IPCV - Shape Detection

November 29, 2024

1 Subtask 1

1.1 Training Performance

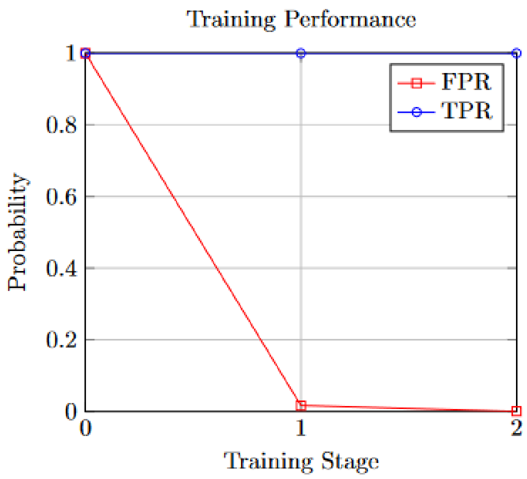

Figure 1: False positive and true positive rates for each stage in training

The False Positive Rate (FPR) decreases from 1 to 0.0005 as training stage increases.

$$FPR = \frac{FP}{FP + TN}$$

For this to occur, the number of false positives decreases and/or the number of true negatives increases.

The True Positive Rate (TPR) remains at 1 in every training stage.

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

We can conclude that there are no false negatives; therefore, the Viola-Jones model successfully detects every no-entry sign in the training stage.

This is due to the boosting algorithm having multiple stages, implementing a form of supervised learning, where results are refined (in this case, false positives are reduced).

$$Precision = \frac{TP}{TP + FP}, \qquad Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

1.1.1 Example Output Images

Figure 2: Example output images

Image Number	TPR	F1
0	1.000	0.286
1	1.000	0.400
10	0.667	0.500
11	1.000	0.267
12	0.500	0.333
13	0.000	0.000
14	1.000	0.200
15	1.000	0.667
2	1.000	0.182
3	1.000	0.500
4	1.000	0.364
5	0.100	0.114
6	0.750	0.545
7	0.000	0.000
8	0.333	0.500
9	0.500	0.400
Average	**0.678**	**0.336**

Table 1: TPR and F1 for all images for Subtask 1

The parameters for 'detectMultiScale' were left as default. This resulted in a high number of false positives detected by the Viola-Jones model.

TPR roughly 67.8% of the signs were detected, which is a fairly high rate compared with the F1 score. This result means that roughly two-thirds of the no-entry signs were detected.

The performance of this is significantly worse than the training performance because the training stage had next to no false positives, whereas the number of false positives for testing is high.

It is expected that testing generally should have worse performance than training, as the testing stage is unsupervised.

2 Subtask 2

2.1 Edge Magnitude Image, Hough Space, Final Detections

Merits Images:

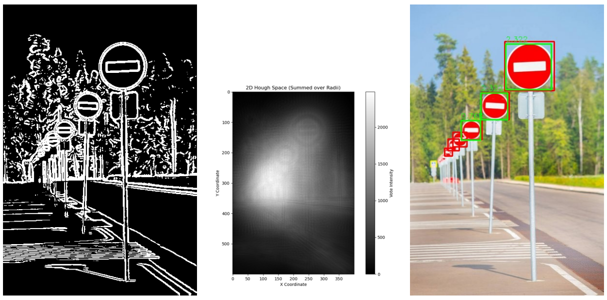

Figure 3: Detection of fully-visible stop signs

Detects all three of the fully-visible stop signs.

Limitations Images:

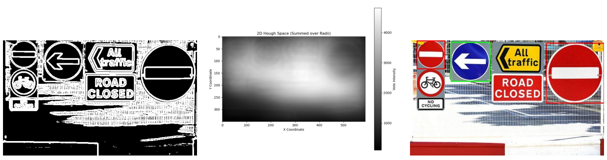

Figure 4: Detection limitations due to false positives

Detects a false positive on the blue sign (due to not checking colours and horizontal bar). Does not detect the large, prominent sign in the top right.

2.2 Performance Evaluation

Image Number	TPR	F1
0	0.500	0.667
1	0.000	0.000
10	0.667	0.800
11	0.500	0.400
12	0.250	0.400
13	0.000	0.000
14	1.000	1.000
15	1.000	1.000
2	1.000	1.000
3	0.500	0.500
4	0.500	0.667
5	0.000	0.000
6	0.250	0.333
7	0.000	0.000
8	0.500	0.500
9	0.500	0.667
Average	**0.448**	**0.517**

Table 2: TPR and F1 for all images for Subtask 2

The circle transformation was used to locate areas of interest in the input image. From the detections, the Viola-Jones model could be run.

