Object Detection & Tracking Summary Report

Objective - To Develop a computer vision and deep learning model capable of detecting and tracking birds and drones from an onboard camera installed on a UAV. The model is expected to effectively differentiate between birds and drones, even when they appear as small objects in the image (low pixel count).

Dataset Description - The chosen data set is retrieved from a roboflow licensed platform. [Dataset-URL](https://universe.roboflow.com/search?q=class%3Abird+-+flying) It has bird and drone images, with 80% train set, 10% validation set and 10% test set.

Data class distribution - The target variable is a classifier having two classes, 0-Bird and 1-Drone

Packages- Ultralytics, torch, torchvision, torchaudio, Yolov8\_expainer, benchmark

Chosen Model - YOLOv8 and YOLOv11 was selected for its balance of accuracy and inference speed, further it best performs image classification

Training Method - Transfer Learning using Yolov8 and Yolov11

GPU Chosen - A100, T4

Model Training Details - The YOLOv8 model was trained for 20 epochs using a batch size of 16. The model provided a good training mAP-score for bird category when compared to drone data accuracy.

Model Evaluation and Metrics - The model achieved 77% score for identifying bird in validation set and 48% score for identifying drone in validation set for the threshold parameter mAP(50-95)

To enhance the prediction for drone , Yolov11 was chosen.

Dataset Description - The final chosen data set is retrieved from a roboflow licensed platform. [Dataset-URL](https://universe.roboflow.com/cv2/birds-2wpg0) It has bird and drone images, with 90% train set, 5% validation set and 5% test set.

Model Evaluation and Metrics - The model achieved 64.5% score for identifying bird in validation set and 64.8% score for identifying drone in validation set for the threshold parameter mAP(50-95)

Object detection: The best fit model differentiates bird or drone for small pixel images. The output file provided in the jupyter notebook - yolo8-11.ipynb

Object detection and tracking is performed for video files, which is saved under yolov11n>runs>detect>predict

Compression and Inference: The best fit model is evaluated by using default parameter pruning values. The model underwent batch normalisation and dropout. Further the tuned model is converted to paddle format as inference.