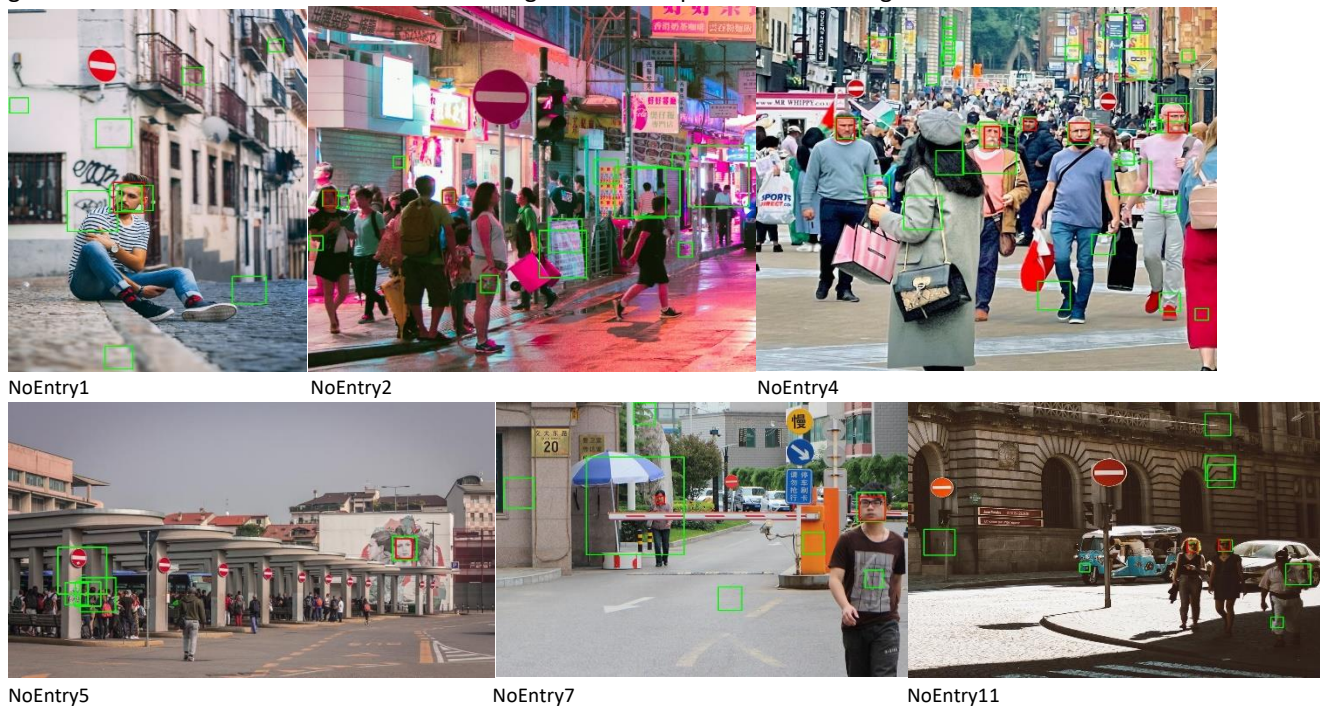


In order to detect all faces that were deemed “front facing” in the 16 example images given, the Viola-Jones detector provided by OpenCV was used in combination with a prebuilt frontal face detector. In order to determine whether the detector successfully identified a face, or identified an area containing similar Haar Features of a face, for each bounding box produced by the detector I determined the IOU(intersection over union), by modifying code from (*‘Intersection over Union (IoU) for object detection’, 2016*), for each ground truth bounding box. If the IOU was greater than a predefined threshold value, 0.5 was chosen to account for the variance between my definition of a frontal face and what was defined when the detector was trained, then a frontal face was successfully detected. A sample of these images, with green bounding boxes containing detected frontal faces and red bounding boxes containing ground truths, as well as a table containing the true positive rate (TPR) and F1 score, using code from (*How the Compute Accuracy For Object Detection tool works—ArcGIS Pro | Documentation*) for all example images given can be seen below. Where there were no ground truths present for an image the TPR and F1 score were set to 0.



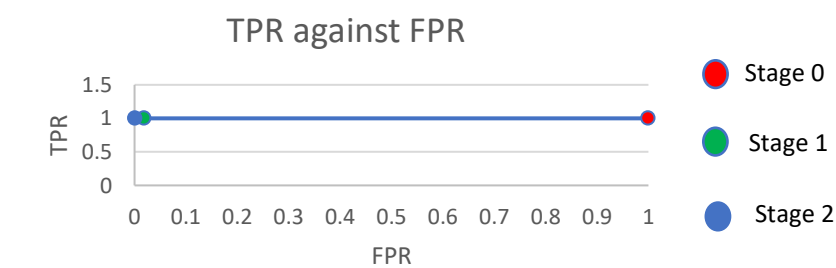
TPR and F1 Scores for example images 1 to 15

