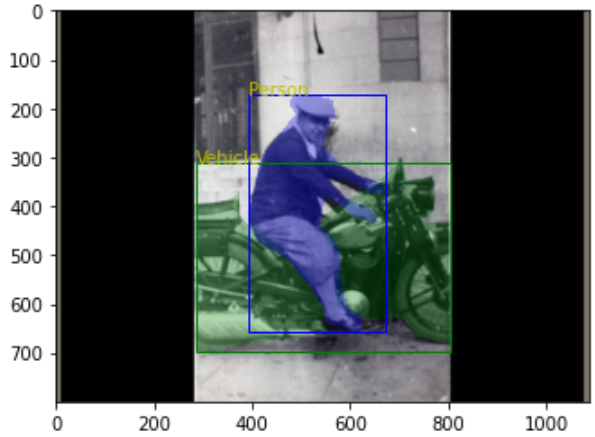
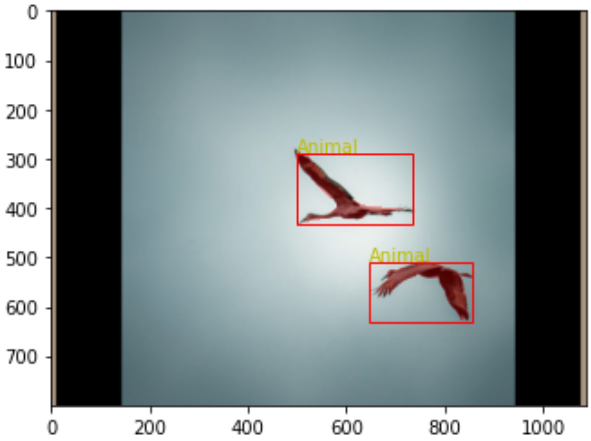
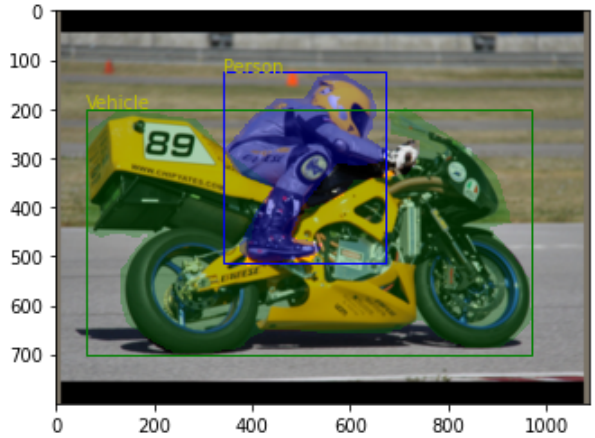
RPN Report

