2021

Tick Object Detection using Deep Learning Algorithms

CONTENTS

INTRODUCTION

Object detection is a critical computer vision task that detects instances of specific visual objects (for example, humans, animals, cars, or buildings) in digital images such as photos or video frames. Object detection seeks to create computational models that provide the most basic information required by computer vision applications: "What objects are where?"

The goal of this research project is to implement various state-of-the-art object detection models on a dataset containing Tick images and compare their results.

We have 2 datasets, one with 4 classes and the other with 11 classes. The datasets are imbalanced in terms of number of samples and hence require data augmentation. We train each model twice for each dataset, once without data augmentation and once with data augmentation.

The model predicts the class of each image, the bounding box of the object and the confidence level (lies between 0 and 1) which indicates the probability of the predicted class begin the actual true class of the object.

DATASET

We will be training each model on 2 datasets:

Tick Dataset 1:

Contains 4 classes:

- Hyalomma Dromedarii Female: 27 images
- Hyalomma Dromedarii Male: 83 images
- Hyalomma Scupense Female: 17 images
- Hyalomma Scupense Male: 47 images

We see that the data is quite imbalanced and hence to tackle imbalance we use **rotation** as the **Data Augmentation** technique, we rotate the images with an angle of **90**, **180** and **270** degrees only for the Hyalomma Scupense Female and Hyalomma Dromedarii Female classes as they are the minority class.

Tick Dataset 2:

Contains 11 classes:

- Hyalomma Dromedarii Female: 107 images
- Hyalomma Dromedarii Male: 211 images
- Hyalomma Scupense Female: 119 images
- Hyalomma Scupense Male: 147 images
- Hyalomma Impeltatum Female: 16 images
- Hyalomma Impeltatum Male: 46 images
- Hyalomma Marginatum Female: 24 images
- Hyalomma Marginatum Male: 60 images
- Other Female: 6 images
- Other Male: 16 images
- Hyalomma Excavatum: 22 images

We see that the data is quite imbalanced and hence we to tackle imbalance we use **Data Augmentation**. We rotate the images with a random angle between **0** and **270** degrees, we **change the contrast** of the image and we **shift the RGB color** of the image by a maximum of **30** units. We carry out Data augmentation for all classes except Hyalomma Dromedarii Female, Hyalomma Dromedarii Male, Hyalomma Scupense Female and Hyalomma Scupense Male as they are in majority.

DATA ANNOTATION

The dataset is not annotated hence we use **makesense.ai** to annotate the images in **YOLO** format.

Different models require the input data to be in different formats hence we convert the data into **COCO** and **Pascal VOC** format.

- **YOLO Format**: In YOLO labelling format, a .txt file with the same name is created for each image file in the same directory. Each .txt file contains the annotations for the corresponding image file, that is object class, object coordinates, height and width.

  ```
  <object-class> <x> <y> <width> <height>
  ```

  For each object, a new line is created.

- **COCO Format**: For object detection, COCO follows the following format:

  ```
  annotation{
  "id" : int,
  "image_id": int,
  "category_id": int,
  "segmentation": RLE or [polygon],
  "area": float,
  "bbox": [x,y,width,height],
  "iscrowd": 0 or 1,
  }
  categories[{
  "id": int,
  "name": str,
  "supercategory": str,
  }]
  ```

- **Pascal VOC Format**: Pascal VOC stores annotation in XML file.

  ```
  <annotation>
  <folder>Train</folder>
  <filename>01.png</filename>
  <path>/path/Train/01.png</path>
  <source>
  <database>Unknown</database>
  </source>
  <size>
  <width>224</width>
  <height>224</height>
  <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
  <name>36</name>
  <pose>Frontal</pose>
  ```

```
<truncated>0</truncated>
<difficult>0</difficult>
<occluded>0</occluded>
<bndbox>
<xmin>90</xmin>
<xmax>190</xmax>
<ymin>54</ymin>
<ymax>70</ymax>
</bndbox>
</object>
</annotation>
```

METRICS AND DEFINITIONS

- **Parameters**
  True positive and true negatives are the observations that are correctly predicted and therefore shown in green. We want to minimize false positives and false negatives, so they are shown in red colour.

		Predicted class	
Actual Class		Class = Yes	Class = No
	Class = Yes	True Positive	False Negative
	Class = No	False Positive	True Negative

  True Positives (TP) - These are the correctly predicted positive values which means that the value of actual class is yes, and the value of predicted class is also yes.
  True Negatives (TN) - These are the correctly predicted negative values which means that the value of actual class is no, and value of predicted class is also no.
  False Positives (FP) – When actual class is no and predicted class is yes.
  False Negatives (FN) – When actual class is yes but predicted class in no.
- **Accuracy**
  It is the ratio of correctly predicted observation to the total observations.
  Accuracy = TP+TN/TP+FP+FN+TN
- **Precision**
  Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.
  Precision = TP/TP+FP
- **Recall (Sensitivity)**
  Recall is the ratio of correctly predicted positive observations to the all observations in actual class - yes.
  Recall = TP/TP+FN
- **F1 Score**
  F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account.
  F1 Score = 2*(Recall * Precision) / (Recall + Precision)

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

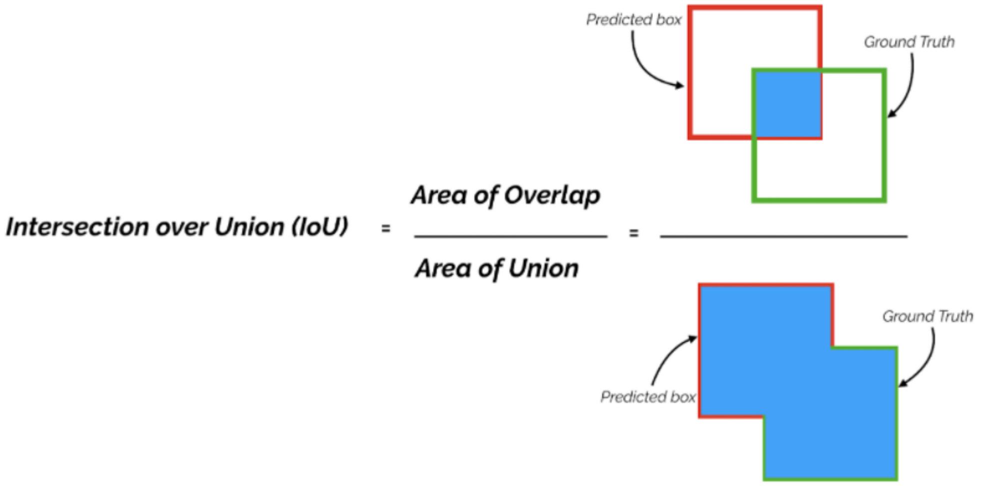
TP = True positive

TN = True negative

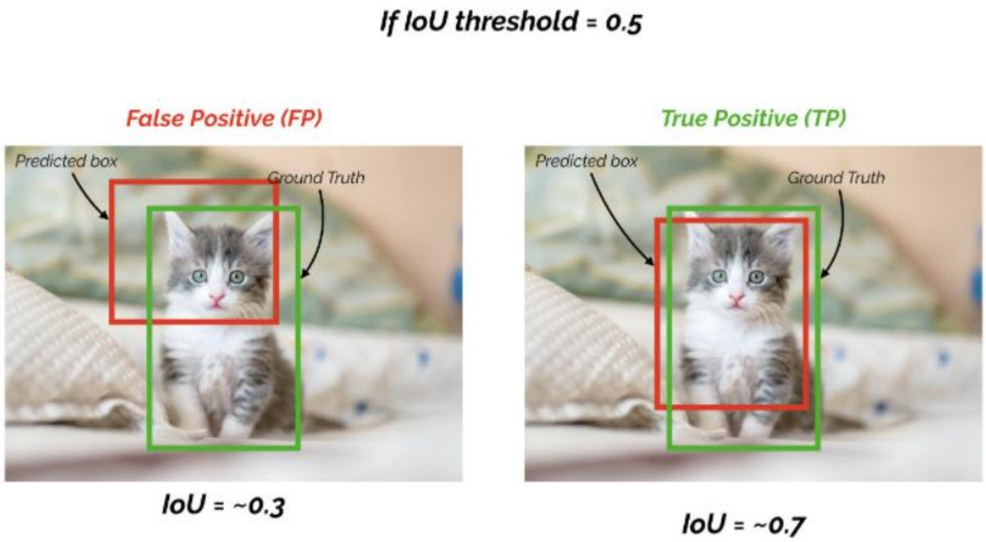
FP = False positive

FN = False negative

- **IoU (Intersection over Union)**
  IoU measures the overlap between two boundaries. We use it to measure how much the predicted boundary overlaps with the ground truth (the real object) boundary.



For object detection tasks, we calculate Precision and Recall using IoU value for a given IoU threshold. For example, if IoU threshold is 0.5, and the IoU value for a prediction is 0.7, then we classify the prediction as True Positive (TP). On the other hand, if IoU is 0.3, we classify it as False Positive (FP).



- **Average Precision (AP)**
  The general definition for the Average Precision (AP) is finding the area under the precision-recall curve.

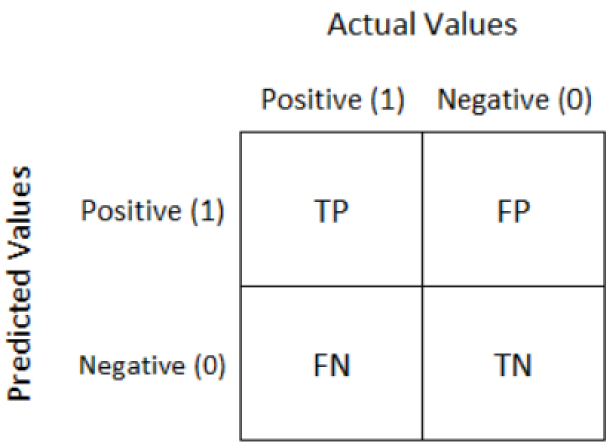
$$\mathbf{AP} = \int_0^1 p(r)dr$$

Precision and recall are always between 0 and 1. Therefore, AP falls within 0 and 1 also.

- **AP@ [0.5]**
  It corresponds to the average precision for IoU threshold of 0.5.
- **AP@ [0.5:0.95]**
  AP@ [.5:.95] corresponds to the average AP for IoU from 0.5 to 0.95 with a step size of 0.05.
- **Mean Average Precision (mAP)**
  The mean Average Precision or mAP score is calculated by taking the mean AP over all classes and/or overall IoU thresholds, depending on different detection challenges that exist.

```
Average Precision (AP):
   AP               % AP at IoU=.50:.05:.95 (primary challenge metric)
   AP^IoU=.50       % AP at IoU=.50 (PASCAL VOC metric)
   AP^IoU=.75       % AP at IoU=.75 (strict metric)
AP Across Scales:
   AP^small         % AP for small objects: area < 32²
   AP^medium        % AP for medium objects: 32² < area < 96²
   AP^large         % AP for large objects: area > 96²
Average Recall (AR):
   AR^max=1         % AR given 1 detection per image
   AR^max=10        % AR given 10 detections per image
   AR^max=100       % AR given 100 detections per image
AR Across Scales:
   AR^small         % AR for small objects: area < 32²
   AR^medium        % AR for medium objects: 32² < area < 96²
   AR^large         % AR for large objects: area > 96²
```

- **Confusion Matrix**
  It is a performance measurement for machine learning classification problem where output can be two or more classes. It is a table with n^2 (where n is the number of classes) different combinations of predicted and actual values.



- **Confidence Level**
  It is the probability of the predicted class being the actual true class.

YOLOv5 Model

YOLO an acronym for **'You only look once'**, is an object detection algorithm that divides images into a grid system. Each cell in the grid is responsible for detecting objects within itself. YOLO is one of the most famous object detection algorithms due to its speed and accuracy.

We train the YOLOv5 model for **150 epochs** on each dataset.

Results:

- **YOLOv5 on Tick Dataset 1 (without augmentation):**

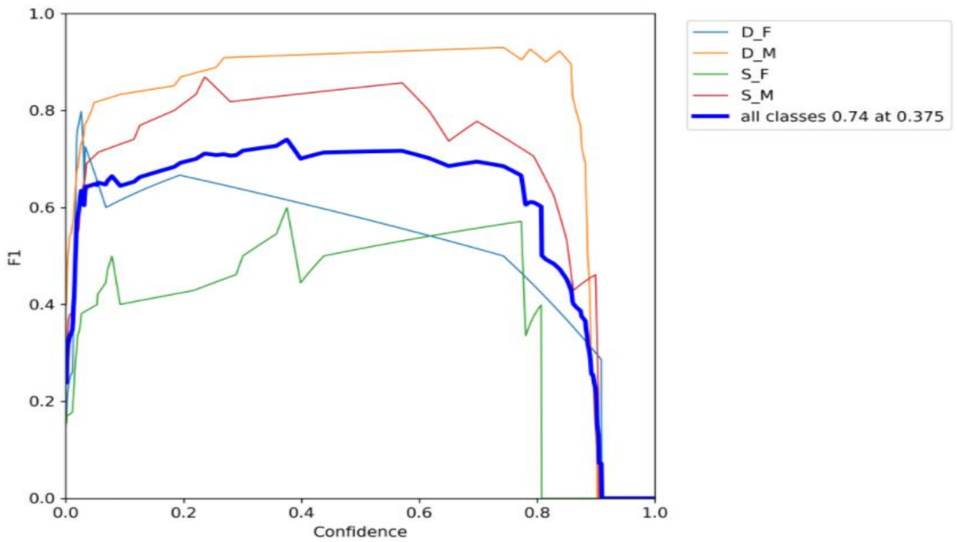
  The mAP@.5 and mAP@.5:.95 for all classes is **0.826** and **0.659** respectively.

Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:
all	40	40	0.778	0.773	0.826	0.659
D_F	40	6	1	0.445	0.851	0.668
D_M	40	20	0.842	1	0.945	0.781
S_F	40	4	0.499	0.749	0.62	0.5
S_M	40	10	0.773	0.9	0.886	0.688

F1 Curve

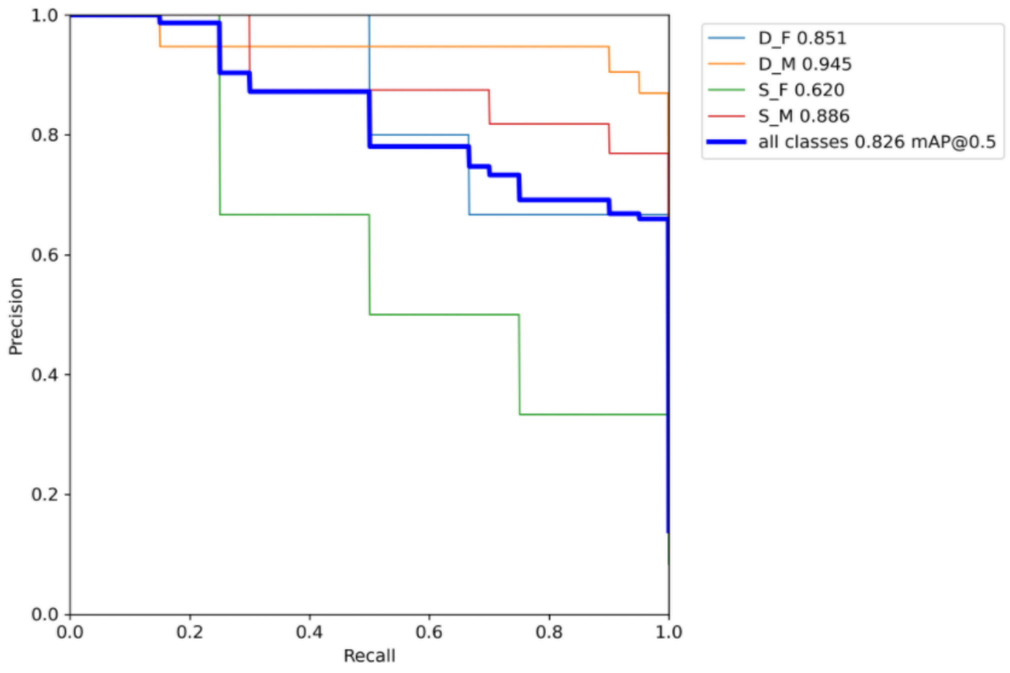
The thick blue line represents the average of all classes. The F1 score for Hyalomma Scupense Female (S_F) and Hyalomma Dromendarii Female (D_F) is below average due to lower number of samples for these classes.

The F1 score drastically drops as we increase the confidence above 0.90. The F1 score for Hyalomma Scupense Female (S_F) drops drastically after 0.80 confidence due to it having only 17 images for training.



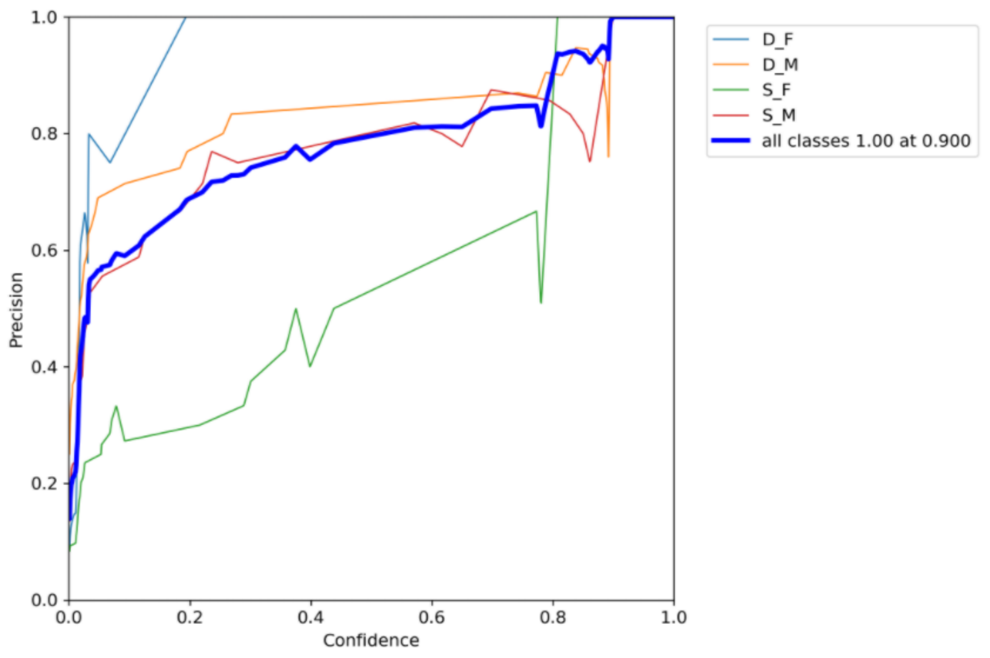
PR Curve

The thick blue line represents the average of all classes. The PR curve for Hyalomma Scupense Female (S_F) is the least due to it having the least number of training samples.



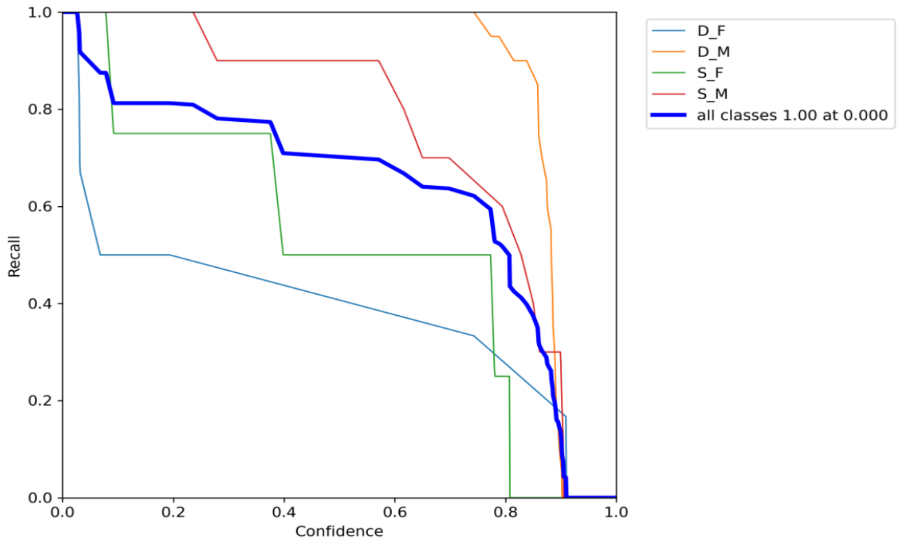
P Curve

As we increase the minimum confidence threshold the precision increases as the model becomes surer of its prediction.



