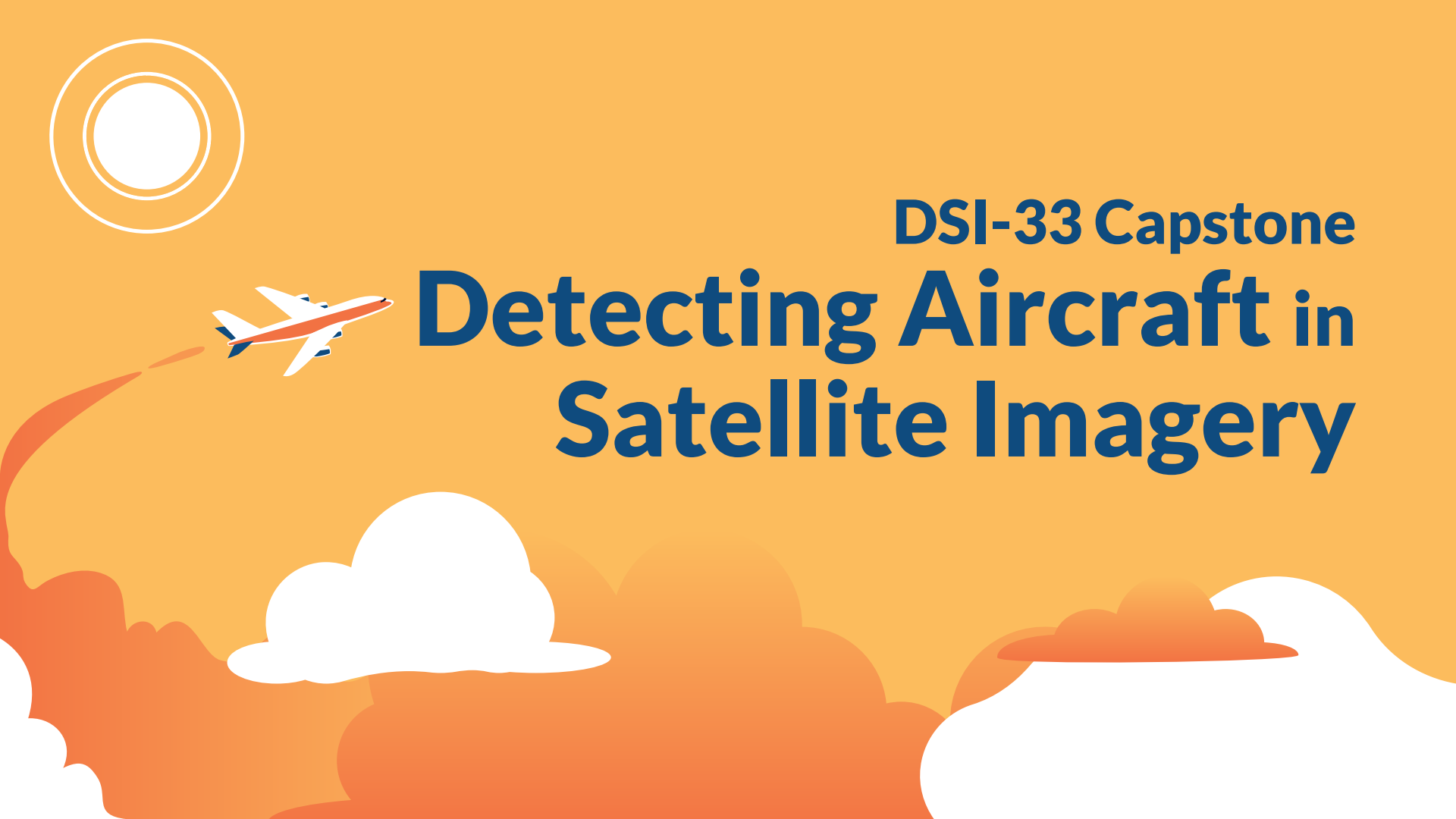

DSI-33 Capstone
Detecting Aircraft in Satellite Imagery

BACKGROUND

Looking at airplanes within satellite images may sound like a weird hobby but it has potential real-world use cases such as:

- Defence intelligence
- Airport capability assessment
- Capacity planning/monitoring
- Search and Rescue

AIRPORTS

01 CAPACITY

Airports are currently facing a capacity crunch.

