Image Processing and Computer Vision

Dartboard Challenge

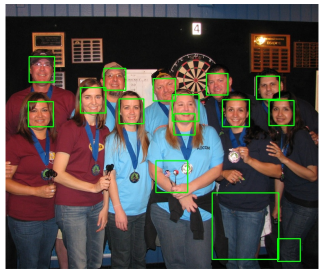
COMS30121 Coursework

Sam Sutherland-Dee, ss15060@my.bristol.ac.uk &
Michael Barlow, mb16329@my.bristol.ac.uk

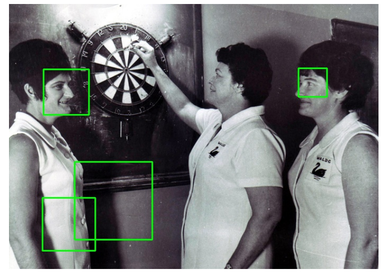
December 3, 2018

1 The Viola-Jones Object Detector

(a) dart4.jpg



(b) dart5.jpg



(c) dart15.jpg

(d) dart13.jpg

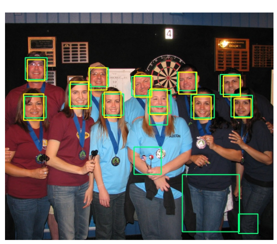
(e) dart14.jpg

Figure 1: Detected face candidates

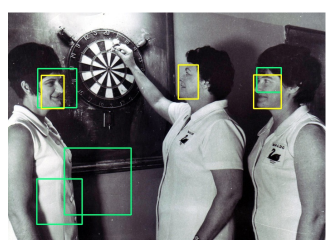
The true positive rates for dart5.jpg and dart15.jpg are 1 and $\frac{1}{3}$ respectively. Practical difficulties in determining the TPR include how to define a face; does it include the chin, ears and forehead? How much hair should be included if it obscures the face? Since there are no quantitative answers to these questions, we implemented our own metric, remaining consistent throughout analysis. We defined a face to be horizontally bounded by sideburns and vertically enclosed by the chin and lowest substantial head hair. A result should be categorised as a true positive if

$$\frac{Ground\ Truth\ Area\ \cap\ Detected\ Area}{Ground\ Truth\ Area\ \cup\ Detected\ Area} > t \tag{1}$$

for some accuracy threshold $0 \leq t \leq 1$. Expressing the true positive rate as an inequality accounts for noise in the image, small inaccuracies in our ground truths and the varying definitions of a face. Working with an accuracy threshold of 0.4, we created a python script implementing this check to determine whether or not each Viola-Jones detection was a true positive. This subsequently allowed us to calculate an F1 score for images dart5.jpg and dart15.jpg, resulting in 0.846 and 0.286 respectively.



(a) dart5.jpg

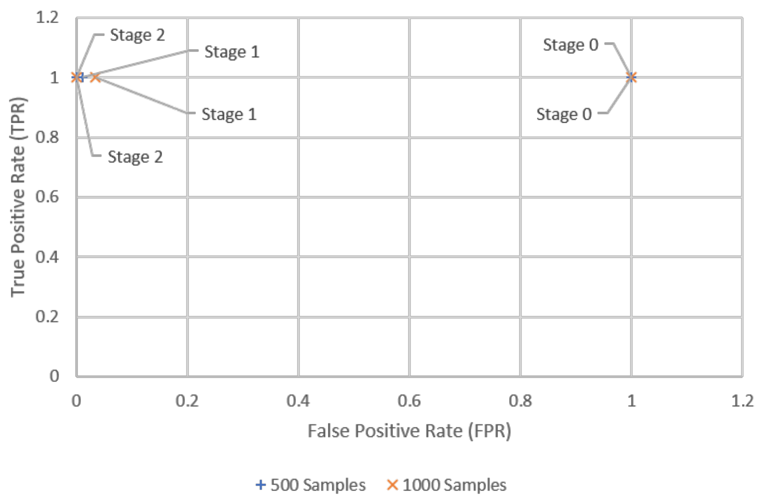


(b) dart15.jpg

Figure 2: Detected faces overlaid with ground truths

A true positive rate of 100% can always be achieved by categorising every pixel area as positive. As the true positive measure is unaffected by false positive values, this model will find every face in an image but the classifications hold no semantic value.

