

Evaluation of Performance of Various OBB Models on UAV Videos

Group Name: Ctrl + Alt + Del

Project Definition: 3

Group Members : Dhruv Patel (AU2444003) , Garima Jain (AU2444017) ,Devanshi Rathod (AU2444023)

This study focuses on evaluating the performance of the OBB (Oriented Bounding Box) family models. The objective is to explore and compare various OBB models, understand their evaluation criteria and dataset, analyze their architecture, and identify the key modules that perform oriented object detection.

Models and Their Descriptions:

The following YOLO11 pretrained OBB models are considered for evaluation. These models are pretrained on the DOTA1 dataset and downloaded automatically from the latest Ultralytics release upon first use.

1. YOLO11n-obb

- **Size:** 1024B
- **Technique:** Lightweight version optimized for real-time performance with lower computational costs. Utilizes depth-wise separable convolutions for efficient feature extraction.
- **Description:** YOLO11n-obb is the smallest model in the YOLO11 series, designed for edge devices and low-powered UAV applications. It trades off some accuracy for significantly faster inference speed.

2. YOLO11s-obb

- **Size:** 1024B
- **Technique:** Small-scale OBB model with improved feature pyramid network (FPN) for better detection of small and rotated objects.
- **Description:** YOLO11s-obb provides a balance between speed and accuracy, making it suitable for UAV-based applications where real-time inference is critical.

3. YOLO11m-obb

- **Size:** 1024B
- **Technique:** Medium-scale model leveraging attention mechanisms and improved anchor-based object detection.

- **Description:** YOLO11m-obb enhances detection capabilities by improving feature extraction layers, making it effective in detecting complex orientations in UAV videos.

4. YOLO11l-obb

- **Size:** 1024B
- **Technique:** Large-scale model with deeper neural network architecture and high-resolution feature maps.
- **Description:** YOLO11l-obb is designed for high-accuracy applications, providing better precision in detecting oriented objects at the cost of increased computational demand.

5. YOLO11x-obb

- **Size:** 1024B
- **Technique:** Extended version with an advanced transformer-based backbone and larger receptive fields for superior object detection.
- **Description:** YOLO11x-obb is the most powerful model in the series, optimized for high-quality detection in aerial imagery, but requiring significant computational resources.

Dataset Used:

- **DOTAv1/DOTAv1.5/DOTAv2:** These datasets provide high-quality aerial images with labeled objects in an oriented bounding box format, making them ideal for evaluating the performance of OBB models.

Next Week's Plan:

- Next week, we will focus on implementing evaluation metrics and running benchmark tests on the selected models to analyze their real-world performance.