

Image Processing and Computer Vision

1st George Lancaster
 dept. of Computer Science
 University of Bristol
 Bristol, United Kingdom
 qv18258@bristol.ac.uk

2nd Ren Jiang
 dept. of Computer Science
 University of Bristol
 Bristol, United Kingdom
 mu18336@bristol.ac.uk

Abstract—This report outlines the tasks completed for Image Processing and Computer Vision assignment one. Our final classifier, which uses template matching and speeded up robust features (SURF), has an F1 score of x.xx when tested against the 16 test images.

I. The Viola-Jones Object Detector

Sixteen test images were annotated for ground truth. Each test image contains either faces, dartboards or a combination of both. In this first task, we use the Viola-Jones object detector to find faces within the test images.



Fig. 1: Five images from the test data set. Green rectangles show where the Viola-Jones classifier has detected a face.

The true positive rate for images darts5 and darts15 when tested using the Viola-Jones face detector are 1 and 0.67 respectively.

Although the true positive rate can be used to indicate a classifiers accuracy, it does not reflect its true performance. It is always possible to get a 100 per cent detection rate on any classification task as we can select all possible areas of an image, regardless of the number of false positives. The key to a good classifier is to get a high true positive rate, whilst keeping the false positive rate minimal. The F1 score measures the relationship between the precision and the recall of the model and can therefore be considered to be a more reliable measure of classifier performance.

Additionally, the true positive rate can be difficult to assess as it requires the definition of a rule to determine what counts as a detection. For this task, a true detection is defined as an area that shares a 68 per cent overlap with a ground truth annotation. A value of 68 per cent was chosen as it is the largest area of overlap before the F1 score starts to decline.

TABLE I: F1 scores, precision and recall for all images, when detecting faces using the Viola-Jones face detector.

Image Name	F1 Score	Precision	Recall
dart0	1.00	1.00	1.00
dart1	1.00	0.00	0.00
dart2	1.00	0.00	0.00
dart3	1.00	0.00	0.00
dart4	1.00	1.00	1.00
dart5	0.79	0.65	1.00
dart6	1.00	0.00	0.00
dart7	1.00	1.00	0.00
dart8	1.00	0.00	0.00
dart9	0.67	0.50	1.00
dart10	1.00	0.00	0.00
dart11	0.67	1.00	0.50
dart12	1.00	0.00	0.00
dart13	1.00	1.00	1.00
dart14	0.36	0.22	1.00
dart15	0.40	0.50	0.33
Average F1 Score			0.87

Using the Viola-Jones face detector gives fair results when applied to the 16 test images. Images that have been correctly classified as containing no faces have been given an F1 score of 1. When excluding these images from the calculation, the F1 score is reduced to 0.81.

II. The Dartboard Detector

To detect dartboards, the Viola-Jones classifier was trained on a data set generated from a single bitmap image of a dartboard, in addition to a set of negative images.

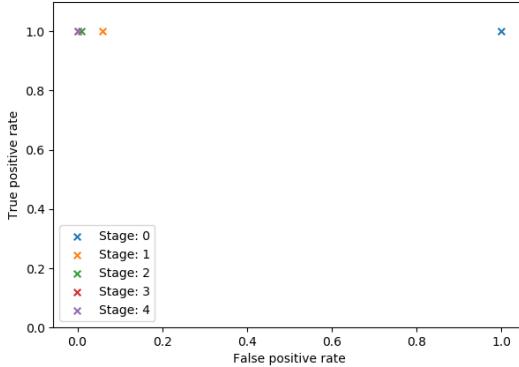


Fig. 2: True positive rate plotted against false positive rate when training the cascade classifier on 500 images of dartboards and 500 negative images. Each stage of training can be identified by its own point.

On the first stage of training, all samples are classified as true. This is reflected by a value of 1 for both the true positive rate and false positive rate. As the classifier progresses through further training stages, the false positive rate decreases, whilst the true positive rate remains the same. This indicates that the classifier improved after each training stage, by decreasing the number of false detections.

To measure how varying the number of training images affects performance, two identical Viola-Jones classifiers were trained on sets of 500 (a), and 1000 (b) true images and negatives. The F1 score of classifier b was 0.06 higher than classifier a. This suggests that the classifiers performance can be improved by increasing the amount of training data.

Observed differences in accuracy between training and testing can be attributed to the training data consisting of a small set of critical features. In contrast, the test images contain a large amount of irrelevant background noise, which needs to be discarded by the cascade.