02 PLANNING

Monitor movements and parking bay utilization to ensure efficiency.

PROBLEM STATEMENT

Airport operators need a way to improve airport capacity planning and monitoring capabilities beyond ground-based methods.

THE PROJECT

Build a computer vision model that can accurately detect aircraft within satellite images, achieving an mAP of at least 0.75

WORKFLOW

DATA

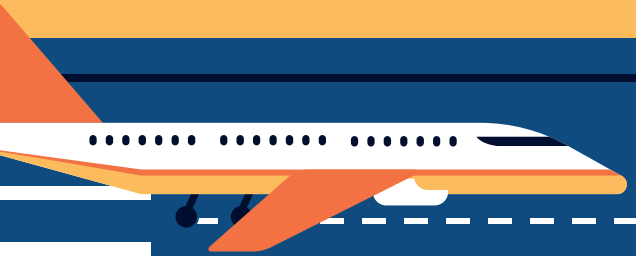
Import, explore, clean, and preprocess

MODEL

Build models and select best performing

EVALUATE

Run test detections and evaluate performance



01

DATA

Import, explore, clean, and preprocess

AIRBUS AIRCRAFT DATASET

EARTH OBSERVATION

Images taken by Airbus' Pleiades twin satellites

103 IMAGES

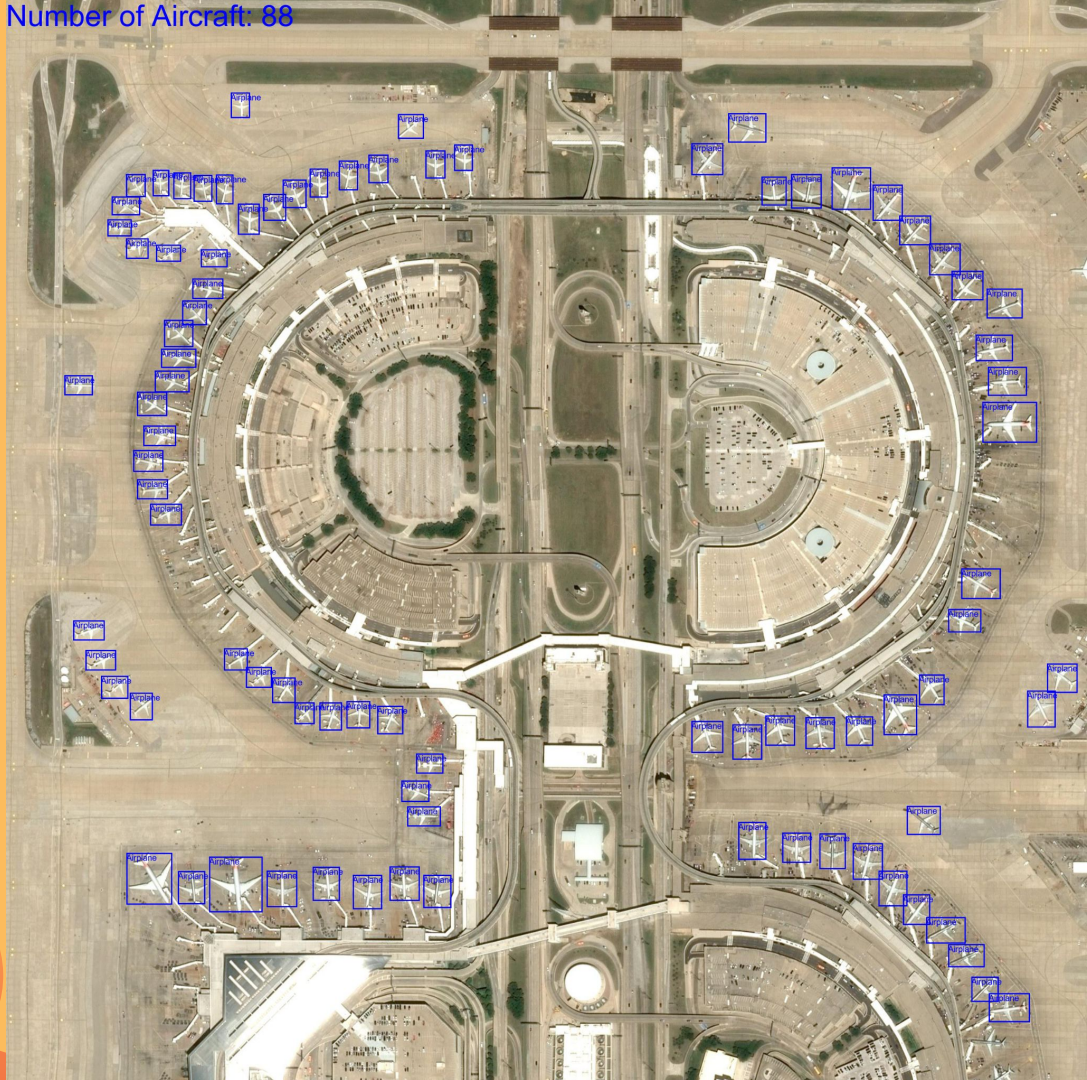
Of various airports worldwide, some appear multiple times at different acquisition dates

