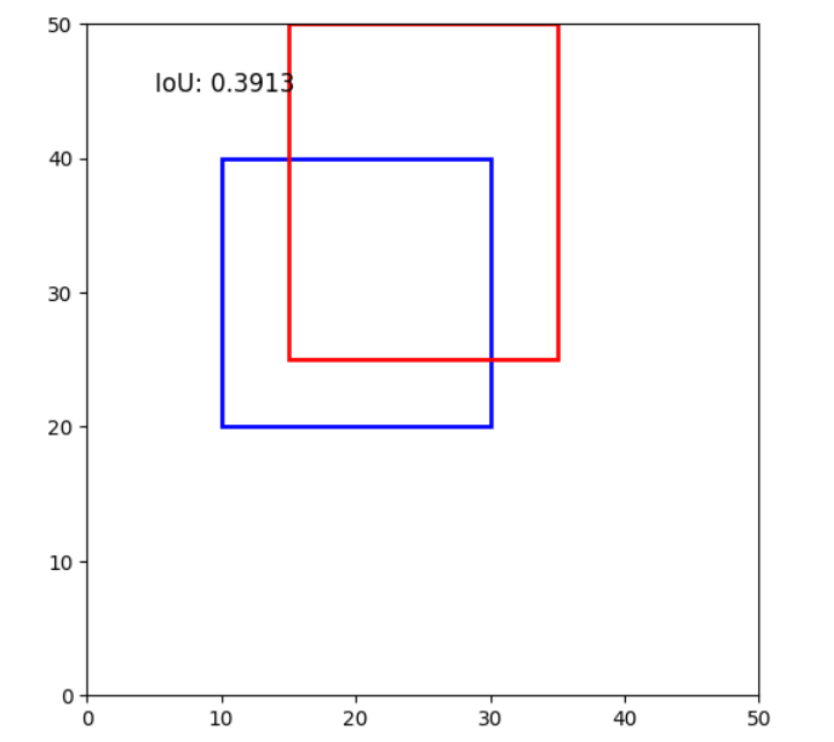
**3. Results**

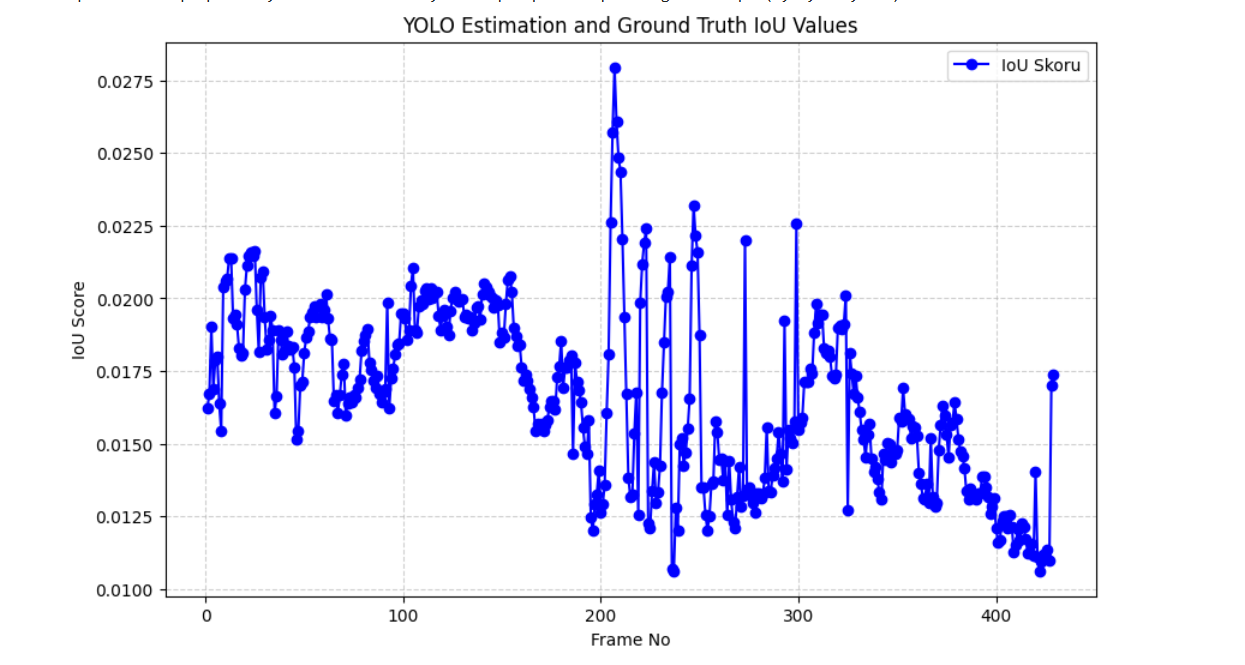
3.1 IoU Calculations

The model's accuracy was evaluated by calculating the IoU for each detection. IoU measures how well the predicted bounding box overlaps with the ground truth. An example IoU score of \*\*0.391\*\* was computed, indicating moderate detection accuracy.

3.2 Visual Comparison

The following comparison shows detected bounding boxes (red) versus ground truth bounding boxes (blue). The closer the overlap, the better the detection performance.





3.3 Output Video

A video showcasing the model’s detections with bounding boxes drawn around detected humans was generated.

**4. Evaluation**

Model Performance: While the model performed well in detecting humans, its performance dropped in highly crowded scenes, where detection errors or missed detections occurred.

The YOLOv8 model's performance on the MOT20 dataset was assessed using IoU metrics:

IoU Distribution: Scores varied across frames, peaking at ~0.0275, with moderate overall accuracy.

Visual Comparison: A sample IoU of 0.391 indicates moderate overlap between predicted (red) and ground truth (blue) bounding boxes.

While the model performs reasonably well, detection accuracy fluctuates, particularly in complex scenes. Further optimization or larger models could enhance performance.

Overall Findings: The visual comparisons and calculated IoU scores demonstrated the model's strengths and limitations in human detection tasks.

**5. Conclusion**

In this project, the YOLOv8 model successfully detected humans in the MOT20 dataset video, achieving moderate accuracy based on IoU calculations. However, there is room for improvement, especially in crowded scenarios.