IPCV report

1 The Viola-Jones Object Detector

Ground Truth and Visualisation

Ground truths help visualise the true values of an object within an image. For the given test images, I manually drew the ground truth bounding boxes and chose to include only frontal faces as the Viola-Jones detector is described as a frontal face detector, and so side faces should not be considered true values. The following images show the manually generated ground truth values for the faces together with the faces detected by the Viola-Jones detector, which are bound by the red and green boxes respectively.

Figure 1a: NoEntry1.bmp

Figure 1b: NoEntry2.bmp

Figure 1c: NoEntry4.bmp

Figure 1d: NoEntry5.bmp

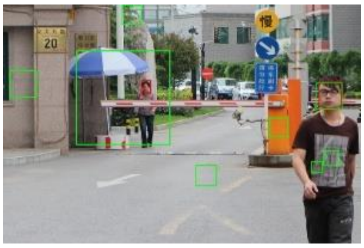

Figure 1e: NoEntry7.bmp

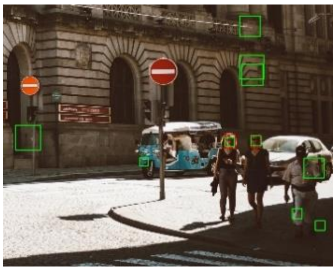

Figure 1f: NoEntry11.bmp

IOU, TPR, F1-Score:

The Intersection Over Union is an evaluation metric used to measure the accuracy of an object detector on a particular dataset by comparing the bounding boxes of the ground truths and the predictions from our model:

$$IOU = \frac{Area\ of\ Overlap}{Area\ of\ Union} = \frac{A_1 \cap A_2}{A_1 \cup A_2}$$

Using the IOU to determine successful detection allows for some ambiguity in defining the manually drawn ground truths.

The True Positive Rate (TPR) is the ratio of correct detections for an image to the number of ground truths for a given image.

$$TPR = \frac{TP}{TP + FN}$$

Another metric to determine the performance of the detector is the F1-Score:

$$F1 = 2 \times \frac{precision\ \times recall}{precision + recall} = = \frac{TP}{TP + {}^{1}/_{2}\,(FP + FN)}$$

Image File:	TPR	F1-Score
NoEntry0.bmp	N/A	N/A
NoEntry1.bmp	1.000	0.250
NoEntry2.bmp	1.000	0.267
NoEntry3.bmp	N/A	N/A
NoEntry4.bmp	1.000	0.333
NoEntry5.bmp	1.000	0.400
NoEntry6.bmp	N/A	N/A
NoEntry7.bmp	0.500	0.200
NoEntry8.bmp	N/A	N/A
NoEntry9.bmp	N/A	N/A
NoEntry10.bmp	N/A	N/A
NoEntry11.bmp	1.000	0.333
NoEntry12.bmp	N/A	N/A
NoEntry13.bmp	N/A	N/A
NoEntry14.bmp	N/A	N/A
NoEntry15.bmp	N/A	N/A
Average	0.917	0.297

Figure 2: Table of TPR and F1-scores for face detector

While TPR is a useful indicator of the performance of a detector, it is unprofitable when using with images where the image does not contain the object you are trying to detect. As exhibited in *figure 2*, images with no ground truths do not yield any useful information. In these cases, metrics which take into account false positives, such as the F1-Score, would be a more useful indicator.

Moreover, TPR relies on the IOU to determine whether the objects detected are to be accepted as the truth (i.e. the true positives). The IOU in turn depends on the manually set threshold value. This means that we can make our IOU reject or accept every image detected by our Viola-Jones detector. It is therefore always possible to achieve a TPR rate of 100% as the IOU threshold is manually set. As the threshold value decreases, more images are accepted and the true positive rate increases. Setting the threshold value to 0 means that the detections do not even need to intersect and will therefore yield a TPR rate of 100%. Consequently, we need to find a suitable threshold value for our dataset. In order to generate the results above, I chose a threshold of 0.4, which I found to give the best results.