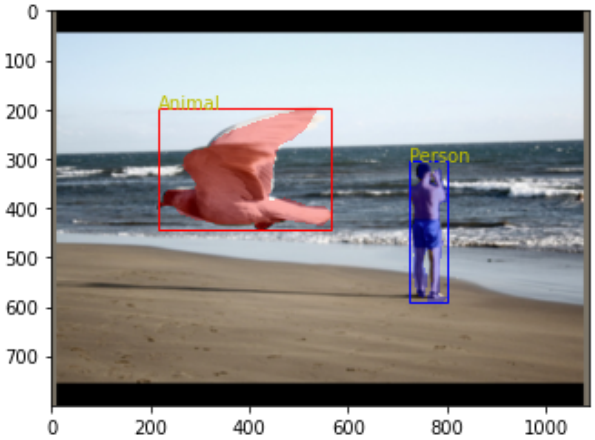
By- Ruchi Gupte, Ishani Mhatre

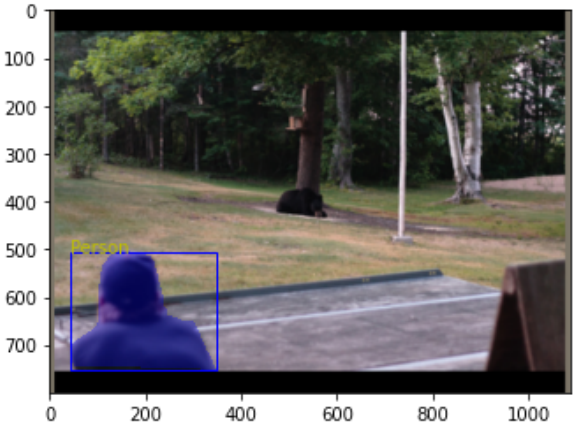
Ground Truth Visualization











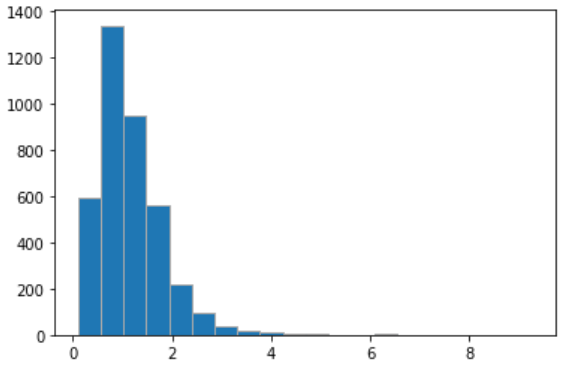
Histograms of scale and aspect ratios:
(Only one scale and aspect ratio used for a basic implementation of RPN)

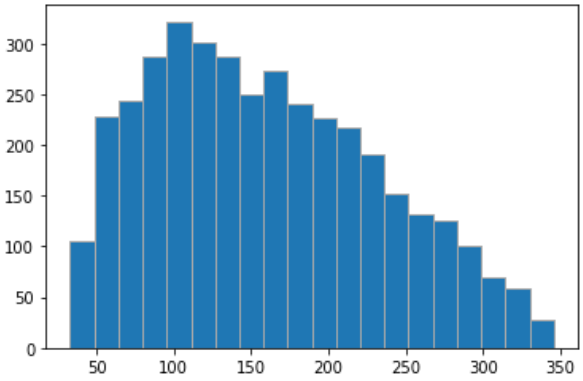
```
Mode of Aspect Ratio:  0.55585647
Mode of Scale:  95.3640061987706
```

Default values of scale and aspect ratio were chosen as it gave us good results. Otherwise we were planning to use the values which occurred with maximum frequency in the histogram.

