The Dartboard Challenge

Cheong Ren Hann & Navya Zaveri

1 Viola-Jones Face Detector

1.1 Visualisation

The following are five images processed using the provided Viola-Jones object detector for human faces. The green boxes below denotes candidate faces given by the detector while the red boxes denotes the manually annotated ground truth.

(a) dart4.jpg (b) dart5.jpg (c) dart13.jpg (d) dart14.jpg (e) dart15.jpg

Figure 1: Predicted faces with ground truth

1.2 Evaluation

To evaluate the detector, we must define (i) what a valid frontal face is and (ii) how well must a bounding box fit the ground truth. For (i), we decided that to qualify a valid 'face' it must have both of the eyes and mouth fully visible (hence no faces in Figure 1e). Next, for (ii) we used a Intersection Over Union (IOU) threshold of 0.3 between the detected bounding boxes and the ground truth to determine if the faces were successfully detected.

We also noted the case where multiple bounding boxes fit the criteria of a true positive (TP) on the same ground truth. In this case, we only count the first such TP while the repeated boxes are ignored (i.e neither TP or false positive (FP)). The rationale behind this is that Viola-Jones trained classifiers tend to return multiple overlapping boxes around the ground truth due to insensitivity over minor changes in translation and scale [1].

This allows us to calculate the True Positive Rate (TPR) and F1-score of the test images as listed in Table 1. With the semantics defined above, we appreciated that the F1-test would evaluate a classification with multiple bounding-boxes that fit the criteria to the *same* score as a perfect classification with one bounding box. We also considered the case of images which do not contain any valid faces in the ground truth (marked with *). As the TPR and F1 would otherwise be undefined, we returned 100% for the TPR and F1 if the detector correctly returned no candidate face and 0% if the detector gave false positives.

Image	**TPR**	**F1**
dart0*	0	0
dart1*	1	1
dart2*	1	1
dart3*	1	1
dart4	1	1
dart5	1	0.88
dart6	0	0
dart7	1	1
dart8*	0	0
dart9	1	0.4
dart10*	0	0
dart11	1	1
dart12*	1	1
dart13	1	0.67
dart14	1	0.5
dart15*	0	0

Table 1: Results of Viola-Jones face detector

We note that the TPR depends on the IOU threshold as it directly affects what is classified as a TP. This is evident in Figure 1d in particular as there is a large bounding box that covers the ground truth but is not sufficiently precise enough to classify it as a TP.

Furthermore, we can always achieve a 100% TPR simply by returning all possible bounding boxes at every pixel and scale in the image. However, this would result in a very high number of FPs and so this approach would not be considered an effective detector. Hence it is difficult to evaluate the detector based on only the TPR as a single measure. However, we can also use the F1 score in addition to the TPR, which gives a better indication of the detector as it also takes into account the rate of FPs. This is evident in Figure 1c as although the TPR is 100%, the F1 is 67% as there is a FP detected.

2 Performance

2.1 Training

Figure 2 shows the plot of the True Positive Rate (TPR) vs the False Positive Rate (FPR) over each stage of the training of the classifier. We can see that both the TPR and FPR are at 1 in the first stage, while continuing through the stages the TPR stays at 1 while the FPR decreases in each stage with a significant drop between stages 1 and 2.

This is characteristic of how the Viola-Jones AdaBoost training works: in the first stage, the Haar features chosen are simple enough that they will match many regions in the image. This explains the initial high TPR and high FPR. In the proceeding stages, more Haar features are selectively added. Features that reduce the TPR are instantly rejected, whilst features that increase the True Negative Rate are accepted into the feature set. This results in a decreased FPR whilst maintaining a high TPR, as evident in the plot.

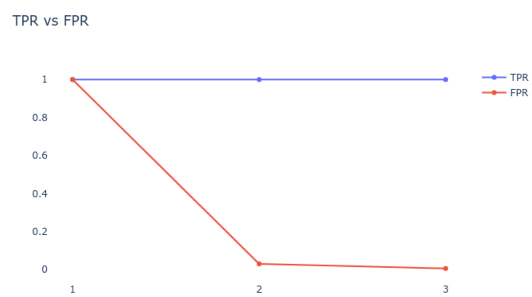


Figure 2: TPR vs FPR in each training stage

2.2 Testing

(a) dart0.jpg (b) dart10.jpg (c) dart11.jpg

Figure 3: Predicted dartboard with ground truth

Table 2 shows the results of the trained cascade classifier. Comparing with Figure 2 above, we can see that the TPR is 100% in both cases.

This is not necessarily the case in general as the TPR of training is usually higher than that of testing. As in the former case the same data used to fit the classifier is used to assess it whilst in the latter the test data is unseen so this is a measure of how well the classifier generalises. In this case, the TPR is high while the F1 is low which suggests that there is a high number of FPs from the detector. However, we can improve on the F1 score in the later sections using Hough detection methods in a manner similar to a cascade classifier.

