Image Processing and Computer Vision CW

1: The Viola-Jones Object Detector

(a) The first task is to use Viola-Jones object detector to find faces in a set of 16 images. Figure 1 below shows the detecting result on 5 example images.

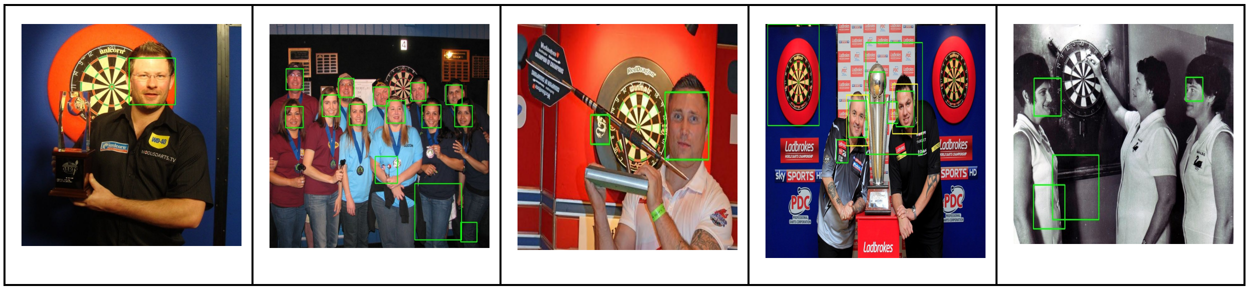


Figure 1. Face classification using Viola-Jones detector

(b) As we can see in figure 1, all the faces in image 2(dart5.jpg) are detected resulting in a True Positive Rate (TPR) of 11/(11+0) = 1. In image 5 (dart15.jpg), 3 there are faces present, but there are 4 detections. Two out of three of the faces present are detected resulting in a TPR of 2/(2 + 1) = 0.67.

(1)Assessing the TPR can prove to be a challenging task, and in many cases does not reflect the true accuracy of a detection algorithm. Facial features can be inconsistent, easily distorted and hard to detect due to lighting, skin color or orientation making it hard to accurately detect a face.
It can also be hard to label what the ground truth of an image is. For example, if an image contains a poster with a face, would we want it to be detected? If a bounding box doesn't fully encapsulate the face (e.g. image 15) is it still a valid detection?
Additionally, the TPR only takes into account positive and negative detection of the object class we are interested in (faces) but does not take into account all other object classes detected. Therefore the TPR does not represent the inaccurate or wrong detections of objects that aren't faces.

(2)This means that as long as all the faces in an image are detected, regardless of other objects detected, we will achieve a TPR of 100% (as seen in image 5). Hence we can always achieve a TPR of 100% for the reason that if we implement an algorithm that considers all objects in an image to be a face all the faces will be detected. Resulting in a TPR of 100%, but also a high False Positive Rate (FPR).

(3)Training different features separately and tuning parameters as optimally as possible can help optimize the accuracy of a classifier.In order to calculate the F1-score, we use the overlapping ratio of the bounding box with the ground truth box. And the threshold has been set to 0.5. For example, in dart 15, the F1-score = $2 * (\frac{1}{4} * \frac{1}{3}) / (\frac{1}{4} + \frac{1}{3})$ = 0.285. Figure 2 shows a image with ground truth in yellow boxes and bounding boxes in green.

Figure 2. Image with ground truth labeled

2: Building & Testing your own Detector

The second task consists of building and training an object detector that recognises dartboards. It uses 500 positive and 500 negative images and OpenCV's boosting tool to construct a detector that uses Haar-like features.

(a) In Figure 3, we can observe that before the first stage of training, both values for the TPR and FPR are 1, or 100%. By then end of the training the value of the TPR stays constant at 1 while the value of the FPR reduced drastically to a very small value. This can be due to our classifier overfitting on the training data. This graph poorly represents the accuracy of our classifier as we can see in figure 4 that only half the images have successful dartboard detections, and there a cases with false dartboard detections. While the classifier always claims a TPR of 1, we can see in the table below that the F1-scores are rather low and correspond better to the detection results found in the images.

