1 Face detection

Figure 1: Viola-Jones face detection (green) with ground truth (red)

Frontal face detection results		
Image name	TPR	F1-SCORE
dart4	1	1
dart5	1	0.88
dart13	1	0.666667
dart14	1	0.5
dart15	1	0

Table 1: Viola-Jones face detection results

The results in Table 1 show the true positivity rate, TPR, and the F1 score of the Viola-Jones face detector on the input images shown in 1. TPR is the fraction of true faces which have been detected and F1 score is given by

$$2 \times \frac{precision \times recall}{precision + recall}$$

where precision is the number of true positives divided by the total number of detected objects and recall is the number of true positives divided by the true number of objects in the image.
The true values used to calculate the above are determined manually for each input image. This can results in some ambiguity as a face does not have a strictly defined bounding region which can result in different "true" values.
The effect of this variation in true faces is mitigated by the use of intersection over union, IOU, to determine if a true value has successfully been detected. As provided by [1] the IOU is calculated by determining

$$\frac{\text{Area of Overlap}}{\text{Area of Union}}$$

for a true value and a given region classified by the detector. In this paper for each true value we use the maximum IOU between that true value and all of the detected values. A threshold is then definied and if the IOU is greater than the required threadhold the true value has been detected. This threshold is what allows for slight variations in the true values.
Similarly the Viola-Jones detector was designed as a frontal face detector meaning a side on face would not be considered a true value. The example in Figure 1 shows that the definition of a side on face can also be ambiguous and therefore for in this paper the definition requires both eyes to be fully visible. The consequence of TPR's ambiguity is evident in the previous figure as the detector did classify one of the side on faces and therefore its F1 would have been was reduced as result of this not being considered a true value.
All the results in Table 1 have a TPR value of 1. This means every valid face in the input images were detected. The Viola-Jones method however can often have a very hight TPR. This is as a result of "cascade" [2] implementation. When a region of the image is evaulated it is repeatedly passed through classifiers in an overall cascade until a classifier rejects the region or it passes thorugh all the classifiers and is therefore on the the detected regions. Each classifier has "very high detection rates" [2]. This means if a classifier does not have many stages then the overall cascade will also have a very high TPR rate by not rejecting many regions. However this will come at the cost of many false positives. In order for a classifier to have a hight F1 score it will need a balance between high TPR (recall) but also need a high precision and thus a low number of false positives.

2 Dart board detection

Having tested the Viola-Jones detector on faces a new cascade was trained. An image of a dart board was used to generate a set of postiive images along side a set of negative images. The detector was then trained on these images and it results are shown below.

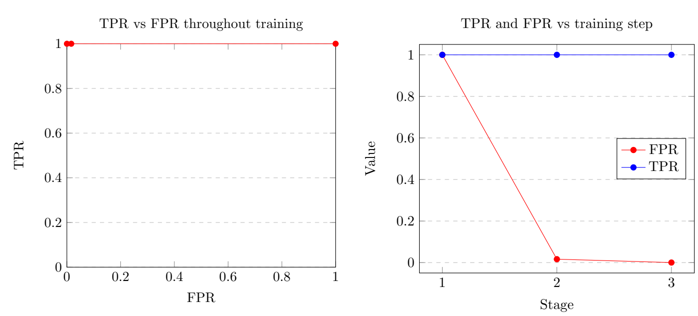


Figure 2: Viola-Jones training results

Figure 2 shows the change in FPR and TPR throughout the training steps. At the beginning of the training process all regions of the image are accepted. As the training goes on more classifiers are added to the cascade which have the oportunity to reject sections. This results in the sharp reduction in the false positivity rates, FPR, seen in 2 as the number of negative regions which are rejected increases. The following stage then shows another reduction in FPR for the same reason however the reduction is far smaller as this classifier now applies to a far smaller set of images (those which are rejected by the first classifier are not passed on to any other classifiers) and it is performing a more "difficult task" [2] as the previous classifiers will have filtered out many of the neagative values.