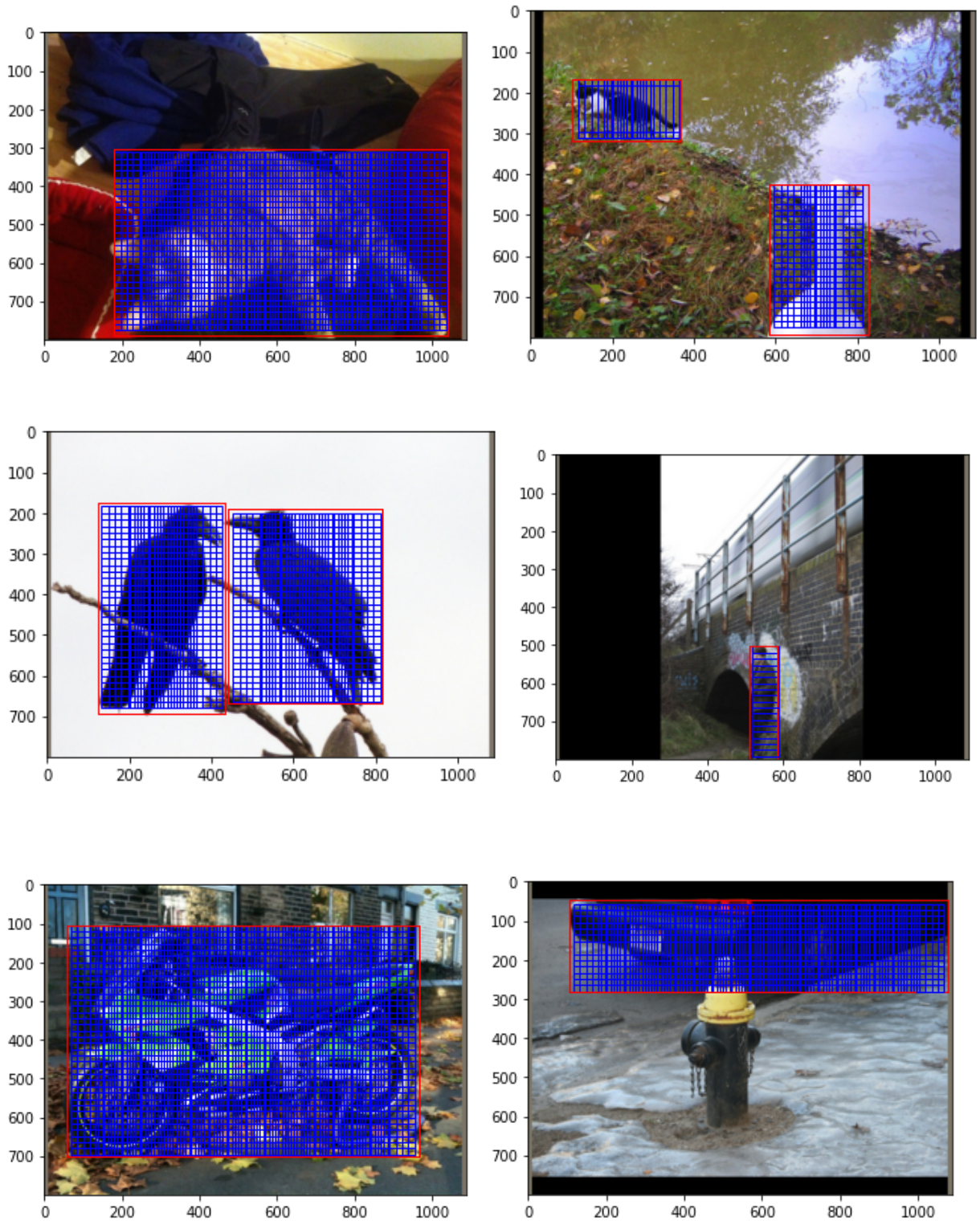
When the mode of the plots were used as the aspect ratio and scale the resulting training took longer and the results were almost the same, hence we went with the default values. Practically using the modes would also give us a good result.

When choosing a scale and aspect ratio it is important to check the distribution of ground truths, since if the scale and ratio are too big, most images would be overestimated in size and the bounding box might be too big for the image.