2 Repurposing Viola-Jones for Dartboards



(a) TPR against FPR for each cascade stage

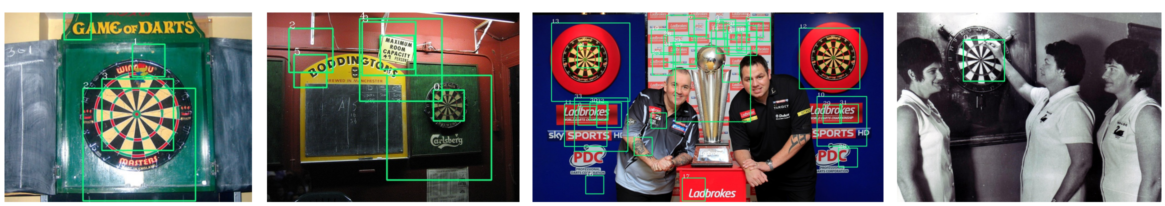
Samples	Stage	TPR	FPR
500	0	1	1
	1	1	0.0108453
	2	1	0.000569413
1000	0	1	1
	1	1	0.033745
	2	1	0.000587049

(b) Raw boosting data

Figure 3: Boosting data

The ROC curve for the Viola-Jones dartboard classifier training is shown in Figure 3a. Throughout boosting, the TPR remains at 1, while the FPR reduces significantly at each training stage. This holds true for both 500 and 1000 sample runs.

After training our classifier, we applied it to each of the provided dartboard images (see figure 4 for four output examples). As can be seen, the results were mixed; some bounding boxes were wrongly sized and positioned while others were very close to the corresponding ground truth.



(a) dart1.jpg (b) dart3.jpg (c) dart14.jpg (d) dart15.jpg

Figure 4: Viola Jones dartboard detector results

Using the same categorisation metric as used for faces with a more lenient threshold of 0.25, we augmented our python script to automatically calculate F1 scores for each of the provided dartboard images. The resulting output is summarised in table 1.

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F1 Score	0.17	0.40	0.22	0.25	0.29	0.25	0	0.14	0.25	0.25	0.20	0	0.67	0.14	0.06	1

Average F1 Score	0.278

Table 1: F1 scores for each dartboard image (Viola Jones)

The average F1 score is fairly poor and performance seems significantly worse than during training. This is likely due to the training process using tiny samples from a single source.

3 Integration with Shape Detectors

Our dartboard detector combines Viola Jones data with both Circle and Line data gathered from two different Hough Spaces. Line data is refined to determine intersection points with a significant angle delta. Figure 5 shows scaled examples of these Hough Spaces, while figure 6a shows the resulting data overlaid on the input image.

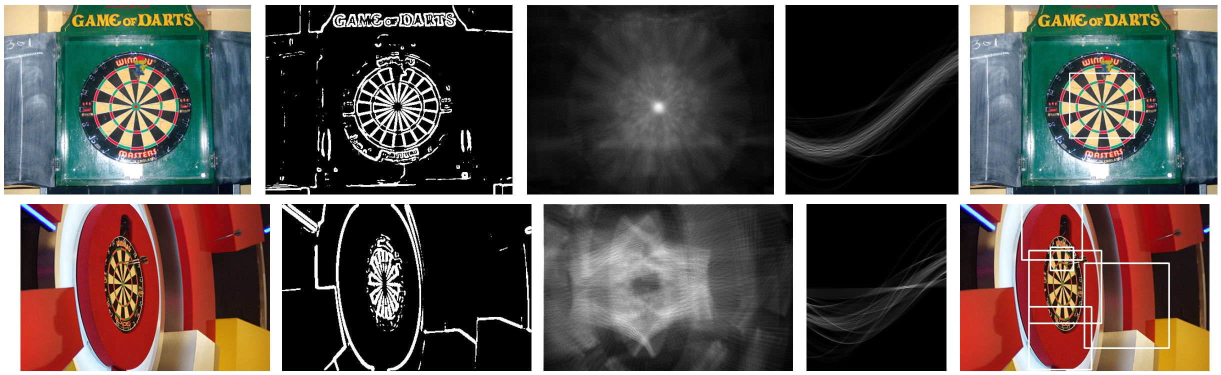
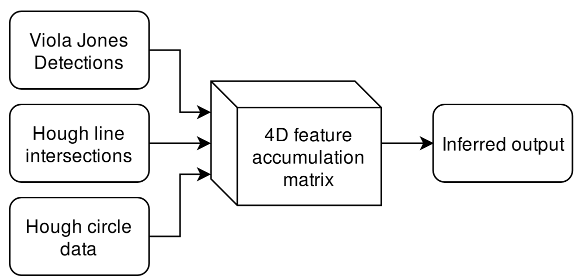


Figure 5: (Left to right) Original Image, Gradient Magnitude, Circle Hough Space, Line Hough Space, detected.jpg

To combine feature data, we use a 4D accumulator matrix representing a width and height space for every (x, y) co-ordinate in the original image. As shown in figure 6b, each detection metric is written to this accumulator with a calibrated weight, after which the matrix is analysed for local maxima which represent our dartboard predictions.

