CIS 680, Vision & Learning, Fall 2018: HW2B

Abhijeet Singh

November 12, 2018

1 Region Proposal Network

1.1 RPN classifier

Final test accuracy: 96.37%
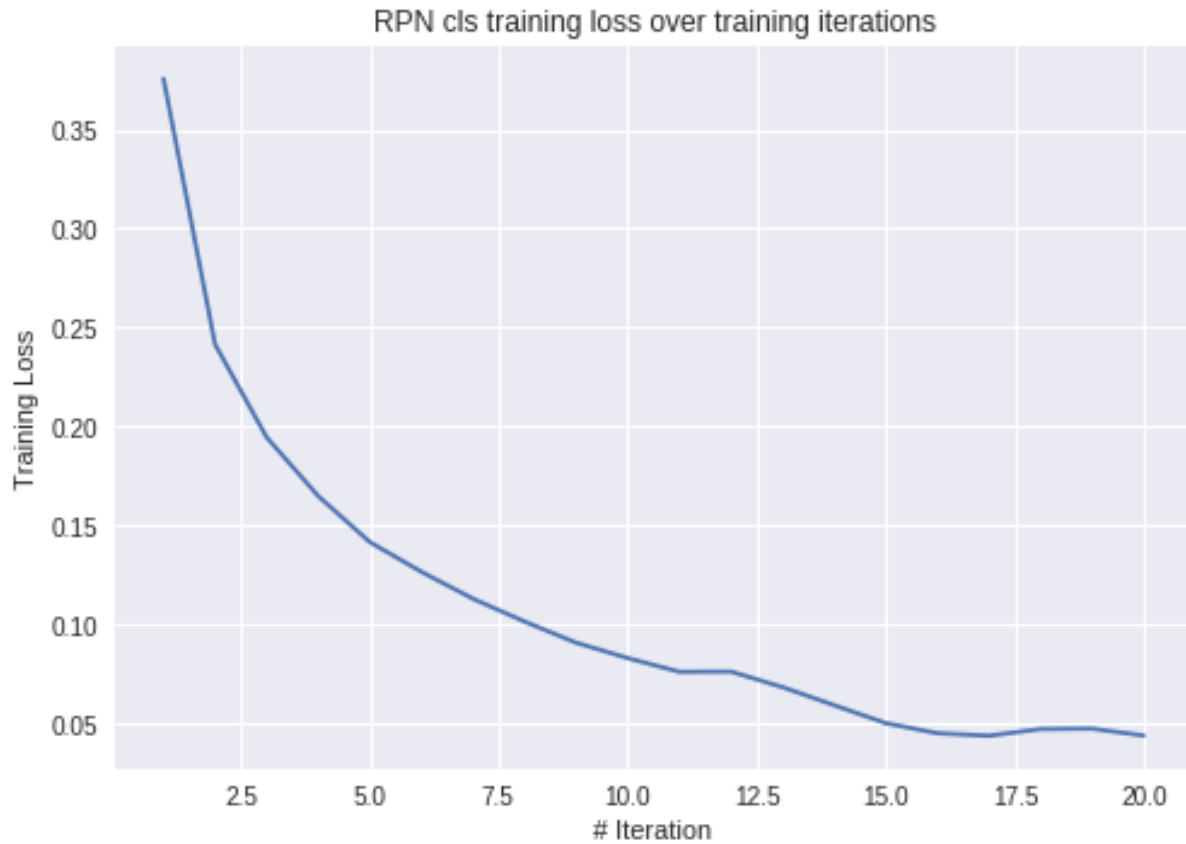

Figure 1: RPN Classifier

1.2 RPN regressor

Final testing regression loss: 0.0298

Figure 2: RPN regressor

2 Faster R-CNN, Mask R-CNN

2.1 Spatial transformer

Sample input and output to the spatial transformer as shown below. θ input to the spatial transformer is given corresponding to the person in the input image.
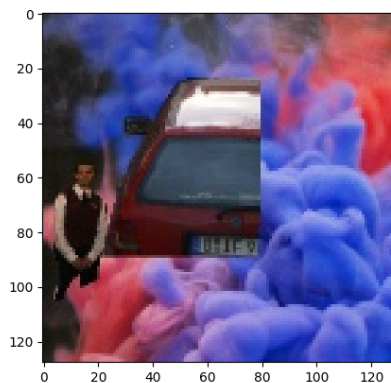