R Curve

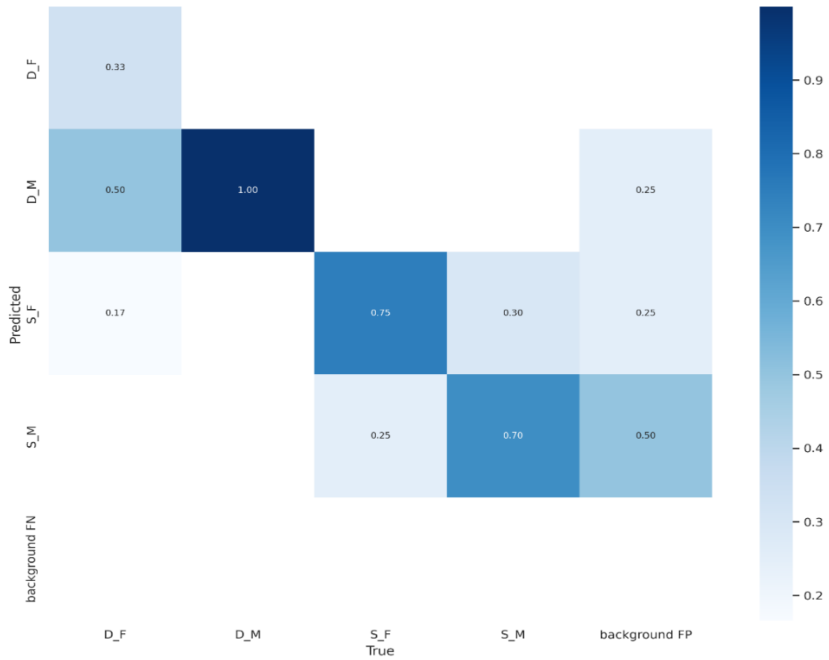
As we increase the confidence level threshold the recall decreases as the model is able to predict less images with higher degree of confidence.



Confusion Matrix

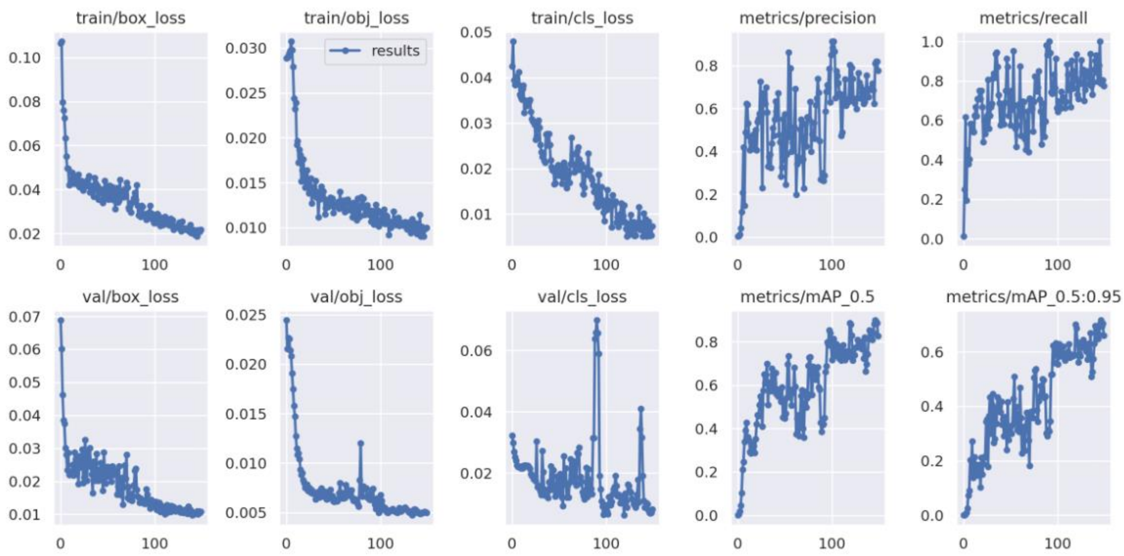
The model predicted all the Hyalomma Dromedarii Male images correctly. It also predicts 75% of the Hyalomma Scupense Female and 70% of the Hyalomma Scupense Male correctly.

The model is only able to predict 33% of the Hyalomma Dromedarii Female images accurately and misclassifies 50% of them as Hyalomma Dromedarii Male.This may be due to gross imbalance in the number of samples in Hyalomma Dromedarii species.



Train and Validation Loss Metrics

The training loss and validation loss decreases as we train for a greater number of epochs. Similarly, the precession, recall and mAP increase as we train for more epochs.



- **YOLOv5 on Tick Dataset 1 (with augmentation):**

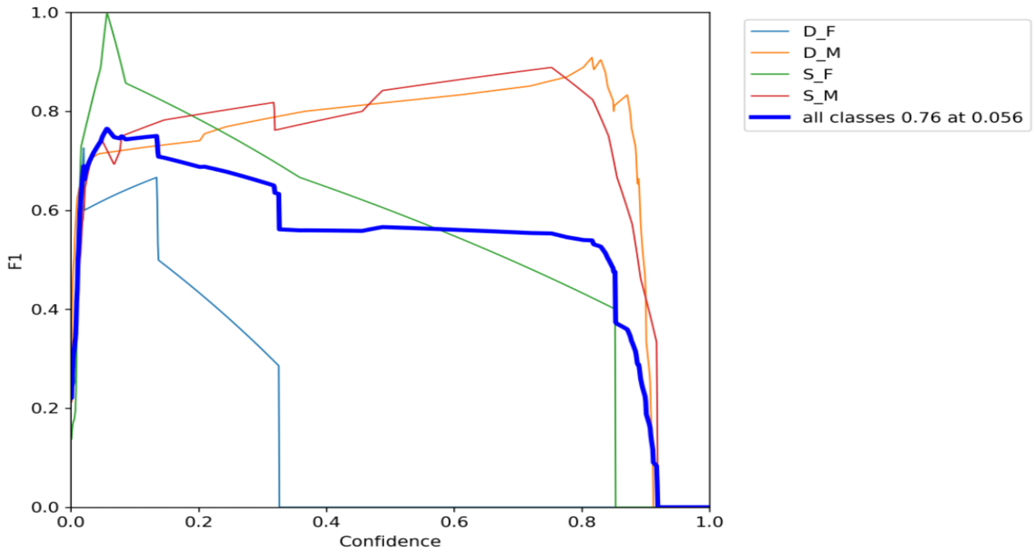
The mAP@.5 and mAP@.5:.95 for all classes is **0.894** and **0.721** respectively.

Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:
all	40	40	0.739	0.866	0.894	0.721
D_F	40	6	0.827	0.5	0.705	0.476
D_M	40	20	0.558	1	0.945	0.799
S_F	40	4	0.994	1	0.995	0.846
S_M	40	10	0.578	0.962	0.929	0.764

F1 Curve

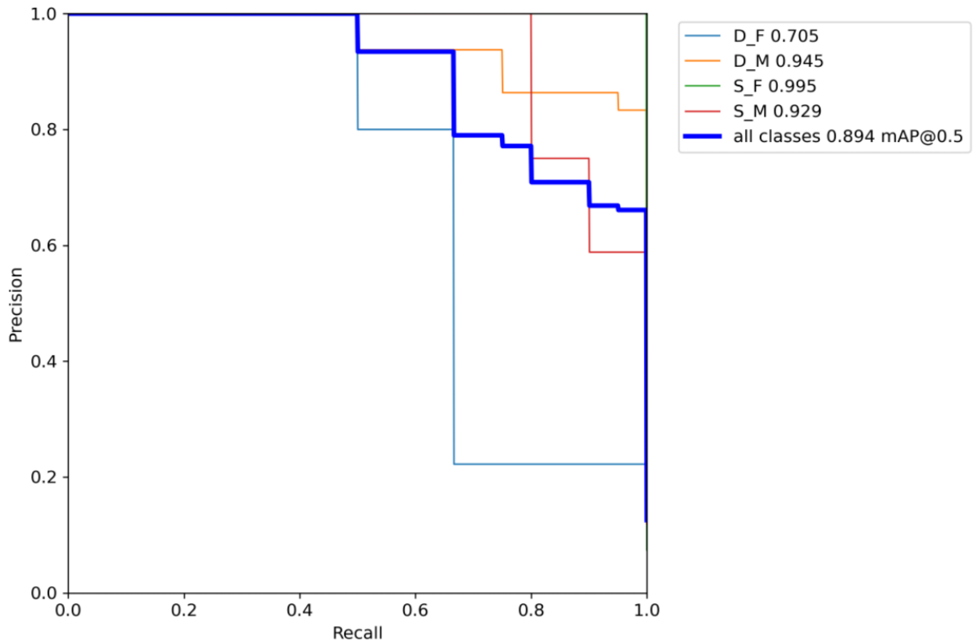
The thick blue line represents the average of all classes. The F1 score for Hyalomma Dromendarii Female (D_F) is below average due to lower number of samples.

The F1 score drastically drops as we increase the confidence above 0.90. The F1 score for Hyalomma Dromendarii Female (D_F) drops drastically after 0.30 confidence.



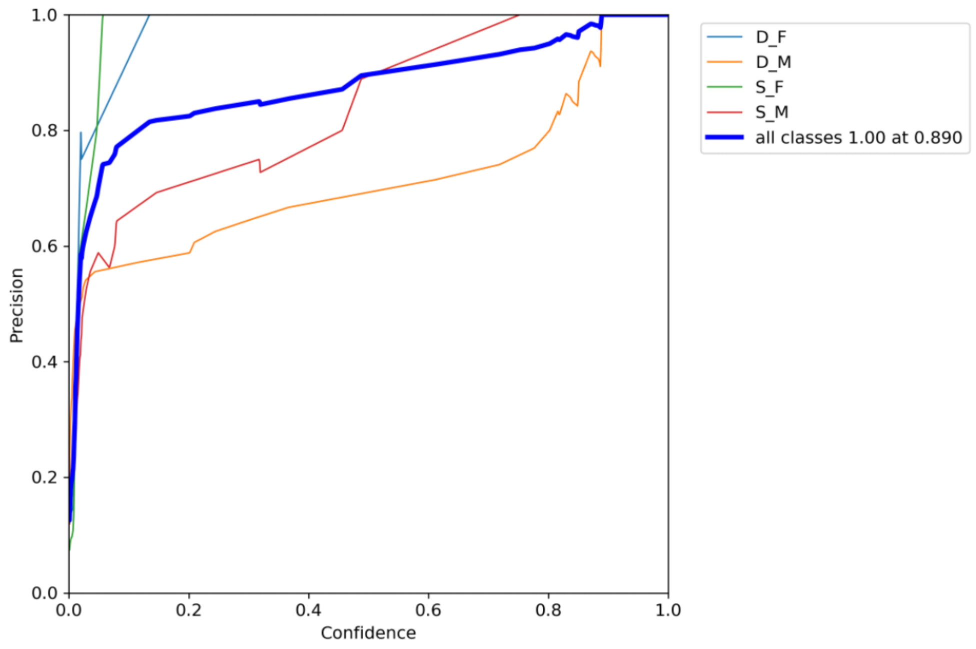
PR Curve

The thick blue line represents the average of all classes. The PR curve for Hyalomma Dromendarii Female (D_F) is the least.



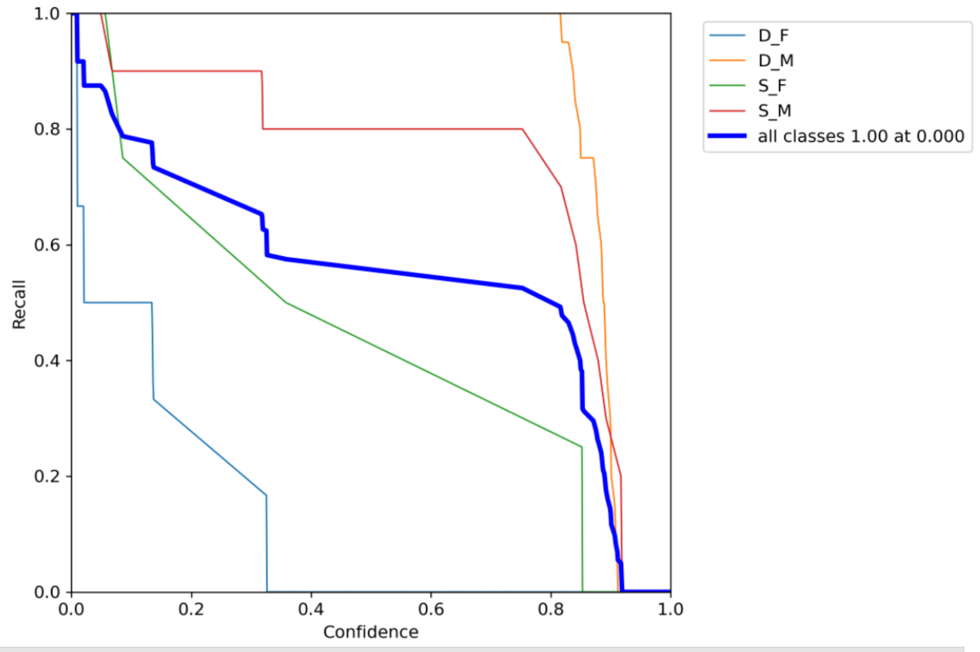
P Curve

As we increase the minimum confidence threshold the precision increases as the model becomes surer of its prediction.



R Curve

As we increase the confidence level threshold the recall decreases as the model is able to predict less images with higher degree of confidence.

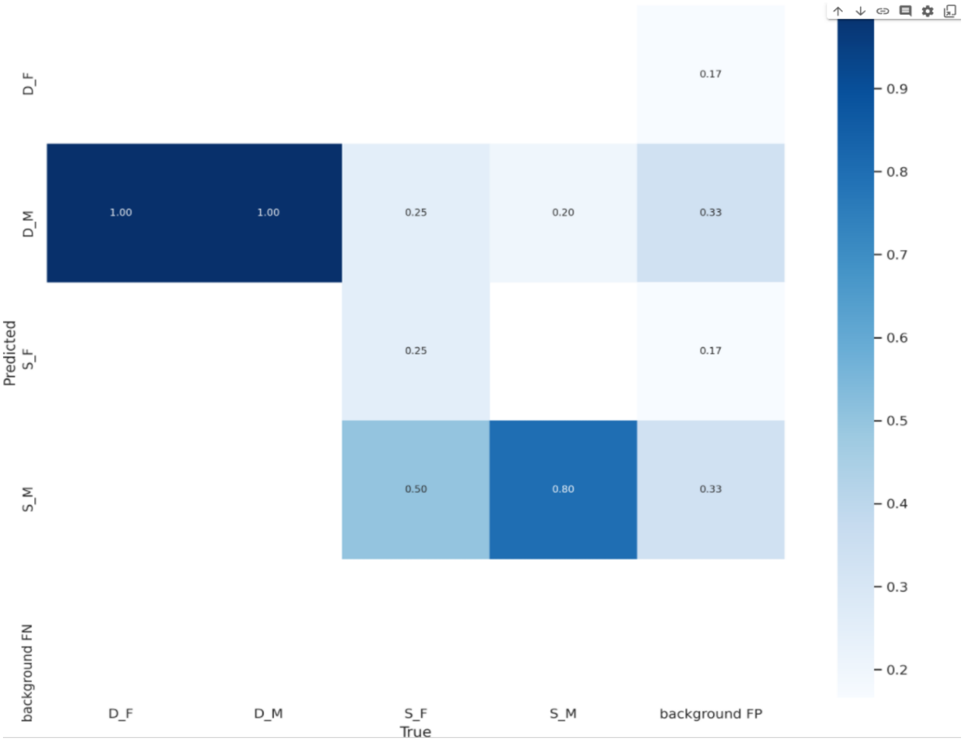


Confusion Matrix

The model predicted all the Hyalomma Dromedarii Male images correctly. It also predicts 80% of the Hyalomma Scupense Male correctly.

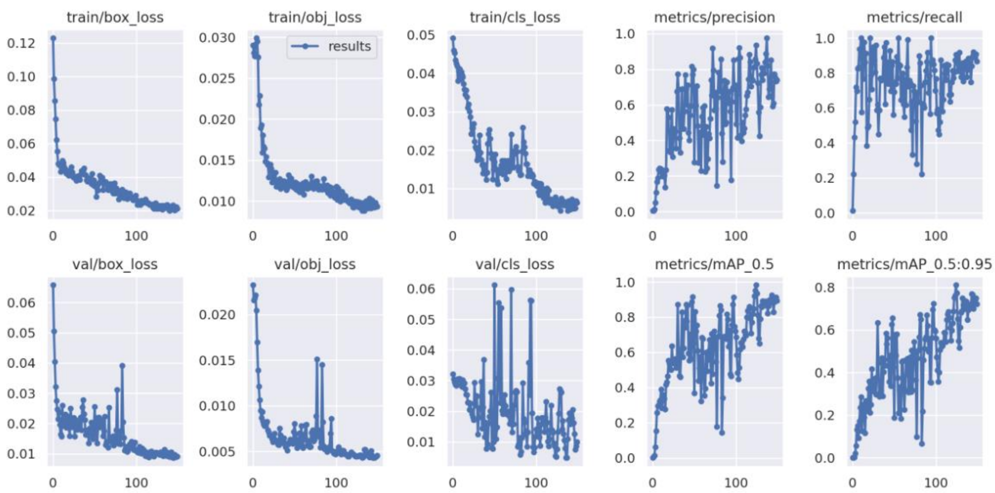
The model is only able to predict 25% of the the Hyalomma Scupense Female images accurately and misclassifies 50% of them the Hyalomma Scupense Male. This may be due to gross imbalance in the number of samples in the Hyalomma Scupense species.

The model is unable to classify Hyalomma Dromedarii Female images and misclassifies them as Hyalomma Dromedarii Male. This may be due to gross imbalance in the number of samples in the Hyalomma Dromedarii species.



Train and Validation Loss Metrics

The training loss and validation loss decreases as we train for a greater number of epochs. Similarly, the precession, recall and mAP increase as we train for more epochs.



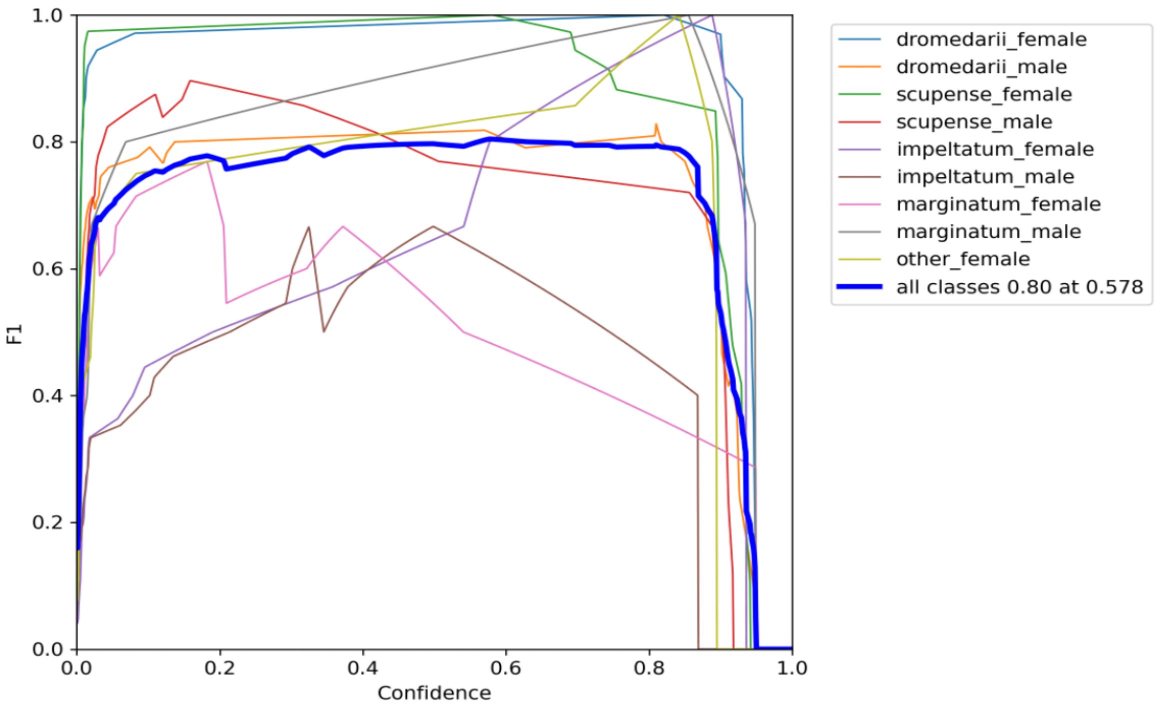
- **YOLOv5 on Tick Dataset 2 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.909** and **0.757** respectively.

Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:
all	88	91	0.897	0.799	0.909	0.757
dromedarii_female	88	17	0.982	1	0.995	0.859
dromedarii_male	88	22	0.817	0.812	0.833	0.694
scupense_female	88	19	1	1	0.995	0.852
scupense_male	88	16	1	0.612	0.885	0.681
impeltatum_female	88	2	0.667	1	0.995	0.945
impeltatum_male	88	4	1	0.446	0.674	0.565
marginatum_male	88	6	1	0.318	0.818	0.317
other_male	88	2	0.882	1	0.995	0.995
excavatum	88	3	0.721	1	0.995	0.907

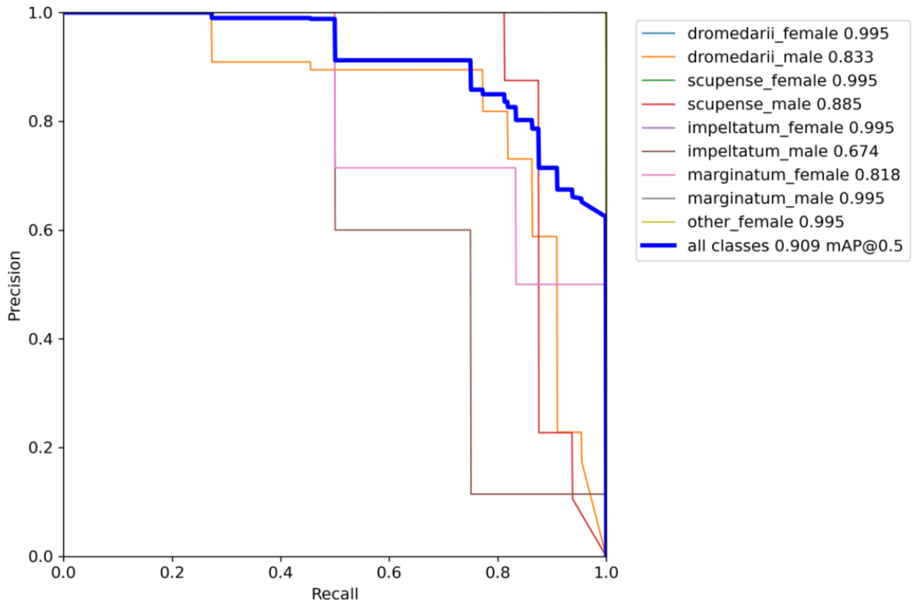
<u>F1 Curve</u>

The thick blue line represents the average of all classes. The F1 score of the classes with lower number of samples begins dropping after 0.80 confidence level and F1 score of the classes with greater number of samples begins dropping after 0.90 confidence level.



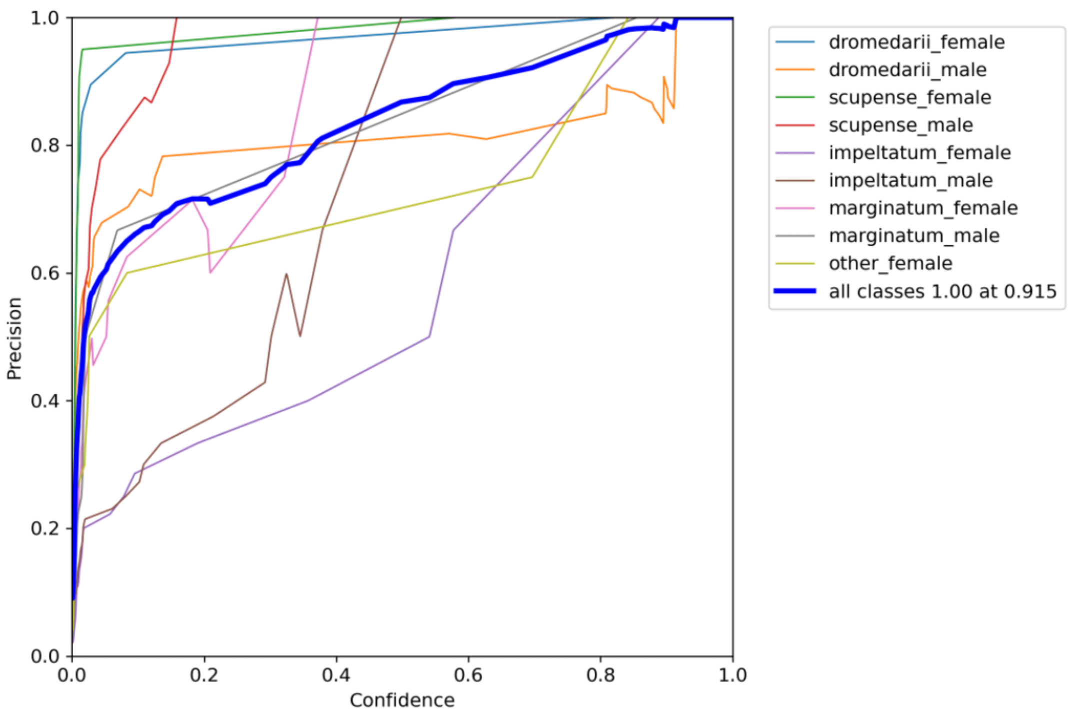
<u>PR Curve</u>

The thick blue line represents the average of all classes. As the recall increases the precision decreases. As precision increases as confidence threshold increases while recall decreases as confidence threshold increases.



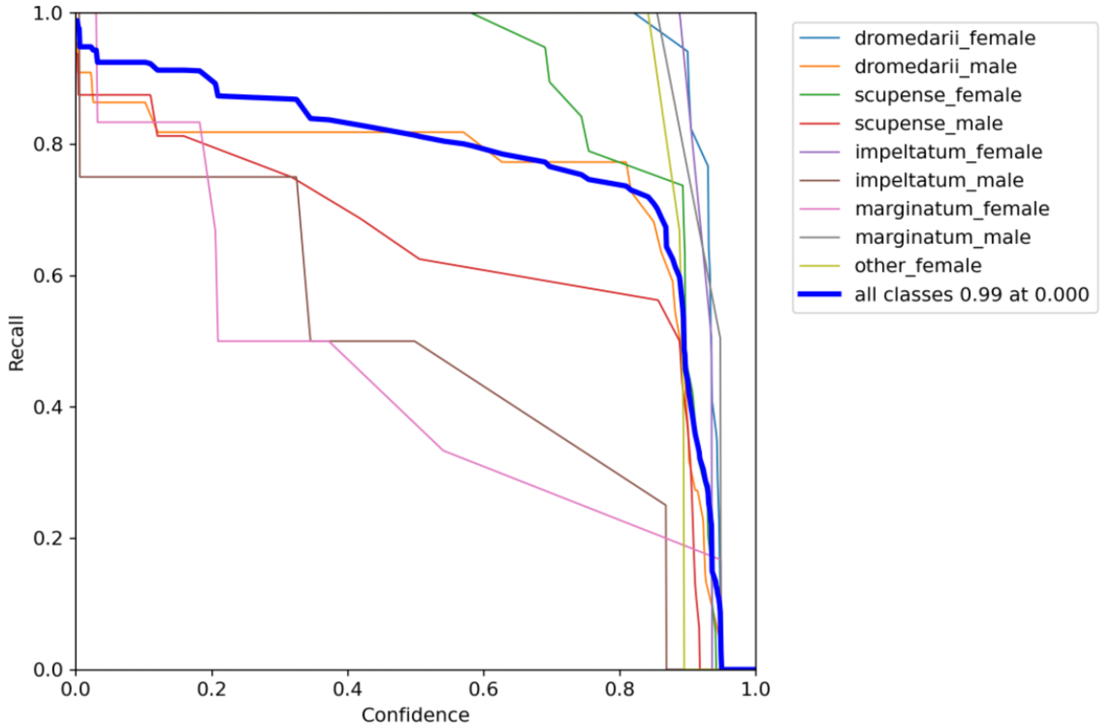
P Curve

As we increase the minimum confidence threshold the precision increases as the model becomes surer of its prediction.



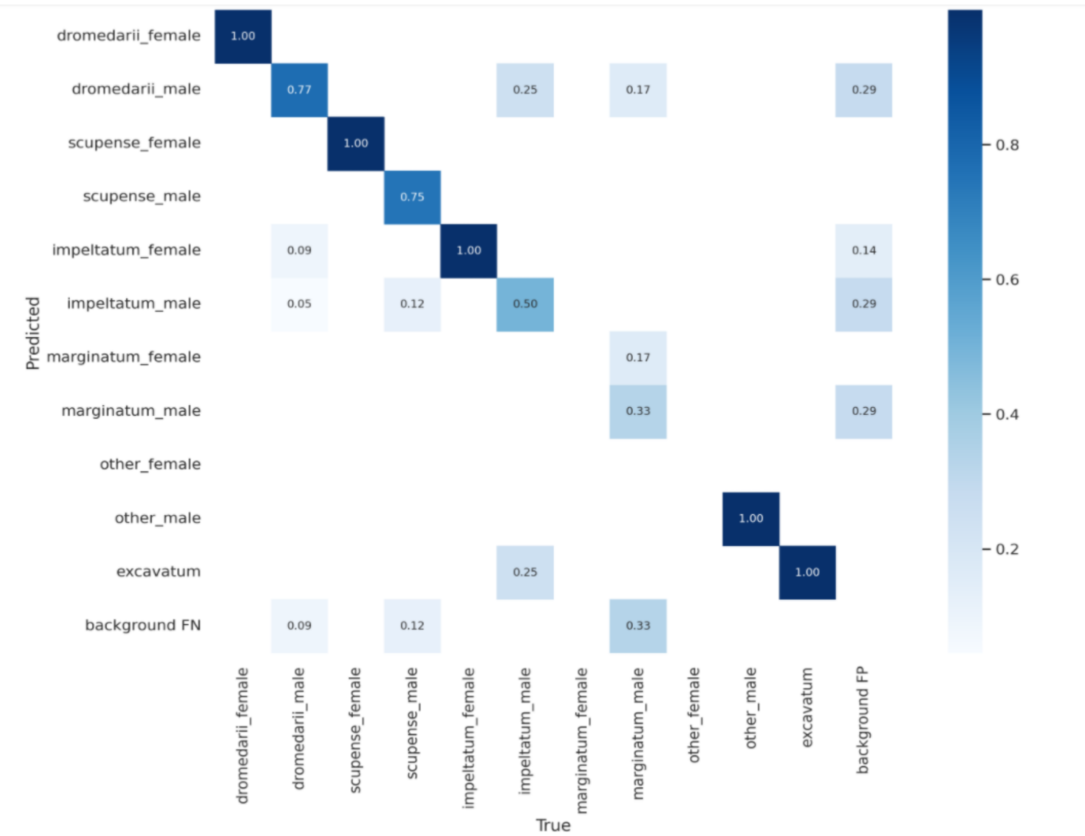
R Curve

As we increase the confidence level threshold the recall decreases as the model is able to predict less images with higher degree of confidence.



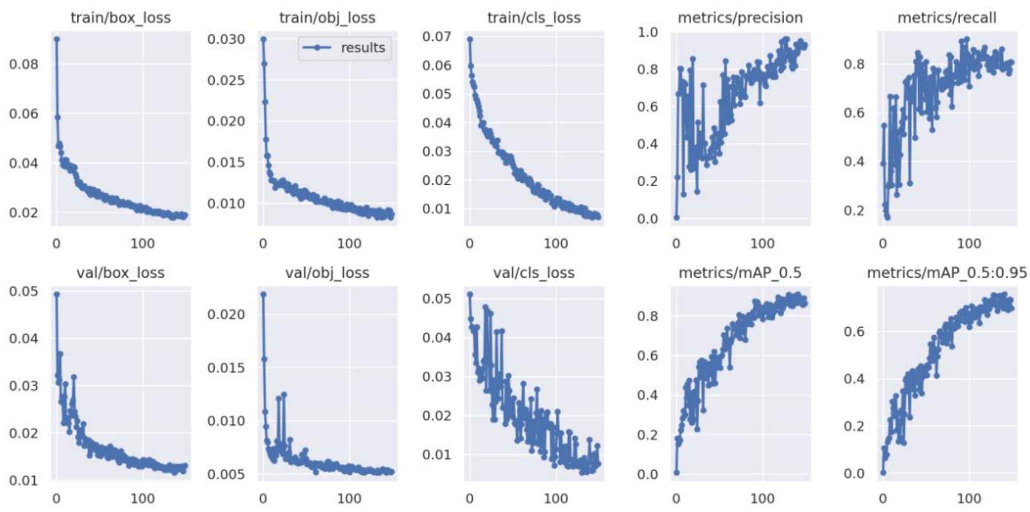
Confusion Matrix

The model classifies all the classes with greater number of training samples accurately while classes with lower number of samples are misclassified.



<u>Loss and Validation Loss Metrics</u>

The training loss and validation loss decreases as we train for a greater number of epochs. Similarly, the precession, recall and mAP increase as we train for more epochs.



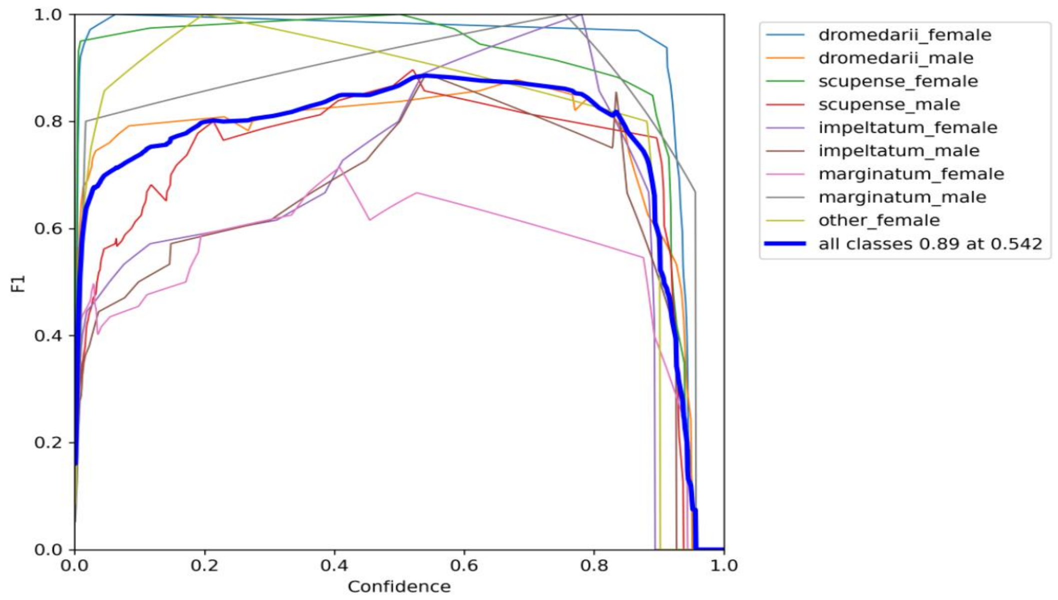
- **YOLOv5 on Tick Dataset 2 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.936** and **0.761** respectively.

Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:
all	92	95	0.931	0.87	0.936	0.761
dromedarii_female	92	17	1	0.965	0.995	0.821
dromedarii_male	92	22	0.871	0.818	0.872	0.737
scupense_female	92	19	1	0.974	0.995	0.829
scupense_male	92	16	1	0.749	0.916	0.626
impeltatum_female	92	4	0.804	1	0.995	0.834
impeltatum_male	92	4	0.8	1	0.945	0.875
marginatum_male	92	8	1	0.495	0.716	0.434
other_male	92	2	0.904	1	0.995	0.796
excavatum	92	3	1	0.833	0.995	0.895

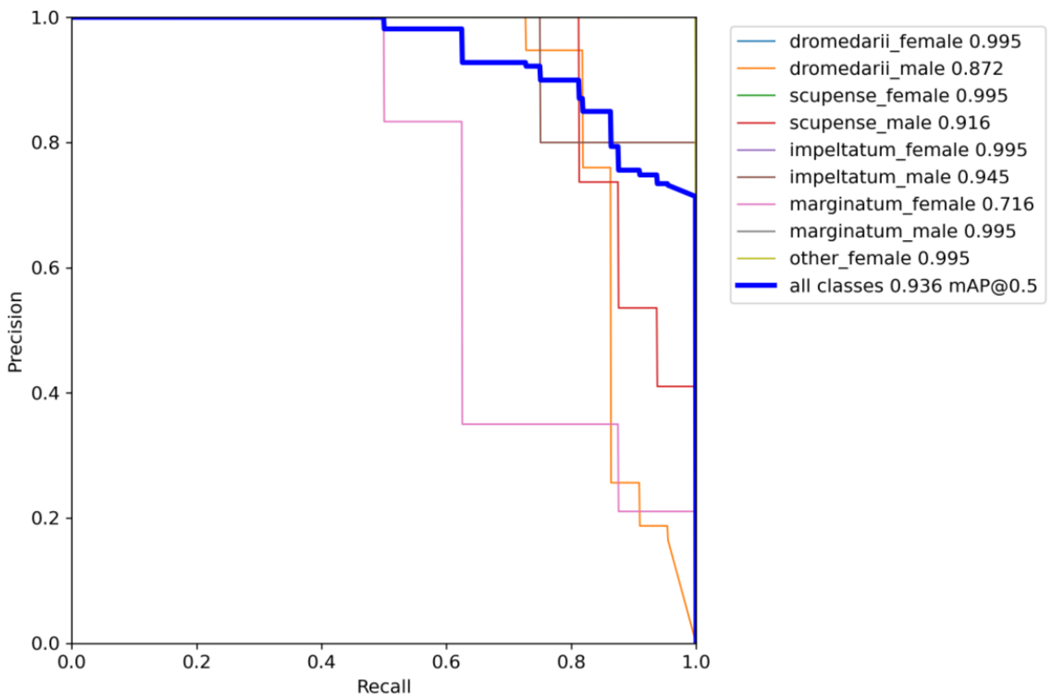
<u>F1 Curve</u>

The thick blue line represents the average of all classes. The F1 score of all the classes with dropping after 0.90 confidence level and hence training has improved due to augmentation.



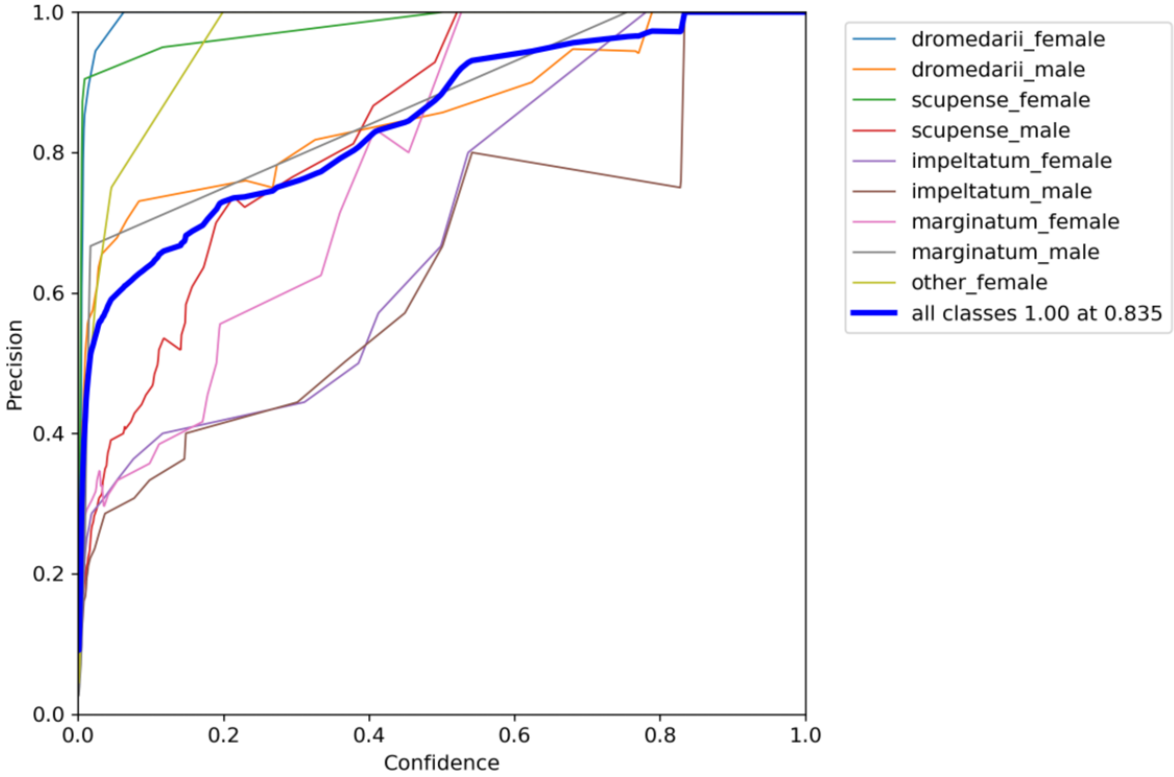
PR Curve

The thick blue line represents the average of all classes. As the recall increases the precision decreases. As precision increases as confidence threshold increases while recall decreases as confidence threshold increases.