ANNOTATIONS

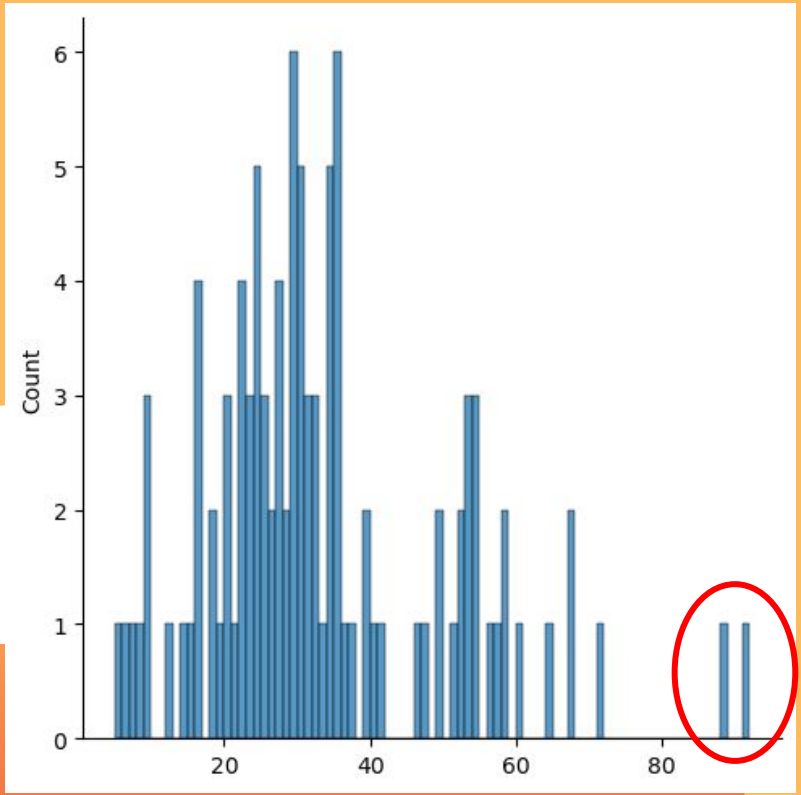
CSV file containing coordinates of the four bounding box corners per target