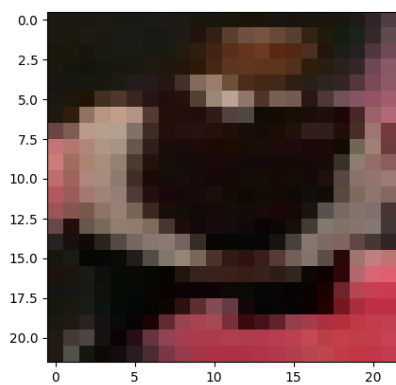Figure 3: Sample input image to spatial transformer

Figure 4: Output of spatial transformer from the above image

2.2 Faster RCNN
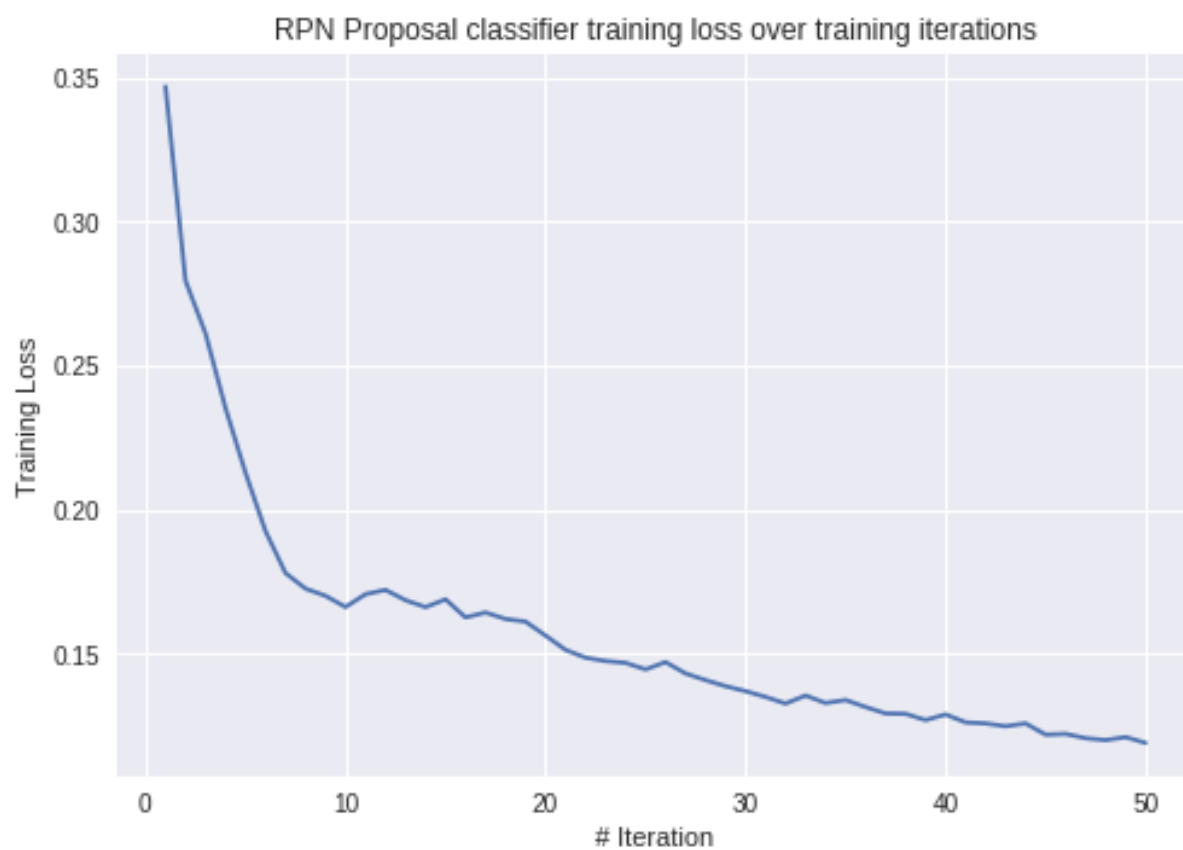