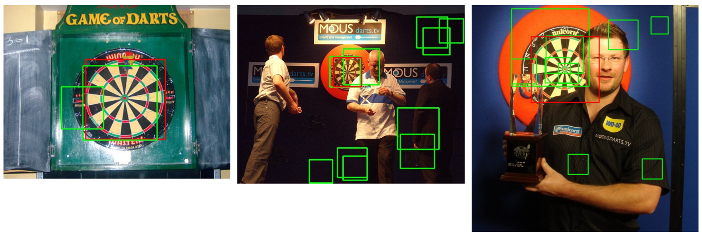


Figure 3: Viola-Jones dartboard detection with true values

Frontal face detection results		
Image name	TPR	F1-SCORE
dart1	1	0.667
dart2	1	0.25
dart3	1	0.4
dart4	0	0
dart5	0	0
dart6	1	0.182
dart7	0	0
dart8	0	0
dart9	0	0
dart10	0	0
dart11	0	0
dart12	0	0
dart13	0	0
dart14	0	0
dart15	1	0.5
Average	0.333333	0.133232

Table 2: Viola-Jones face detection (green) with ground truth (red)

The Viola-Jones algorithm as described in [2] uses very simple feature (2, 3 or 4 rectange features). This allows the Viola-Jones algorithm to be very efficient at detecting certain features in objects such as faces but can be difficult to scale up to more complex objects such as dartboards. As seen in 3 the classifier often finds a subsection of the true dartboards, this confirms the assumption that the Viola-Jones algorithm is not suited to detecting a full dart board but is effective at detecting sections of repeating dartboard pattern.
The results in Table 2 also show that the detector has many false positives through its low F1 score. This suggests that the dartboard features are also being detected elsewhere in the image and show that additional steps will be required to filter the Viola-Jones results in order to improve the detector.

3 Combining Viola-Jones with hough circle detection

From the above section it is clear that the Viola Jones detector is not well suited to detecting full dart boards. Although it is effective at detection the dart boards repeating pattern it is rarely detecting the full circular shape of the dart board and has many false positives. In order to overcome these limitation the results from the Viola-Jones detector can be combined with a circle hough transform. This circle hough transform can then be used to shrink the results set and increase the F1 score of the detector. Moreover the bounding region from the circle hough transform can be used as the resulting region as it is well suited to detecting the full dart board. The combination step is as follows:

- Run Viola-Jones and hough circles.
- For every Viola-Jones results find the maximum IOU with the circles from hough.
- If the IOU is greater than a specified threshold (0.25 in this case) add the item with the largest area to the set of final results.

The use of IOU with hough resulted in a large reduction in the set of results. For the required IOU a relatively low threshold was chosen (0.25) as often the Viola-Jones detector would find a sub section of the dart board (matching the expected pattern) but not the circle. For this reason the final result used the maximum of the bounding boxes from the two detectors rather than the Viola-Jones detector as this was more likely to encompass the full dartboard and not just a sub section or inner circle. This removed many of the original Viola-Jones false positives and therefore increased the F1-score dramatically. However the limitation of this approach are clear in **??** which shows the circle detector failing to detect circles with large angles to the camera or with objects blocking sections of the circle. This can result in a reduction in the TPR rate of the overall detector in exchange for the reduction in the number of false positives.

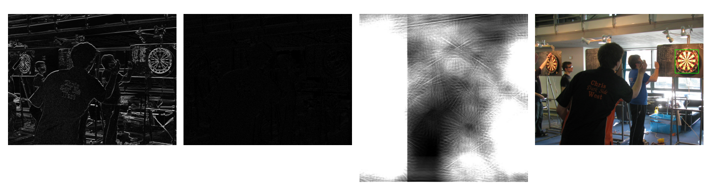


Figure 4: Hough circle detection on dart8.jpg

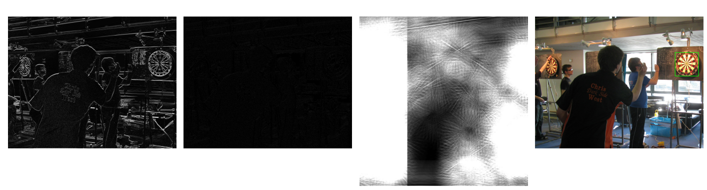
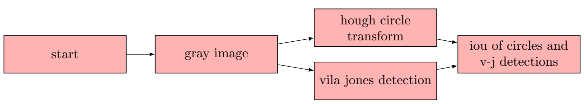