Raw image from dataset

What the images look like with annotations



Distribution of Aircraft Counts per Image

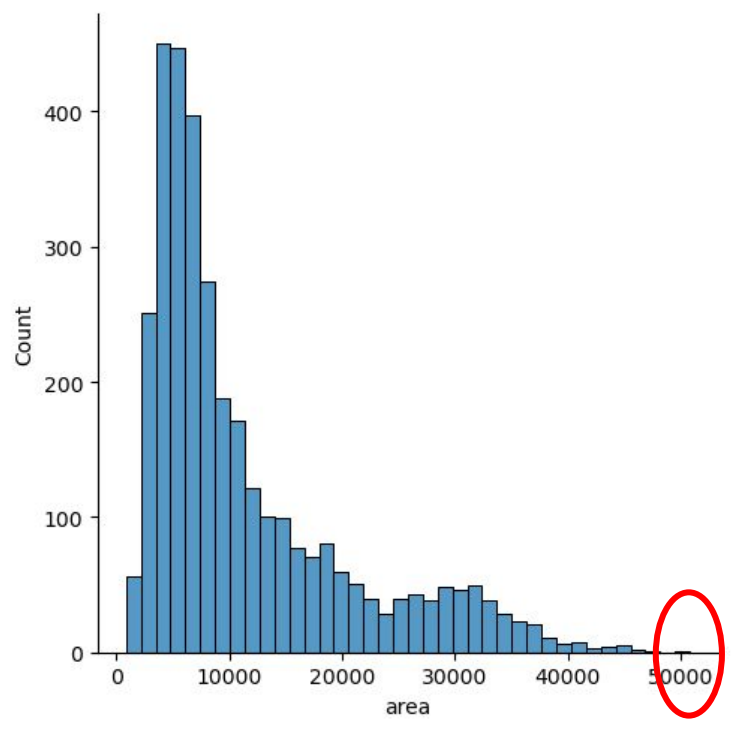


Bulk of the images have between 10-40 aircraft

2 images with particularly high number of aircraft

Extra set of annotations

Distribution of Bounding Box Size



Investigate bounding boxes with area >50,000 pixels

Only one bounding box of size

Just a jumbo A380

IMAGES & ANNOTATIONS

103

IMAGES

W:2560xH:2560

3379

AIRCRAFT

Annotated targets

TRAIN-VAL-TEST SPLIT

test_size = 0.2

TRAIN

65 Images

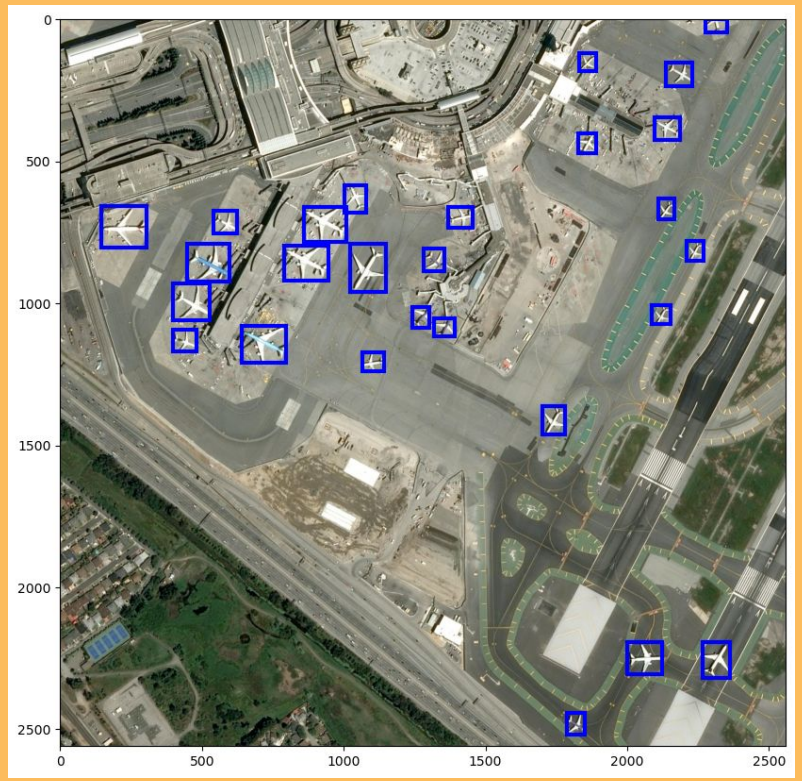
VAL

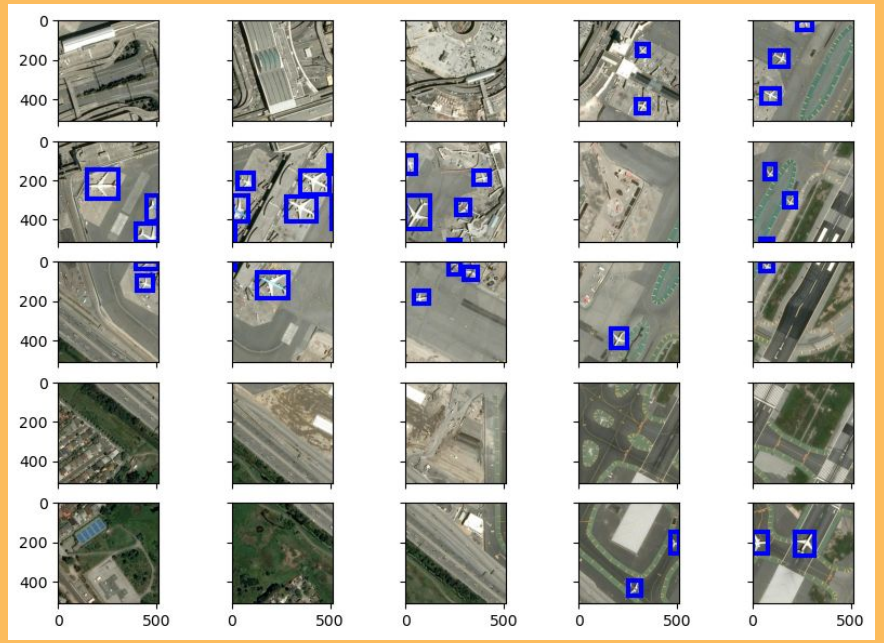
