Group 2 - Model Training & Performance Comparison

Objective:

Train **Faster R-CNN** (a two-stage object detection model) and **YOLOv5** (a single-stage object detection model) on the same dataset and compare their performance based on key evaluation metrics. This will help analyze the trade-offs between detection accuracy and inference speed in object detection models.

Task Breakdown:

1. Dataset Preparation

- Select an appropriate object detection dataset such as **COCO, Pascal VOC, Open Images, or a custom dataset**.
- Split the dataset into **training (70%), validation (20%), and testing (10%)**.
- Ensure that the dataset is correctly formatted:
    - Faster R-CNN requires **COCO JSON or Pascal VOC XML format**.
    - YOLOv5 requires **YOLO TXT annotation format**.
- If using a custom dataset, perform annotation using a tool like **LabelImg or Roboflow**.

---

2. Model Selection

Train and evaluate two different object detection models on the same dataset:

A. Faster R-CNN (Region-Based CNN Family)

- Use a pre-trained **Faster R-CNN model** and fine-tune it on the dataset.
- Understand the two-stage approach:
    1. **Region Proposal Network (RPN)** generates potential object locations.
    2. **Classifier & Bounding Box Regression** refine and classify objects.

B. YOLOv5 (You Only Look Once - Single Stage Detector)

- Use a pre-trained **YOLOv5 model** and fine-tune it on the dataset.
- Understand the single-stage approach:

- YOLO directly predicts **bounding box coordinates, object confidence scores, and class probabilities** in one forward pass.

---

3. Model Training

- Implement training pipelines for both models with appropriate **hyperparameters** such as:
  - **Batch size, learning rate, optimizer, and number of epochs**.
- Train both models on the **same dataset** to ensure a fair comparison.
- Monitor training loss and performance metrics across epochs.

---

4. Performance Evaluation & Comparison

After training, evaluate the models on the **same test dataset** using the following criteria:

A. Mean Average Precision (mAP)

- Compute **mAP@0.5 (Pascal VOC standard)** and **mAP@0.5:0.95 (COCO standard)**.
- Analyze the impact of **IoU thresholds** on detection accuracy.

B. Inference Time (Speed)

- Measure the **inference time per image** for both models.
- Compute **Frames Per Second (FPS)** for real-time detection scenarios.

C. Detection Accuracy

- Compare the **number of false positives and false negatives**.
- Analyze the **bounding box precision and recall** for different object classes.
- Visualize sample detection results to compare **bounding box quality**.

---

5. Final Report & Analysis

Each group must submit a report including:

1. **Comparison of mAP scores** for Faster R-CNN and YOLOv5.
2. **Inference time analysis** – Which model is faster?
3. **Bounding box precision** – Which model provides better object localization?
4. **Accuracy vs. Speed Trade-off** – Discussion on when to prefer Faster R-CNN vs. YOLOv5.

5. **Graphical analysis**:
    - **mAP vs. IoU threshold plots**
    - **Inference time comparison**
    - **Sample detected images with bounding boxes**

Deliverables

✓ **Code Implementation**: Training and evaluation scripts for both models.
✓ **Model Weights**: Saved trained models for further testing.
✓ **Evaluation Metrics**: mAP, inference time, and accuracy comparison results.
✓ **Performance Analysis Report**: Documented findings with observations.
✓ **Visualizations**: Charts and sample detections for comparative study.