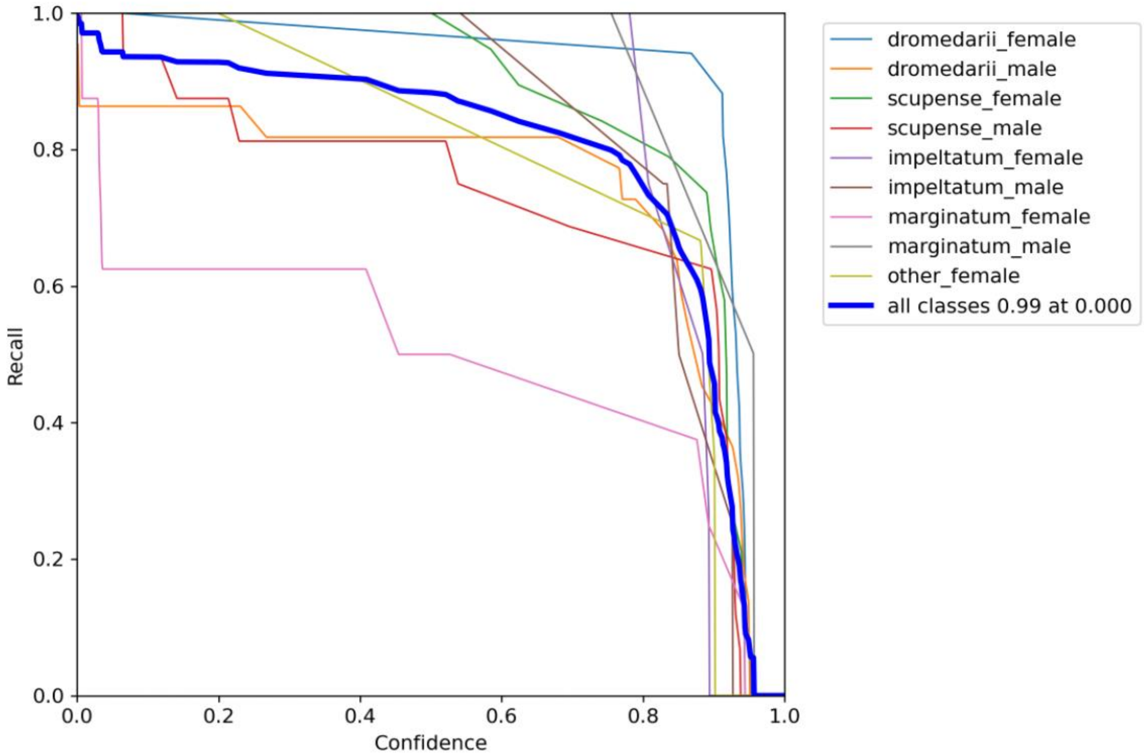
P Curve

As we increase the minimum confidence threshold the precision increases as the model becomes surer of its prediction.



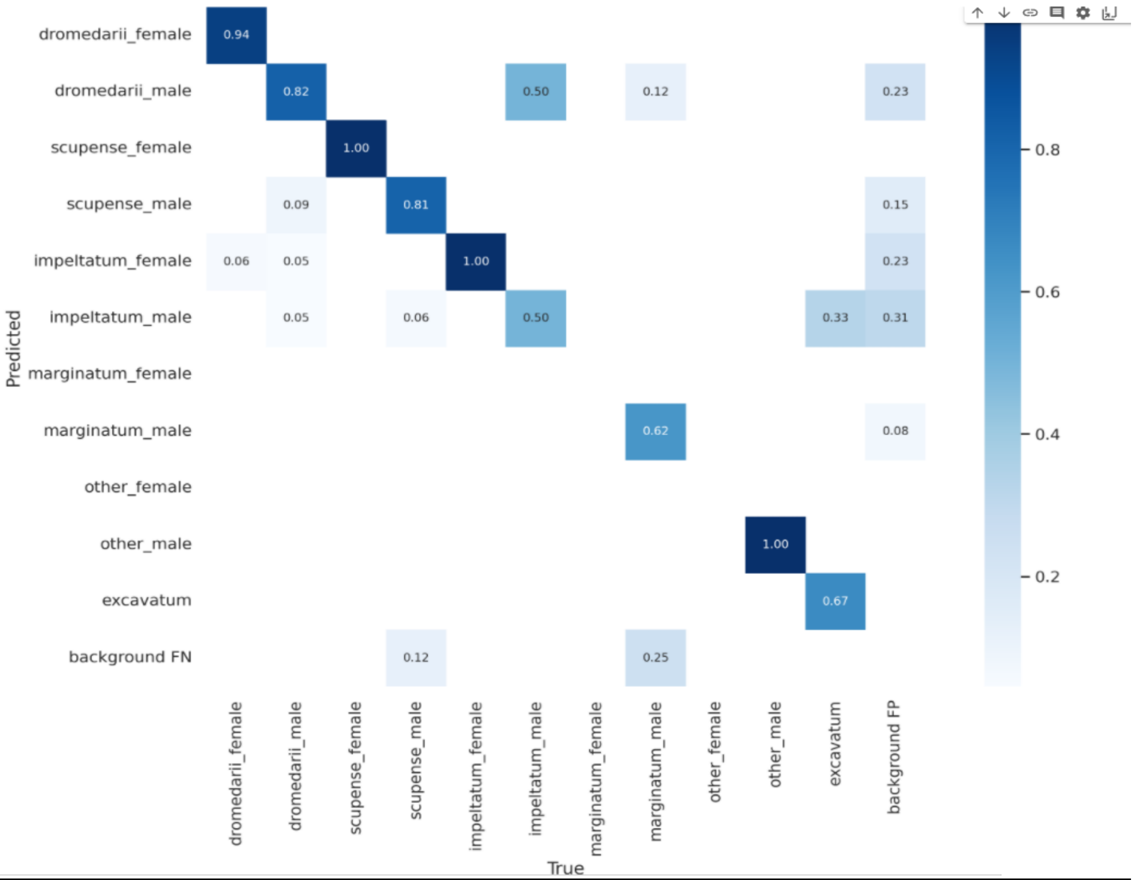
R Curve

As we increase the confidence level threshold the recall decreases as the model is able to predict less images with higher degree of confidence.



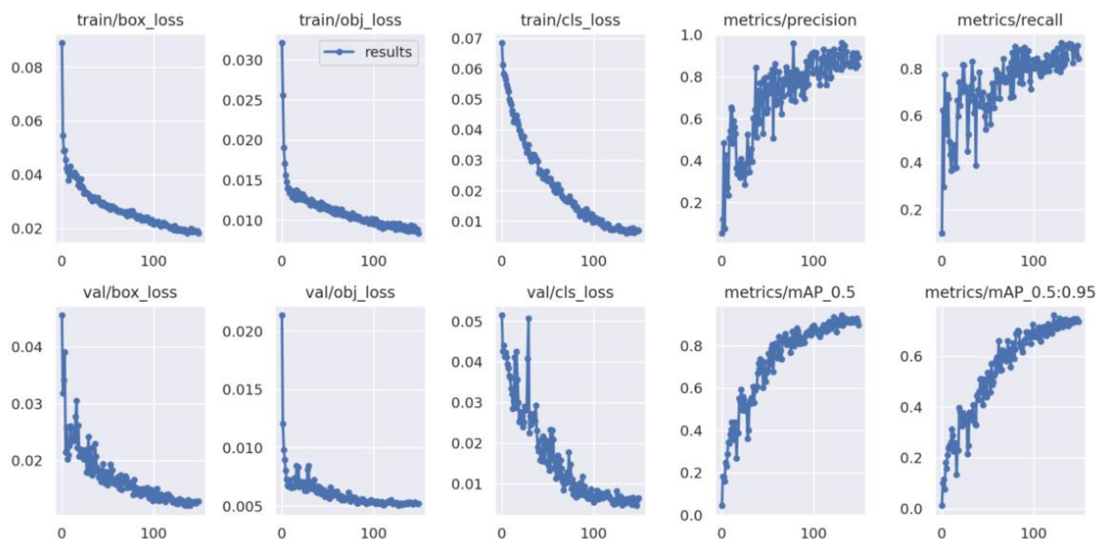
Confusion Matrix

The model classifies all the classes with greater number of training samples accurately while classes with lower number of samples are misclassified.



Loss and Validation Loss Metrics

The training loss and validation loss decreases as we train for a greater number of epochs. Similarly, the precession, recall and mAP increase as we train for more epochs.



DETR Model

DETR (DEtection TRansformer) approaches object detection as a direct set prediction problem. It consists of a set-based global loss, which forces unique predictions via bipartite matching, and a Transformer encoder-decoder architecture. Given a fixed small set of learned object queries, DETR reasons about the relations of the objects and the global image context to directly output the final set of predictions in parallel. Due to this parallel nature, DETR is very fast and efficient.

We train the DETR model for **100 epochs** on each dataset.

Results:

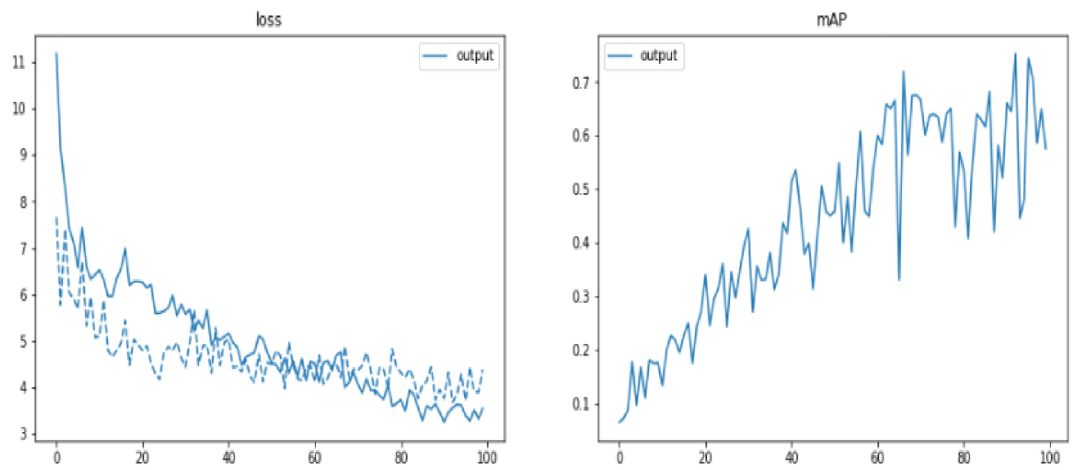
- **DETR on Tick Dataset 1 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.576** and **0.445** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.445
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.576
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.496
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.445
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.510
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.562
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.602
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.602
```

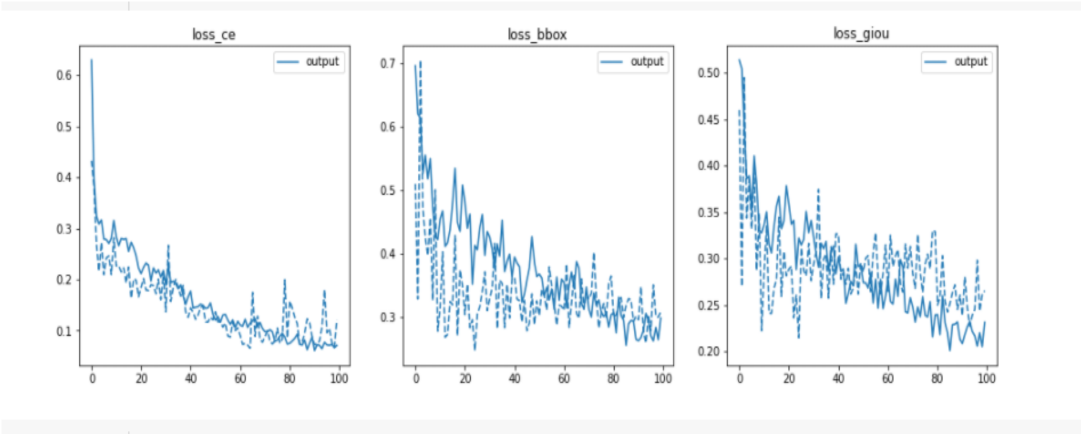
Learning Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases as mAP increases.



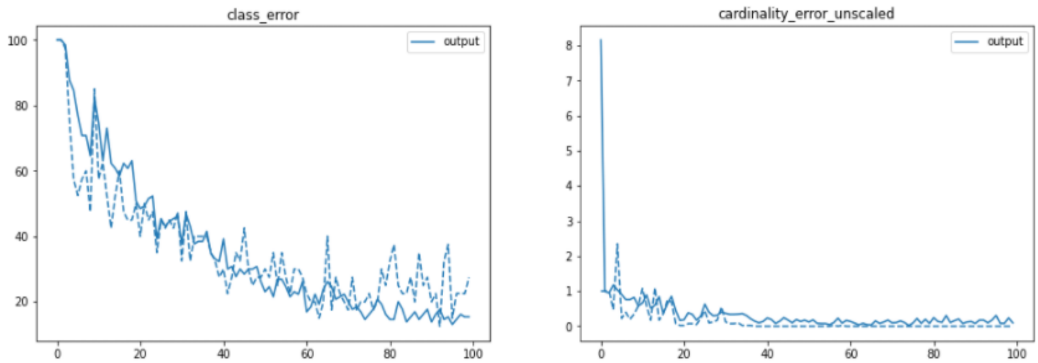
Loss Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases.



Error Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the error decreases.



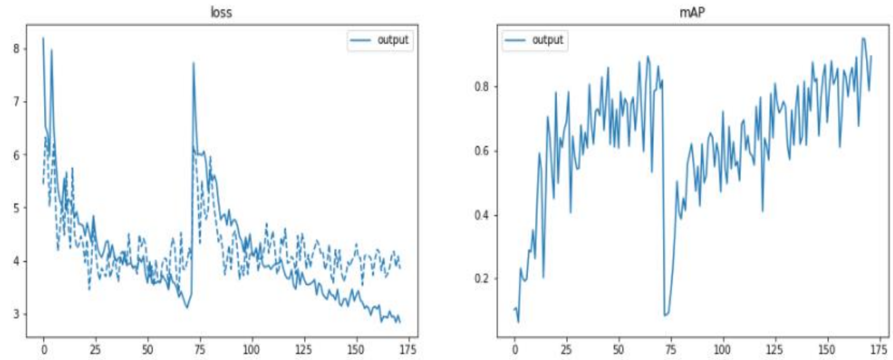
- **DETR on Tick Dataset 1 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.895** and **0.744** respectively.

```
IoU metric: bbox
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.744
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.895
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.880
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.744
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.767
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.767
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.767
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.767
```

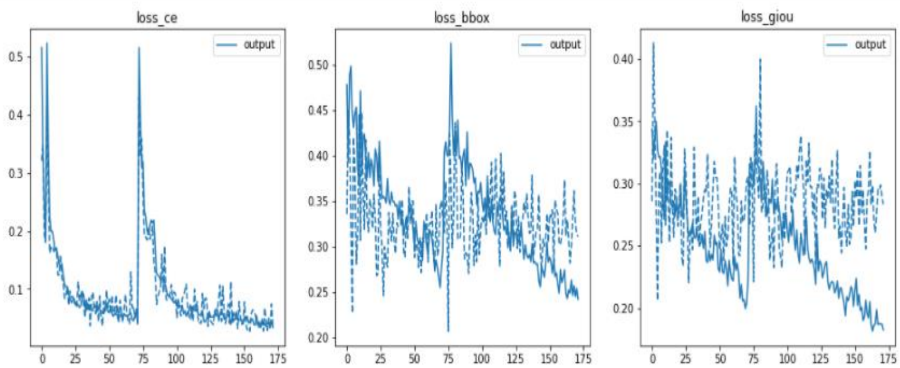
Learning Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases as mAP increases.



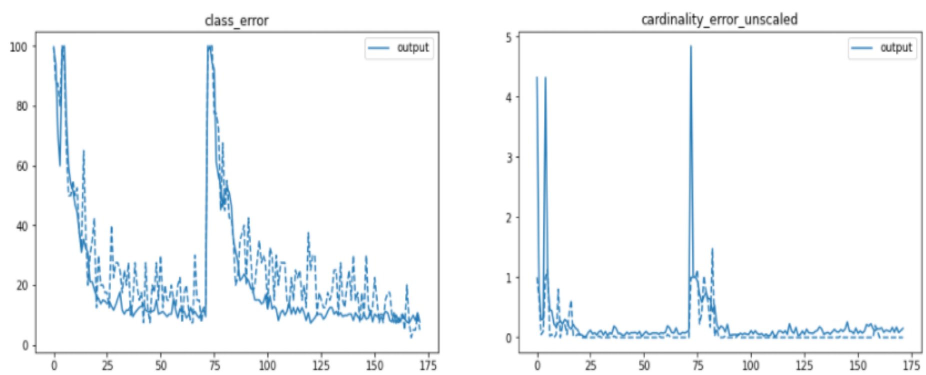
Loss Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases.



Error Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the error decreases.



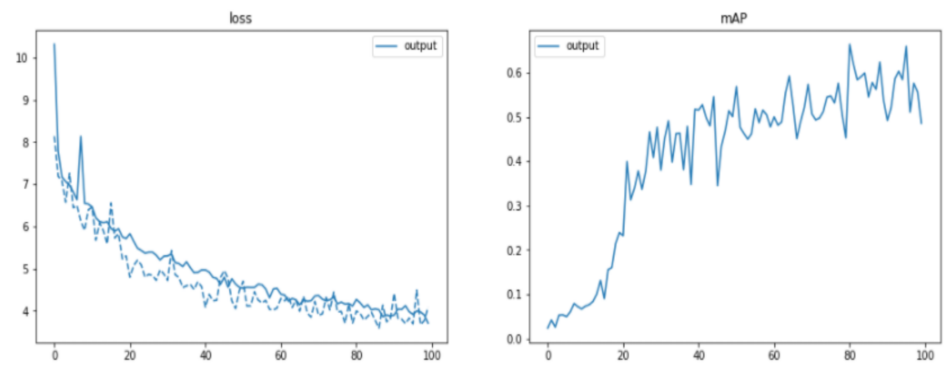
- **DETR on Tick Dataset 2 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.486** and **0.361** respectively.

```
IoU metric: bbox
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.361
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.486
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.442
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.361
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.514
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.623
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.699
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.699
```

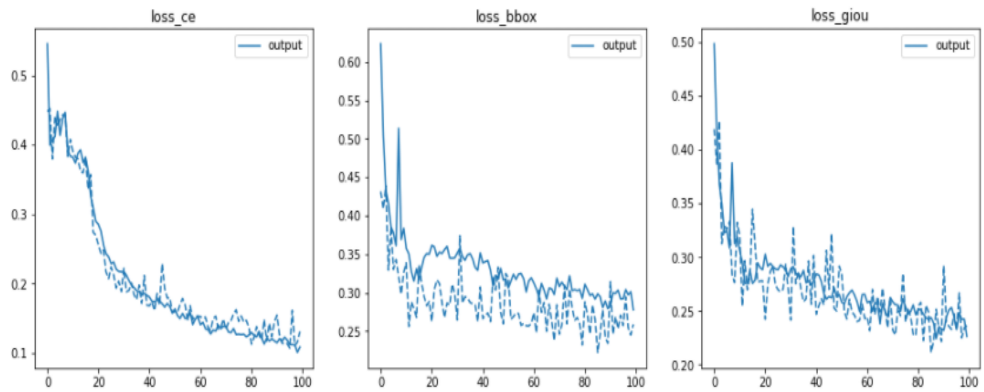
<u>Learning Curves</u>

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases as mAP increases.



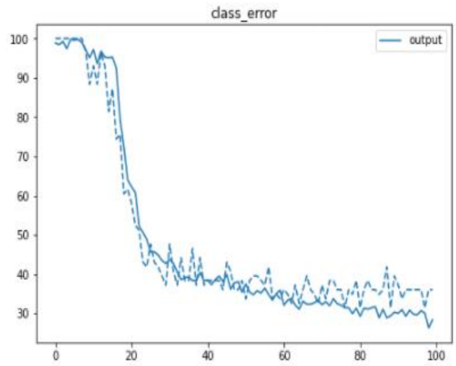
<u>Loss Curves</u>

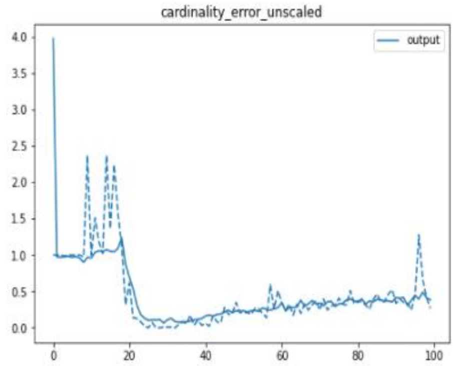
The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases.



Error Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the error decreases.





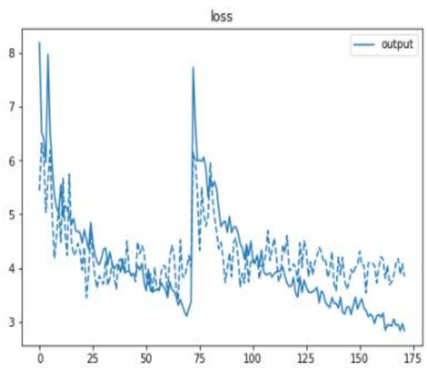
- **DETR on Tick Dataset 2 (with augmentation):**

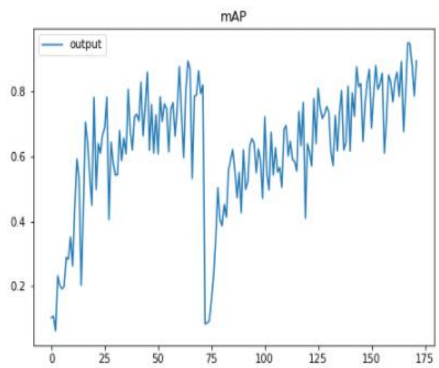
The mAP@.5 and mAP@.5:.95 for all classes is **0.629** and **0.483** respectively.