17 Images

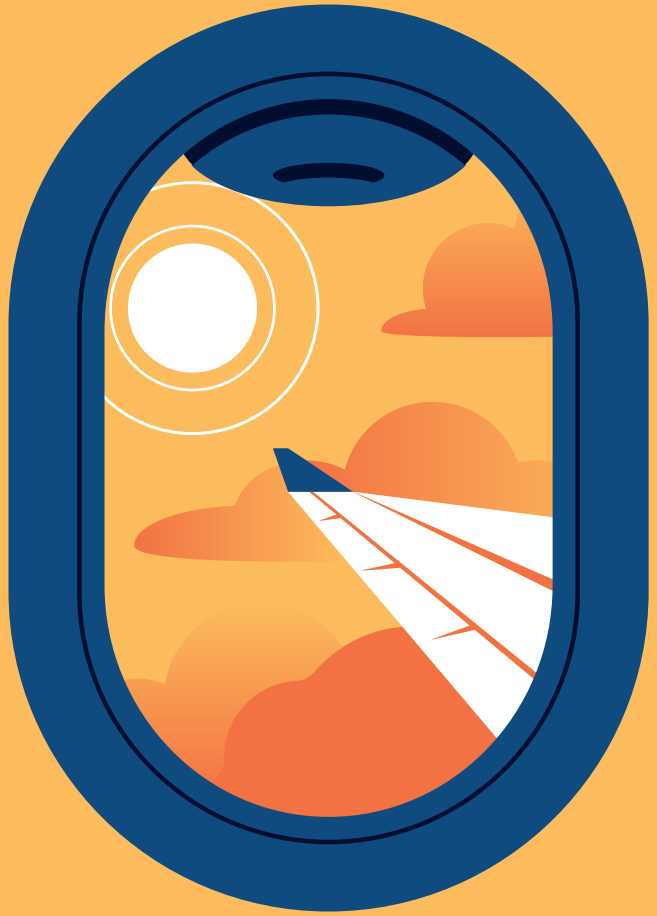
TEST

21 Images

TILING


0
500
1000
1500
2000
2500
0
500
1000
1500
2000
2500


0
200
400
0
500



WHY TILE?

LARGE IMAGE SIZE

Training on 2560x2560 is taxing on GPU resources. Also, YOLO is trained on lower resolution images.

SMALL TARGETS

Down-sampling the image is not ideal since it'll reduce the resolution and turn targets into blobs

FINAL DATASET

SET	IMAGE TILES	ANNOTATIONS
TRAIN	1625	762
VAL	425	492

02

MODEL

YOLOv7
YOLOv8

RECAP

Build a computer vision model that can accurately detect aircraft within satellite images, achieving an **mAP** of at least **0.75**