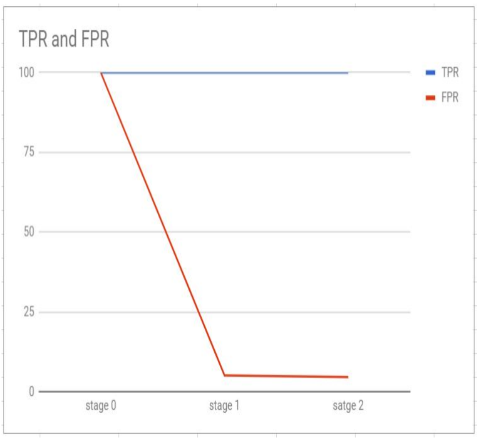
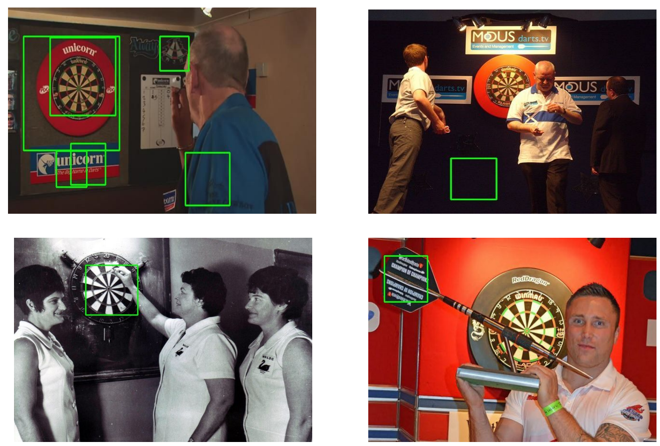


Figure 3. True positive and False negative change in three stages



Figue 4. Dartboard detection with cascade on dart 2, dart 6, dart 13 and dart 15

(b) Based on the F1-scores calculated for each test image we can see that the graph plotted above does not reflect the accuracy of our classifier. While the training stages of our classifier reveal a TPR of 100%, the F1-scores reveal the inaccuracy of our classifier in detecting dartboards. This shows us the limitations of the TPR in assessing the accuracy of a classifier, while showcasing the F1-score as a much more reliable representation.

Image #	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
TPR	1	1	1	1	1	1	0	1	1	1	0,7	0	1	0	1	1
F1-score	0,5	1	0,5	0,7	0,2	0,5	0	0,8	0,7	0,5	0,5	0	1	0	0,4	1

Table 1. Result of dartboard detector based on Viola-Jones method

3: Integration with Shape Detectors

In the third subtask, in addition to the object detection, we use a combined Hough transform. Dartboards have two distinct features: concentric circles and several lines intersecting with each other. So we decided to use both a line and circle Hough transform for the task. Figure 5 below shows the flow diagram of our implementation.

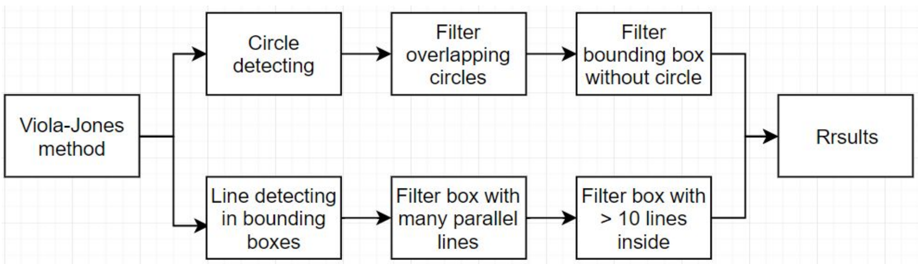


Figure 5. The flow chart for our combined shape detector

The rationales behind our combination are as follows:

- First we combine the Hough circle detection with Viola-Jones detector. Viola-Jones provides potential dartboards, but not all are valid, so we use a Hough circle to filter most of the invalid bounding boxes by checking if they have a circle center inside them.
- Secondly the Hough line has been combined with Viola-Jones. Since Hough circle can only deal with perfect circles, in order to find tilted circles (ellipses), we used the Hough line detection. In this case, if a bounding box has no circle, the line detection process will run. After filtering most of the parallel lines (as dartboards contain intersected lines), we check if the box contain more than 10 lines. Then the decision is made of whether or not it is a dartboard.

Figure 6 below shows the result of applying our combined circle and line detector, we should mention that both the circle and line detection have been processed on the threshold gradient magnitude image. We check for circles in the whole image, but only within bounding boxes for lines. Overall the combined object, circle and line detector has precision 0.93 and recall 0.7, and the final F1-score is 0.79. With TP = 14, FP = 1 and FN = 6.

