Project Report on Object Detection using Yolov8 for Aerial Maritime Dataset

Submitted by –
Rohan Sachidanand Chopade

Presented to:
Aethrone Aerospace

Index

1. Dataset Overview

The dataset provided for this problem statement consisted of 3 folders test, train and valid respectively along with a yaml file. This dataset contained images of Aerial Maritime photographs taken via a Mavic Air 2 drone, consisting of 4 objects namely **docks, boats, lifts, jetskis, and cars**.

Data Distribution	Number of Images
Train	371
Test	32
Valid	105

2. Annotation Process

The dataset annotation process was performed using **Roboflow**, a tool specifically built for creating and managing labeled datasets for computer vision applications. Initially, bounding boxes were manually drawn around key objects in each image, such as *boats, cars, docks, jet skis, and lifts*. Each bounding box was then assigned the appropriate class label to clearly define the objects in the images. After annotation, the dataset was exported in **YOLOv8** format, containing the bounding box coordinates and class labels, ready for training the model.

3. Model Training

For the object detection task, I chose to implement **YOLOv8s** (YOLOv8 Small), a cutting-edge deep learning model optimized for real-time object detection in computer vision tasks. This model was selected for its impressive combination of accuracy and speed.

The model was trained in **Google Colab**, leveraging its GPU resources to accelerate the training process.

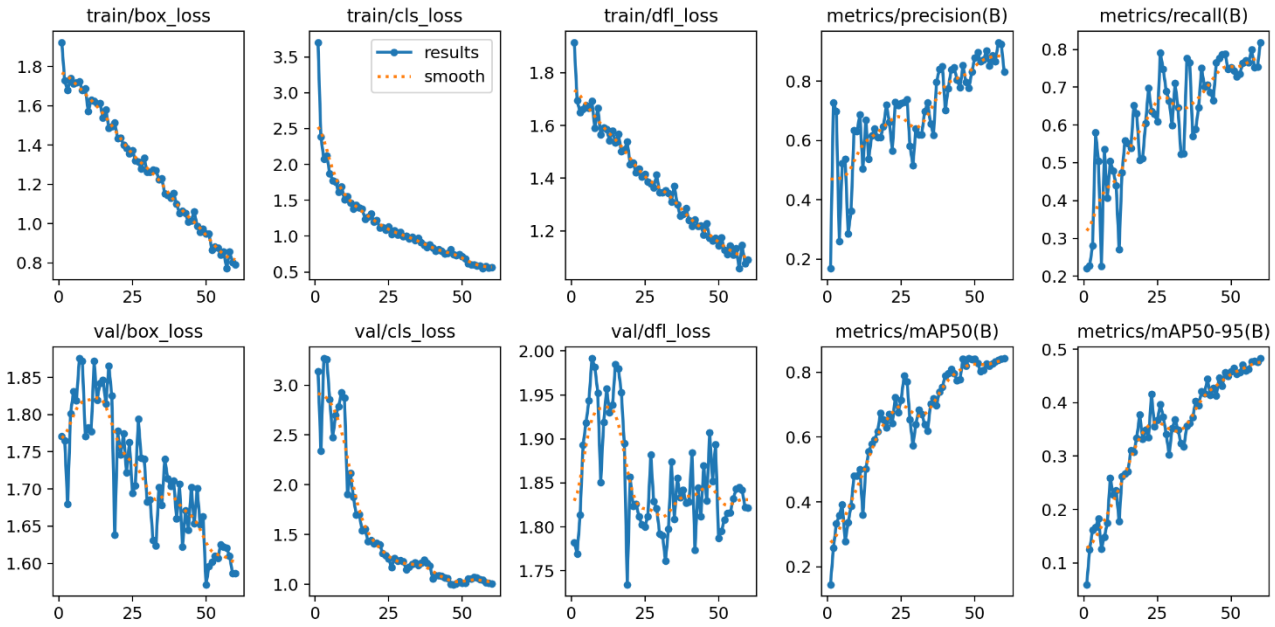
Parameters for training –

- Epochs: 60
- Imgsz (image size): 640

4. Evaluation

The model's performance was evaluated on the **validation dataset**. The evaluation process involved the use of several standard object detection metrics to assess the model's ability to detect and classify the objects accurately.

Classes	Precision (in %)	Recall (in %)	Mean Average Precision (mAP)(in %)
All	83.2	81.8	**84.3**
Boat	72.6	71.4	**72.4**
Car	100	81.9	**85.1**
Dock	74.8	79.2	**81.6**
Jetski	78	96.8	**95.9**
Lift	**90.7**	**79.6**	**86.6**



4.1 Output


DJI_0272_JPG.rf.aee02e49573c
lift 0.9
lift 0.8
dock 0.4
boat 0.6
dock 0.5
DJI_0282_JPG.rf.9daae03614262
dock 0.9
jetski 0.8
dock 0.8
dock 0.5
lift 0.9
dock 0.6
dock 0.8
jetski 0.8
DJI_0288_JPG.rf.073380258bfe
dock 0.9
lift 0.8
dock 0.6
lift 0.9
DJI_0288_JPG.rf.13c2d7
dock 0.5
dock 0.8
dock 0.5
boat 0.8


DJI_0259_JPG.rf.0b45fb1a66b0
dock 0.9
car 0.9
dock 0.8
dock 0.3
lift 0.9
dock 0.5
lift 0.7
jetski 0.8
dock 0.8
dock 0.3
lift 0.8
dock 0.9
dock 0.7
dock 0.9
dock 0.4
car 0.8
lift 0.6
car 0.7
dock 0.9
DJI_0259_JPG.rf.60
dock 0.9
lift 0.9
dock 0.3
dock 0.4
dock
lift 0.8
dock 0.9
boat 0.9

5. Object Count Analysis

The object count analysis was carried out using Roboflow's built-in class count feature, which offered a summary of the object distribution throughout the dataset.

6. Comparative Analysis

The comparative analysis was done between two models. **Model 1**, trained on five distinct classes and **Model 2**, which combines *boat* and *jetski* into *watercraft*.

The validation results for the two YOLOv8 models—one trained with 5 classes and the other with 4 classes (where the *boat* and *jetski* classes were merged into a single watercraft class)—are summarized and compared below.

Metric	**Model 1 (5 classes)**	**Model 2 (4 classes)**
Classes	Boat, Car, Dock, Jetski, Lift	Car, Dock, Lift, Watervehicle
Box (P)	0.832	0.893
Box (R)	0.818	0.699
mAP50	0.843	0.828
Inference Speed	9.9 ms	14.4 ms
Postprocess Speed	11.7 ms	10.7 ms

**Postprocess speed is the time taken per image to perform operations like non-maximum suppression, formatting outputs, visualizing results, and saving detections after the model's inference.*

The comparison between the 5-class and 4-class YOLOv8 models highlights key trade-offs. The 4-class model achieves higher precision (0.893) but lower recall (0.699), indicating more confident detections but more missed instances compared to the 5-class model, which balances precision (0.832) and recall (0.818) for a higher overall mAP50. Combining boat and jetski into watercraft improves detection performance for the merged category (mAP50: 0.915), simplifying the task. Both models have similar complexity, with the 4-class model having slightly faster postprocessing (10.7 ms vs. 11.7 ms) but slower inference (14.4 ms vs. 9.9 ms). While the 5-class model excels in detecting smaller, specialized classes like jetski, larger classes like dock and lift perform comparably in both models.