```
IoU metric: bbox
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.483
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.629
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.556
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.483
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.526
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.621
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.693
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.693
```

Learning Curves

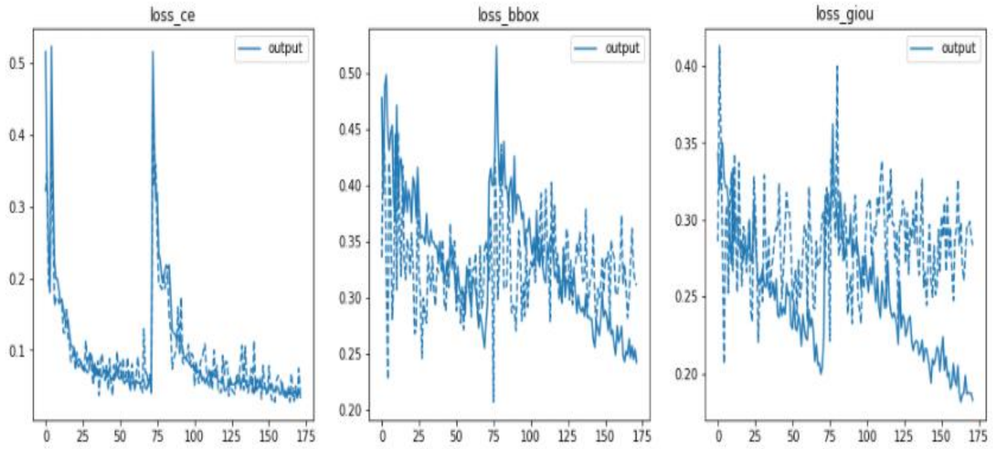
The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases as mAP increases.





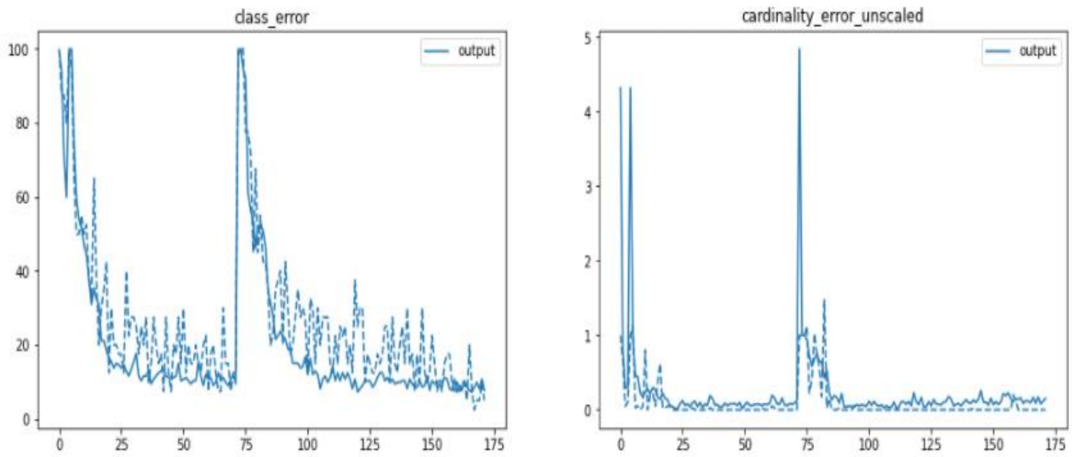
Loss Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the loss decreases.



Error Curves

The solid line represents the training curve, and the dotted line represents the validation curve. As the number of epochs increase, the error decreases.



FASTER RCNN MODEL

Faster RCNN is a deep convolutional network used for object detection, that appears to the user as a single, end-to-end, unified network. The network can accurately and quickly predict the locations of different objects.

We train the Faster RCNN model for **4000 iterations** on each dataset.

Results:

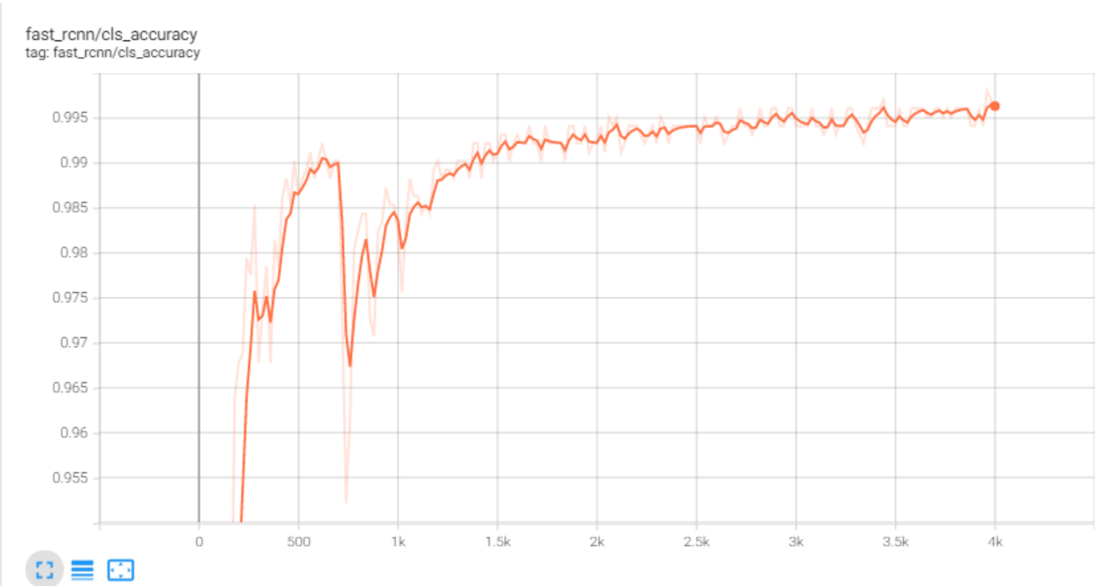
- **Faster RCNN on Tick Dataset 1 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.819** and **0.662** respectively.

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.662
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.819
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.819
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.662
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.700
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.700
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.700
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.700
[10/24 12:05:08 d2.evaluation.coco_evaluation]: Evaluation results for bbox:
|   AP   |  AP50  |  AP75  |  APs  |  APm  |  APl   |
|:------:|:------:|:------:|:-----:|:-----:|:------:|
| 66.214 | 81.858 | 81.858 |  nan  |  nan  | 66.214 |
[10/24 12:05:08 d2.evaluation.coco_evaluation]: Some metrics cannot be computed and is shown as NaN.
[10/24 12:05:08 d2.evaluation.coco_evaluation]: Per-category bbox AP:
| category          | AP     | category        | AP     | category        | AP     |
|:------------------|:-------|:----------------|:-------|:----------------|:-------|
| dromedarii_female | 62.955 | dromedarii_male | 80.399 | scupense_female | 46.733 |
| scupense_male     | 74.768 |                 |        |                 |        |
```

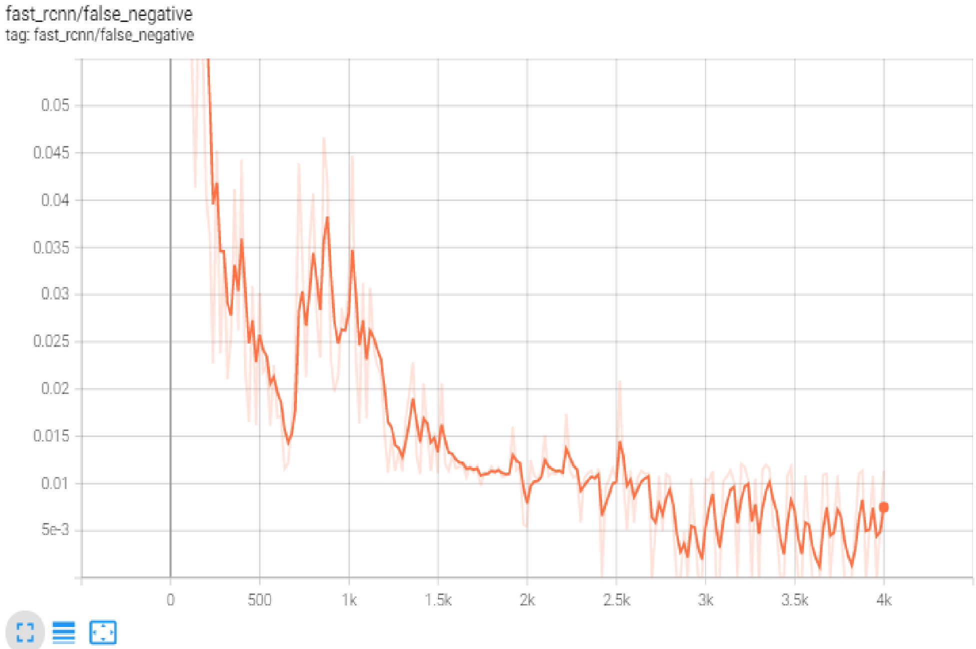
Training Class Accuracy Curve

As the number of iterations increases, the class accuracy increases.



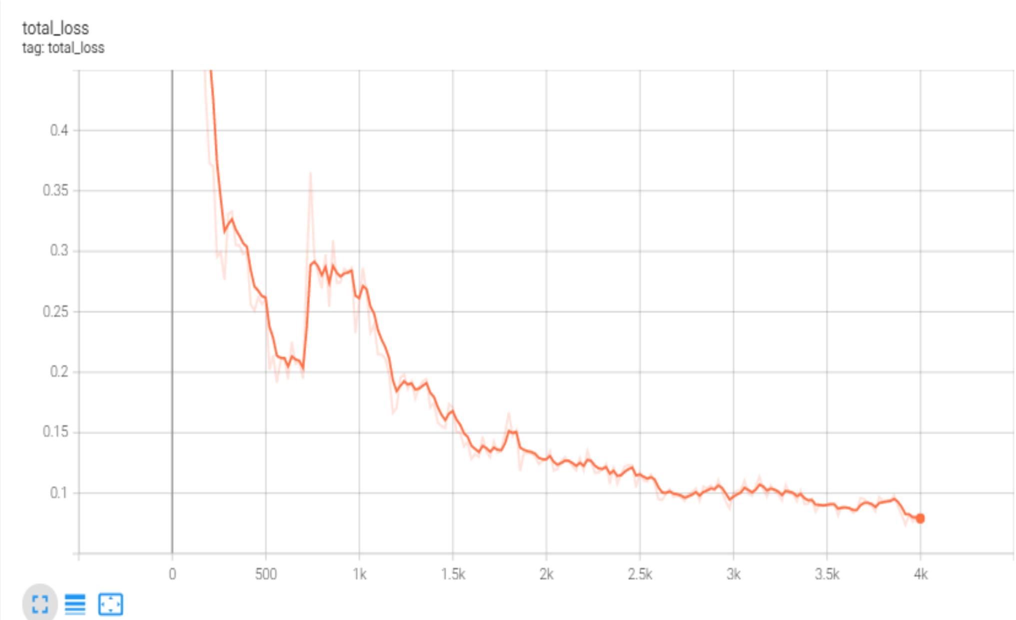
Training False Negative Curve

As the number of iterations increases, the number of False Negative decreases.



Training Total Loss Curve

As the number of iterations increases, the total loss decreases.



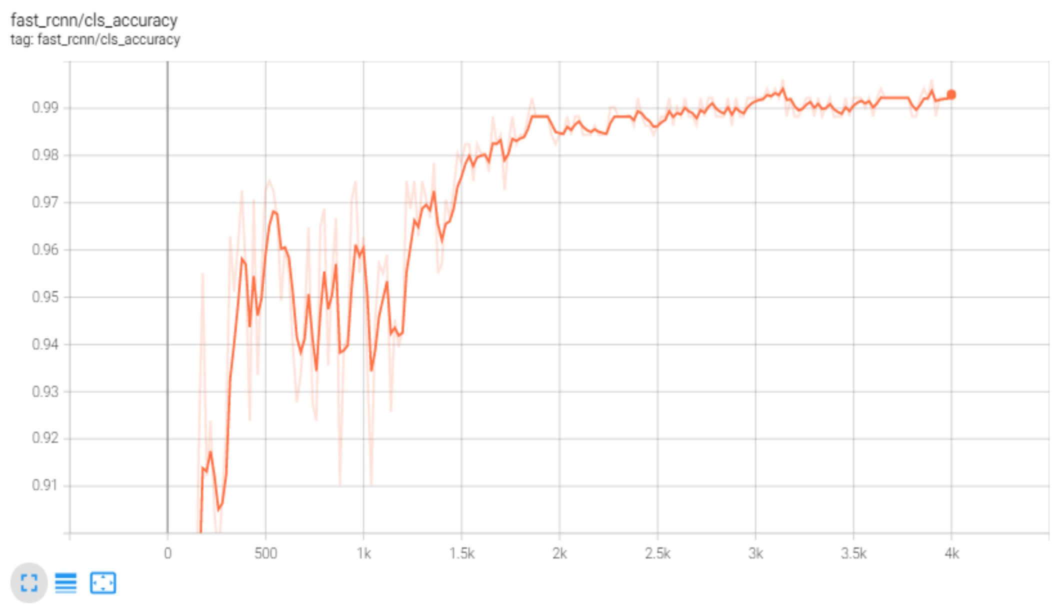
- **Faster RCNN on Tick Dataset 1 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.809** and **0.598** respectively.

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.598
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.809
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.767
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.598
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.626
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.626
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.626
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.626
[10/24 10:47:41 d2.evaluation.coco_evaluation]: Evaluation results for bbox:
|   AP   |  AP50  |  AP75  |  APs  |  APm  |  APl   |
|:------:|:------:|:------:|:-----:|:-----:|:------:|
| 59.843 | 80.941 | 76.733 |  nan  |  nan  | 59.843 |
[10/24 10:47:41 d2.evaluation.coco_evaluation]: Some metrics cannot be computed and is shown as NaN.
[10/24 10:47:41 d2.evaluation.coco_evaluation]: Per-category bbox AP:
| category          | AP     | category        | AP     | category        | AP     |
|:------------------|:-------|:----------------|:-------|:----------------|:-------|
| dromedarii_female | 61.584 | dromedarii_male | 75.213 | scupense_female | 72.574 |
| scupense_male     | 30.000 |                 |        |                 |        |
```

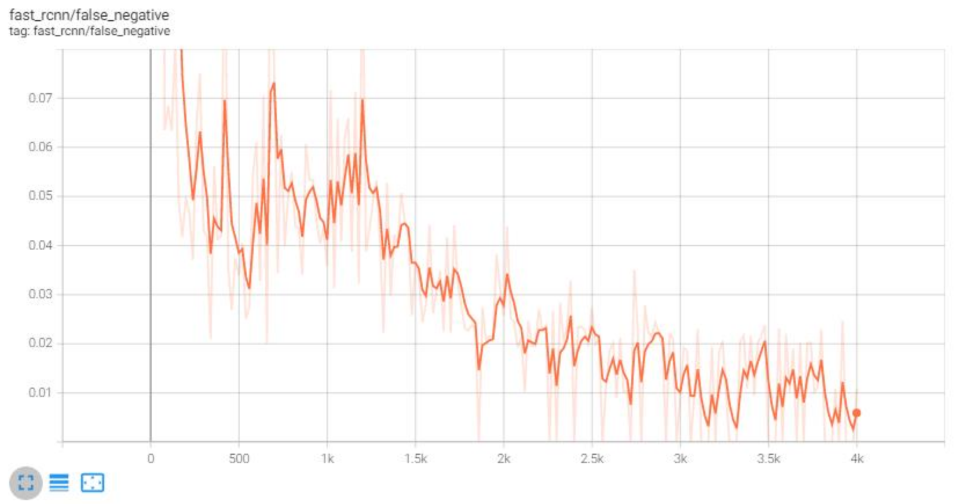
Training Class Accuracy Curve

As the number of iterations increases, the class accuracy increases.



Training False Negative Curve

As the number of iterations increases, the number of False Negative decreases.



Training Total Loss Curve

As the number of iterations increases, the total loss decreases.

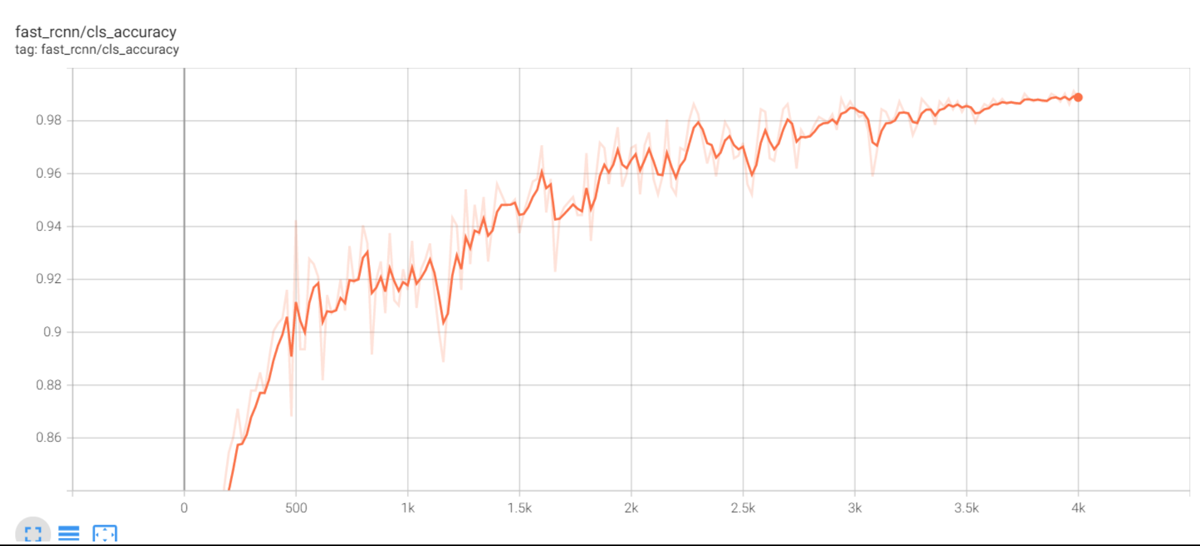
- **Faster RCNN on Tick Dataset 2 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.916** and **0.690** respectively.

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.690
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.916
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.862
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.690
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.718
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.718
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.718
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.718
[10/19 04:10:54 d2.evaluation.coco_evaluation]: Evaluation results for bbox:
|   AP   |  AP50  |  AP75  |  APs  |  APm  |  APl   |
|:------:|:------:|:------:|:-----:|:-----:|:------:|
| 69.021 | 91.639 | 86.202 |  nan  |  nan  | 69.021 |
[10/19 04:10:54 d2.evaluation.coco_evaluation]: Some metrics cannot be computed and is shown as NaN.
[10/19 04:10:54 d2.evaluation.coco_evaluation]: Per-category bbox AP:
| category          | AP     | category          | AP     | category         | AP     |
|:------------------|:-------|:------------------|:-------|:-----------------|:-------|
| dromedarii_female | 80.150 | dromedarii_male   | 68.662 | scupense_female  | 66.838 |
| scupense_male     | 54.180 | impeltatum_female | 90.000 | impeltatum_male  | 66.931 |
| marginatum_female | nan    | marginatum_male   | 54.431 | other_female     | nan    |
| other_male        | 70.000 | excavatum         | 70.000 |                  |        |
```

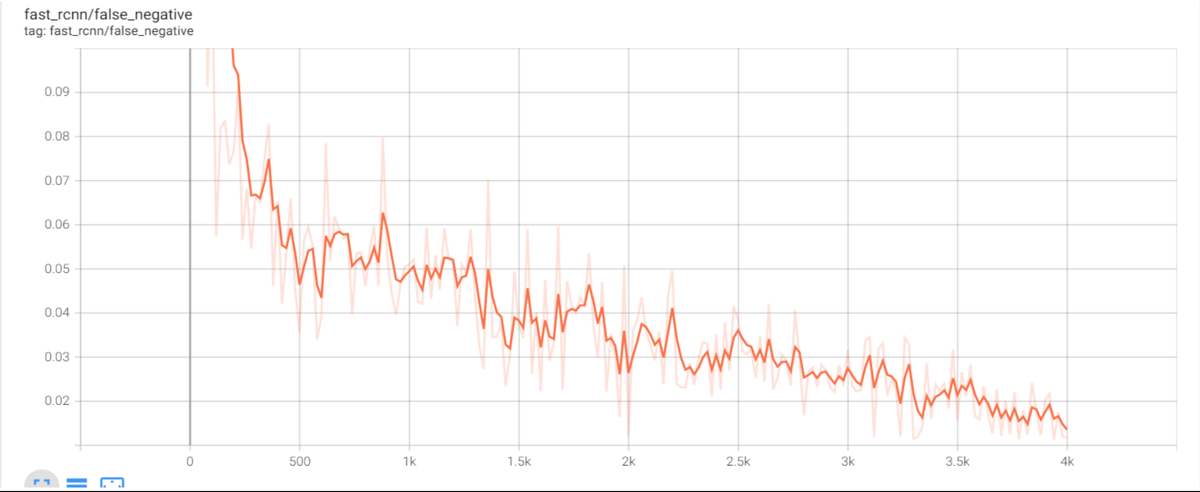
<u>Training Class Accuracy Curve</u>

As the number of iterations increases, the class accuracy increases.