The performance of the model at this stage is remarkably worse than when detecting faces. The face detection classifier has been trained on a larger number of stages, so the two systems are not comparable.

TABLE II: F1 scores for all images, when detecting dartboards using the Viola-Jones cascade classifier trained on dartboards.

Image Name	F1 Score	
	500 Training Images (a)	1000 Training images (b)
dart0.jpg	0.50	0.25
dart1.jpg	0.00	0.50
dart2.jpg	0.25	0.40
dart3.jpg	0.54	0.33
dart4.jpg	0.40	0.25
dart5.jpg	0.15	0.20
dart6.jpg	0.33	0.50
dart7.jpg	0.00	0.25
dart8.jpg	0.16	0.25
dart9.jpg	0.22	0.20
dart10.jpg	0.40	0.55
dart11.jpg	0.22	0.22
dart12.jpg	0.67	0.50
dart13.jpg	0.20	0.33
dart14.jpg	0.09	0.14
dart15.jpg	0.67	0.50
Average F1 score	0.28	0.34

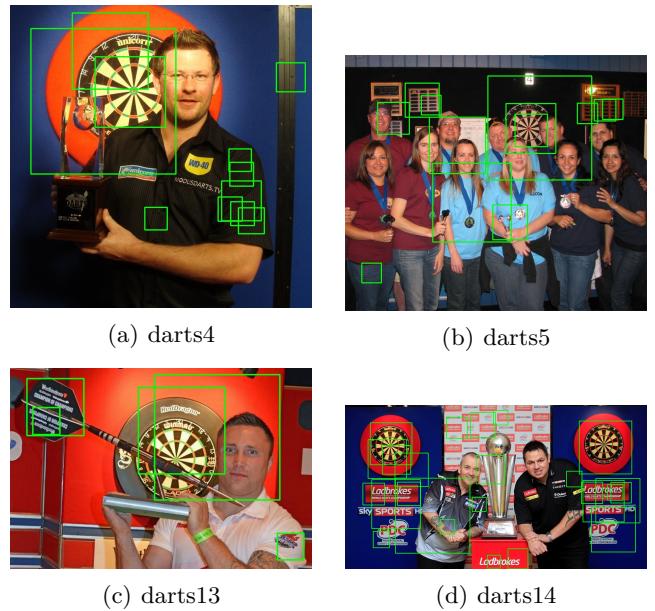


Fig. 3: Four images from the test data set. Green rectangles have been drawn where the Viola-Jones classifier has detected a dartboard.

III. Integration with Shape Detectors

Both linear and circular hough transforms have been used in conjunction with the Viola-Jones classifier to improve the precision and recall of the system.

Reasons for building the classifier in this way include:

- A circular hough transform is used to find circular features in the image;
- The Viola-Jones object detector then determines if there are dartboards in the important area of the image;

TABLE III: F1 scores, precision and recall for all images, when detecting for dartboards using the Viola-Jones classifier combined with shape detection techniques.

Image Name	F1 Score	Precision	Recall
dart0	1.00	1.00	1.00
dart1	1.00	1.00	1.00
dart2	1.00	1.00	1.00
dart3	1.00	1.00	1.00
dart4	1.00	1.00	1.00
dart5	1.00	1.00	1.00
dart6	1.00	1.00	1.00
dart7	1.00	1.00	1.00
dart8	1.00	1.00	1.00
dart9	0.67	0.50	1.00
dart10	1.00	1.00	1.00
dart11	1.00	1.00	1.00
dart12	1.00	1.00	1.00
dart13	1.00	1.00	1.00
dart14	0.67	0.50	1.00
dart15	1.00	1.00	1.00
Average	0.96	0.94	1.00

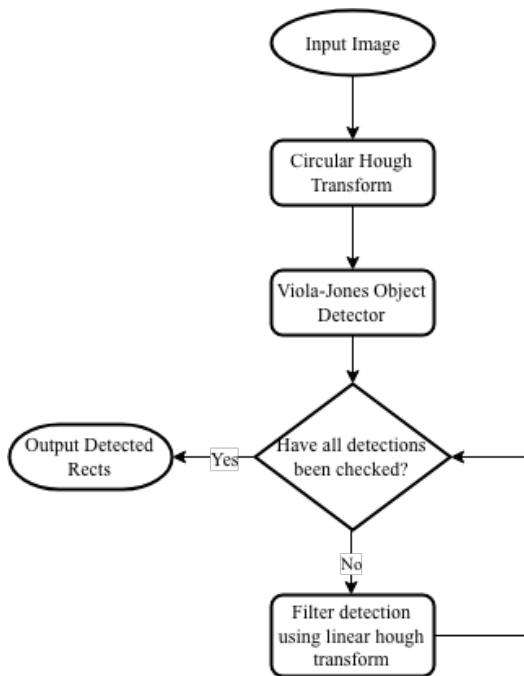


Fig. 4: Flow diagram of image detection algorithm.

