COMS30121 – IMAGE PROCESSING AND COMPUTER VISION

THE DARTBOARD CHALLENGE

Introduction:

The task introduces us to the Viola-Jones object detector which provides real time object detection rates. Through this assignment we explore the Viola-Jones object detector and implement further functionalities to improve it.

Below are displayed the annotated test images displaying ground truth.

Figure 1: dart4.jpg

Figure 2: dart5.jpg

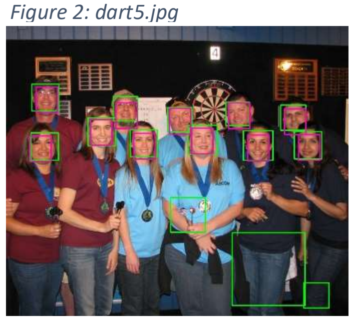
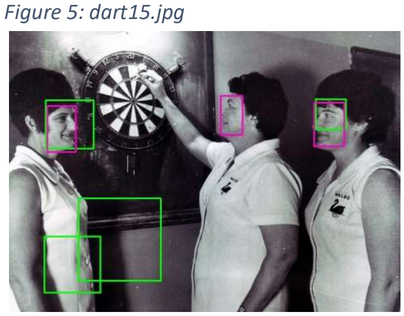


Figure 3: dart13.jpg

Figure 4: dart14.jpg

Figure 5: dart15.jpg



Assessing the Performance of the detector:

Given the two sample images **dart5.jpg** and **dart15.jpg**, we found a true positive rate of 100% and 67% respectively. After using the detector and discussing, we figured out some practical difficulties in assessing the true positive rate accurately. One is that the ground truth boxes must be entered manually, which makes them susceptible to human error. As seen in figure 5, faces turned to profile may not register to the detector. Another difficulty in assessing TPR is that the area detected might be smaller than the actual face and not actually on the face, which would vary on opinion whether it should be included in the true positive results. It is always possible to achieve 100% TPR as the number of possible detections is infinite, and the detector could detect as many detections to achieve that 100% TPR but would end up increasing the false positive rate. A function was implemented to calculate the F1-score of the images. The following table displays the results.

Image	Ground truth	Detected	True positive	Missed	F1 score
dart4.jpg	1	1	1	0	1
dart5.jpg	11	14	11	3	0.88
dart13.jpg	1	2	1	1	0.67
dart14.jpg	2	6	2	4	0.5
dart15.jpg	3	4	2	2	0.5

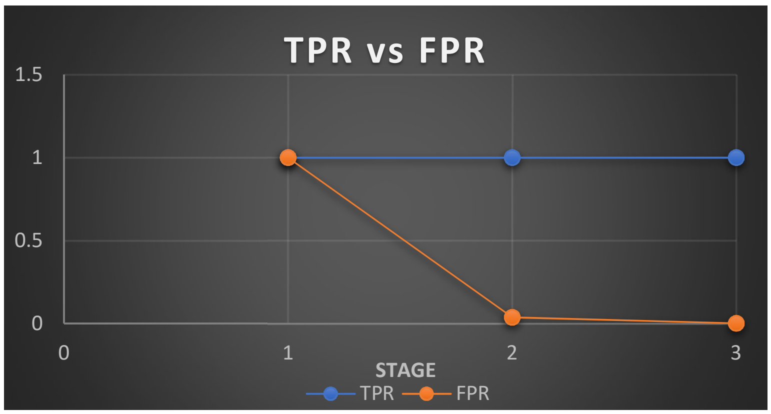
Table 1: F1-Score

DARTBOARD DETECTOR

Interpretation of the TPR vs FPR graph:

To begin, a sample of 500 positive pictures was created. To optimise the detector, we trained it with the 1000 negative pictures provided to us and the 500 positive pictures created. This provided a positive to negative ratio of 500:1000. As we can see from the graph below, the detector begins by detecting all possible dartboards, producing a high false positive rate, and at each stage gradually reduces its number of detections to minimise the false positive rate. By looking at the graph we can clearly see that the detector reduces the FPR at each stage, which means the detector is improving. As the TPR begins and remains at one, we can interpret that the detector detects all dartboards at all stages.

Figure 6: TPR vs FPR graph



Detector's Performance:

Figure 7: dart0.jpg

Figure 8: dart3.jpg

Figure 9: dart7.jpg

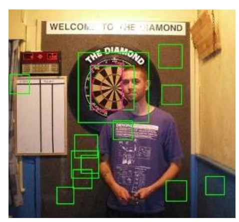
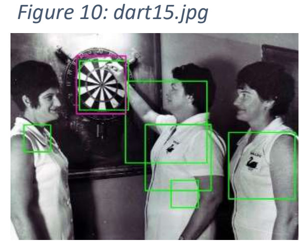


Figure 10: dart15.jpg



Below you will find the F1-scores of all the images:

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Average
F1-score	0.4	0.4	0.2	0.33	0.4	0.18	0.15	0.14	0.125	0.18	0.06	0.4	0	0.25	0.15	0.29	0.23

The F1-score is a measure of the detector's accuracy. By using the precision or recall individually, we often cannot determine superiority between algorithms. The F1-score is a combined metric of the two. It considers the precision and the recall in order to compute a value between 0 and 1, 1 being the best (perfect precision and recall) and 0 being the worst. The average F1-score of the images is 0.23 which is a relatively low score. This is because there are a lot of false positive detections. Looking at the graph and F1-scores, we can understand that the detector is not very successful. At stage 3 of the training, the graph displays a minimal number of FP, while in reality the F1-scores show a very high FPR. The results of the detector can be improved by combining it with other detectors.

