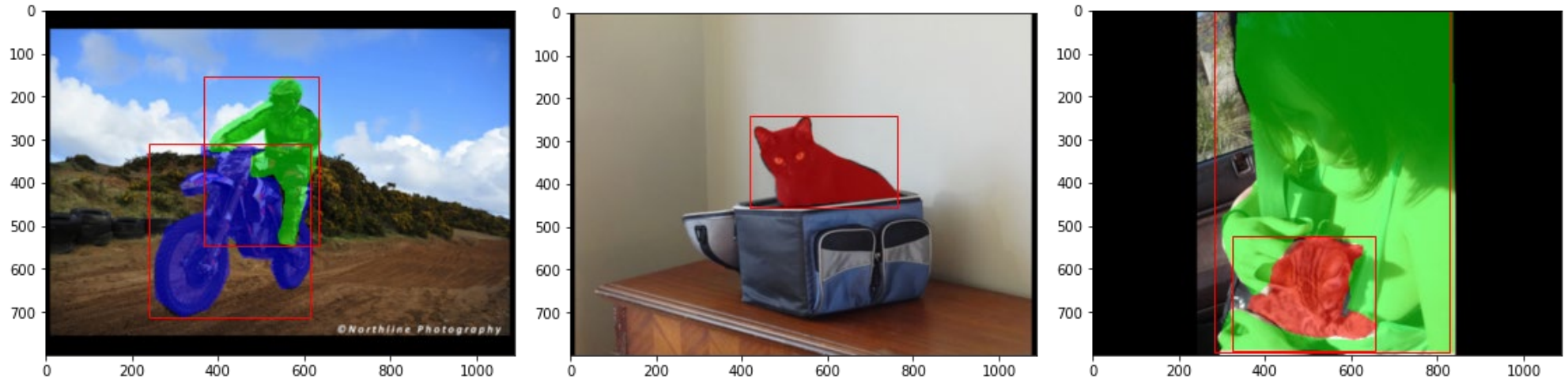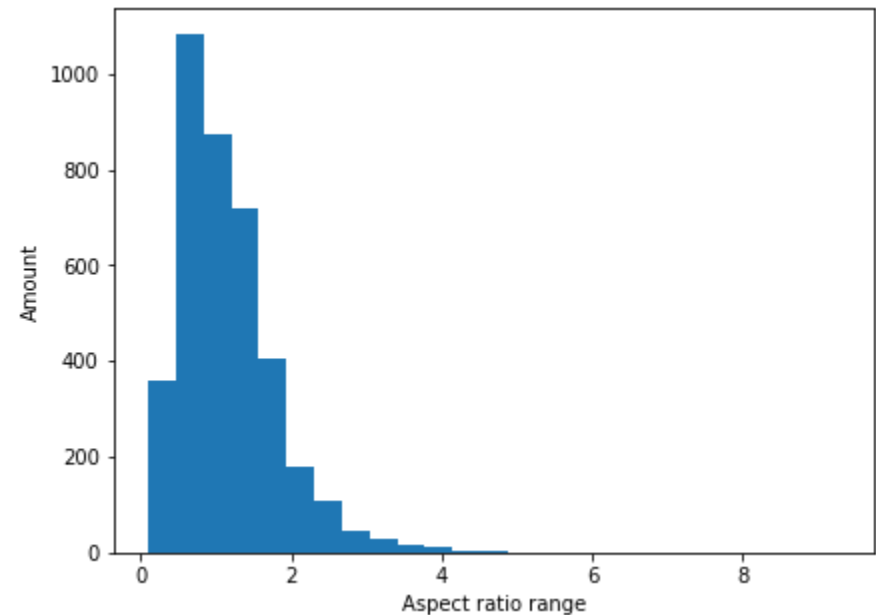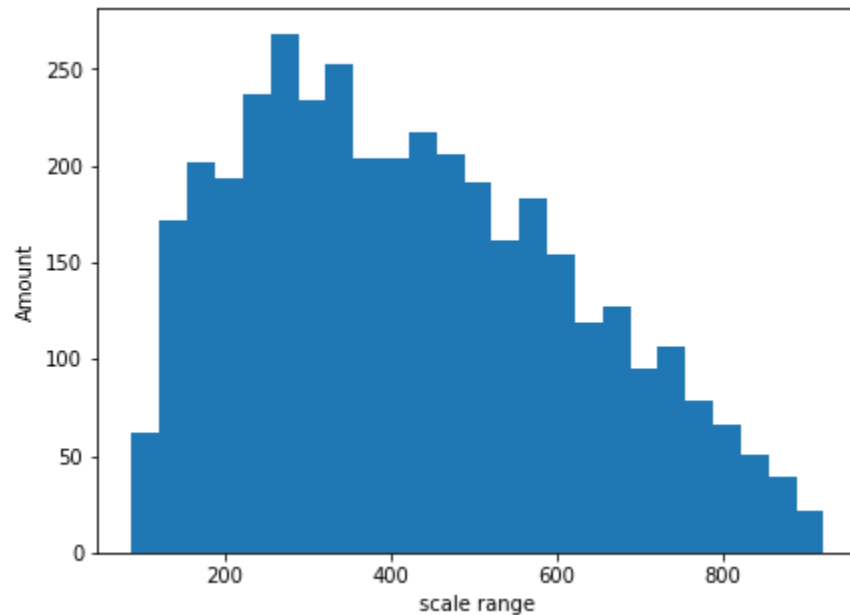# 1. Images with the visual correctness check (Same as SOLO)


# 2. Histograms of scales and aspect ratios

- **Choice of aspect ratio and scale:**

It can be observed from the histograms that among all the bounding