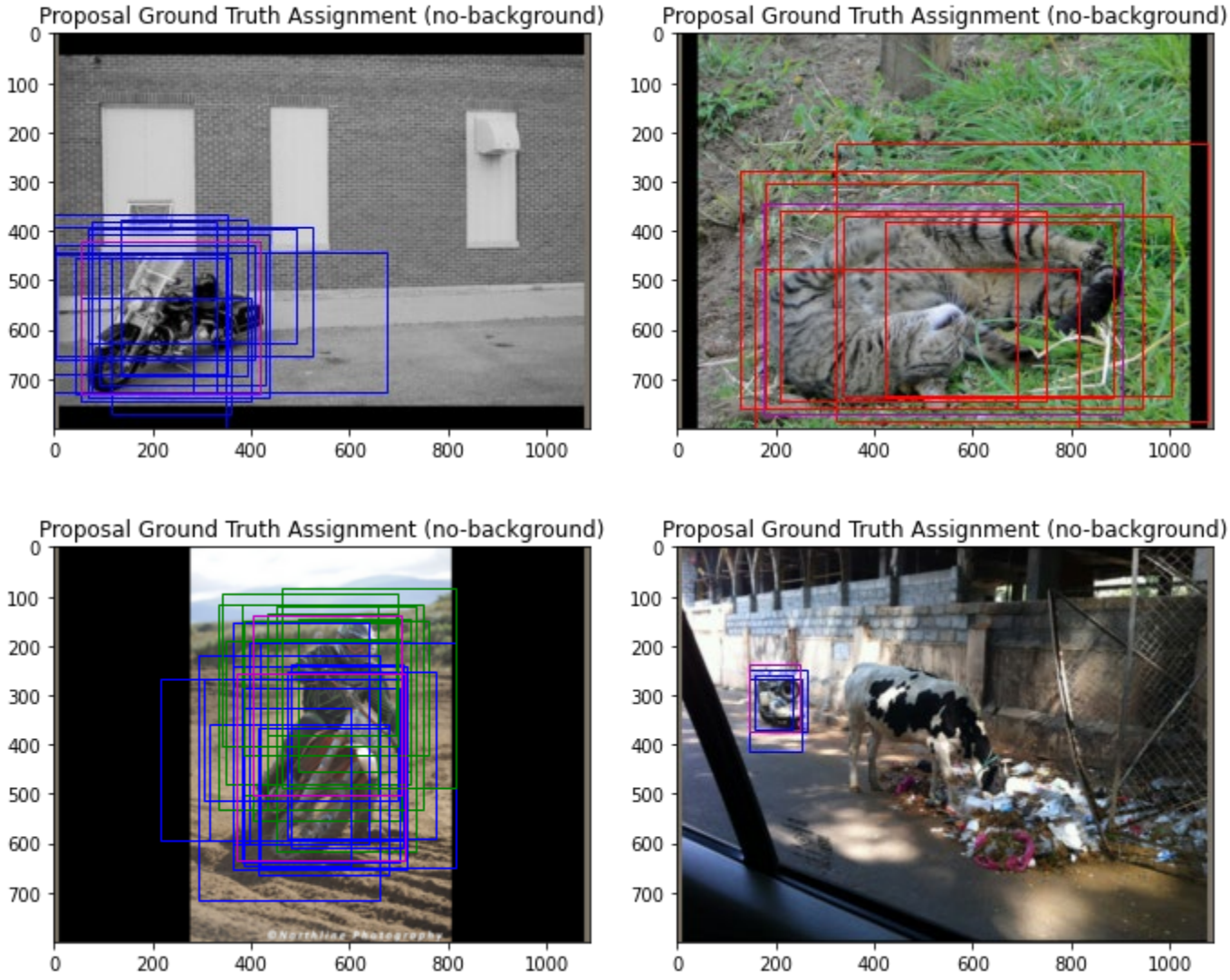
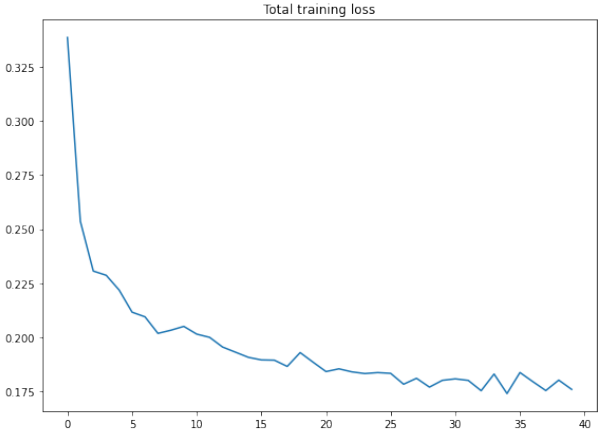
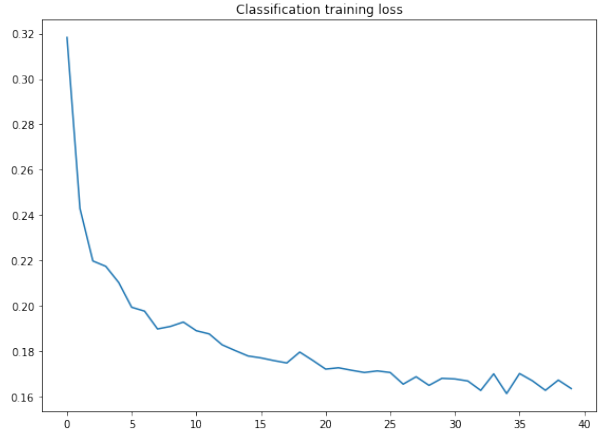
1. Creation of the Ground Truth from the Proposal