Figure 5: Hough circle detection on dart8.jpg

Frontal face detection results				
Image name	TPR	F1	Δ TPR	Δ F1
dart1	1	1	0	+ 0.333
dart2	1	0.5	0	+0.25
dart3	1	0.5	0	+0.1
dart4	0	0	0	0
dart5	1	1	0	0
dart6	0	0	-1	-0.182
dart7	1	0.5	+1	+0.5
dart8	0.5	0.667	+0.5	+0.667
dart9	0.5	0.4	+0.5	+0.4
dart10	0	0	0	0
dart11	0	0	0	0
dart12	0	0	0	0
dart13	1	0.4	+1	+0.4
dart14	1	0.5	+1	+0.5
dart15	1	1	+1	+1
Average	0.333333	0.133232	+0.289	+0.298

Table 3: Viola-Jones plus hough circles detection with difference in F1 score and TPR between results in Table 2

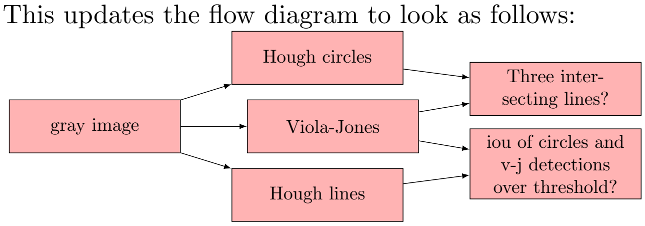


4 Extending the Viola-Jones and Hough Circles detector with Hough Lines

From the previous results it is clear that the hough circles implementation was not effective when there was object blocking the dartboard or the dartboard was at an angle meaing the shape of the dartboard in the gradient image was an ellipse. In order to improve the detector an alternative method would need to be added which could detect dartboards in these specific conditions. The porposed improvement is:

- Run the Viola-Jones detector and Hough Circles and filter as before
- Run Hough lines
- For the detected regions from Viola-Jones which are rejected by the comparison with hough circles check if there are 3 or more lines which intersect inside this region. If so accept the region.

This updates the flow diagram to look as follows:



Frontal face detection results				
Image name	TPR	F1	Δ TPR	Δ F1
dart1	1	1	0	0
dart2	1	0.667	0	+0.167
dart3	1	0.667	0	+0.167
dart4	1	0.286	+1	+0.286
dart5	1	0.5	0	-0.5
dart6	1	1	+1	+1
dart7	1	0.5	0	0
dart8	0.5	0.667	0	0
dart9	0.5	0.286	0	-0.114
dart10	0.333	0.182	+0.333	+0.182
dart11	0	0	0	0
dart12	0	0	0	0
dart13	1	0.4	+1	+0.4
dart14	1	0.5	+1	+0.5
dart15	1	1	+1	+1
Average	0.333333	0.133232	+0.289	+0.298

Table 4: Viola-Jones plus hough circles detection with difference in F1 score and TPR between results in Table 2

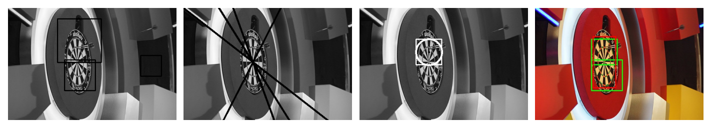


Figure 6: Example out of detector using hough lines

The images in Figure 6 show the main limitation of this approach. In many cases the output from Viola-Jones detects subsection of the dartboard. When using hough circles this was overcome by using the output of the hough circle instead of the original Viola-Jones region. However when incorporating Hough lines, there is no mechanism to find the center of the region. 2 potential solutions to this short coming would be:

4.1 Probabilistic Hough Line Transform

Using a probabilistic Hough line transform, instead of the generic Hough line transform would also return the end points of the detected lines. These line could then both be used to check for the intersecting lines found on the dartboard as well as determine the size and shape of the dartboard.

4.2 Hough Ellipse Transform

Alternatively an Hough ellipse transform could be used. This would allow the detector to find a greater number of the dartboards and an angle and also indicate the true shape of the board in the image. However a hough ellipse requires a 5D Hough space which can be expensive to create, store and use.

References

[1] Adrian Rosebrock. Intersection over union (iou) for object detection. 2016.

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