boxes from the dataset, the dominant aspect ratio is around 0.8 and the dominant scale is around 256. In order to choose a proper value for both scale and aspect ratio that accommodates most of the cases, we choose the aspect ratio to be 0.8 and scale to be 256.

3. Images with positive anchors and corresponding ground truth boxes
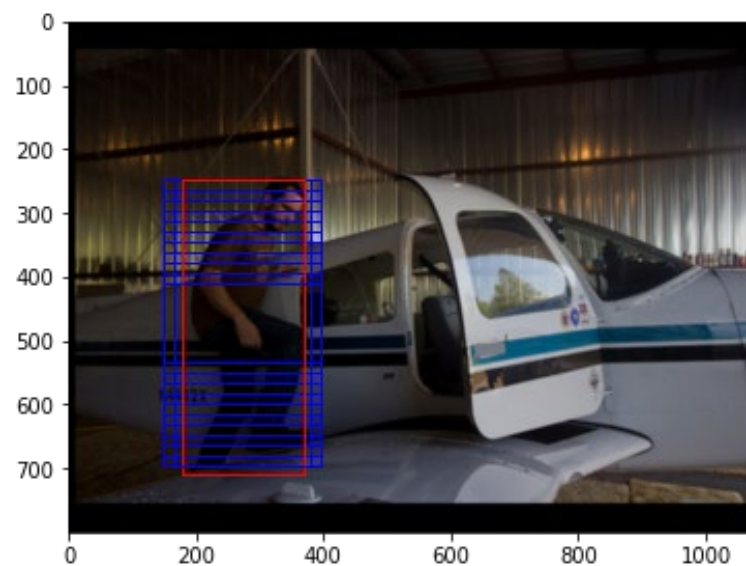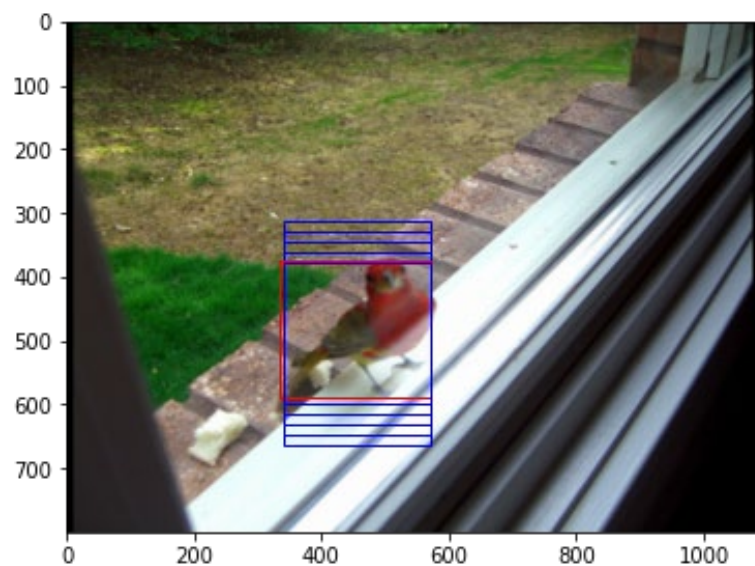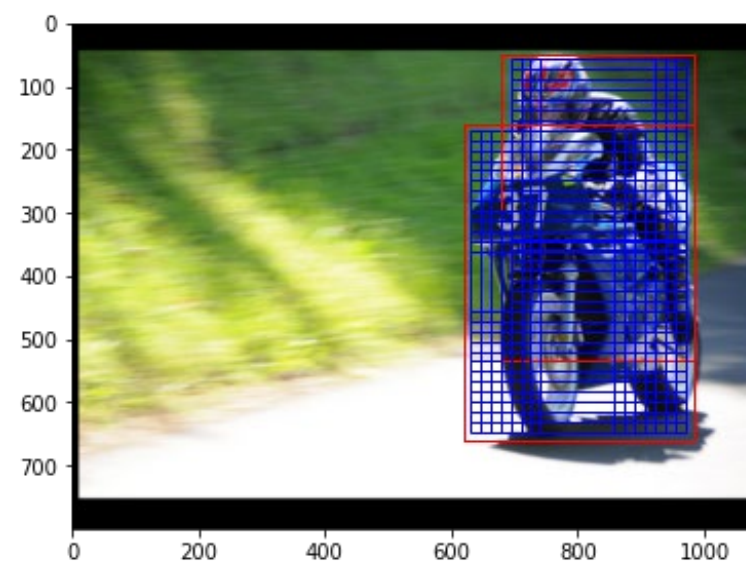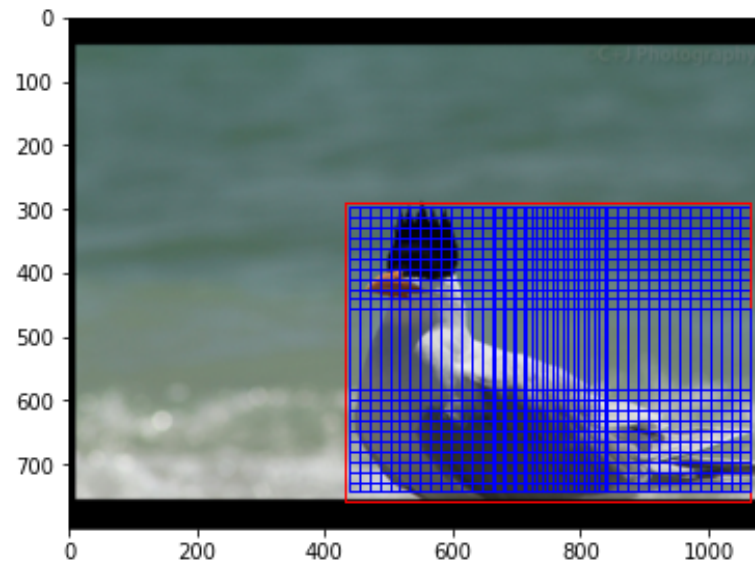
4. Training and Validation curves
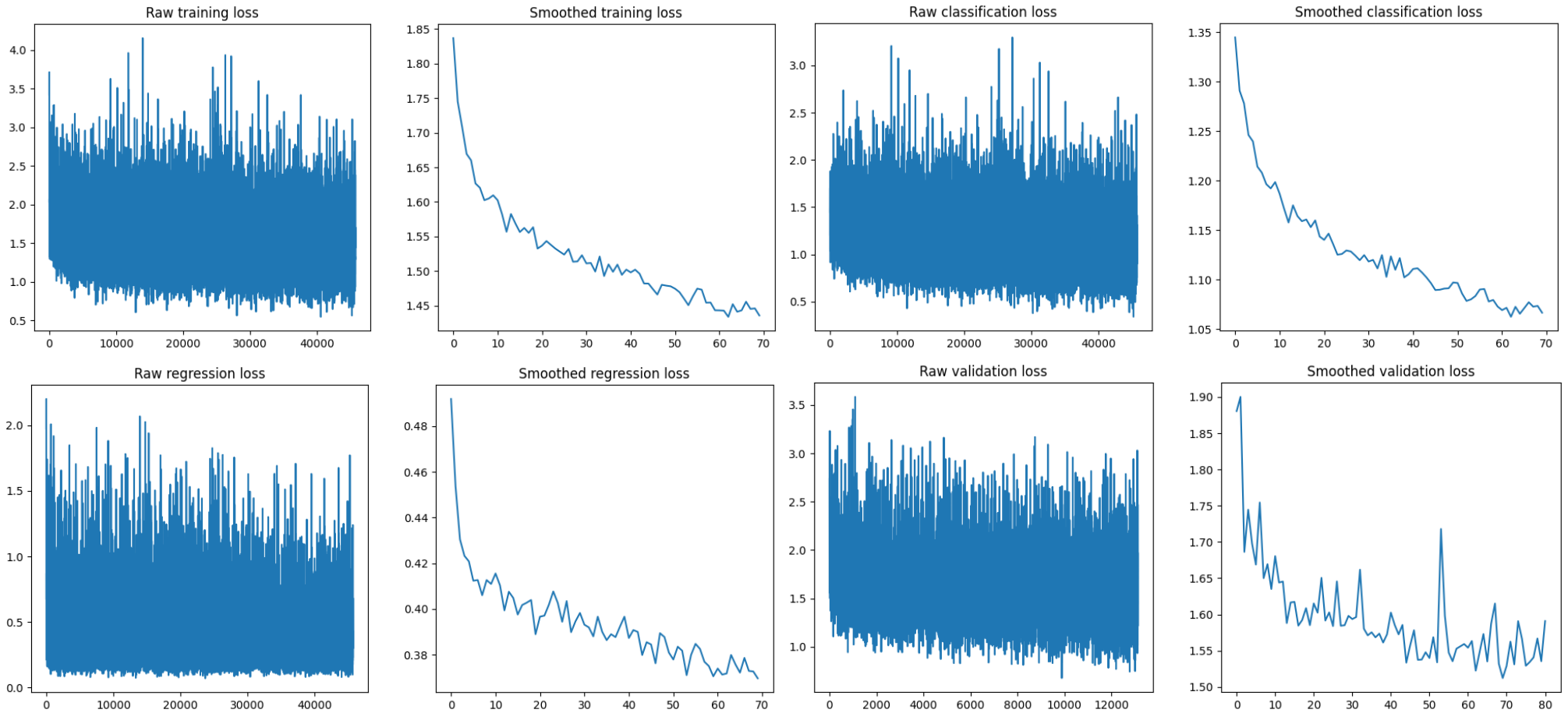
- **Training the RPN**

We follow the pipeline provided in the homework document and train the network for 70 epochs. The total training loss is steadily decreasing, and so are classification and regression loss. The validation loss also decreases with slight fluctuations. The performance of the network would be analyzed with the result in the following sections.

For other training details, see README.md and the code.

# 5. Point-wise Accuracy and Plots of the top 20 proposals for test images
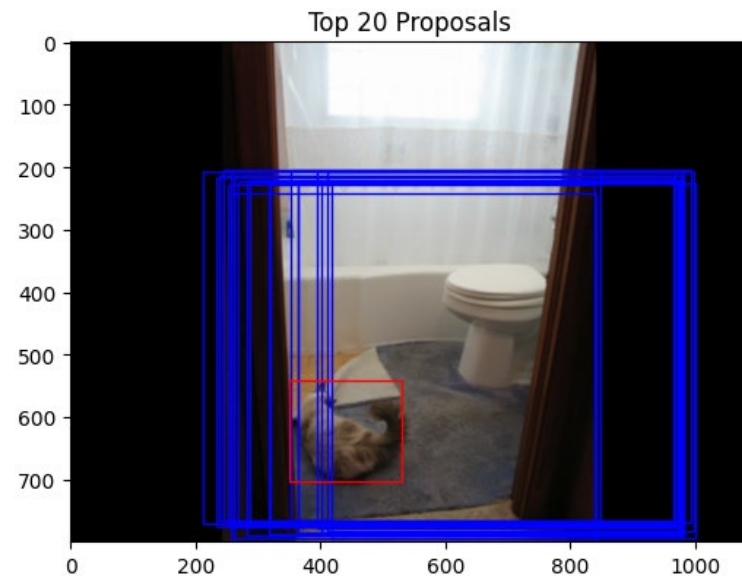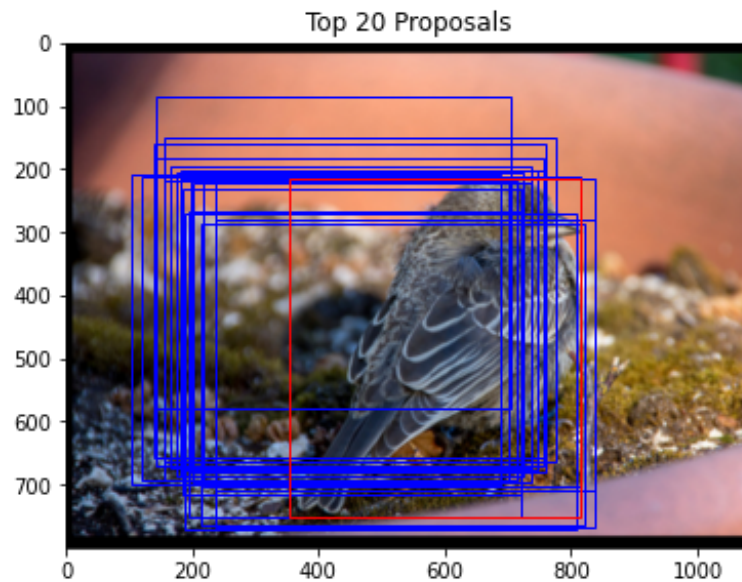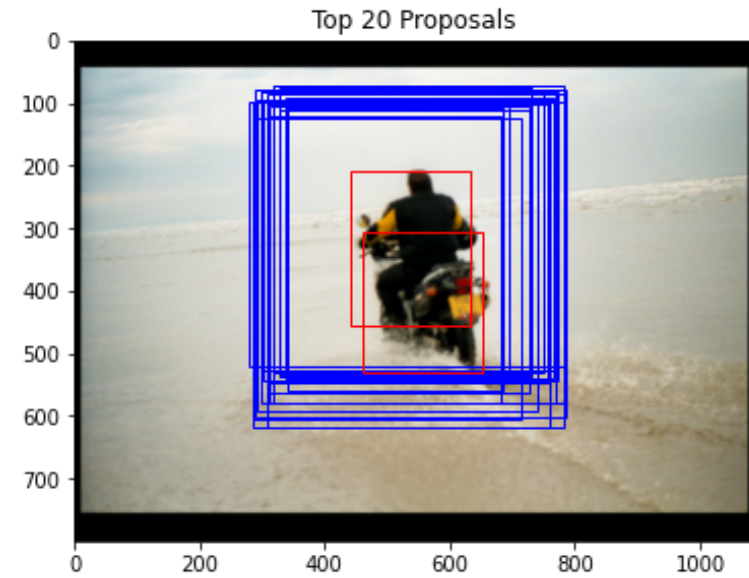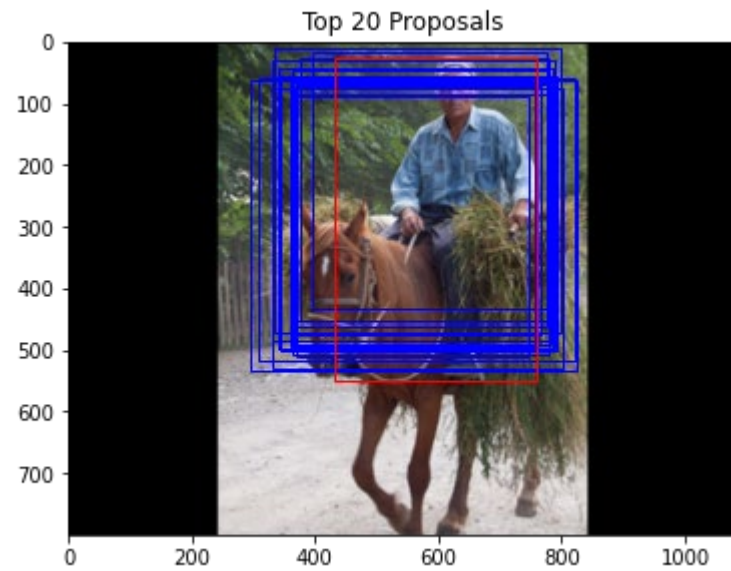
- **Point-Wise Accuracy**

According to our calculation, the point-wise accuracy for classification is 70.15%.

The details of the calculation are included in the inference.ipynb. (Note: Since each time we do the train-test split randomly, so the point-wise accuracy on the test set might vary, but the result mostly is around 70%)

6. Visualization of the proposed boxes before and after the NMS
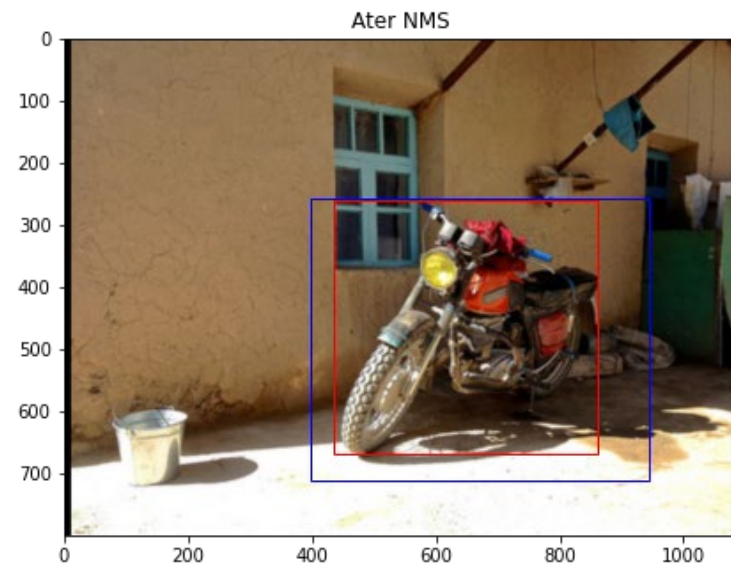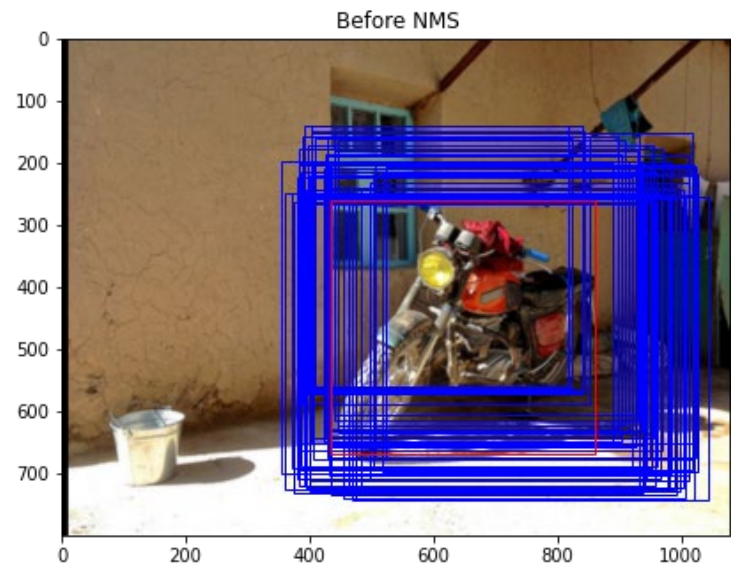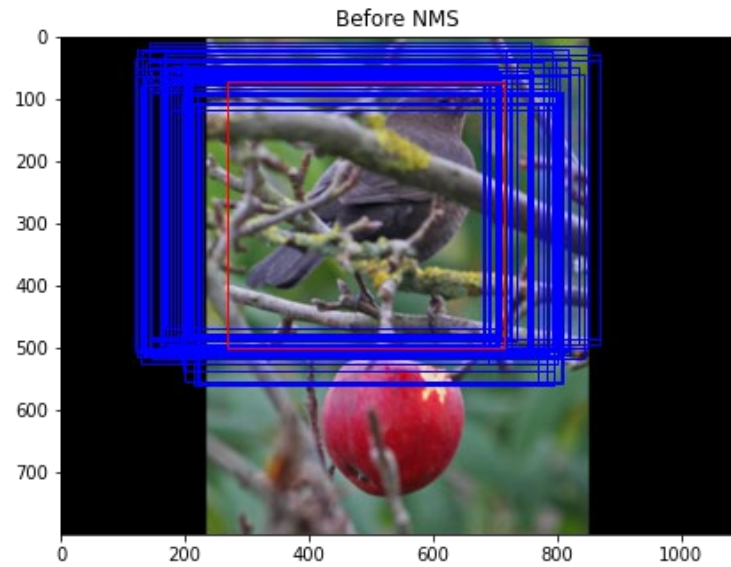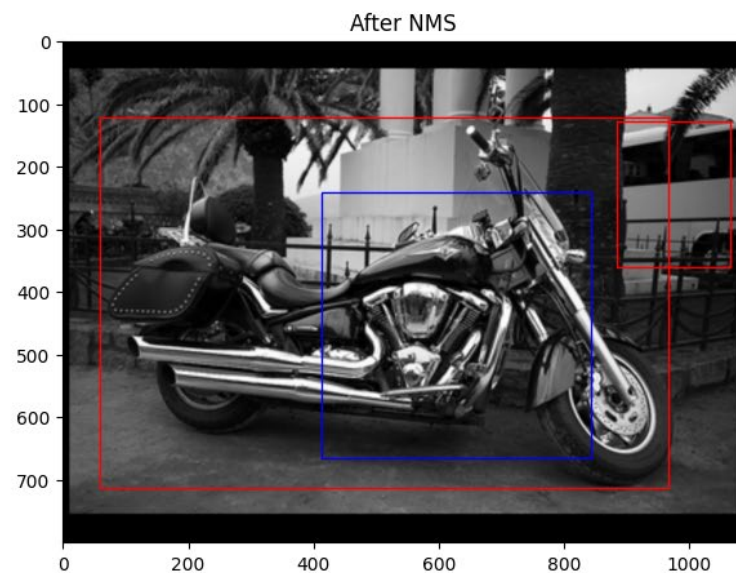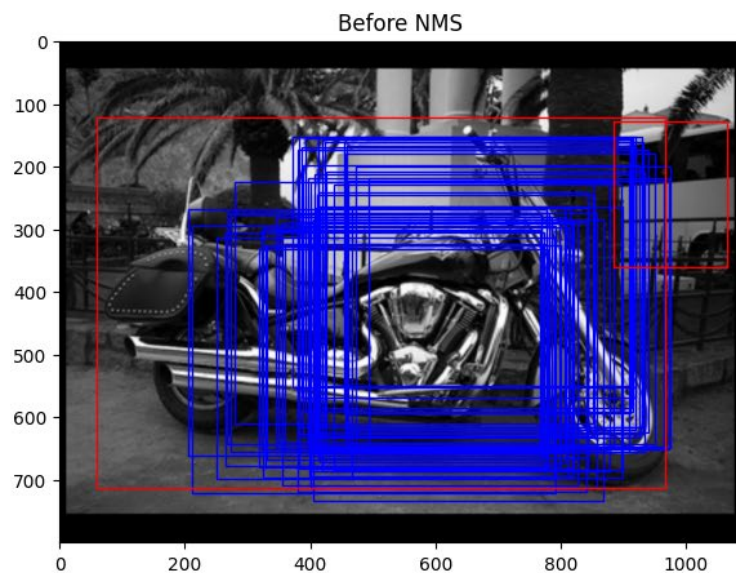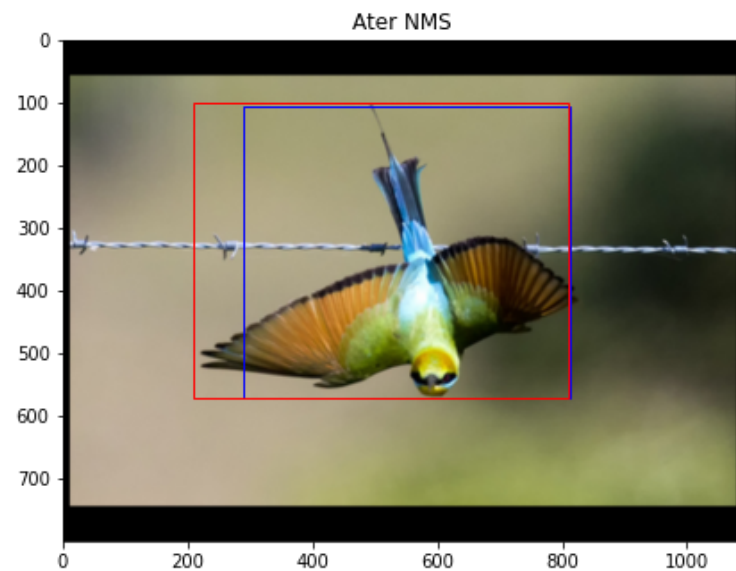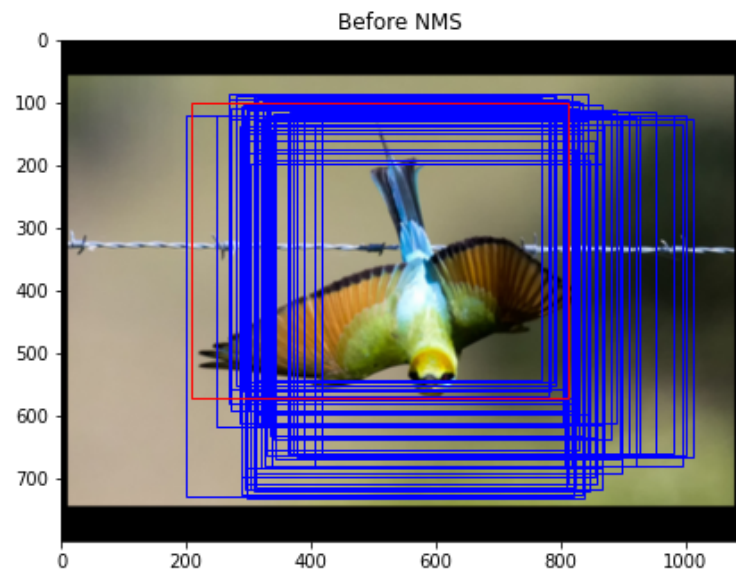
- **RPN Performance and Comments**

We can see from the result displayed that, all in all, the RPN implemented in this project has relatively good performances in both classification and regression (locating the bounding boxes). However, some problems are still worth mentioning. In this implementation, we use **only one** anchor box per grid cell and all of the anchor boxes will have **the same predefined scale and aspect ratio**. We set the predefined scale and ratio value (256 and 0.8) according to our inspection and belief for the dominant. Unluckily, as the last examples in section 5 and 6 show, such a setting would make the region proposal ambiguous for the ground truths that are slender(minor aspect ratio) or smaller(minor scale). The proposed boxes would either be too small to cover the whole ground truth, or too large to include unnecessary backgrounds, which is not ideal.