Final classification test accuracy: 74.75%

Figure 5: RPN classifier training loss during Faster-RCNN training

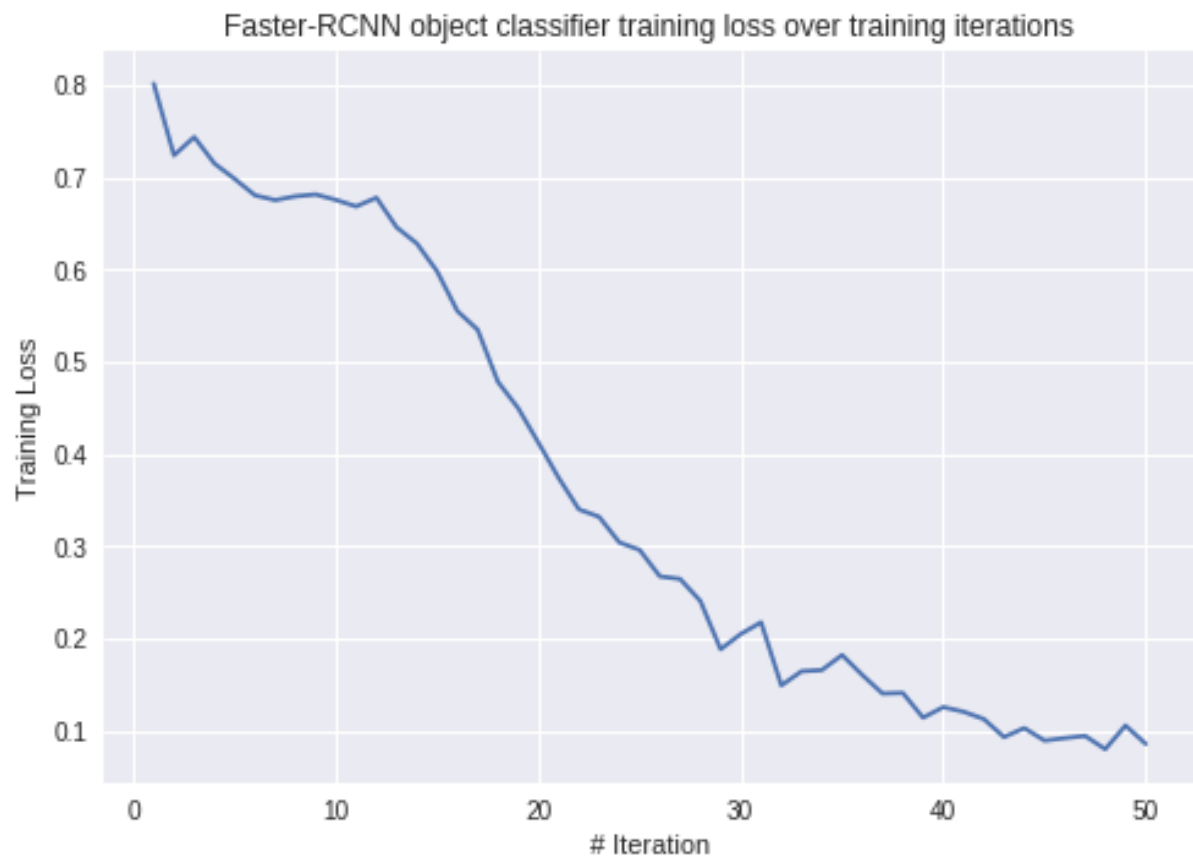Figure 6: RPN regressor training loss during Faster-RCNN training