(a) All features overlaid on dart4.jpg



(b) Feature combination flowchart

Figure 6: An example of our features and how they are combined

Combining data in this manner allows for the easy addition of further metrics. The voting system also means resulting bounds do not always come from a single metric so non-circular dartboards are bounded accordingly.

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F1 Score	1	1	1	0	1	1	0.33	1	0.67	1	0.50	0	0.33	1	1	1

Average F1 Score	0.740

Table 2: F1 scores for each dartboard image (Our detector)

Table 2 shows the F1 scores for our detector when applied on all images. As can be seen, our detector offers a significant improvement over the plain Viola Jones detector analysed in table 1. The merits and shortcomings of our detector are summarised below.

Merits

- Good at detecting partly obscured dartboards
- Combines data metrics very well
- Considerably fewer false positives than Viola Jones

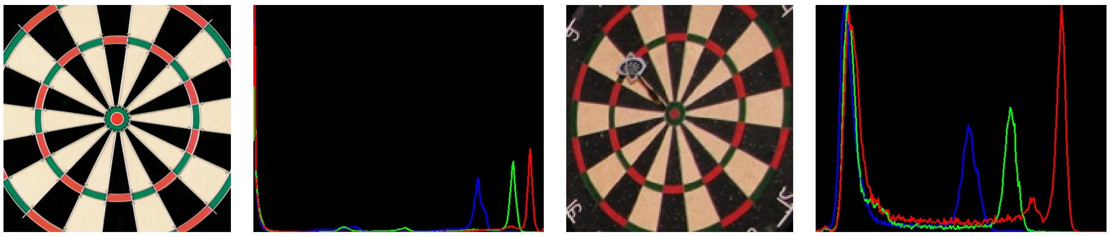
Shortcomings

- No better than Viola Jones at detecting ellipse shaped dartboards
- Very sensitive to weighting and thresholding
- Slow to execute and uses a large amount of memory

4 Further Improvements

To this point, we have only considered the shape of a dartboard and not other features such as colour. Since a dartboard has an equal count of white and black segments, we experimented with a range of colour analysis techniques to determine whether a region contains a dartboard. Due to the high true positive rate of Viola Jones, we used the detected windows as our regions of interest (ROIs).

The first technique which we tried was histogram analysis. We compared RGB histograms (shown in figure 7) using the opencv *compareHist* function, achieving mixed results due to the varying lighting conditions.



(a) dart.bmp (b) Histogram of dart.bmp (c) Viola Jones ROI (d) Histogram of ROI

Figure 7: Comparing the dart.bmp histogram with a Viola Jones region of interest

Since the histogram analysis was highly sensitive to variations in brightness and contrast, we had to employ an alternative solution. We preprocessed each region by first converting it to greyscale and subsequently applying a binary threshold. Figures 8b, 8c, 8d and 8e illustrate this, showing the thresholded regions of interest obtained from the application of Viola Jones on dart1.jpg. We then calculated the ratio of black pixels to white pixels; if this ratio is close to 1, it is a strong indicator that the region contains a dartboard.

(a) Viola Jones on dart1.jpg (b) ROI 1 (c) ROI 2 (d) ROI 3 (e) ROI 4

Figure 8: Obtaining and thresholding the regions of interest from dart1.jpg

To combine this new metric with our current detector, we added an extra weight to Viola Jones regions which had a black/white pixel ratio close to 1.

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F1 Score	1	1	1	0	1	1	0	1	0.67	1	0.50	0	0.67	1	1	1

Average F1 Score	0.740

Table 3: F1 scores for each dartboard image (Our final detector)

Table 3 shows the resulting F1 scores obtained by this version of our detector. While the average score is equal to the previous detector, this improvement leads to fewer false positives. Given a larger data set we are confident that an improvement would be seen.

5 Conclusion

Overall, we are very pleased with our detector and proud of the significant improvement that it offers over plain Viola Jones. Given more time, we would add additional metrics to our detector, further calibrate the thresholds and optimise the code.