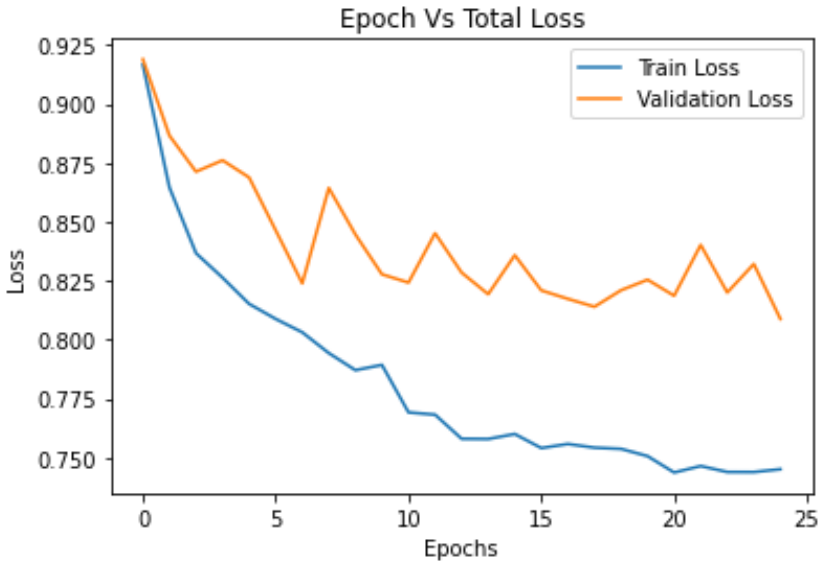


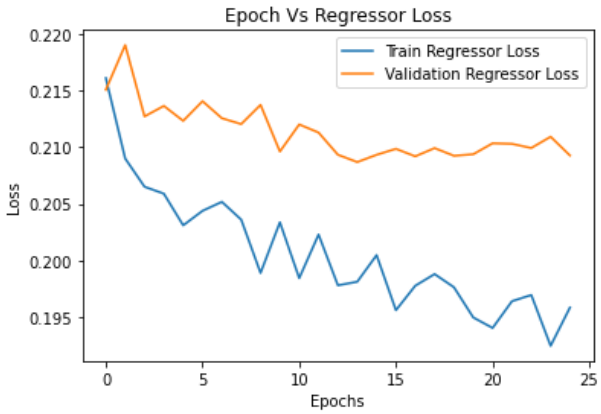


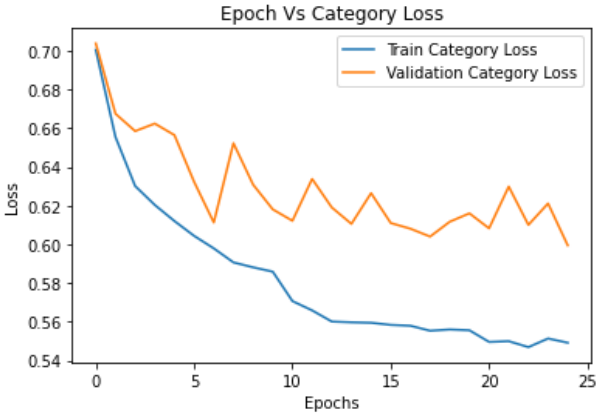
Images with positive anchor boxes and corresponding ground truth box



Training, validation loss curves



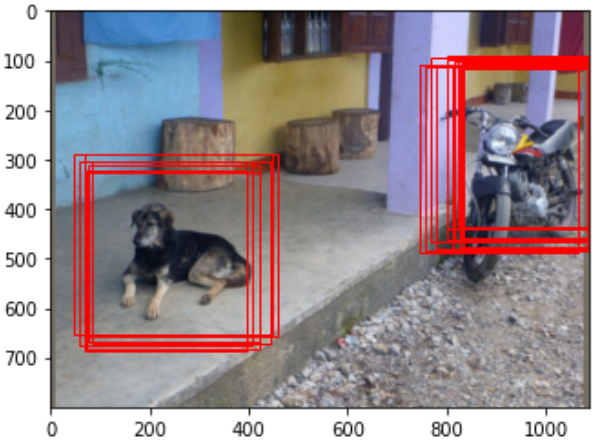


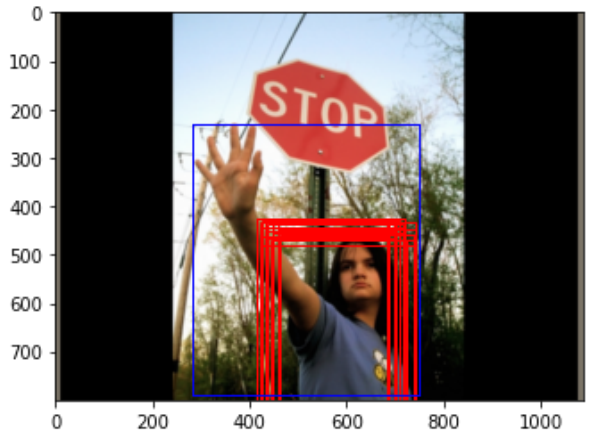


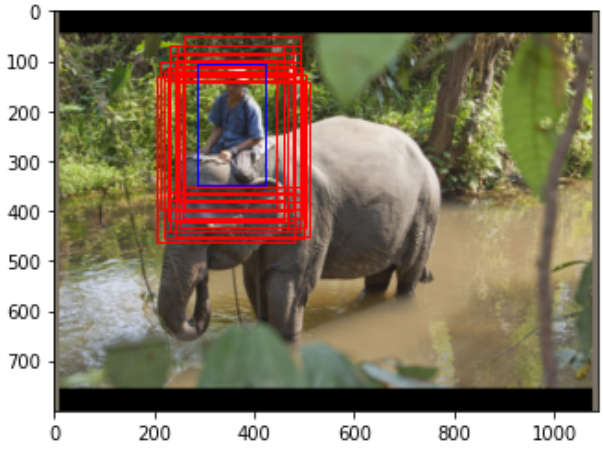

```
Point-wise accuracy of the proposal classifier 0.770325779914856
```