Integration with Shape Detectors

Below are displayed the resulting images from the improved detector, using the circle Hough transform.

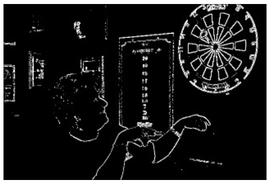


Figure 11: dart0.jpg thresholded gradient magnitude

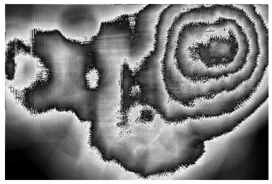


Figure 12: dart0.jpg Hough Space

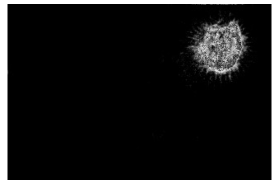


Figure 13: dart0.jpg threshholded Hough Space

Figure 14: dart0.jpg Result

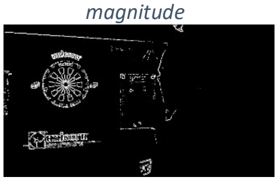


Figure 15: dart2.jpg thresholded gradient magnitude

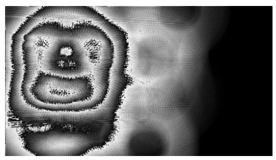


Figure 16: dart2.jpg Hough Space

Figure 17: dart2.jpg threshholded Hough Space

Figure 18: dart2.jpg Result

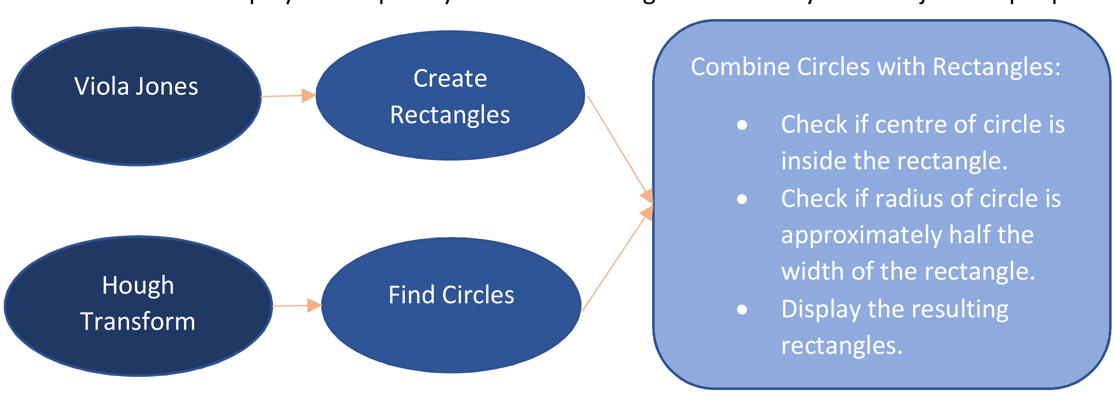
The following table displays the F1-scores achieved on all the test images provided, by using our implementation of the improved detector using a circle Hough transform:

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Average
F1-score	1	1	1	0	1	0.67	0.67	0.67	0.67	1	0.33	0	0	0	0.57	0.67	0.58

Merits and Shortcomings:

The detector performed significantly better with the Hough transform function implemented. This can be seen by comparing the F1-score with the Hough transform and without.

- The detector failed to detect some of the dartboards, for example at picture 13 as there are other parts of the picture with circular forms. At picture 13, the letters on the dartboard are arranged in a circular way, which produces a false result, as the detector detects a dartboard where the letters are, and our implementation of the Hough transform finds the circles there. The dartboard is not detected with ease in this picture, as the circle is broken by the huge dart in the middle. The detection of this picture could be improved by decreasing the threshold on the Hough transform, but that would interfere with our overall results.
- The detector did exceptionally well on most of the images, where it only displays the bounding boxes for the correct dartboards. Our implementation works exceptionally well when the dartboard is displayed completely rather than being obstructed by other objects or people.



Improving the Detector

- In our effort to improve the detector, we chose to use Local Binary Pattern features when training the Viola-Jones classifier instead of the default Haar-like features.
- This was done as the original detector had some problems when detecting dartboards with varying illumination across itself, and LBP features are better able to handle variations in object pose and illumination [1].
- The LBP operator works by thresholding the 3x3 neighbourhood of a center pixel using the center pixel as the threshold and producing the resulting pixel values as one binary number. By then segmenting the image into bins and taking both local and global histograms of these numbers, this approach would describe the local texture and global shape of dartboards.

Figure 19: dart13.jpg

Figure 20: dart15.jpg

We calculated the F1-scores across all test images with the new detector:

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Average
F1-score	0	1	0.67	0	1	0.4	0	0.67	0.5	0	0.5	0	0	1	0	0.1	0.42

Merits and Shortcomings

It turns out that our average F1-score across all images actually decreased when switching to LBP features.

- Our new detector performed really well for some images that it previously struggled with, as can be seen in the two images above, showing that it can actually deal with images where the illumination across the dartboard varies.
- For some of the images however, our new detector would actually accurately determine the position of the dartboards, but the bounding box it detected was much smaller than the ground truth, leading to an evaluation that it was a false detection.
- One good thing about the LBP features is that it is extremely quick to train the Viola-Jones classifier, which makes it likely to be viable for real-time object detection.
- Due to LBP features being computationally simpler than Haar-like features, this results in a lower detection rate as the accuracy of the detector is decreased for some dartboard images.

References

1. https://medium.com/@ckyrkou/object-detection-using-local-binary-patterns-50b165658368