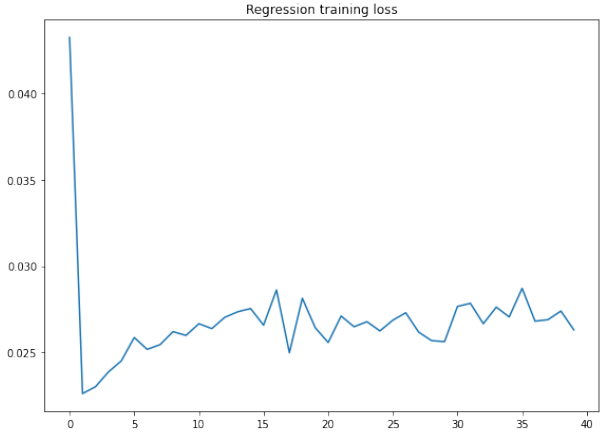


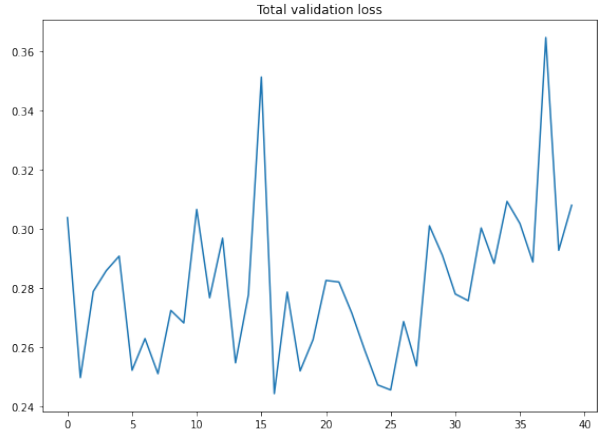
NOTE: Here the **magenta** box represents the ground truth bounding box. The **red, blue** and **green** box corresponds to the RPN proposals in non-background classes in *Animal, Vehicle and Human* category respectively.

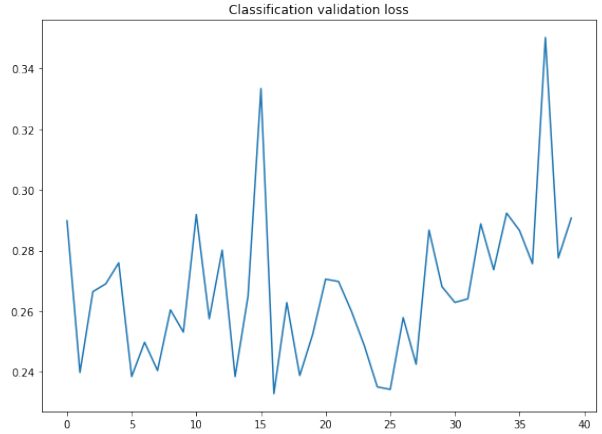
2. Training and Validation curves

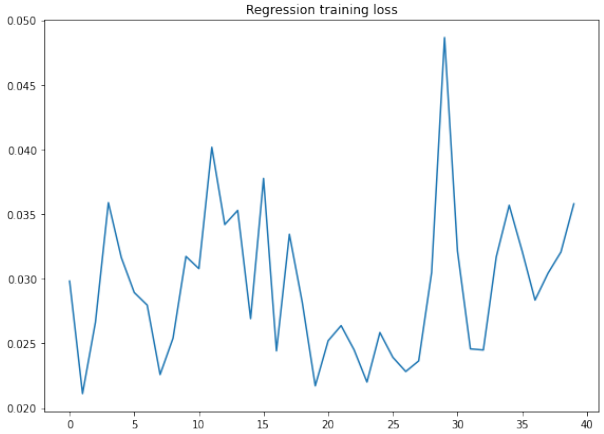










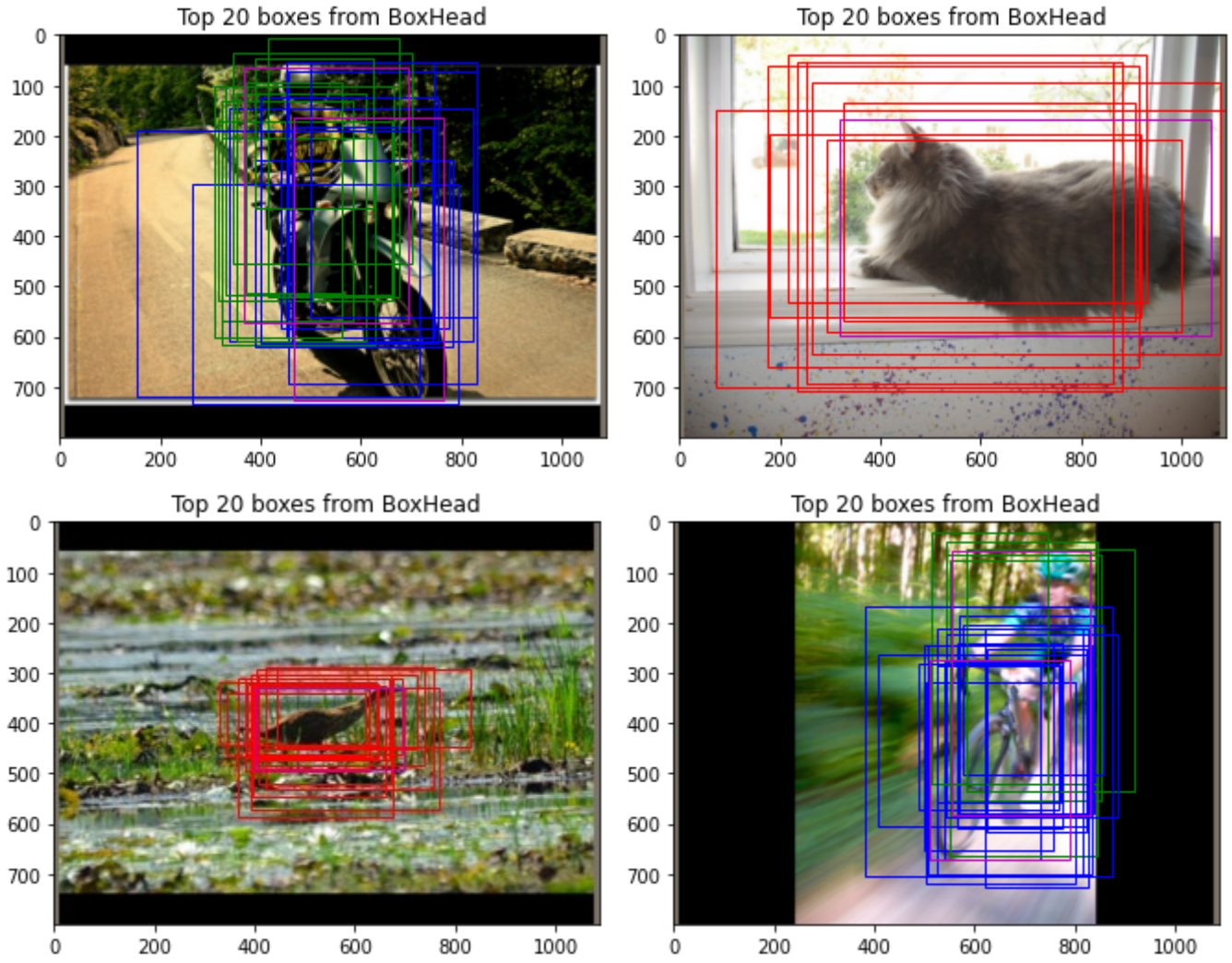


- **Training the Model**

We follow the pipeline provided in the homework document and train the network for 40 epochs. The total training loss is steadily decreasing, and so are classification and regression loss. The validation losses just fluctuate up and down in a small range and we can hardly make them decrease. According to the result displayed below, we believe the training is successful though validation losses almost do not decrease.

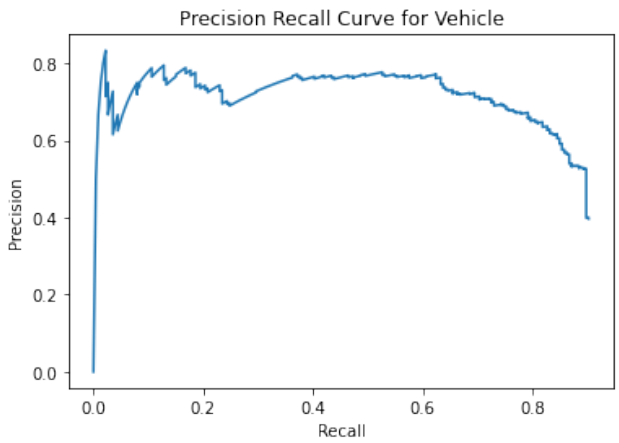
For other training details, see README.md and the code.

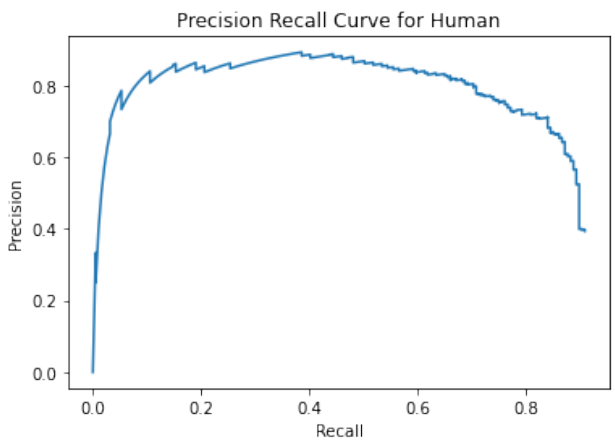
3. Top 20 boxes produced by the Box Head

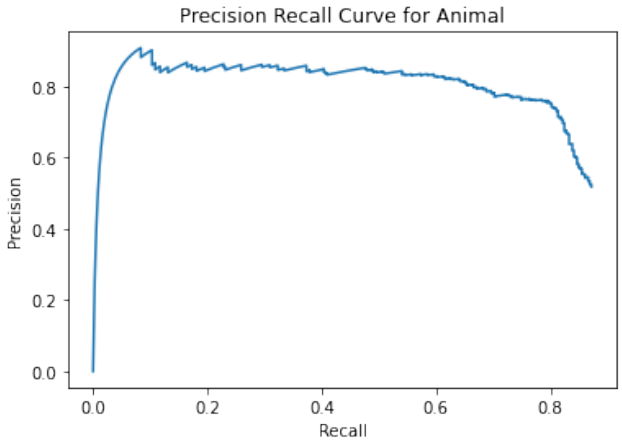


NOTE: Here the **magenta** box represents the ground truth bounding box. The **red, blue** and **green** box corresponds to the boxHead top-20 box output in *Animal, Vehicle and Human* category respectively.

4. PR-curve and Report of the APs and the mAP



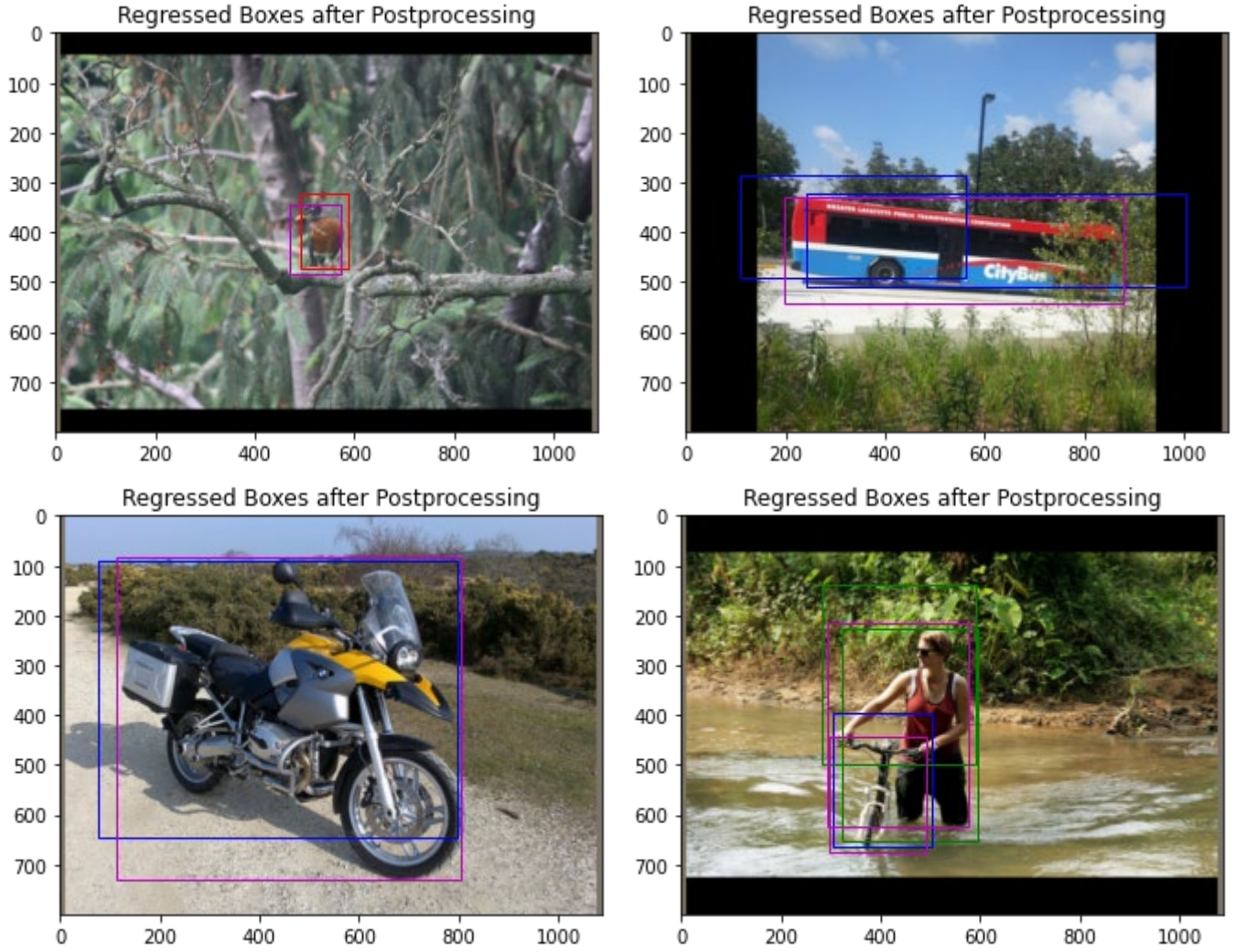




- **Box Head Performance and Comments**

1. **Average Precision for each class:** Vehicle 65.2%, Human 72.9%, Animal 70.7%
   Mean Average Precision: 69.6%
2. **Choice of top K and top M:** According to trial and error, if the top K boxes to keep after the cross-boundary cross and the top M boxes to keep after NMS are larger, higher APs and mAP could be achieved. Therefore, we take K as 100 and M as 50 in this assignment. In the next section, it could be observed that our model produces relatively ideal outcomes.
3. **Other Comments**:
   We can see that the PR curve is somewhat different from the typical type (i.e. starting from the left up (0,1) and going downwards). We believe that may be due to the high variance of the distribution of negative scores or the overlapping of the distribution of negative and positive scores. It shows that our model still has some potential to be improved. However, the overall performance is satisfactory.

5. Image plots of the regressed boxes after the postprocessing



NOTE: Here the **magenta** box represents the ground truth bounding box. The **red, blue** and **green** box corresponds to the boxHead top-20 box output in *Animal, Vehicle and Human* category respectively.