We noticed that trained classifiers typically have lower TPR than suggested in Figure 2. This is likely due to the training is done on samples from a single image which the Harr-like features learnt may not necessarily generalise to the larger variety of test samples. In general, having a high training TPR may suggest over-fitting especially if the testing error is considerably low. Also, the fact that the training was done with outdoor scenes as negative examples is not ideal considering that our test image are indoors. In a few cases, there were partial occlusion of the dartboard (e.g. Figure 3c) or the dartboards were small and hard to see (e.g. Figure 3b) which may further lower the detection rate. Although our detector managed to identify these points, the bounding boxes drawn were not as precise.

Image	**TPR**	**F1**
dart0	1	0.18
dart1	1	0.4
dart2	1	0.18
dart3	1	0.25
dart4	1	0.2
dart5	1	0.2
dart6	1	0.33
dart7	1	0.11
dart8	1	0.24
dart9	1	0.22
dart10	1	0.18
dart11	1	0.8
dart12	1	1
dart13	1	0.25
dart14	1	0.1
dart15	1	0.5
Average	1	0.32

Table 2: Results of trained Viola-Jones dartboard detector

3 Hough Transform

As discussed in the previous section, the Viola-Jones trained detector is far from perfect. We needed a more precise technique to confirm the presence of a dartboard. To that end, we decided to exploit hough transforms to discern if a given test image has the characteristic shapes of a dartboard, such as concentric circles and intersecting lines.

3.1 Detection Pipeline

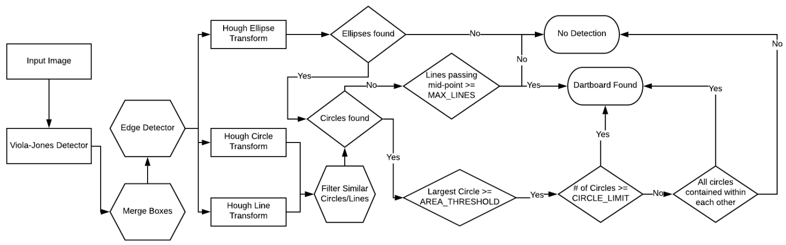


Figure 4: Flow diagram of the detection pipeline

The above pipeline is a strict-layered system that consists of a Viola Jones pass to capture all potential detections, followed by a hough transform for each detection, concluded by a greedy heuristic that confirms the presence of a dartboard based on the circles and lines found on the detection patch.

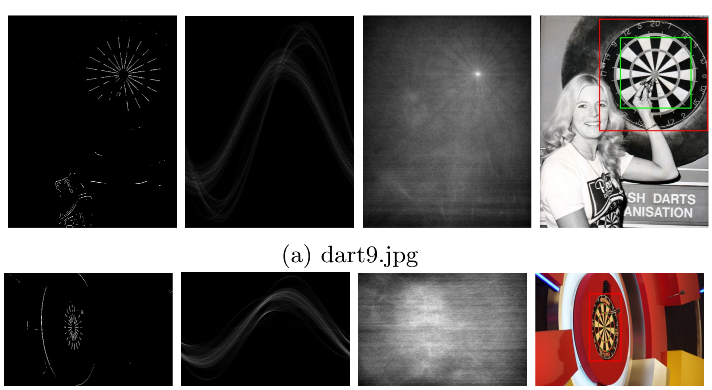
3.2 Evaluation

The merits of the classifier are the following:

- Low False Positives: The pipeline selectively examines patches (from the Viola-Jones pass) that are likely to contain dartboards. This makes it easier to filter out non-dartboard images.
- Efficient: Because the pipeline will only run on each detection patch, it is feasible to run computationally intensive algorithms (such as hough transforms for circles). This also means we have faster iteration cycles. In Figure 5a, there were only few candidates boxes to process and the FPs were quickly eliminated.

Some of the drawbacks are the following:

- Missed Detection: If the Viola-Jones classifier doesn't capture a region containing a ground truth, the pipeline will not be able to detect it. This is evident in Figure 5b as the detection box given was particularly small which is not sufficient to capture dartboard features.
- Overlapping Detections: Sometimes the Viola-Jones classifier returns overlapping bounding boxes that cover the same ground truth. In such cases, it is hard to detach such overlapping detections using hough transforms alone, which can result in over-counting.



(a) dart9.jpg

(b) dart12.jpg

Figure 5: The thresholded gradient magnitude, 2D representation of the Hough Spaces (line and circle) and the final detection results

Image	TPR	ΔTPR	F1	ΔF1
dart0	1	+ 0	0.67	+ 0.49
dart1	1	+ 0	1	+ 0.6
dart2	1	+ 0	1	+ 0.82
dart3	1	+ 0	1	+ 0.75
dart4	1	+ 0	1	+ 0.8
dart5	1	+ 0	1	+ 0.8
dart6	1	+ 0	1	+ 0.67
dart7	0	- 1	0	- 0.11
dart8	0.5	- 0.5	0.5	+ 0.26
dart9	1	+ 0	1	+ 0.78
dart10	0.67	- 0.33	0.8	+ 0.62
dart11	0.5	- 0.5	0.67	- 0.13
dart12	0	- 1	0	- 1
dart13	1	+ 0	0.5	+ 0.25
dart14	1	+ 0	0.8	+ 0.7
dart15	1	+ 0	1	+ 0.5
Average	0.79	- 0.21	0.75	+ 0.43

