# Subtask 1: The No-Entry Sign Detector

## Training Performance

Training performance improved rapidly over the three stages. The initial stage had only 100% rate of true and false positives by evaluating every image to be positive. The two next stages quickly reduced the false positive rate to 1.65% and 0.05% respectively. The true positive rate remained at a constant 100% throughout, exposing the model’s bias towards positive images on the training data. This is because the positive images are very straightforward and are tested on the same data that it is trained on.

## Testing Performance

A picture containing text, tree, outdoor, sign

Description automatically generatedCompared to the training, the model performed considerably worse on testing data than training data. The true positive rate averaged 58.5% and an F1 score of 0.477. Due to the testing positives being partially obscured (Figure 2) while varying much in shape and proportion (Figure 3), there is a considerable drop of the true positive rate from the training performance.

Figure : True Positive Rate and False Positive Rate for each stage

The model is unable to spot signs without the high contrast borders that are present in the positive training images and is also unable to distinguish between brighter areas in images and the main white stripe on the stop sign. This is most likely due to the regularity and lack of variety of the positive training images or the model being underfitted.

Adjusting the model’s ‘minNeighbours’ attribute resulted in a worse true positive rate and overall F1 score, however much increased the model’s reliability (Figure 5). This could be argued to perform better depending on the context the model is used for.

Figure : Obscured signs

A picture containing text, sign

Description automatically generatedPerformance could be improved with a wider variety of training images, both positive and negative. A further stage in the model’s creation process could theoretically improve performance by reducing underfitting but experimentation with 3 stages resulted in a massively reduced true positive rate and therefore overfitting.

Figure 5: Min neighbours against various scores

Figure : Mistaken white stripes

Figure 3: Obscured and difficult orientations

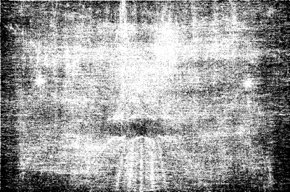
# Diagram Description automatically generatedSubtask 2: Integration with Shape Detectors

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Shape Filtering | |  |  |  |  |  |  |
| Name | TP | FP | FN | TPR | F1 | TP Difference | FP Difference |
| NoEntry0 | 2 | 0 | 0 | 1 | 1 | 0 | -5 |
| NoEntry1 | 1 | 0 | 0 | 1 | 1 | 0 | -1 |
| NoEntry2 | 1 | 1 | 0 | 1 | 0.6667 | 0 | 0 |
| NoEntry3 | 2 | 1 | 0 | 1 | 0.8 | 0 | -4 |
| NoEntry4 | 2 | 0 | 0 | 1 | 1 | 0 | -2 |
| NoEntry5 | 1 | 0 | 10 | 0.0909 | 0.1667 | -1 | 0 |
| NoEntry6 | 3 | 0 | 1 | 0.75 | 0.8571 | 0 | 0 |
| NoEntry7 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| NoEntry8 | 3 | 0 | 3 | 0.5 | 0.6667 | 0 | 0 |
| NoEntry9 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| NoEntry10 | 2 | 0 | 1 | 0.6667 | 0.8 | 0 | -2 |
| NoEntry11 | 1 | 0 | 1 | 0.5 | 0.6667 | 0 | -1 |
| NoEntry12 | 6 | 0 | 1 | 0.8571 | 0.9231 | 0 | 0 |
| NoEntry13 | 0 | 0 | 1 | 0 | 0 | 0 | -1 |
| NoEntry14 | 1 | 1 | 0 | 0.6667 | 0.6667 | 0 | -1 |
| NoEntry15 | 2 | 1 | 0 | 0.8 | 0.8 | 0 | -1 |

A picture containing text

Description automatically generatedA picture containing text, street, sign

Description automatically generatedA picture containing text, building, altar

Description automatically generatedA picture containing text, tree, outdoor, sign

Description automatically generated

A close-up of a person's chest

Description automatically generated with low confidence

A picture containing night sky

Description automatically generatedA picture containing tree, outdoor, night

Description automatically generated

Figure 7: NoEntry12

Figure 6: NoEntry8

* Viola-Jones detection drastically lowered false positive rate, leaving only 4 total false positives.
* Detection kept the true positive rate high however still filtered out a true positive.
* Further analysis could be done to reduce false positive rate further, and a less strict initial detector could allow for more true positives to be detected while filtering away the false positives.
* This however has not contributed to high number of false negatives as only positives have filtered out.

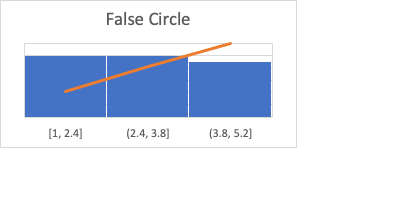
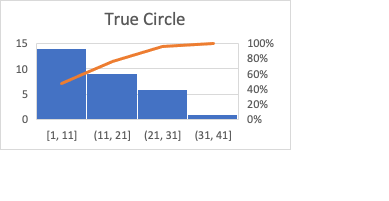
To avoid introducing more false positives into the initial detector, a filtering approach was taken to the shape detector. Detections were either accepted or discarding depending on the model’s certainty of the detection being in the shape of a stop sign. This is more performant as Hough spaces only need to be calculated for an individual detection rather than an entire image while being flexible as the acceptance threshold or importance of individual elements can be easily adjusted.

## Subtask 3: Further improvements

### Chart, line chart Description automatically generatedParameter Analysis

Each element of the above shape filtering created a probability based on the input value from the Hough space, a minimum, maximum, range of uncertainty and importance.

To optimise these parameters, I generated these values for true detections and false detections.



For example, in the graph given for the Hough circles element, it is shown that although a large proportion of true values are 11 and above while all false values are between 1 and 5.2 shown in the False Circle graph. From this information, setting a minimum greater than 5.2 and a very high maximum will filter out all these false positives. To avoid filtering out the true positives between 1 and 11, a higher minimum and a falloff value is used to decrease the certainty of the model when the value is between 1 and 10.

### Segmentation/Colour analysis

A picture containing text

Description automatically generatedGraphical user interface

Description automatically generated with medium confidenceSegmentation was used to further filter the detections. The same probabilistic approach and analytics was used above, except the value was the proportion of pixels in the detection which were above a red or white threshold. The white threshold was calculated as the sum of the pixel values divided by the average image brightness to account for brighter/darker images. The red threshold was if the difference of the red channel with the other channels was above a threshold. This led to an increase in F1 score as more false positives were filtered. Red filtering shown in left image and white filtering shown in right image. The correct white threshold is difficult to find for the as even with normalisation the brightness and contrast varies greatly across different images.