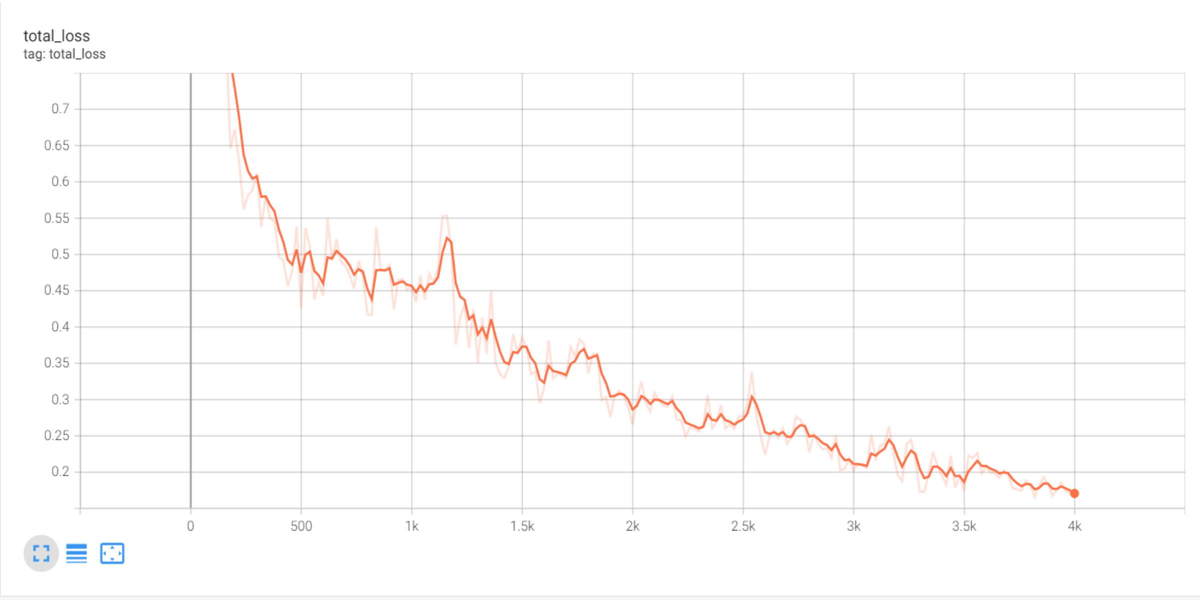
Training False Negative Curve

As the number of iterations increases, the number of False Negative decreases.



Training Total Loss Curve

As the number of iterations increases, the total loss decreases.



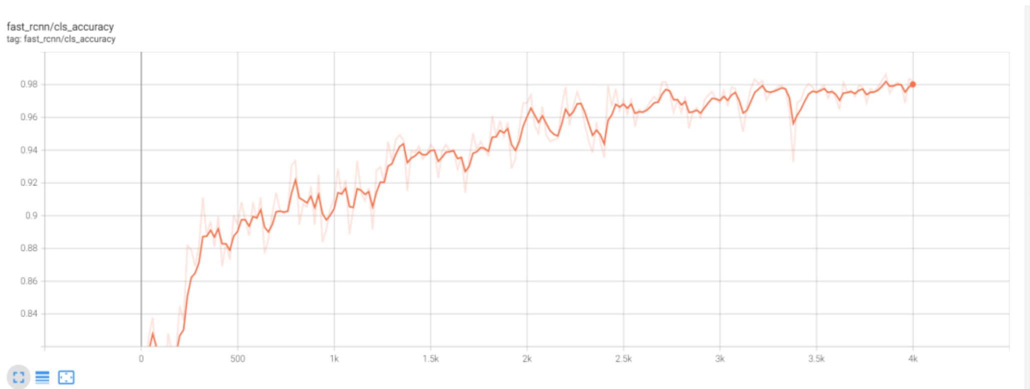
- **Faster RCNN on Tick Dataset 2 (with augmentation)**

The mAP@.5 and mAP@.5:.95 for all classes is **0.750** and **0.572** respectively.

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.572
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.750
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.709
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.572
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.598
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.598
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.598
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.598
[10/19 17:28:06 d2.evaluation.coco_evaluation]: Evaluation results for bbox:
|   AP   |  AP50  |  AP75  |  APs  |  APm  |  APl   |
|:------:|:------:|:------:|:-----:|:-----:|:------:|
| 57.173 | 75.024 | 70.881 |  nan  |  nan  | 57.173 |
[10/19 17:28:06 d2.evaluation.coco_evaluation]: Some metrics cannot be computed and is shown as NaN.
[10/19 17:28:06 d2.evaluation.coco_evaluation]: Per-category bbox AP:
| category          | AP     | category          | AP     | category        | AP     |
|:------------------|:-------|:------------------|:-------|:----------------|:-------|
| dromedarii_female | 76.344 | dromedarii_male   | 77.459 | scupense_female | 76.583 |
| scupense_male     | 62.221 | impeltatum_female | 82.525 | impeltatum_male | 0.000  |
| marginatum_female | nan    | marginatum_male   | 49.224 | other_female    | nan    |
| other_male        | 40.396 | excavatum         | 49.802 |                 |        |
```

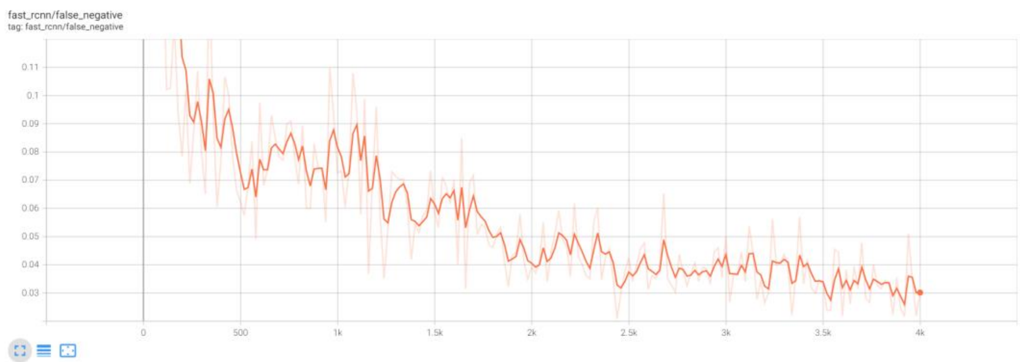
Training Class Accuracy Curve

As the number of iterations increases, the class accuracy increases.



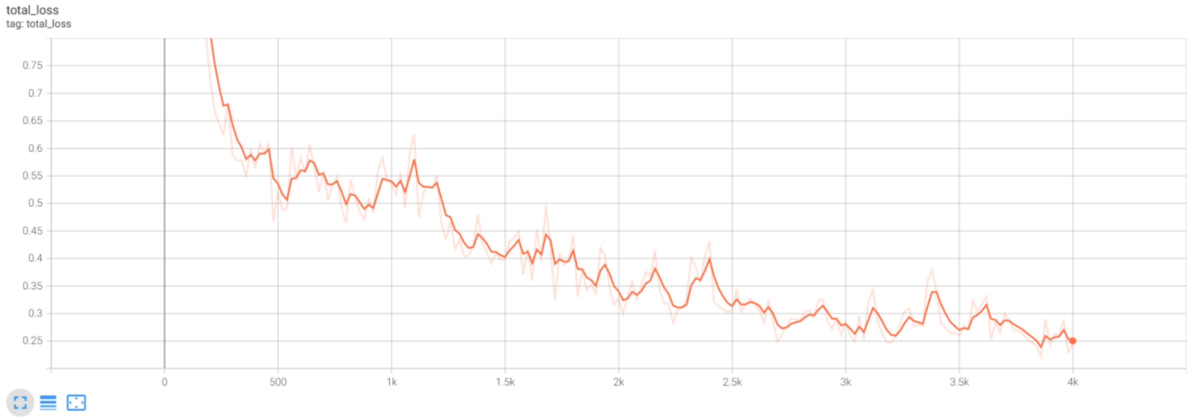
Training False Negative Curve

As the number of iterations increases, the number of False Negative decreases.



Training Total Loss Curve

As the number of iterations increases, the total loss decreases.



EFFICIENTDET MODEL

EfficientDet detectors are single-shot detectors much like SSD and RetinaNet. The backbone networks are ImageNet pretrained EfficientNets. The proposed BiFPN serves as the feature network, which takes level 3–7 features {P3, P4, P5, P6, P7} from the backbone network and repeatedly applies top-down and bottom-up bidirectional feature fusion. These fused features are fed to a class and box network to produce object class and bounding box predictions respectively. The class and box network weights are shared across all levels of features.

We train the EfficientDet model for **20000 steps** on each dataset.

Results:

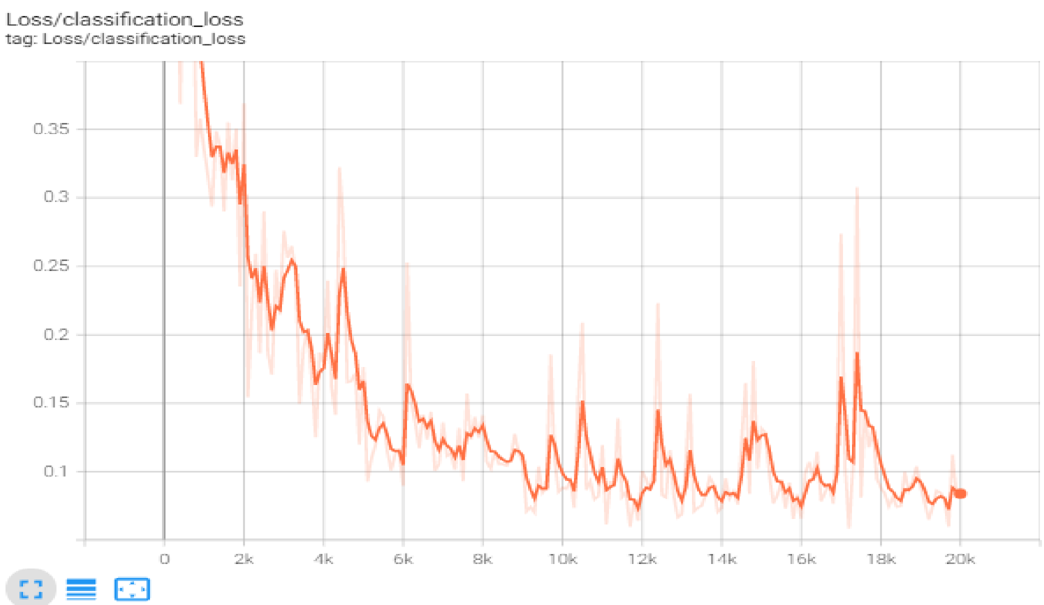
- **EfficientDet on Tick Dataset 1 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.739** and **0.543** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.543
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.739
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.651
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.543
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.711
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.738
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.738
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.738
```

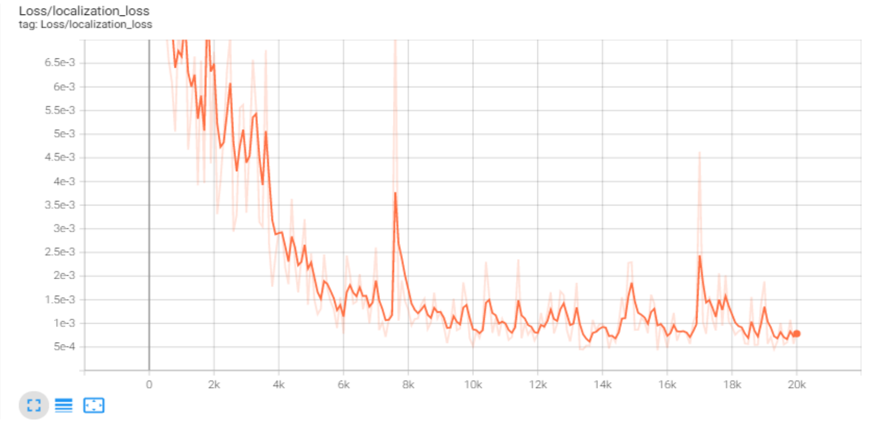
<u>Training Classification Loss Curve</u>

As the number of steps increases, the classification loss decreases.



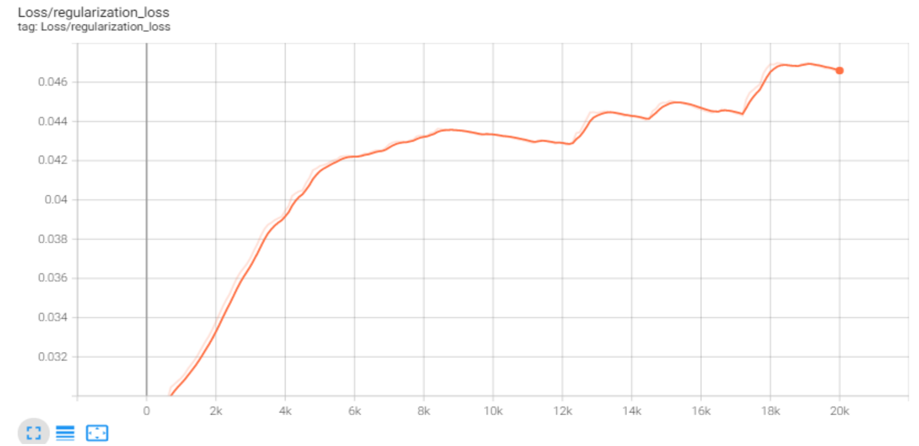
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



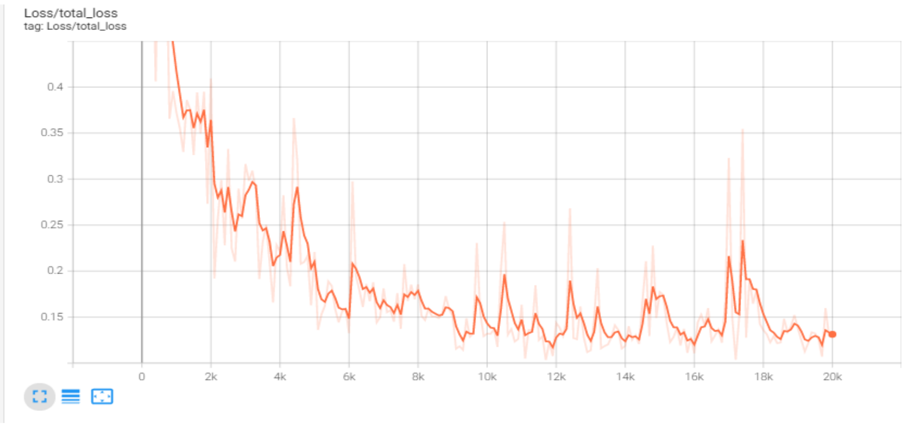
Training Regularization Loss Curve

As the number of steps increases, the regularization loss should decrease but it increases. This indicates that the model is not training well on our data.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



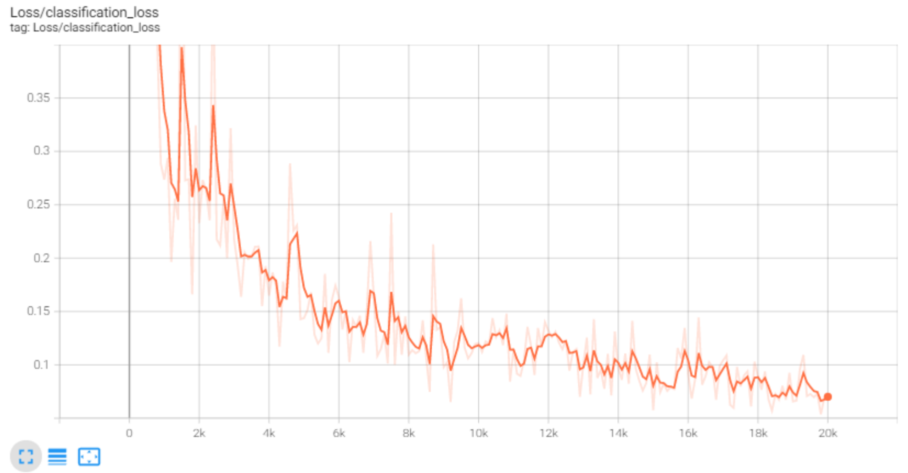
- **EfficientDet on Tick Dataset 1 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.793** and **0.618** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.618
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.793
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.740
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.618
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.753
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.786
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.786
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.786
```

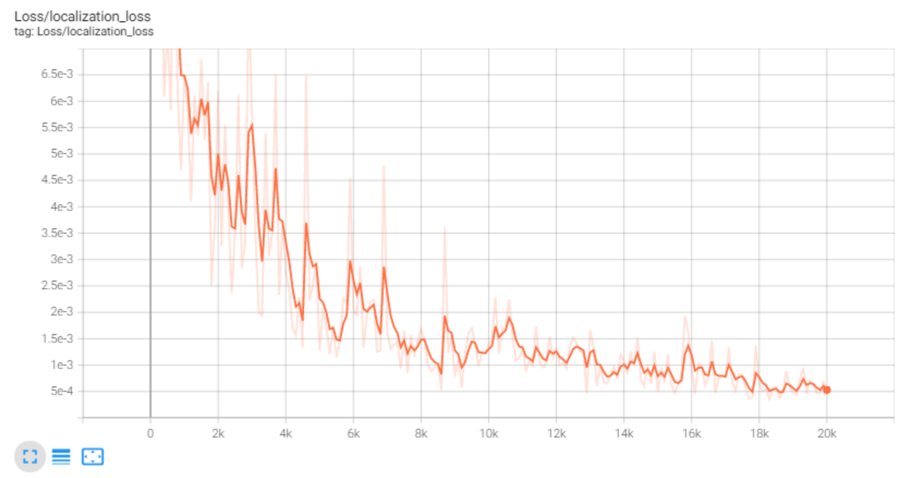
Training Classification Loss Curve

As the number of steps increases, the classification loss decreases.



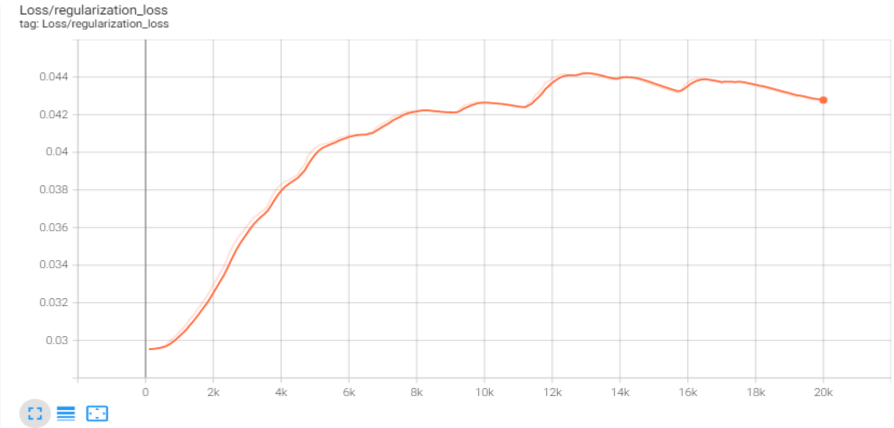
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



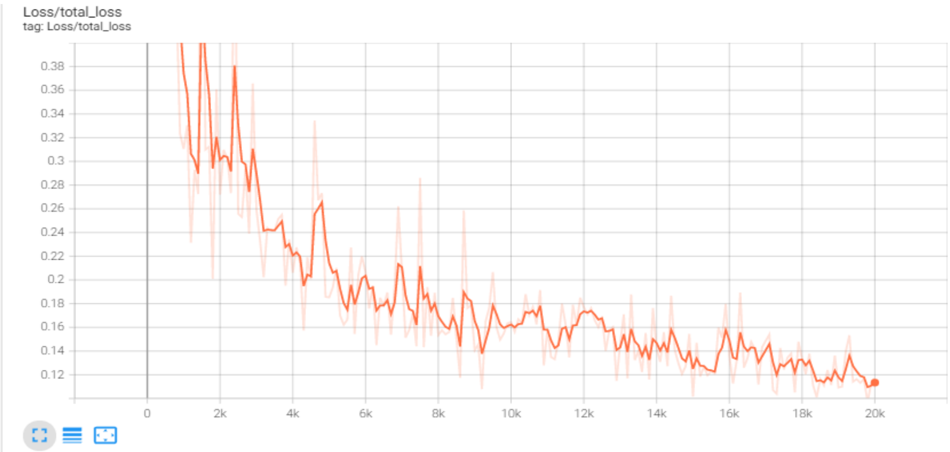
Training Regularization Loss Curve

As the number of steps increases, the regularization loss should decrease but it increases. This indicates that the model is not training well on our data.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



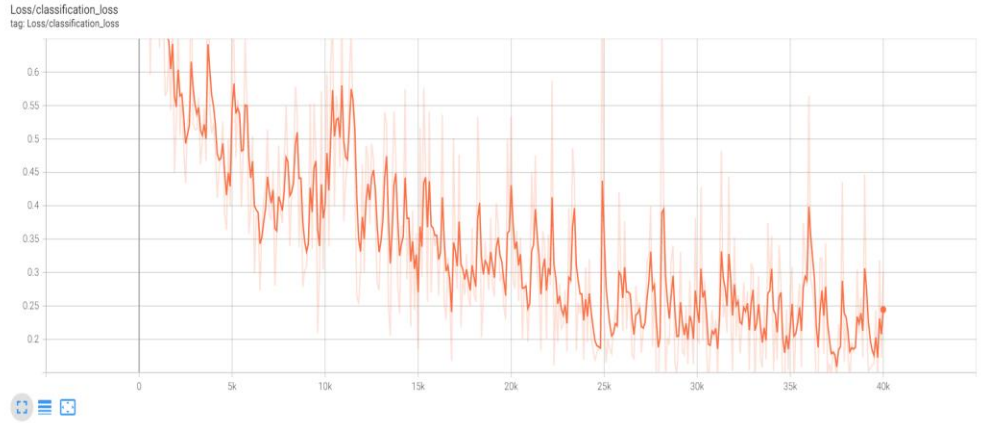
- **EfficientDet on Tick Dataset 2 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.437** and **0.295** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.295
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.437
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.354
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.295
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.489
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.565
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.593
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.593
```

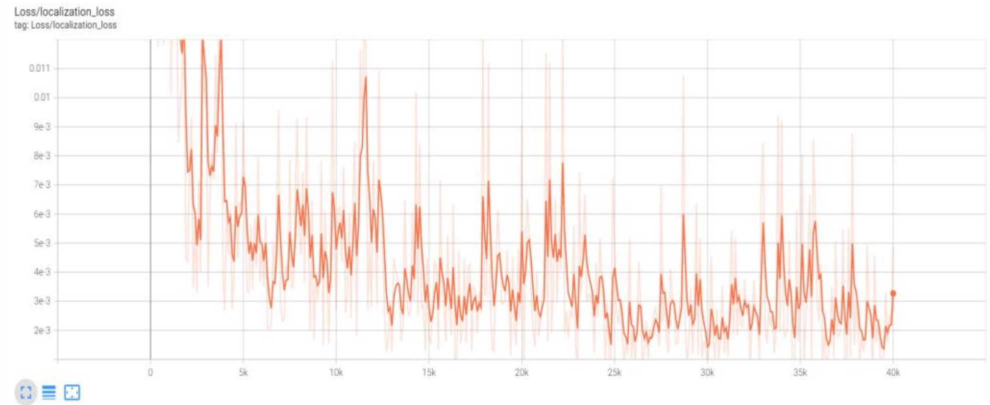
Training Classification Loss Curve

As the number of steps increases, the classification loss decreases.



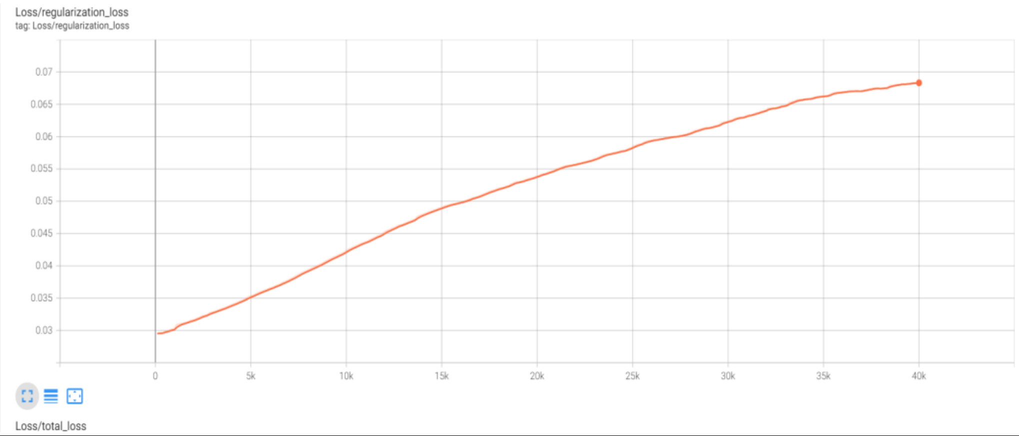
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



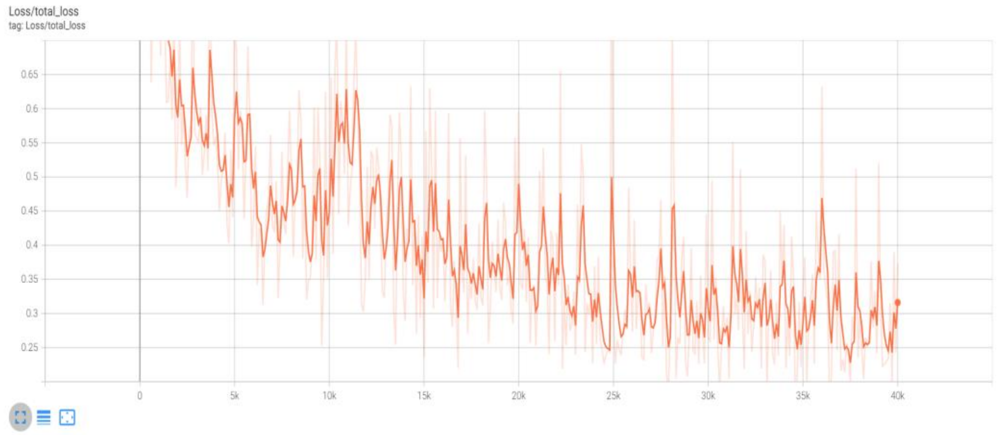
Training Regularization Loss Curve

As the number of steps increases, the regularization loss should decrease but it increases. This indicates that the model is not training well on our data.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



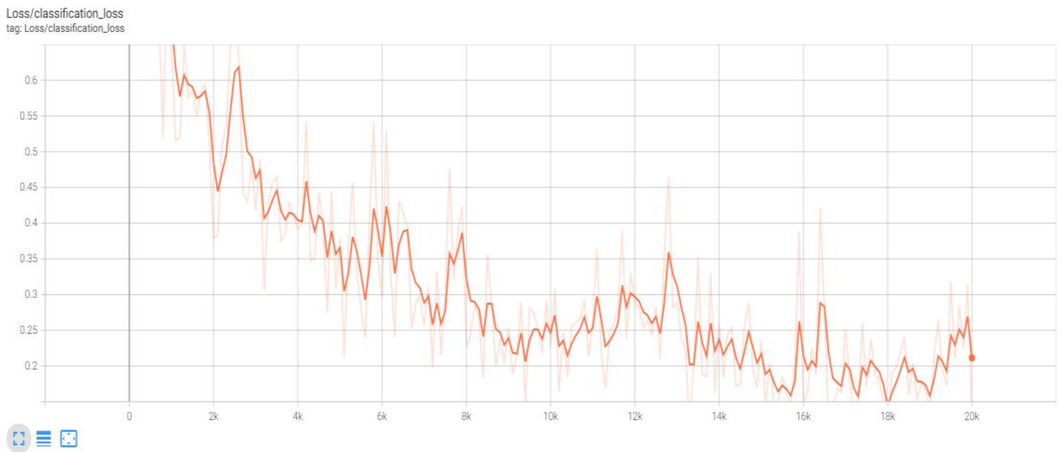
- **EfficientDet on Tick Dataset 2 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.703** and **0.429** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.429
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.703
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.500
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.429
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.519
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.569
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.602
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.602
```

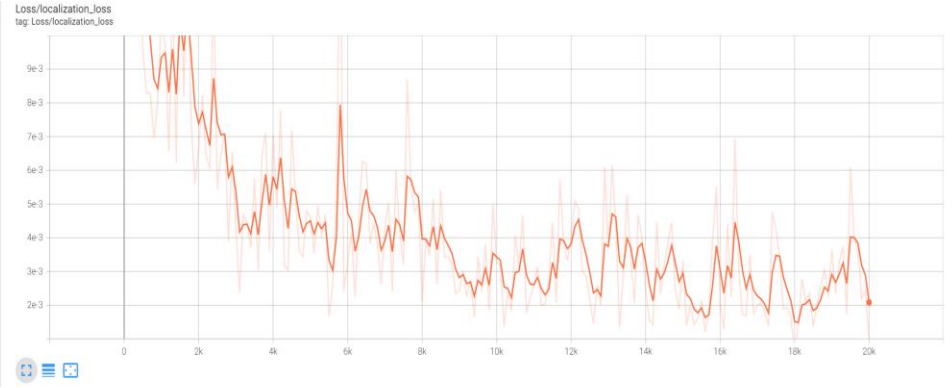
Training Classification Loss Curve

As the number of steps increases, the classification loss decreases.



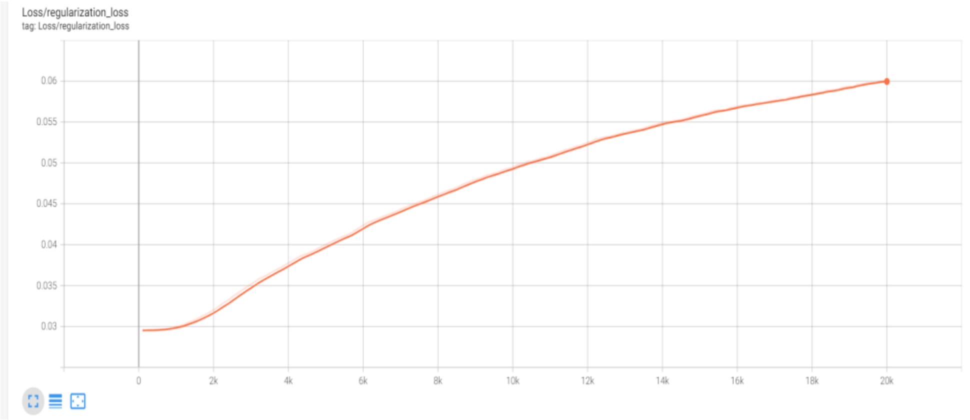
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



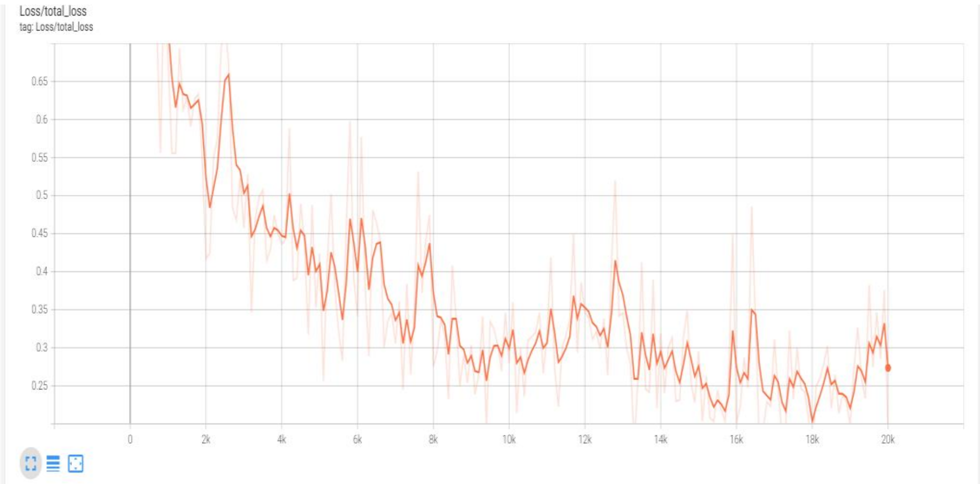
Training Regularization Loss Curve

As the number of steps increases, the regularization loss should decrease but it increases. This indicates that the model is not training well on our data.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



SSD MODEL

SSD (Single-Shot Detector) has two components: a **backbone** model and **SSD head**. *Backbone* model usually is a pre-trained image classification network as a feature extractor. This is typically a network like ResNet trained on ImageNet from which the final fully connected classification layer has been removed. We are thus left with a deep neural network that is able to extract semantic meaning from the input image while preserving the spatial structure of the image albeit at a lower resolution.

We train the SSD model for **40000 steps** on each dataset.

Results:

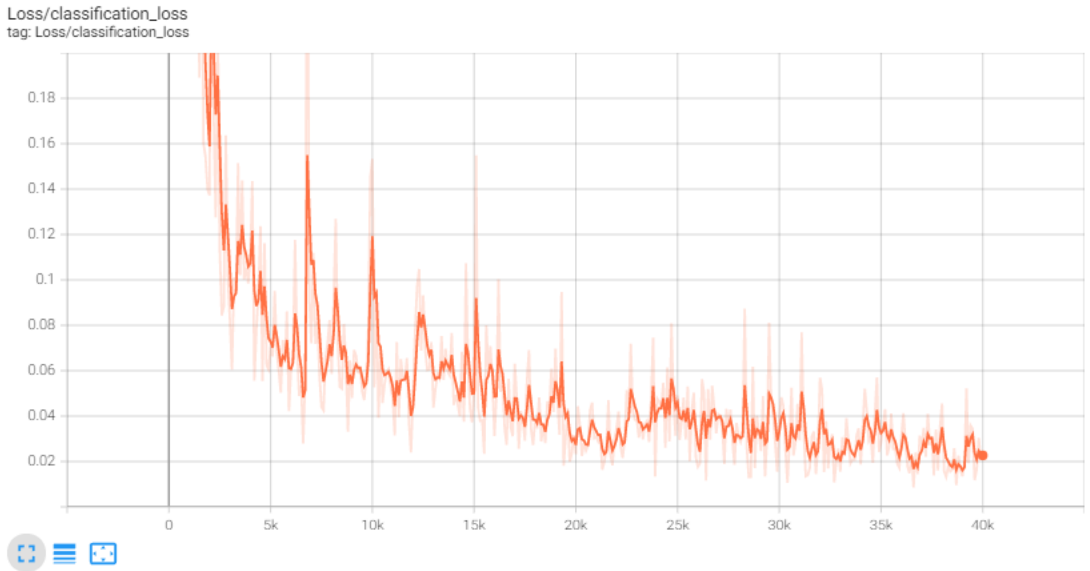
- **SSD on Tick Dataset 1 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.639** and **0.520** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.520
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.639
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.615
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.520
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.804
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.804
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.804
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.804
```

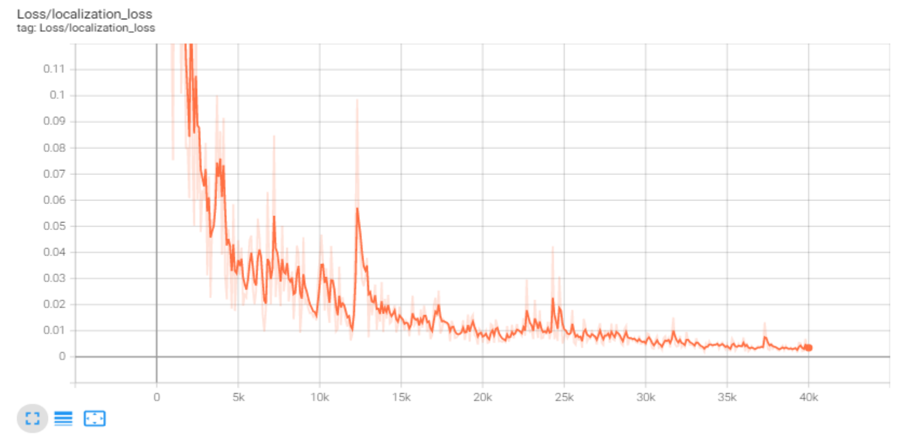
<u>Training Classification Loss Curve</u>

As the number of steps increases, the classification loss decreases.



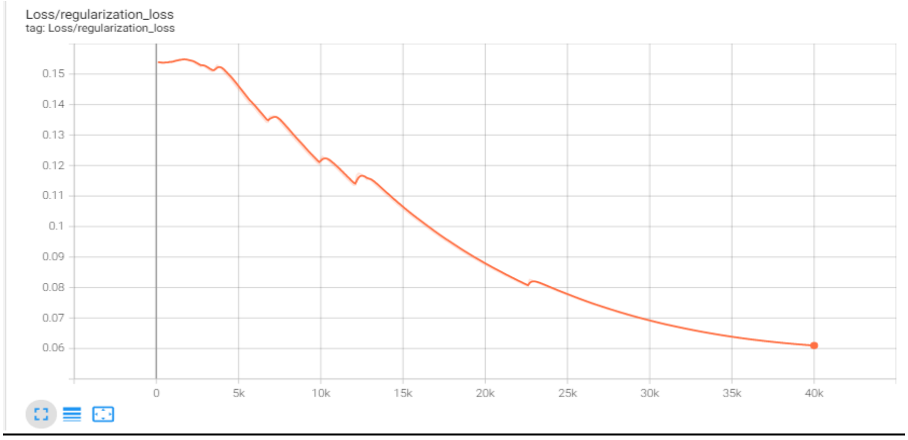
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



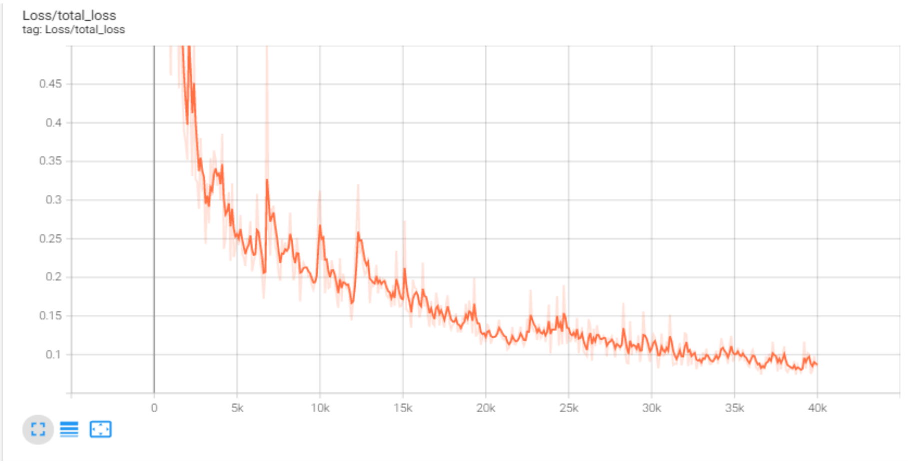
Training Regularization Loss Curve

As the number of steps increases, the regularization loss decreases.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



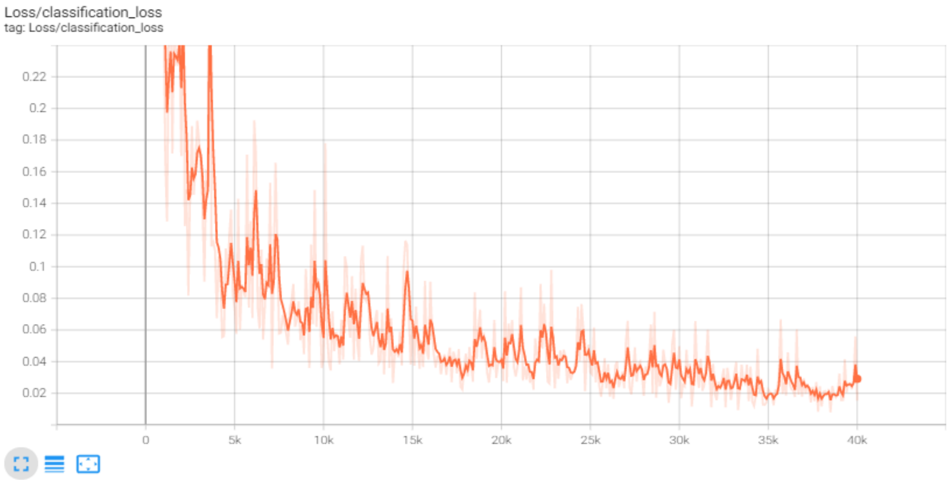
- **SSD on Tick Dataset 1 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.886** and **0.706** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.706
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.886
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.855
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.706
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.800
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.800
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.800
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.800
```

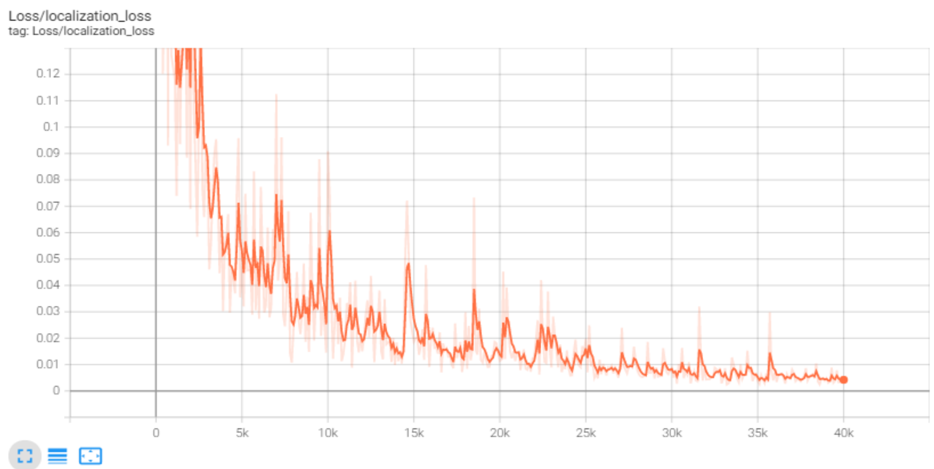
<u>Training Classification Loss Curve</u>

As the number of steps increases, the classification loss decreases.



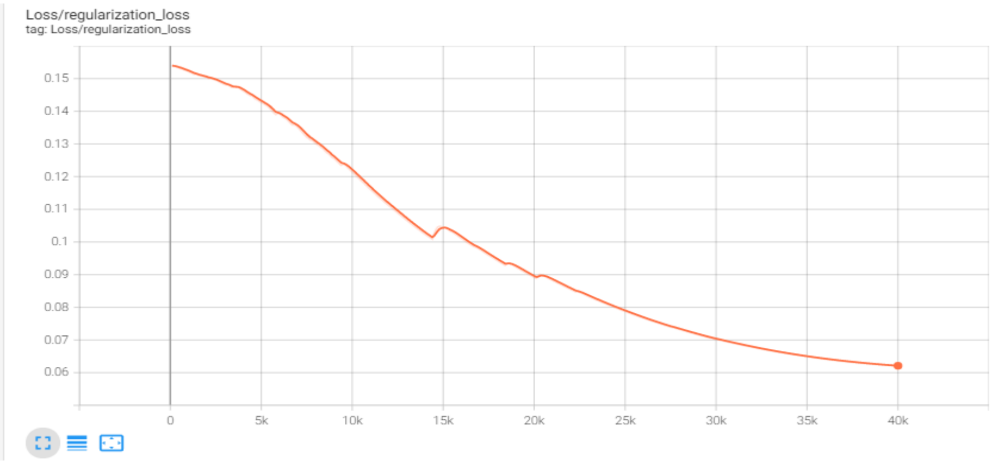
<u>Training Localization Loss Curve</u>

As the number of steps increases, the localization loss decreases.



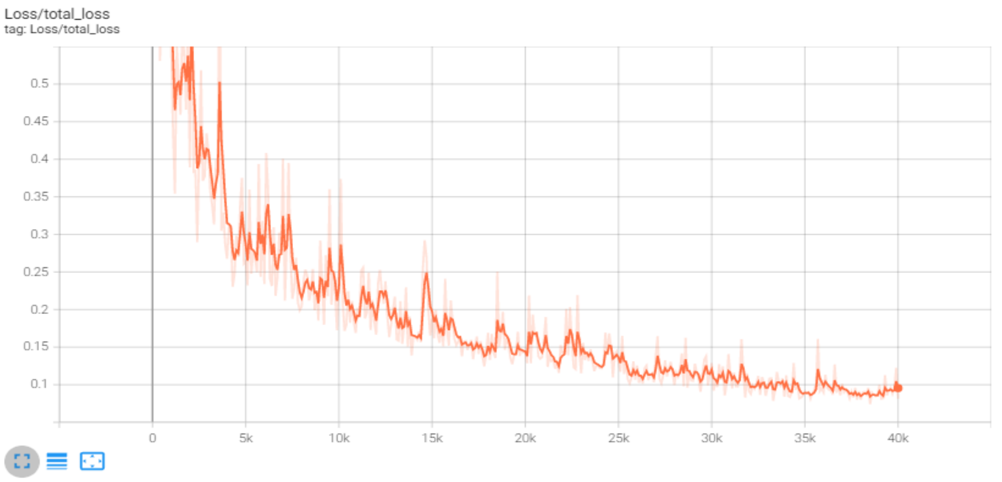
Training Regularization Loss Curve

As the number of steps increases, the regularization loss decreases.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



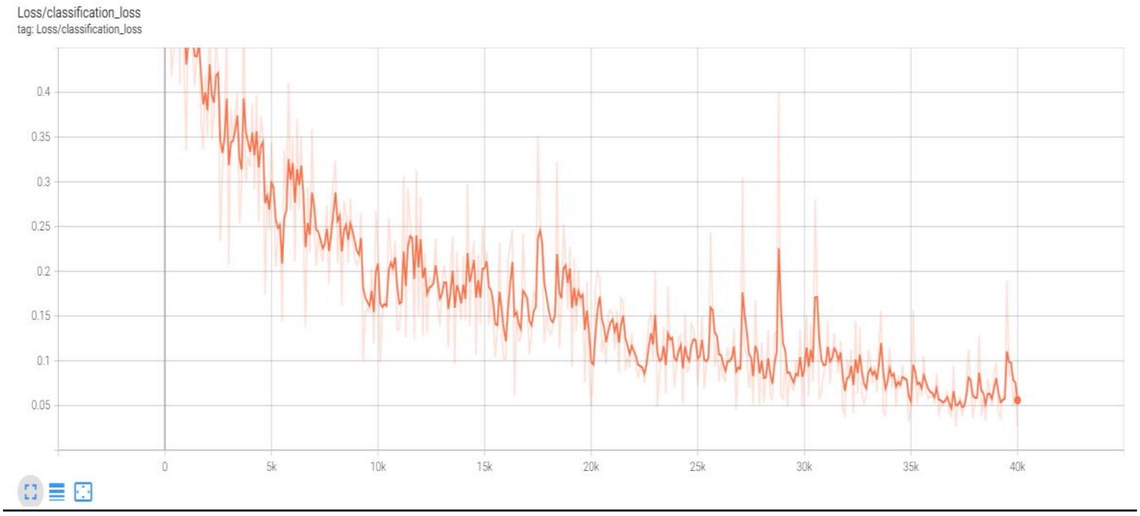
- **SSD on Tick Dataset 2 (without augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.701** and **0.479** respectively.

```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.479
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.701
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.500
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.479
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.598
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.648
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.679
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.679
```

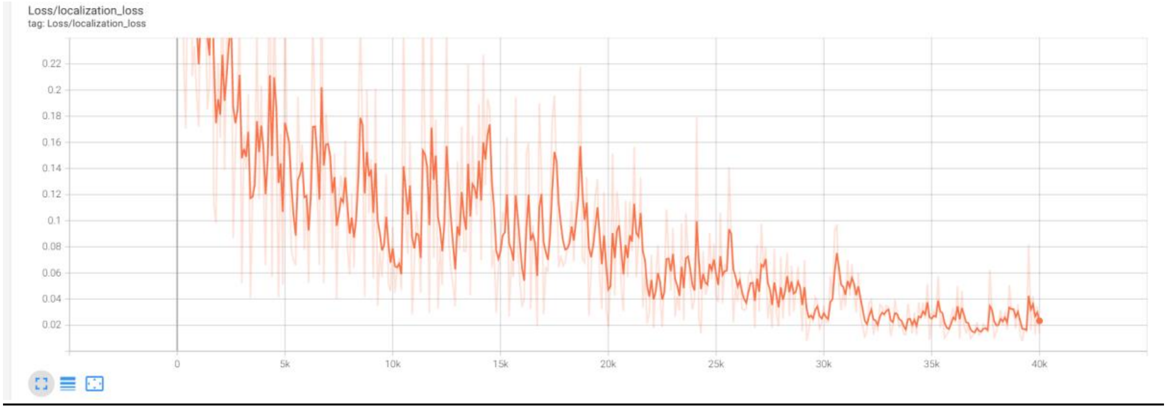
Training Classification Loss Curve

As the number of steps increases, the classification loss decreases.



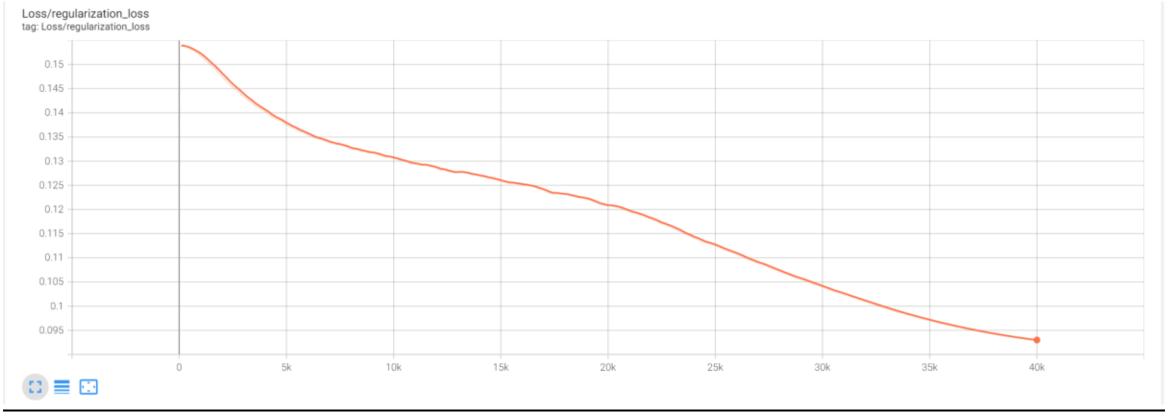
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



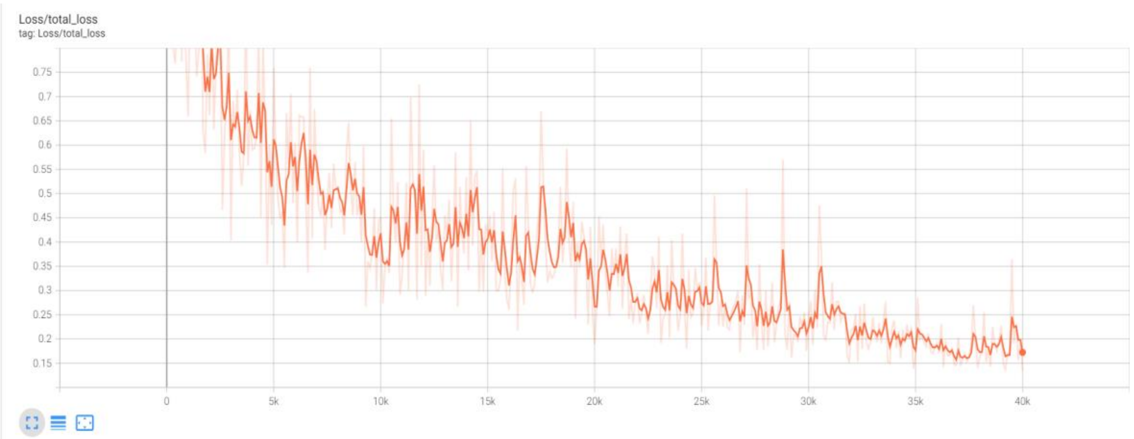
Training Regularization Loss Curve

As the number of steps increases, the regularization loss decreases.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



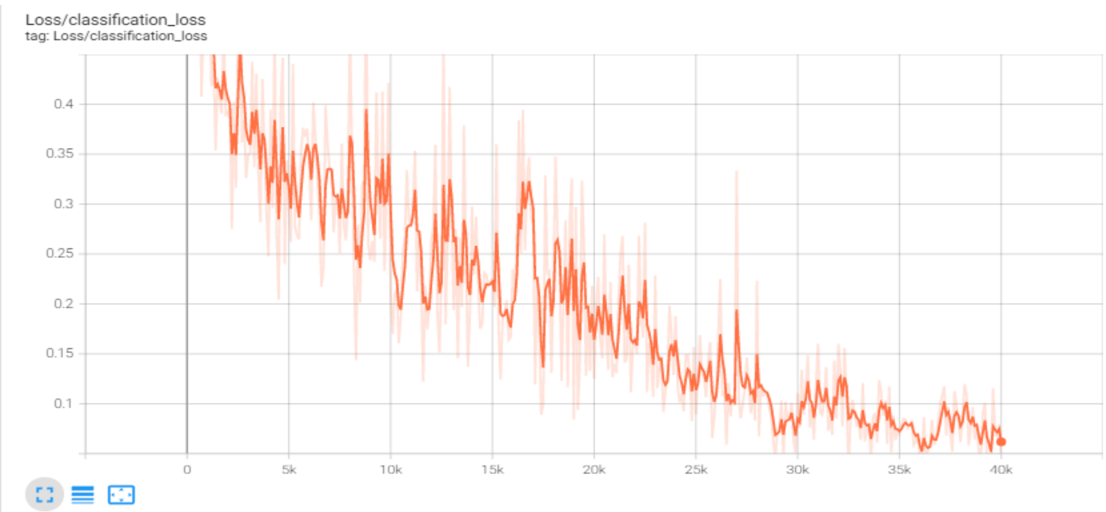
- **SSD on Tick Dataset 2 (with augmentation):**

The mAP@.5 and mAP@.5:.95 for all classes is **0.763** and **0.492** respectively.

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.492
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.763
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.508
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.492
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.580
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.652
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.682
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = 0.682
INFO:tensorflow:Eval metrics at step 40000
```

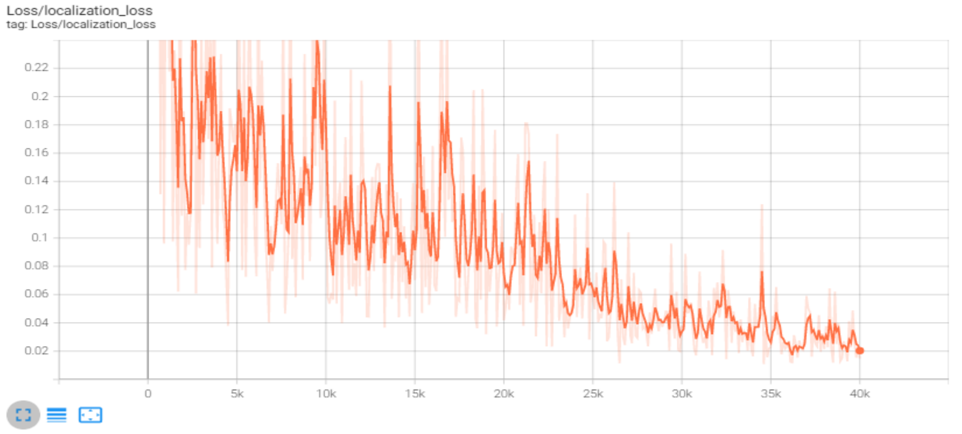
Training Classification Loss Curve

As the number of steps increases, the classification loss decreases.



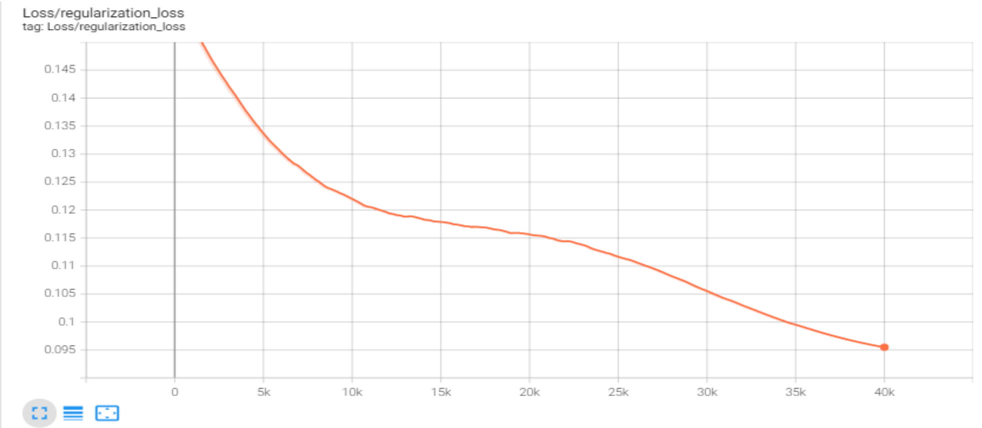
Training Localization Loss Curve

As the number of steps increases, the localization loss decreases.



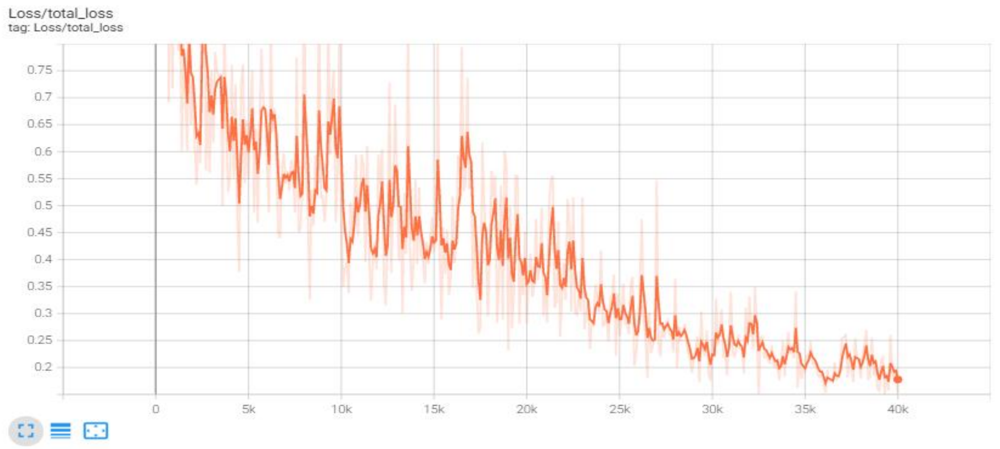
Training Regularization Loss Curve

As the number of steps increases, the regularization loss decreases.



Training Total Loss Curve

As the number of steps increases, the total loss decreases.



CONCLUSION

The results of the research are as follows:

Model	Tick Dataset 1				Tick Dataset 2			
	Without Augmentation		With Augmentation		Without Augmentation		With Augmentation	
	AP@ 0.5	AP@0.5: 0.95	AP@ 0.5	AP@0.5: 0.95	AP@ 0.5	AP@0.5: 0.95	AP@ 0.5	AP@0.5: 0.95
YOLOv 5	0.826	0.659	0.894	0.721	0.909	0.757	0.936	0.761
Faster RCNN	0.819	0.662	0.809	0.598	0.916	0.690	0.750	0.572
DETR	0.576	0.445	0.895	0.744	0.486	0.361	0.629	0.483
Efficient Det	0.793	0.543	0.793	0.618	0.437	0.295	0.703	0.429
SSD	0.639	0.520	0.886	0.706	0.701	0.479	0.763	0.492

The **YOLOv5** Model performs well with both the datasets with and without augmentation.

The **Faster RCNN** Model performs well with both the datasets with and without augmentation.

The **DETR Mode**l performs well with Dataset 1 but performs poorly with Dataset 2.

The **EfficientDet Model** and the **SSD Model** perform the worst among all the models.

Hence, based on the AP scores above the **most accurate** model is the **YOLOv5 Model** and the **least accurate** model is the **EfficientDet Model.**

REFERENCES

1. YOLOv5
2. DETR
3. Faster RCNN
4. EfficientDet
5. SSD