Table 3: Results of the dartboard detector pipeline

4 Extension

4.1 Idea

4.1.1 Pre-processing Merge

As mentioned earlier, the output of the Viola-Jones classifier tends to return many overlapping boxes [1]. This increases the processing time in subsequent steps as we have to process many similar regions. We can mitigate this by merging similar boxes as suggested in [1], which we implemented using a fixed IOU threshold to determine similarity. Then we can partition the boxes and return a single box per partition with the new parameters being the average of the parameters of all boxes in the partition.

4.1.2 Canny Edge Detector

Results from the Hough transforms are highly dependent on the input of the thresholded magnitude image. Using Sobel we noticed that the magnitude images tend to contain many noise-like isolated edges. Canny, on the other hand, is not very susceptible to such noise [2] as it able to reduces the number of "false" edges by thinning edges from Sobel and by applying Hysteresis thresholding to preserve edge continuity. This results in a more defined magnitude image which allows us to better identify shapes using Hough in the next stage.

4.1.3 Ellipse Detector

Another idea we attempted was Hough Ellipse detection methods. The goal was to be able to find dartboards that were tilted at an angle, which the circle detection would not likely identify. A key problem was the complexity of $O(n^5)$ of the Hough space as there are 5 parameters of an ellipse to keep track of. We implemented a more efficient $O(n^3)$ algorithm, [3] but this still proved to be computationally intractable. To optimise this further, we attempted to use a Randomized Hough transform approach [4] to reduce the complexity to $O(kn^2)$ with $k = 20$. While we were able to identify ellipses on well-defined magnitudes such as from Canny, this approach detects also detects many false positives. As such, we were only able to make limited use of it in the final pipeline.

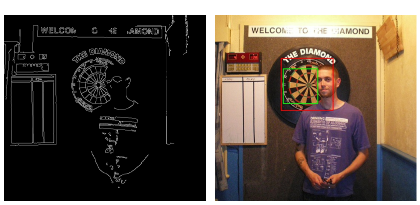
4.2 Evaluation

The merits of the extended classifier include:

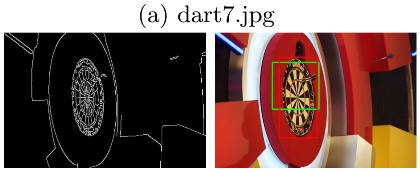
- Feature Preservation: Due to the Hysteresis thresholding employed by the Canny Edge detection, we can identify edges that are more consistent and continuous which allows us to identify shapes and patterns more easily. In both Figure 6a and 6b we can see both dartboards patterns clearly even with partial occlusion and rotation.
- Less Noise: In addition to the merge which helps to consolidates candidate boxes, a key advantage of Canny is that it tends to reduce the number of isolated edges as well as provide sharper edges compared to Sobel. This allows for a more precise Hough Transform for detection.

However, some of the drawbacks are:

- Excessive reduction of search space: While the merge simplifies our pipeline and removes duplicate candidates, there is a risk that the merged box may not fit the ground truth as well as before. This could lead to missing TPs due to the boxes being realigned by the merge.



(a) dart7.jpg



(b) dart12.jpg

Figure 6: The gradient magnitude using Canny Edge Detector and the final detection results

Image	**TPR**	**ΔTPR**	**F1**	**ΔF1**
dart0	1	+ 0	1	+ 0.33
dart1	1	+ 0	1	+ 0
dart2	1	+ 0	1	+ 0
dart3	1	+ 0	1	+ 0
dart4	1	+ 0	1	+ 0
dart5	1	+ 0	1	+ 0
dart6	1	+ 0	1	+ 0
dart7	1	+ 1	1	+ 1
dart8	0.5	+ 0	0.67	+ 0.17
dart9	1	+ 0	1	+ 0
dart10	0.67	+ 0	0.8	+ 0
dart11	0	- 0.5	0	- 0.67
dart12	1	+ 1	1	+ 1
dart13	1	+ 0	1	+ 0.5
dart14	1	+ 0	1	+ 0.2
dart15	1	+ 0	1	+ 0
Average	0.89	+ 0.1	0.90	+ 0.15

Table 4: Results of extended dartboard detector pipeline

References

[1] Paul Viola and Michael Jones. Robust real-time object detection. volume 57, 01 2001.

[2] Radhika Chandwadkar. Comparison of edge detection techniques. 08 2013.

[3] Yonghong Xie Qiang and Qiang Ji. A new efficient ellipse detection method. In *in International Conference on Pattern Recognition 2002*, page 20957, 2002.

[4] Samuel A. Inverso. Ellipse detection using randomized hough transform. 2002.