Image Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TPR	0.00	1.00	1.00	0.00	0.60	1.00	0.00	0.50	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
F1 Score	0.00	0.20	0.25	0.00	0.17	0.33	0.00	0.22	0.00	0.00	0.00	0.36	0.00	0.00	0.00	0.00

The TPR values clearly show that the detector can successfully identify a majority of all frontal faces in almost every example image given, no frontal faces were detected in NoEntry12.bmp where one was present, however the TPR does not have any information on how much of a valid frontal face was actually detected; therefore, making a detector’s level of accuracy appear to be lower than it could be argued to be. This can clearly be seen with the bounding boxes generated for image NoEntry7 as all predefined frontal faces are contained in green bounding boxes, however only one of them is deemed valid. This is due to the threshold value used to calculate the IOU between bounding boxes being too high when determining whether a face has been detected. This resulted in a green bounding box being treated as a false negative, even though it surrounded a red bounding box. In order to prevent this, the threshold value can be set to 0 and a green bounding box can be created around the borders of the image. This will result in all ground truths being detected, if duplicate detections are ignored, since an IOU value greater than zero will be generated for each ground truth. This means that it is possible to achieve a TPR of 100% for any image detection task as long as there is a ground truth present, which can be guaranteed by only giving the detector images containing at least 1 ground truth. However, this would seriously impact the accuracy and suitability of the detector since even though it could detect frontal faces, it would not be able to determine their position in the image since as the ratio of bounding box sizes would be completely ignored. Therefore, the largest green bounding box in NoEntry7 was treated as an invalid detection, since it does not show the position of the red bounding box to a high enough level of accuracy.

Furthermore, since the TPR does not contain any information about how many false positives were generated, which can make a detector appear more accurate than it actually is, if only the TPR is used to assess the suitability of a detector. This can be seen with the image NoEntry2 as even though it has a TPR of 1, many false positives were generated across the image. Therefore, it is not always practical to assess a detection system based on TPR values alone, since it does not contain enough useful information to give a meaningful and accurate result. This is why the F1 score was also generated for each image, as it provides much greater detail of the accuracy of the detector by including the number of false positives and false negatives generated.

In order to further test the performance of the Viola Jones detector, the detector proved was changed to detect no entry signs instead of frontal faces. This was done by creating a strong classifier trained using a set of no entry signs generated from the image provided. The classifier was built in stages, and the TPR and false positive rate(FPR) was calculated for each stage. A graph showing these values can be seen below:



The graph shows that the classifier does not improve on detecting valid no entry signs as it is built, as all points have a TPR of one, However the graph does show a dramatic decrease of FPR values between stages 0 1. This means that the features added to the classifier in stage one greatly increased its ability to not detect a no entry sign where one was not present. This increase was very small between stages 1 and 2, which shows that the features added in stage 2 did not have a great impact on the FPR rate of the classifier.

The detector used previously was then given this classifier and tested on the same example images. A sample of result images with bounding boxes as well as a table containing TPR and F1 scores for all example images can be seen below:



TPR and F1 Scores for example images 1 to 15

Image Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TPR	0.50	1.00	1.00	1.00	1.00	0.40	0.00	1.00	1.00	0.50	0.66	0.00	0.29	0.00	0.00	1.00
F1	0.06	0.10	0.06	0.29	0.11	0.21	0	0.01	0.89	0.17	0.36	0.00	0.17	0.00	0.00	0.80

The data shows that the detector can also successfully identify a majority of all no entry signs for all images, but it is less successful at doing so than when it is detecting frontal faces, when there are ground truths present, as there are more TPR values in the range $0 < \text{TPR} < 1$. This is likely due to the different set of Haar Features used for detecting no entry signs as only one is really necessary: a white horizontal rectangle sandwiched between two black horizontal rectangles. An increase in images with no valid detections, that still included ground truths occurred due to this, which can be seen in NoEntry6. This is because the central white rectangle in each no entry sign is either slanted at a large angle or has a similar level of brightness to the red of the sign, and therefore would not match the Haar Feature used for detection. Furthermore, the simpler set of Haar Features used led to a far greater increase in false positives, which can be seen when comparing the bounding box images for NoEntry4, as the image is filled with many more regions that match the Haar Feature used; for example an outwards crease in a shirt or the sole or a group of differently coloured paving stones. Overall, the F1 scores show that the detector can be very accurate, as seen with NoEntry8, but only when the input image is less cluttered and the signs aren't slanted by too great of an angle. When this is not the case, the F1 scores are lower when compared to the F1 scores of the frontal face detector, as the frontal face detector used a more complex set of Haar features to detect a frontal face. This issue can be resolved by combining the Viola Jones detector with a Hough Transform, as a no entry sign is composed of a clearly defined circle or ellipse enclosing two parallel lines

References

How the Compute Accuracy For Object Detection tool works—ArcGIS Pro | Documentation (no date). Available at: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/how-compute-accuracy-for-object-detection-works.htm> (Accessed: 8 December 2021).

‘Intersection over Union (IoU) for object detection’ (2016) *PyImageSearch*, 7 November. Available at: <https://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/> (Accessed: 8 December 2021).