NTNU
Norwegian University of
Science and Technology

TDT4265 - Computer Vision and Deep Learning

Assignment 4

Jakob Løver
Einar Henriksen

March 17, 2019

Contents

1 Object Detection Metrics

1.1 Task a

Explain what the Intersection over Union is and how we can find it for two bounding boxes. Illustrate it with a drawing.

Intersection over union is a metric that is used to evaluate the performance of object detectors. It works by dividing the "area of overlap" by the "area of union." This will give us a metric that is mapped between [0,1] that tells us how well the labelled bounding box overlaps with the predicted bounding box.

$$\text{Intersection of Union} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{1}$$

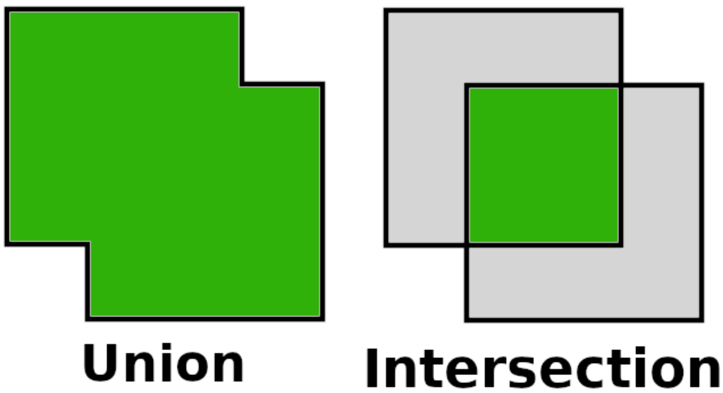

Figure 1.1: Illustration of Intersection and Union

1.2 Task b

What is a true positive? What is a false positive?

First, we can define two classes; a positive class and a negative class. If we want to detect a traffic sign for example, the classes would be

Positive class Sign

Negative class No sign

We can then define true positive and false positive as

True positive Object detector detected correctly the positive class

False positive Object detector detected falsely the positive class

In other words, a true positive is when the object detector detects a sign correctly, and a false positive is when it detects a sign that isn't there.

1.3 Task c

Write down the equation of precision and recall

$$\text{Precision} = \frac{\#\text{ true positives}}{\#\text{ true positives} + \#\text{ false positives}} \tag{2}$$

$$\text{Recall} = \frac{\#\text{ true positives}}{\#\text{ true positives} + \#\text{ false negatives}} \tag{3}$$

1.4 Task d

Given the following precision and recall curve for the two classes, what is the mean average precision?

$$\begin{aligned}
\text{Precision}_1 &= [1.0, 1.0, 1.0, 0.5, 0.20] \\
\text{Recall}_1 &= [0.05, 0.1, 0.4, 0.7, 1.0] \\
\text{Precision}_2 &= [1.0, 0.80, 0.60, 0.5, 0.20] \\
\text{Recall}_2 &= [0.3, 0.4, 0, 5, 0.7, 1.0]
\end{aligned}$$

The way this is calculated can be found in the zipped folder under `task1_d.py`. We ended up with the following values

Class 1 Average precision	0.6455
Class 2 Average precision	0.5182
Mean average precision	0.5818

2 Implementing Mean Average Precision

The mean average precision was as expected 0.9066. The precision recall curve is shown in figure 2.1.

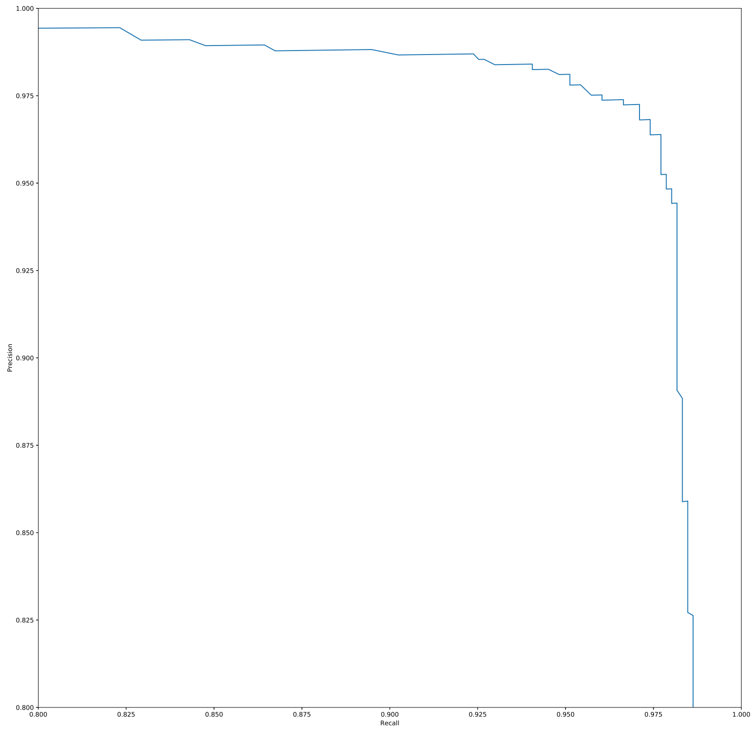

Figure 2.1: Precision recall curve for assignment 2

3 You Only Look Once

3.1 Task a

YOLO divides the image into a grid of cells. Each of these cells can only be assigned one class and can predict a limited number of bounding boxes. This means the resolution of the grid constrains the size and amount of objects YOLO is able to detect. It especially struggles with small objects appearing in groups.
The model uses data to learn how to predict bounding boxes, so it struggles to generalize in order to identify objects with unusual aspect ratios etc.

3.2 Task b

FALSE: Like the name implies, YOLO sees the entire image at once. It does not use sliding-window.

3.3 Task c

Fast YOLO's neural network has 9 convolutional layers instead of 24, and each layer has fewer filters. Other than that they are identical, the training and testing parameters are the same.

3.4 Task d

Faster R-CNN has improved speed and accuracy over standard R-CNN by using neural networks instead of Selective Search. Much effort has been put into speeding up the DPM pipeline with for example HOG computation and using GPUs for heavy computation. However, even though it's faster than standard R-CNN, Faster R-CNN is too slow for real time. YOLO on the other hand skips the pipeline altogether, and manages to be much faster. The fastest R-CNN at the time the paper was written was still 2.5 times slower than YOLO, and also less accurate.

4 You Only Look Once

This assignment was implemented using Jupyter Notebook. See the attached notebook. Notice that the only change in starter code is the line where we import the `drawing_utils.py` file. The Jupyter notebook was started with the yolo folder as the notebook directory, such that the necessary files can be imported using its relative path.