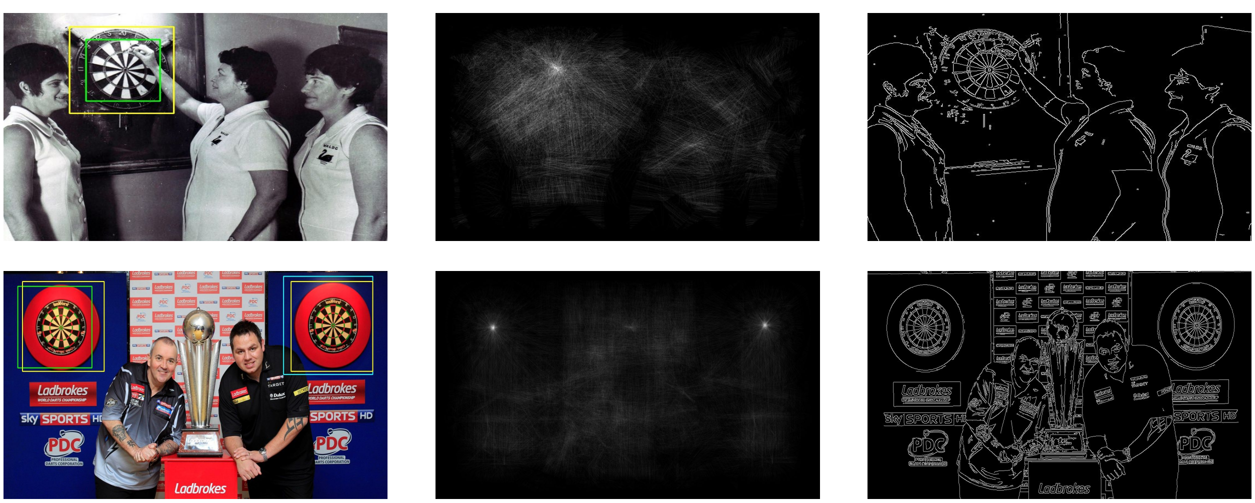


Figure 6. Test images, dart 14 and dart 15, with 2D Hough circle space and the edge detection image

4: Improving your Detector

4.1 Detector improvement

In order to improve the performance, we first try to increase the training data of the classifier. And it provides a more accurate detecting result as shown in Table 2 below. And as we can see in figure 7 below, we do get improved results comparing with the old classifier. But after we combine the new classifier with the Hough transform detector we built, the final result does not improve a lot.

Stage 0:TPR: 100%, FPR: 100 %	Stage 0:TPR: 100%, FPR: 100%
Stage 1:TPR: 100%, FPR: 5.07%	Stage 1:TPR: 100%, FPR: 6.90%
Stage 2:TPR: 100%, FPR: 4.59%	Stage 2:TPR: 100%, FPR: 0.95%

Table 2. Results after increase training sample from 500 negative(left) to 1000(right)

Because the combined detector we have built really depends on the Hough circle detector and the Hough line detector, although we have a slightly improved the output of bounding boxes, the circle and line detector does not improve at this time, which give us a quite similar results at the end.

A problem we encountered was the running time of our implementation. The Hough circle detection we have created uses a large 3D matrix. Operations such as finding the index of a maximum value is quite time consuming. One of the improvements we have done is that for some of the operations we convert the 3D matrix to a 1D array, which significantly speeds them up. The final running time has been improve from 10s per image to 5s.

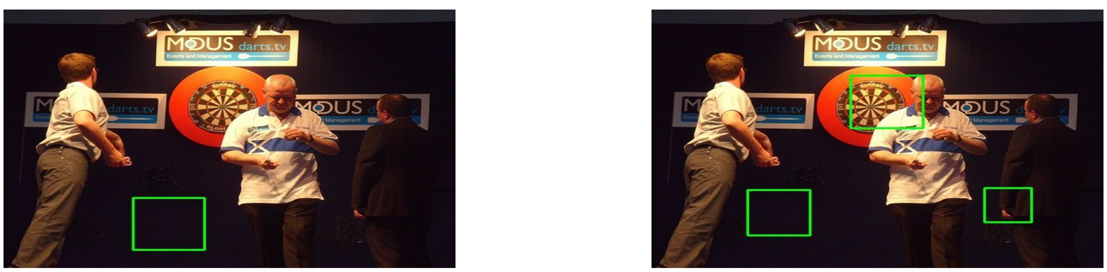


Figure 7. Object detection result after increase sample from 500(left) to 1000(right)

4.2 Evaluation and future improvement

Merits and the shortcomings of our detector:

Performs well on perfect dartboards
Can detect partially circle dartboards
Can detect tilted dartboards
* Large Hough space
* False negatives exist

Currently we are using line feature to detect tilted dartboards. It is quite limited when the line is not clear in the image, one way to solve this is to implement hough ellipse detection.

And another aspect we can use to improve is the color pattern of the dartboard. Noticed the face that a dartboard has various color inside, we can use it to improve the performance of our detector.