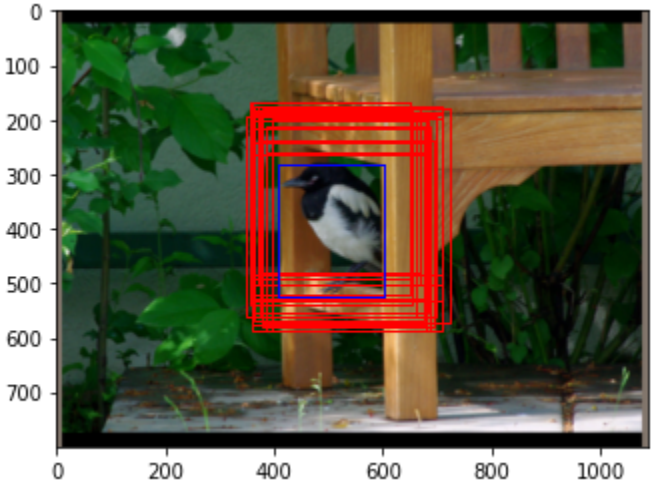
Results


0
100
200
300
400
500
600
700
0
200
400
600
800
1000

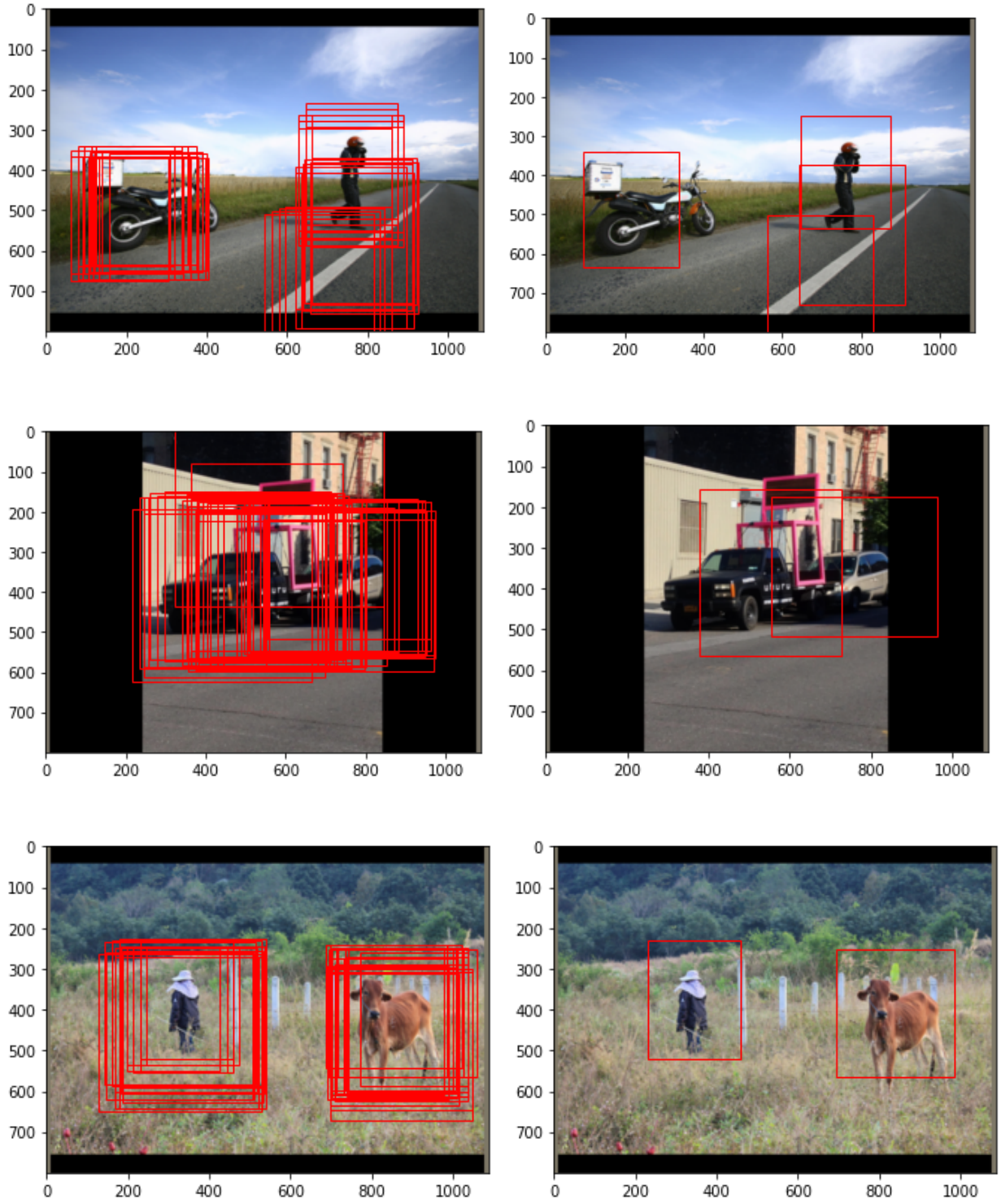

0
100
200
300
400
500
600
700
0
200
400
600
800
1000

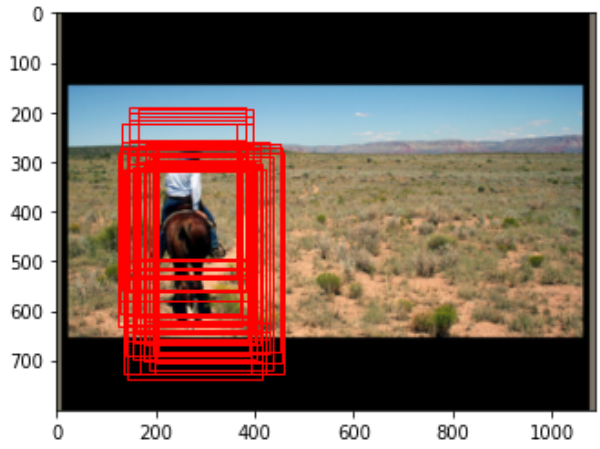

STOP
0
100
200
300
400
500
600
700
0
200
400
600
800
1000

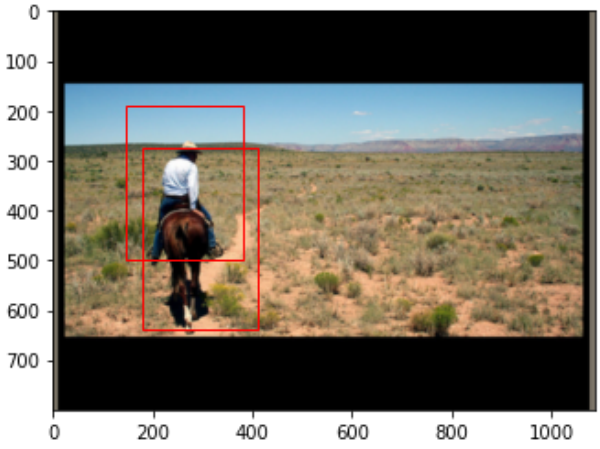

0
100
200
300
400
500
600
700
0
200
400
600
800
1000

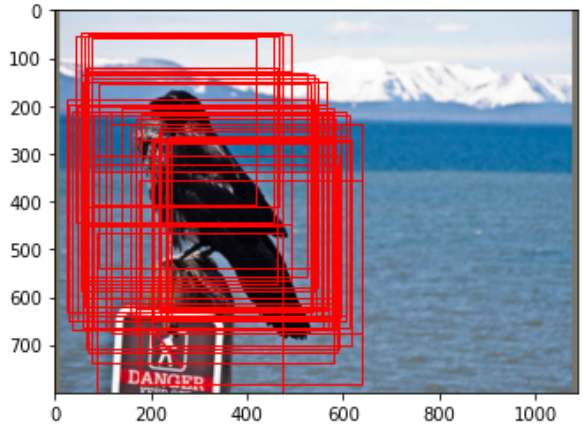

0
100
200
300
400
500
600
700
0
200
400
600
800
1000

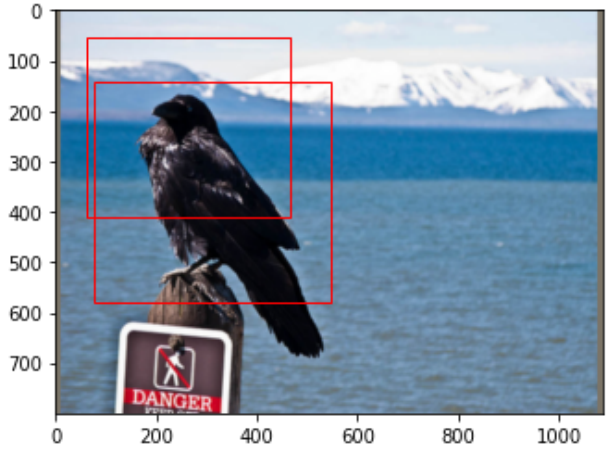
Visualization of the proposed boxes before and after the NMS

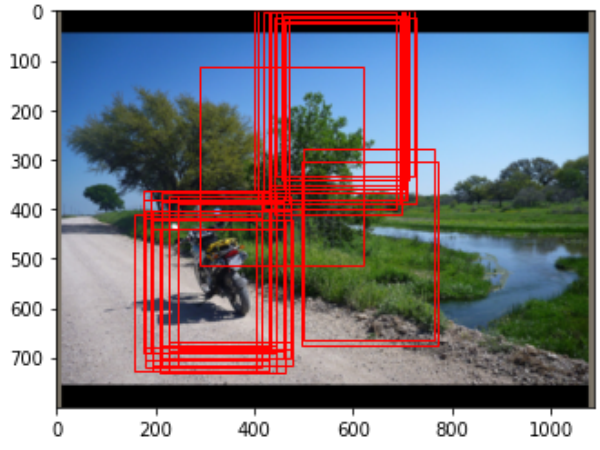


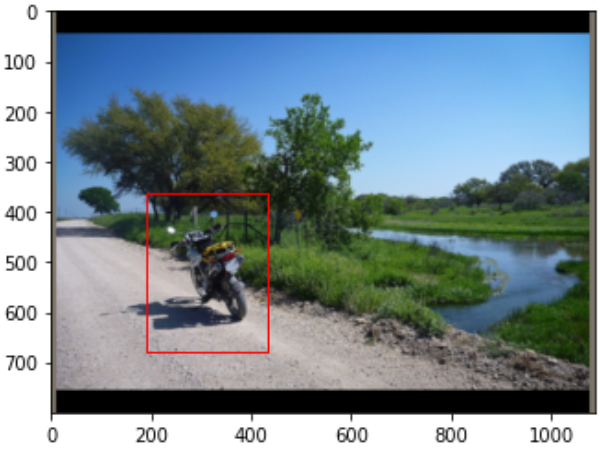

0