YOLO ADVANTAGES

Speed

YOLO is faster than other state-of-the-art object detectors

Performance

High detection performance, few background errors

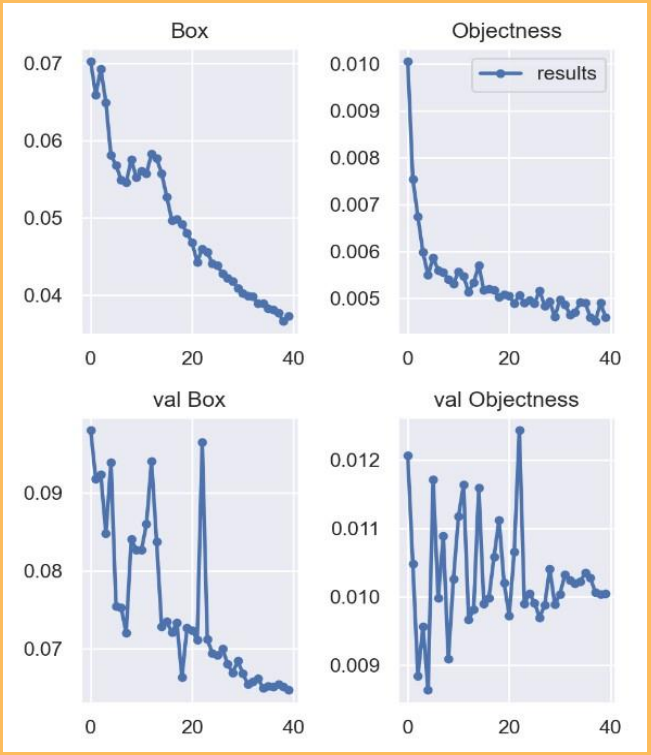
Generalization

Able to generalize from natural images to domains like art

Open Source

Community driven development has led to major improvements

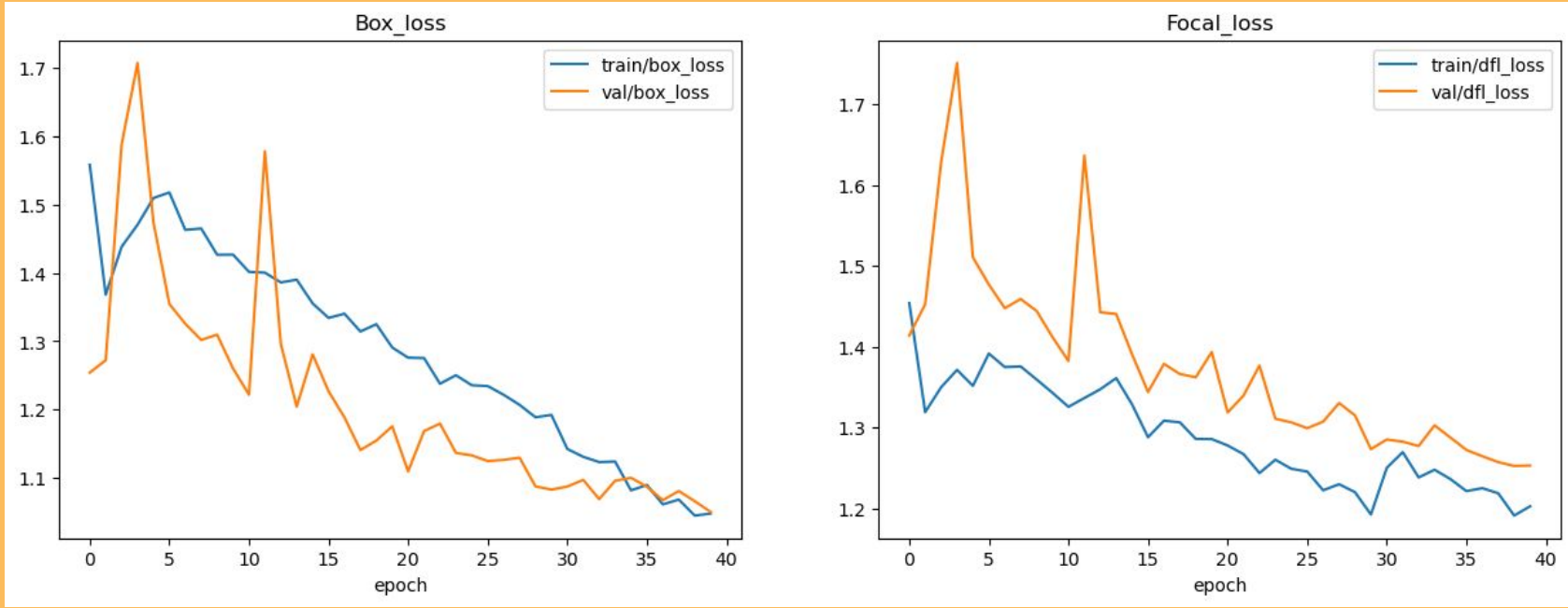
YOLOv7 TRAIN LOSS



Train and validation loss functions generally decrease across the 40 epochs, except for validation obj_loss. validation obj_loss looks very turbulent but the value is actly quite stable around 0.0105.

