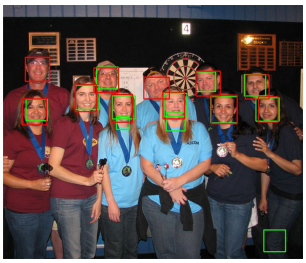
Dylan Holden-Sim DH17830 & Alexander Carolan AC17588

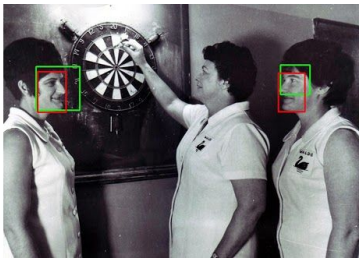
The Dartboard Challenge

1. Viola Jones Face Detector

1.1 Visualisation

Detections using a scale factor 1.1, minimum neighbour 3 and an IOU threshold 0.3 for images [4, 5, 13, 14, 15].



1.2 Evaluation

Image	4	5	13	14	15	Mean
TPR	1.00	0.81	1.00	1.00	0.50	**0.86**
F1	0.67	0.78	0.67	0.57	0.50	**0.64**

In assessing the TPR we must first define what should be classed as a face in order to generate the ground truth values, which is an inherently subjective decision. Ultimately we decided that the face should be rotated through an angle of less than 90 degrees, thus making more than half the face visible to the camera. This is conveniently expressed in image 15 where the centremost face is excluded but the leftmost and rightmost faces are included.

Furthermore we must define an IOU threshold for deciding which detections should be considered a true detection, another subjective decision. A detection could be defined by capturing the centre of a face, capturing all the essential features of a face or any number of possible definitions. After some testing we found that an IOU threshold of 0.3 gave an ideal balance between overdetection and underdetection. This too is expressed in image 15 where the rightmost face is not considered a detection due to excluding the mouth and chin.

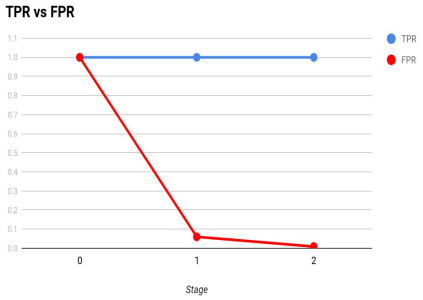
On its own TPR isn't a perfect metric, for example, it is always possible to achieve a TPR of 100% by simply classifying every possible subsection of the image as a detection and therefore never producing any false negatives.

Contributions: Dylan 1.0, Alex 1.0
Signed:

2. Viola Jones Dartboard Detector

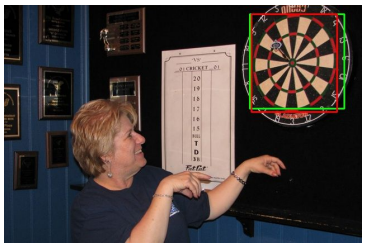
2.1 Training

The production of a cascaded classifier involves combining strong classifiers in stages. As false negative classifications of previous stages exclude them from the following stages, the initial stage is designed to have a high TPR by allowing a high FPR, in order to not limit the overall TPR of the cascaded classifier. Stage 0 utilizes only 2 Haar features in order to execute execute quickly and exclude most true negatives, and the following stages use progressively more Haar features in order to exclude more true negatives. This process is reflected in the trend in FPR from 1 to 0 in the accompanying graph.



2.2 Visualisation

Detections using a scale factor 1.1, minimum neighbour 8 and an IOU threshold 0.3 for Images [0, 1, 5].



2.3 Evaluation

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean
TPR	1.00	1.00	1.00	1.00	0.00	1.00	0.00	0.00	0.50	1.00	0.34	0.00	0.00	0.00	1.00	1.00	**0.55**
F1	1.00	0.67	0.67	0.34	0.00	0.67	0.00	0.00	0.25	1.00	0.20	0.00	0.00	0.00	0.34	1.00	**0.38**

The mean TPR during testing was significantly lower than during training, this is however to be expected as the model is trained on a different data set to the one that it is tested on, allowing for discrepancies between the two such as the variation in the angle of the object, the lighting of the object including exposure, colour, shadows etc. and the optical properties of the camera such as focal length.

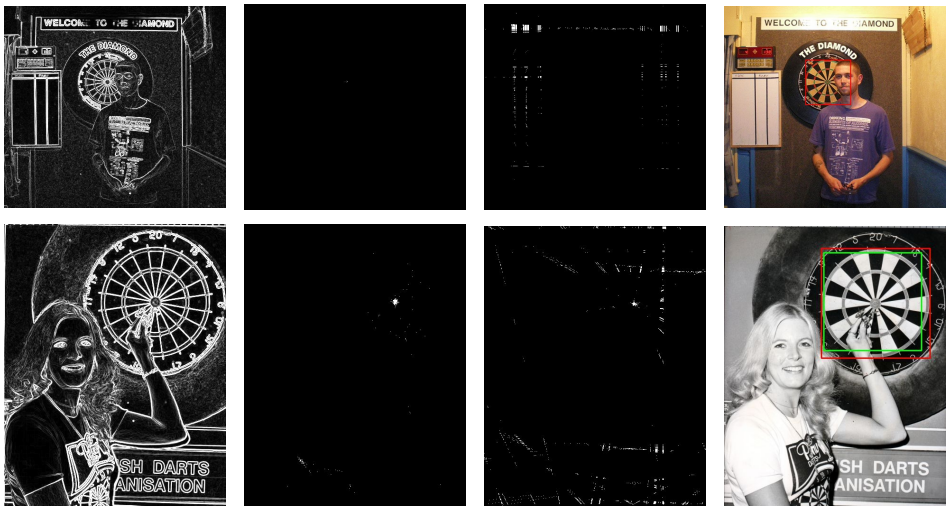
The mean TPR and F1 scores for the dartboard detector are also significantly lower than the face detector, by 0.21 and 0.26 respectively. There are a number of factors that could explain the difference in performance, such as the inherent complexity of the object being detected, the quality in the positive and negative training data sets, the parameters used during training and testing such as the number of stages, window scaling and minimum neighbour requirement.