100
200
300
400
500
600
700
0
200
400
600
800
1000


0
100
200
300
400
500
600
700
0
200
400
600
800
1000


0
100
200
300
400
500
600
700
0
200
400
600
800
1000


0
100
200
300
400
500
600
700
0
200
400
600
800
1000
DANGER


0
100
200
300
400
500
600
700
0
200
400
600
800
1000


0
100
200
300
400
500
600
700
0
200
400
600
800
1000

Q7. Brief explanation about specific choices in the implementation that were not mentioned in the rest of the report.

We trained our model with multiple hyperparameter choices and chose the best working model. For optimizers we used Adam with learning rate = 0.01 and SGD with learning rate = 0.01, weight decay = 0.0001 and momentum of 0.9. We saw that SGD optimizer with lr_scheduler led to gradual decrease in both the training and validation losses and worked best in our case.

When choosing a scale and aspect ratio it is important to check the distribution of ground truths, since if the scale and ratio are too big, most images would be overestimated in size and the bounding box might be too big for the image.

For the classifier head, we used Binary Cross Entropy loss and SmoothL1Loss for the regression head which was given to us. Instead of training on all the anchor boxes which had a lot of negative boxes without any matching ground truth box, we selected only 50 such proposals with 25 positive anchors and 25 negative.