Not much significant improvements seen after epoch 30 or so, implying that more training is unlikely to yield better performance.

YOLOv8 TRAIN LOSS

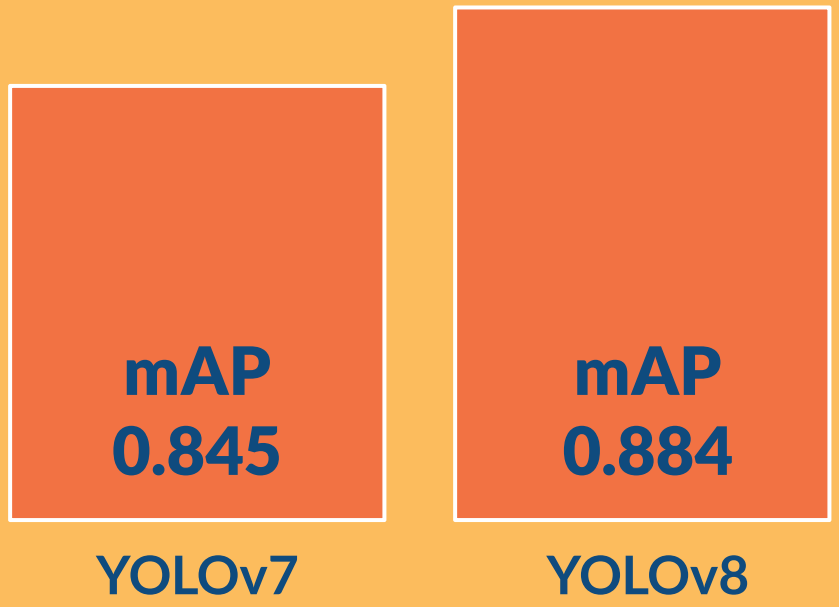


Losses are both pretty turbulent in the initial epochs but gradually tapers off in a downward trend.
Train and validation losses (for both box and focal loss) reduces as the number of epoch increases, implying good fit up till epoch 40.

OBJECT DETECTION METRICS

Precision	True Positive / (True Positive + False Negative)
Recall	True Positive / (True Positive + False Positive)
Intersection over Union	Area of Overlap / Area of Union
Mean Average Precision	Weighted mean of precisions at each IoU threshold

COMPARE MODELS - mAP



YOLOv8 edges its predecessor out by a small margin.