Contributions: Dylan 1.0, Alex 1.0
Signed:

3. Combining Viola Jones & Hough

3.1 Visualisation

Gradient magnitudes using a threshold 200, circle centres using a threshold of 8, line intersections using a threshold of 40 and detections using a scale factor 1.1, minimum neighbour 8 and an IOU threshold 0.3 for images [7, 9].



3.2 Evaluation

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	Diff
TPR	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	0.50	1.00	0.33	0.00	0.00	0.00	1.00	1.00	**0.49**	**-0.06**
F1	1.00	0.67	1.00	0.00	0.00	1.00	0.00	0.00	0.40	1.00	0.40	0.00	0.00	0.00	0.44	1.00	**0.44**	**+0.06**

Pros:

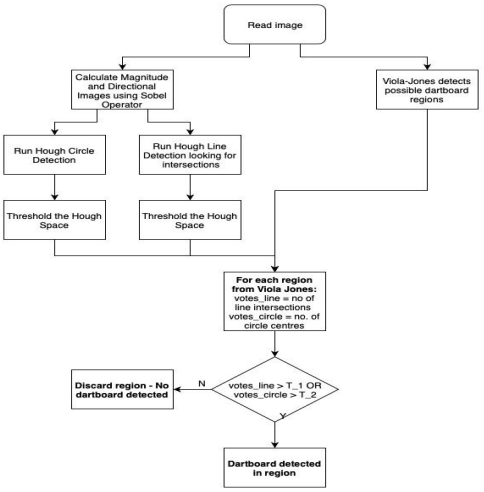
- Increased F1 through removal of false positives with Hough spaces.
- Hough spaces for circles centres is very clean due to small amount of circles in images.

Cons:

- Decreased TPR through accidental removal of true positives with Hough spaces.
- Hough spaces for line intersections is very noisy due to large amount of horizontal and vertical lines in images.

3.3 Pipeline

- The rationale behind our combined detector was that in order to get a high TPR on our combined detector we could afford a high FPR on all of our individual detectors and let their false positives cancel out.
- Our intuition here was that the combination of circle centres and line intersections captures the essence of a dartboard and few other objects would have these features, making it unlikely for a false positive to be simultaneously detected by the Viola Jones, Hough circle centres and Hough lines intersections.



Contributions: Dylan 1.0, Alex 1.0
Signed:

4. Improving Viola Jones & Hough

4.1 Idea

After analysing the detector effectiveness and efficiency we found both to be massively limited by the calculation of line intersections. There was a large variance in the number of lines detected in each image, which not only made it incredibly inefficient but also made it impossible to find an effective threshold which maintained the line intersections on a dartboard but not elsewhere. In order to overcome this we implemented a variety of improvements (speedup calculated for image [8]):

- Stopped checking for horizontal and vertical lines as these aren't unique to dartboards, this gave a speedup of 4.8X.
- Stopped checking for intersections between lines with very similar angles, this gave a speedup of 1.07X (~10% of lines detected were similar).
- Dynamic thresholding where we increment the threshold for lines until we have a specified number of lines remain in the space, this gave a speedup of 4.9X for 150 lines.

4.2 Visualisation

Line intersections using a static threshold of 40 and a dynamic threshold along with the other optimisations for image [7]. Whilst the left image is dominated by intersections from horizontal and vertical lines the right image consists only of diagonal intersections at the dartboard centre.

4.3 Evaluation

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	Diff
TPR	1.00	1.00	1.00	1.00	0.00	1.00	0.00	0.00	0.50	1.00	0.33	0.00	0.00	0.00	1.00	1.00	**0.55**	**0.00**
F1	1.00	1.00	1.00	0.50	0.00	1.00	0.00	0.00	0.33	1.00	0.33	0.00	0.00	0.00	0.57	1.00	**0.51**	**+0.13**

Pros:

- TPR score maintained and F1 greatly improved from Viola Jones
- Average runtime of 95s with standard deviation of 60s compared to a previous average runtime of 2024s with standard deviation of 4980s, this gives a speedup of 21.3X.
- Presence of only diagonal lines removed many undesired line intersections that would otherwise maintain false positives.
- Voting threshold for line intersections is more meaningful across images as the dynamic line threshold results in roughly the same number of line intersections per image.
- Hough line intersections is able to detect dart boards at angles

Cons:

- TPR is limited by the TPR of Viola Jones and despite tuning it to overdetect, the model wasn't trained using particularly wide viewing angles and so produces false negatives, which can't be gained at subsequent stages.
- Dynamic thresholding is weighted towards longer lines.
- Hough circle centres not able to detect dart boards at angles as it is not modified for ellipses.

Contributions: Dylan 1.0, Alex 1.0
Signed: