CS783 Assignment 3

Aman Tiwari(160094), Ankush Tiwari(160130)

March 2019

1 Problem Statement

The problem is to implement object detection algorithm which can detect bounding boxes of objects in images and their classes.

2 Methods Used

Our approach was to first train a resnet18 model with the images of objects belonging to the 4 classes (aeroplane, bottle, chair, background). The images were extracted from the training images with the help of the ground truth bounding boxes. The images for background class were extracted by taking randomly cropped patches from the image and selecting those with minimum intersection over union. Then during testing, we use a sliding window to slide through the image and find the class of the patch returned by the sliding window.

3 Challenges Faced

1. One of the problems was that the patches of the images that we generated by the sliding window approach were not able to cover all objects(since some objects took up large space within the image and some were very small). This affected the prediction where a lot of patches were predicted as background. Instead of keeping constant values for the scale, we tried to use scale values based on the image size, which resulted in improving the predictions.

2. We also faced some problems with implementing the detection using two layers of the resnet18. We tried two ideas. First we tried to use two different models, one using all the layers of resnet18, and other only upto the second last layer of resnet18(convolutional layer of shape 512x7x7). Then we used the predictions from both these models to predict bounding boxes in the image. So, the final predictions contained the union of the bounding boxes predicted by both the models. We also tried to use a single model by flatenning the convolutional layer and concatenating it with the last layer, then using a linear layer to get the predictions.

4 Results

The results for the single layer detection method are:
AP score for aeroplane class = 0.0227
AP score for chair class = 0.013
AP score for bottle class = 0.003
The mAP score = 0.012

The results for two layer detection method are:
AP score for aeroplane class = 0.0257
AP score for chair class = 0.017
AP score for bottle class = 0.0041
The mAP score = 0.0156

The codes that we used as a reference for this assignment are:

(a) The code for Non-Maximum Supression from the following blog Faster Non Maximum Supression

(b) The code to calculate average precision from the following blog mAP Calculation