03

EVALUATE

Run test detections and evaluate performance

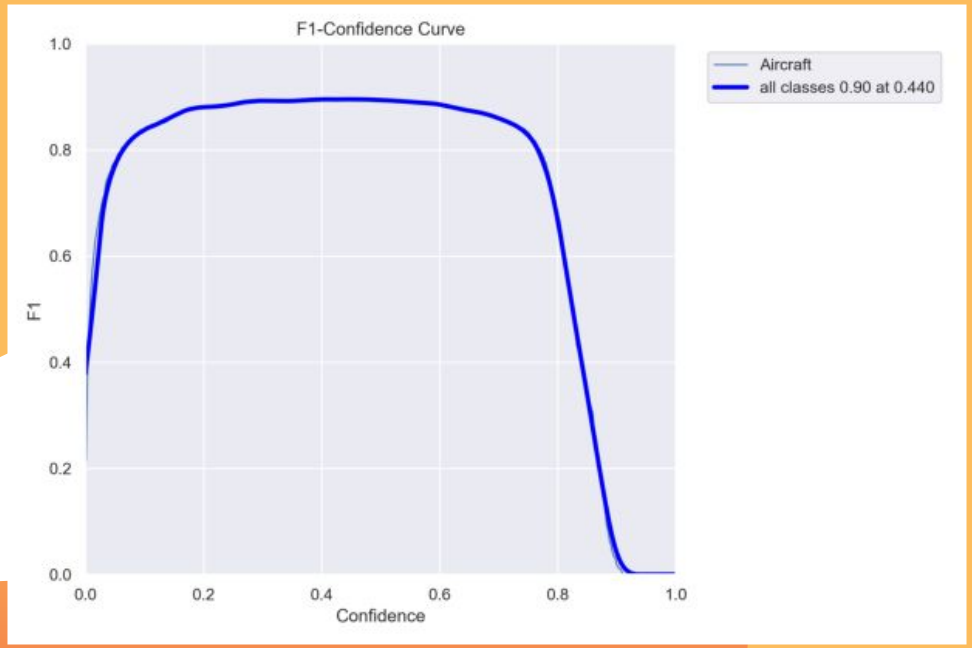
YOLOv8 Test Scores

PRECISION	RECALL	MEAN AVERAGE PRECISION
0.969	0.923	0.922

DETECTION ON TEST SET IMAGE

- Confidence threshold set at 0.001 as per YOLO test parameters.
- Model has a lot of erroneous detections, although most if not all of these have very low confidence scores

SETTING AN APPROPRIATE CONFIDENCE THRESHOLD



Use F1 score to inform choice of confidence threshold for detection using the final model, since the F1 score strikes a balance between both precision and recall.

Selected confidence threshold will be 0.440 or 44% for the maximum F1 of 0.90.

DETECTION ON UNSEEN IMAGE

Model is able to detect all visible aircraft in the picture while managing to avoid non-aircraft detections.


Aircraft 0.75
Aircraft 0.77
Aircraft 0.84
Aircraft 0.78
Aircraft 0.80
Singapore Changi Airport
Aircraft 0.83
Aircraft 0.83
Aircraft 0.80
Aircraft 0.79
Aircraft 0.82
Aircraft 0.81
Aircraft 0.80
Aircraft 0.84
Aircraft 0.83
Aircraft 0.84
Aircraft 0.80
Google Earth
Image © 2023 Maxar Technologies

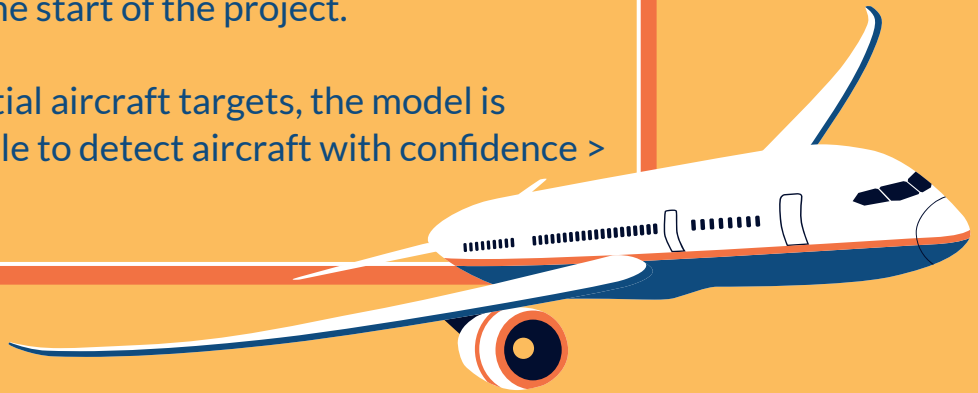
04

CONCLUSIONS

Key takeaways, limitations, and future directions

CONCLUSIONS

- The final YOLOv8 custom model is capable of detecting airplanes from satellite images with little to no noise using the right confidence threshold.
- The selected model has also met the target mAP stipulated the start of the project.
- Besides partial aircraft targets, the model is generally able to detect aircraft with confidence > 0.75



FUTURE DIRECTIONS

- Collect train images across a greater variety of backgrounds, e.g. over water, to train and test on.
- Expand dataset to include aircraft that are not commercial jets. Classification of aircraft into finer grained subcategories would no doubt prove to be useful to a wider range of stakeholders.
- Experiment with the use of synthetic training data.
- Deploy the model on a platform such as UP42 that is able to provide "live" satellite feeds.

QnA