If both of these detect the no-entry sign, then we say that is a true positive detection.

Implementing IOU evaluation with the predicted no-entry signs and ground truths made the false positive rate (which was previously high) become zero.

Non-Maximum Suppression was used to prevent multiple detections of the same object, allowing the true positive rate calculated to be the number of distinct ground truths detected.

Performance could be marginally improved by manually adjusting parameters and thresholds.

This will vary from the training performance due to:

- Some signs are too small to detect. - Some signs are obscured by other objects. - Signs vary slightly from country to country (thicker middle line, no white border, etc.). - Signs have noise in real images.

The pre-trained classifier runs on grayscale images; hence, in the example output, the blue sign was detected.

Increased number of false positives and true positives. F1 score decreased.

2.3 Pipeline Flowchart

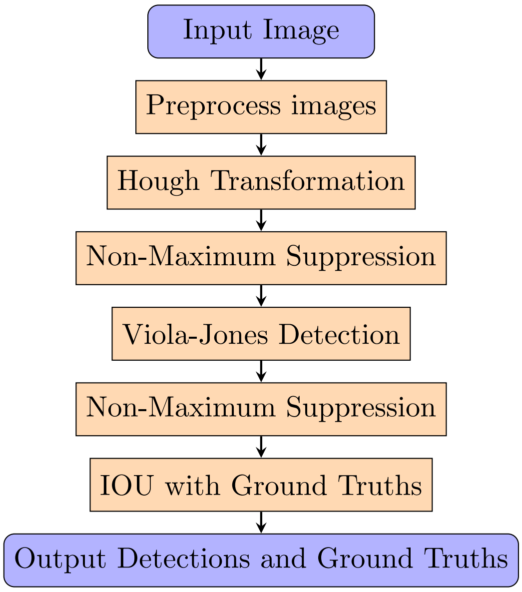

Figure 5: Flowchart of the detection pipeline used

3 Subtask 3

3.1 Improved Detection Algorithm

Idea:

- Normalise the images so that different brightness levels do not affect model performance. - Threshold the image by the intensity of red pixels. - Use OpenCV's adaptive thresholding on the blurred red data. - Perform morphological operations (opening and closing) to remove salt and pepper noise on the thresholded image. - Use OpenCV's built-in Canny edge detection to get clearer edges. - Generate contours of the gradient image and then fit ellipses and rectangles to them. This helps find the middle bar (rectangles) and perspective-warped signs (ellipses) without using the computationally expensive Hough ellipse function. - Pass these detections into the Viola-Jones classifier to get final detections. Compute Non-Maximum Suppression on this combined model.

Decreased the minimum size of the detections of the cascade classifier now that we have more confidence in the accuracy of the preliminary detector, causing an increase in TPR.

3.2 Performance Evaluation

Image Number	TPR	F1
0	0.000	0.000
1	0.000	0.000
10	0.667	0.800
11	1.000	1.000
12	0.500	0.571
13	0.000	0.000
14	1.000	1.000
15	1.000	1.000
2	1.000	0.667
3	1.000	0.800
4	1.000	1.000
5	0.000	0.000
6	0.750	0.857
7	1.000	0.667
8	0.667	0.800
9	0.500	0.500
Average	**0.733**	**0.670**

Table 3: TPR and F1 for all images for Subtask 3

This has a significantly higher TPR and F1 score, meaning the model is performing better on the test set. The fact that the TPR increased much more than the F1 score suggests that the model has a relatively high false positive rate.

3.3 Merits and Limitations

Merits:

- It can detect no-entry signs that are not facing directly forward, where the shape appears as an ellipse. - The processing time of calculating the Hough circles is reduced due to the red filter stage because the image has less data. - Overall detection quality is improved due to advanced preprocessing and edge detection methods.

Limitations:

- Some of the ground truths are nearly impossible to detect without introducing false positives. This is due to them being far away (see ‘NoEntry8‘ or ‘NoEntry5‘). - The ground truth in image 13 is not detected by the Viola-Jones classifier when it is passed in. This means the preprocessing that was implemented was not successful enough at reducing lighting effects (‘NoEntry13‘ has a lamp on top which makes the sign not have uniform lighting). - When there are two ellipses close together, they are merged. While this works well when the ellipses represent the top and bottom half of a sign, as seen in ‘NoEntry10‘, it can also merge two stop signs into one region, resulting in an incorrect classification.