2 Building & Testing your own Detector

Training performance

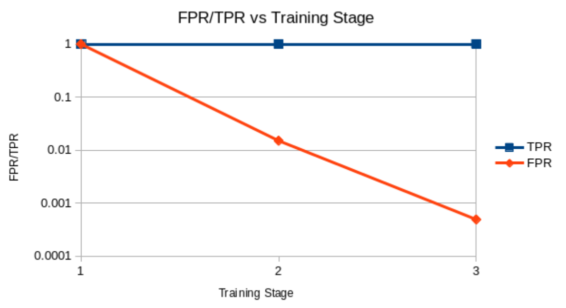

Figure 3: Graph of FRP vs TPR during training stages

A cascading classifier for no-entry signs was built by first generating a set of positive sample images as training data. These images together with the set of given negative images were then used to train the detector in 3 stages. Figure 2 shows the changes in the True Positive Rate and the False Negative Rate during the 3 stages.

In the initial stage of training, all regions of the image are accepted, and so both the FPR and TPR start off at 100%. TPR remains at 100% throughout the training process. As training proceeds, more classifiers are added to the cascade and so more images are rejected, resulting in a sharp drop in the FPR from 1 to 0.015 in the subsequent stage. The FPR continues to decrease in the 3rd stage to a final rate of 0.00049. This decrease is a lot less significant. This is likely because subsequent classifiers have harder examples to train as they only use examples which have not yet been rejected in prior stages. Moreover, features selected in later stages yield higher error rates.

Testing performance:

The following images are a subset of the images produces by running the Viola Jones detector with the newly trained classifier on the test images, together with the ground truths, which are bound by the green and red boxes respectively:

Figure 4a: NoEntry0.bmp

Figure 4b: NoEntry5.bmp

Figure 4c: NoEntry8.bmp

Image Name	TPR	F1 Score
NoEntry0.bmp	1.000	0.182
NoEntry1.bmp	1.000	0.333
NoEntry2.bmp	1.000	0.200
NoEntry3.bmp	1.000	0.364
NoEntry4.bmp	1.000	0.190
NoEntry5.bmp	0.100	0.071
NoEntry6.bmp	0.250	0.167
NoEntry7.bmp	0.000	0.000
NoEntry8.bmp	0.500	0.600
NoEntry9.bmp	0.500	0.286
NoEntry10.bmp	0.667	0.500
NoEntry11.bmp	0.500	0.154
NoEntry12.bmp	0.625	0.476
NoEntry13.bmp	0.000	0.000
NoEntry14.bmp	1.000	0.250
NoEntry15.bmp	1.000	0.500
Average	0.634	0.267

Figure 5: Table of TPR and F1-Scores for no-entry detector

In the table above, you can see the TPR and F1 score of the no-entry detector for each of the images. On average, this detector fared worse than the face detector with a TPR of 63.4%, which is 28.3% lower than that of the face detector. Compared to the face detector, the no-entry detector struggles to correctly identify no-entry signs. This is likely due to faces being easier to detect and distinguish as they have a lot more distinctive features which sets them apart from other objects. Another reason may be because no entry signs are often in the background of an image, so the image is often unclear, as opposed to faces which are often at the forefront.

The classifier performs best when detecting images face on, as in *figure 4a*, where it successfully detects all no-entry signs. The detector was also trained to detect no-entry signs at different angles and is therefore able to successfully detect signs that are slightly titled. However, it struggles to do so at higher angles (*figure 4b*). It is also unable to detect signs which are highly obscured (*figure 4c*).

There was also a 3% drop in the F1-Score to 26.7%. This is likely due to the vast number of false positives (*figure 4*).

3: Integration with Shape Detectors

Hough Details

Clearly, the Viola-Jones detector is not very good at detecting no-entry signs. We can therefore combine it with a circle Hough transform in order to filter out the vast number of false positives.