Figure 7: Faster-RCNN classifier training loss

2.3   Mask RCNN

Final classification test accuracy (pixel-wise): 81.04%
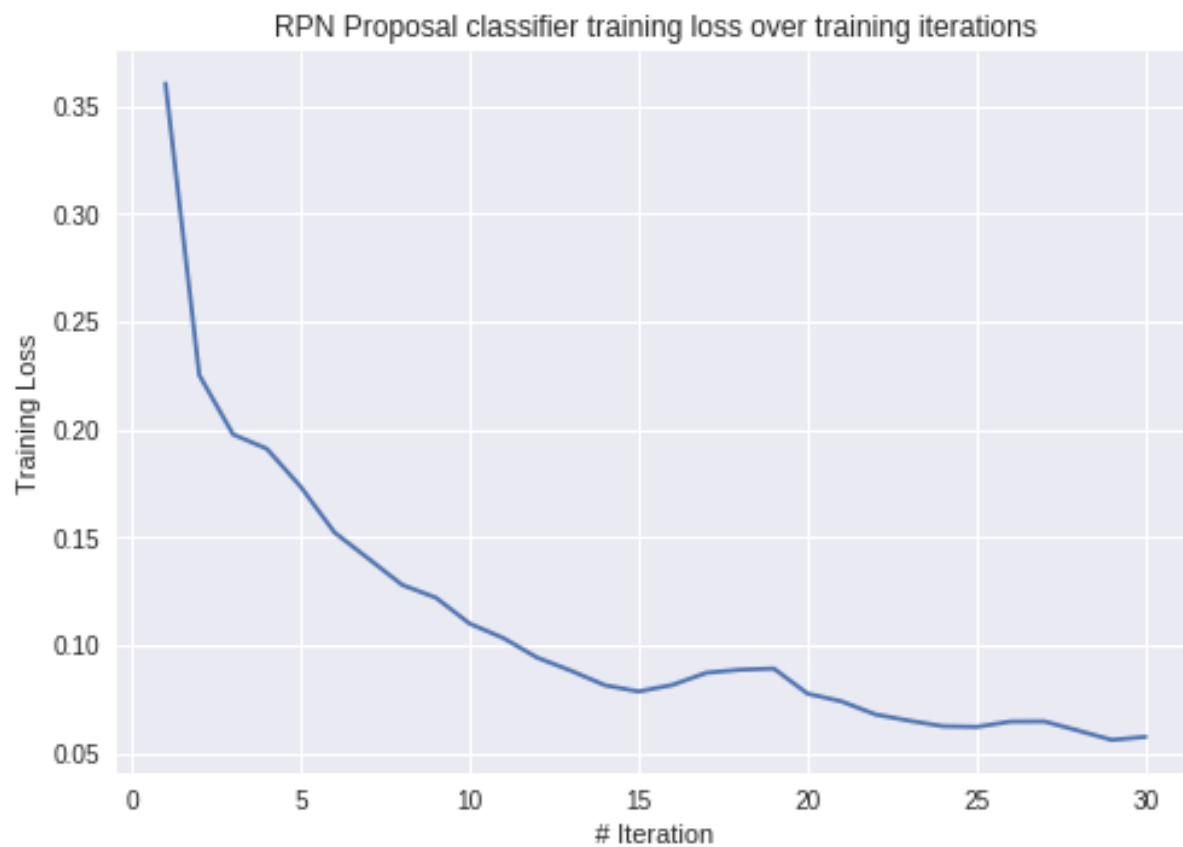Final classification test accuracy (IoU-wise): 78.41%

Figure 8: RPN classifier training loss during Mask-RCNN training

Figure 9: RPN regressor training loss during Mask-RCNN training

Figure 10: Mask-RCNN training loss (combined training)

I also tried alternate training for Mask-RCNN where the RPN was first trained and then the Mask-RCNN layers. The results were -

Final classification test accuracy (pixel-wise): 79.55%

Final classification test accuracy (IoU-wise): 76.61% (Results from combined training are better, maybe because the base layers are updated during combined training which is not the case in the alternate training)
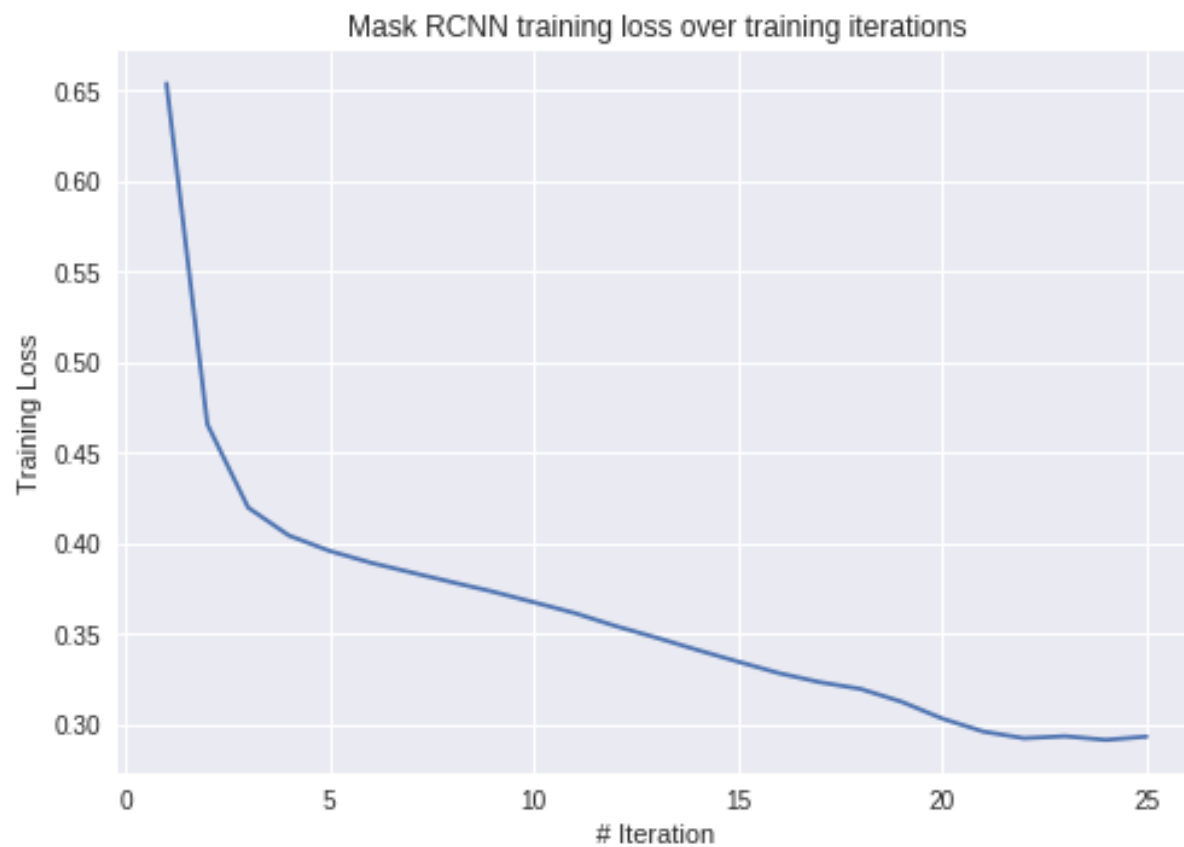
Figure 11: Mask-RCNN training loss (alternate training)