Figure 6a: Threshold magnitude NoEntry11.bmp

Figure 6b: Hough Transform NoEntry11.bmp

Figure 6c: Hough Circles NoEntry11.bmp

Figure 6d: Filtered detected signs NoEntry11.bmp

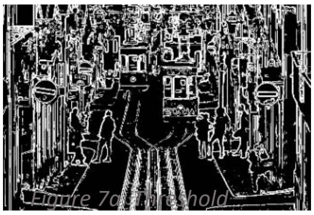

Figure 7a: Threshold Magnitude NoEntry12.bmp

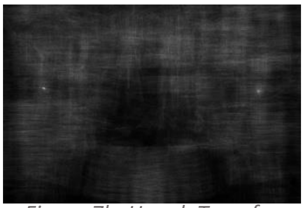

Figure 7b: Hough Transform NoEntry12.bmp

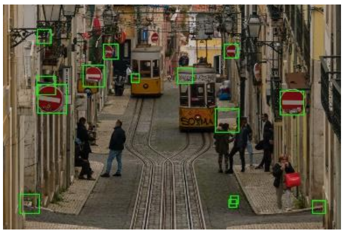

Figure 7c: Unfiltered detected signs NoEntry12.bmp

Figure 7d: Filtered detected signs NoEntry12.bmp

Image Name	TPR	ΔTPR	F1	Δ F1
NoEntry0.bmp	1.000	0.000	1.000	+0.818
NoEntry1.bmp	1.000	0.000	1.000	+0.667
NoEntry2.bmp	1.000	0.000	1.000	+0.800
NoEntry3.bmp	1.000	0.000	0.800	+0.436
NoEntry4.bmp	1.000	0.000	1.000	+0.810
NoEntry5.bmp	0.100	0.000	0.182	+0.111
NoEntry6.bmp	0.250	0.000	0.400	+0.233
NoEntry7.bmp	0.000	0.000	0.000	0.000
NoEntry8.bmp	0.500	0.000	0.667	+0.067
NoEntry9.bmp	0.500	0.000	0.667	+0.381
NoEntry10.bmp	0.667	0.000	0.800	+0.300
NoEntry11.bmp	0.500	0.000	0.667	+0.513
NoEntry12.bmp	0.500	-0.125	0.667	+0.191
NoEntry13.bmp	0.000	0.000	0.000	0.000
NoEntry14.bmp	1.000	0.000	1.000	+0.750
NoEntry15.bmp	1.000	0.000	1.000	+0.500
Average	0.626	-0.008	0.678	0.411

Figure 8: TPR and F1 table for filtered no-entry signs

Key Merits:

- Significant increase in precision as majority of false positives filtered out
- Major increase in F1 Score, increasing accuracy of detector

Key Shortcomings:

- Slight decrease in TPR due to simplicity of shape detector used *(figure 7c, 7d* – one less sign detected) so false negative increases but only 1 sign affected
- Limited by the performance of previous detector as new bounding boxes cannot be created. Detector can only be as good as the Viola-Jones detector even if signs are detected by circle detector (*figures 6c, 6d*), so number of true positives cannot increase
- Some false positives still remain so more advanced filtering may be required

Detection Pipeline:

The steps taken to integrate the Viola-Jones detector together with the circle detector are depicted below. At this stage, more complex shape detectors are unnecessary as only NoEntry12.bmp will be affected. Even if implemented, due to the obscured nature of that particular sign, the shape detector may fail to detect it, and there will be no change in TPR.

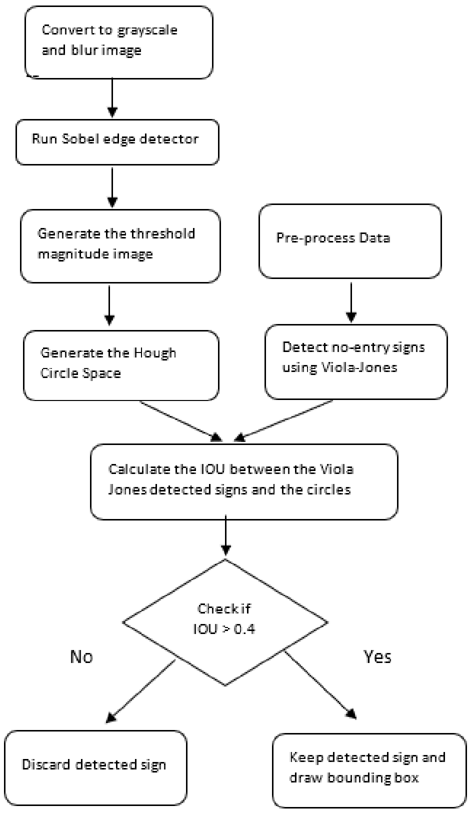

Figure 9: Detection Pipeline to filter detected faces

4: Improving Detector

Idea:

One approach to improve the shape detector is by detecting more complex shapes, such as ellipses and parallel lines. This means that the shape detector will be able to filter out more false positives and detect tilted no-entry signs. However, as seen in *figure 12*, this approach would have minimal effect on our results as it would only improve the TPR of NoEntry12.bmp. This is because the final detector is constrained by the accuracy of the Viola-Jones detector. Even if the shape detector is improved, the final detector will continue to be limited by the accuracy of the Viola-Jones detector and results will vary by little. [1] mentions averaging overlapping bounding regions. This would have little effect for the same reasons mentioned above, and also because the Viola-Jones detector generates little overlapping regions over the no-entry signs. As a result, I sought to improve the Viola-Jones detector instead. At the moment, the detector struggles to correctly identify no-entry signs. Surprisingly, some signs which are frontal facing and in the foreground of the images are not detected (*figure 9a*). This can be attributed to noise in the image, so I tried to perform some extra pre-processing, specifically median blur. Additionally, the detector performs poorly with signs which are tilted at higher angles. As such, I retrained the detector with more sample images and a higher y-rotation angle.

Visualise:

With the median blur, there was a slight increase in the TPR of the viola-Jones detector with 2 newly detected no-entry sigs from all 16 images.

Figure 10a: Previous Viola-Jones detection NoEntry10.bmp

Figure 10b: New Viola-Jones detection NoEntry10.bmp

However, neither of these 2 images are detected by the Hough detector as both no-entry signs are partially obscured. Consequently, there is no improvement in the final detector, and so the TPR and f1-scores remain the same.

Retraining the cascade with the new parameters with no blur yields better results, with several new true positives. However, we are once again restricted by the circle detector. After combining the new classifier with the Hough detector, we yield similar results.

Figure 11a: Previous final detector NoEntry6.bmp

Figure 11b: New final detector NoEntry6.bmp

Combining the median blur with the new cascade yields worse results.

Evaluate:

Image Name	TPR	ΔTPR	F1	Δ F1
NoEntry0.bmp	1.000	0.000	1.000	0.000
NoEntry1.bmp	1.000	0.000	1.000	0.000
NoEntry2.bmp	1.000	0.000	1.000	0.000
NoEntry3.bmp	1.000	0.000	0.800	0.000
NoEntry4.bmp	0.500	-0.500	0.667	-0.333
NoEntry5.bmp	0.100	0.000	0.182	0.000
NoEntry6.bmp	0.500	+0.250	0.667	+0.267
NoEntry7.bmp	0.000	0.000	0.000	0.000
NoEntry8.bmp	0.500	0.000	0.667	0.000
NoEntry9.bmp	0.500	0.000	0.667	0.000
NoEntry10.bmp	0.667	0.000	0.800	0.000
NoEntry11.bmp	0.500	0.000	0.667	0.000
NoEntry12.bmp	0.500	0.000	0.667	0.000
NoEntry13.bmp	0.000	0.000	0.000	0.000
NoEntry14.bmp	0.000	-1.000	0.000	-1.000
NoEntry15.bmp	0.500	-0.500	0.667	-0.333
Average	0.517	-0.109	0.591	- 0.087

Figure 12: TPR and F1 table for newly trained detector

With the new cascade, only 4 images were affected. The only image on which the new detector improved was NoEntry6.bmp (*figure 11*) while results worsened for the other 3 images.

Key Merits:

- Circles tilted at higher angles now able to be detected by Viola-Jones detector

Key shortcomings:

- Although able to detect more no-entry signs, due to simplicity of Hough, majority of these signs still unable to be detected in final detector
- Some previously detected no-entry signs now unable to be detected, leading to worse TPR and F1-score overall
- False negatives continue to exist

Next steps:

In order for the detector to accept tilted no-entry signs, the Hough detector should be able to detect ellipses. However, as this requires a 5D Hough Space, this brings the question of whether the slight increase in precision is worth the significant increase in computational power. Another potential improvement would be to detect parallel lines as well as the circles. However, this would only filter out false positives, bringing minimal effect as only 1 false positive remains in all test images. It is clear by now that in order to significantly improve the performance of the final detector, both detectors will have to be improved. Another detection algorithm in place of the Viola-Jones can be used.

References:

[1] Paul Viola and Michael J.Jones. Robust real-time face detection. International Journal of Computer Vision, 57(2):137–154, 2004