- Using the linear version of the hough transform allow for false detections to be filtered out.

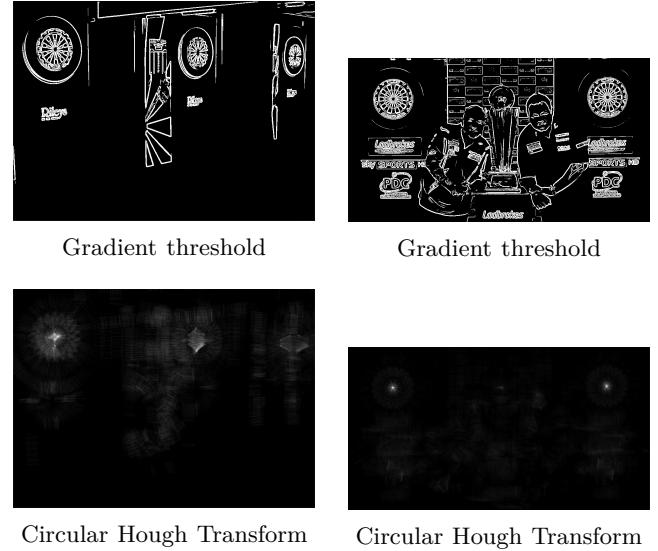


Fig. 5: Best and worst cases for the dartboard detector. Column a shows the best case, and has a F1 score of 1. Column b shows the worst case is the worst case, and has an F1 score of 0.667

The solution at this level gives very good results, achieving a true positive rate of 1 for all images. Only images 9 and 14 have imperfect F1 scores, caused by false positives. Image 9 could be considered to be a true positive as the classifier has detected a dartboard on a t-shirt.

Merits of the classifier include:

- It has a perfect recall value for all images;
- The false positive rate is very low, with only three false positives in the entire data set;

And the shortcomings include:

- The classifiers parameters have been tuned to perform well on the test data set, it may not generalise well to other images;
- Classification speed is slow due to the hough transform. It could not be applied to data in real-time.

IV. Further Improvements

As the dartboard detector from part three has a perfect true positive rate, improvement can only be made to the false positive rate. Two filters, using SURF and template matching have been used as final stages to reduce the number of false positives, and achieve a perfect F1 score for the test data set.

A. SURF

SURF aims to map the relationship between a descriptor and detector to find similarities in images. (Will write more about this later).

Only areas with more than two significant points of interest have been kept in the list of detected dartboards.



Fig. 6: SURF, applied to the test data set.

B. Template Matching

Template matching is used to match parts of an image to a template. Because the viola-jones detector only has 3 levels, it lacks features which can further reduce false positive rate.

After experiments of using different templates, one of them is decided because it perfectly helps to reject false positives.

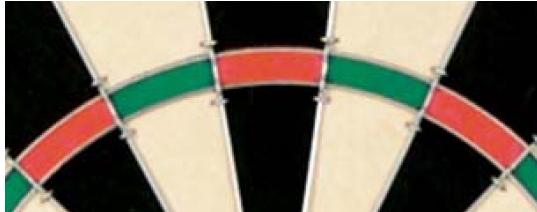


Fig. 7: template matching, final template chosen.



Fig. 8: template matching rejects 2 remaining false positives.

C. Results

The key to successful template matching is to choose a template with enough feature to reject false areas. However, it shouldn't contain all the features of a dartboard because there may shelters, irregular shapes or light in the test images that prevent them from getting high enough scores in template matching. So many different templates should be tried before a suitable one can be found.

TABLE IV: F1 scores, precision and recall for all images, when detecting for dartboards using the Viola-Jones classifier combined with shape detection techniques, SURF and template matching.

Image Name	F1 Score	Precision	Recall
dart0	1.00	1.00	1.00
dart1	1.00	1.00	1.00
dart2	1.00	1.00	1.00
dart3	1.00	1.00	1.00
dart4	1.00	1.00	1.00
dart5	1.00	1.00	1.00
dart6	1.00	1.00	1.00
dart7	1.00	1.00	1.00
dart8	1.00	1.00	1.00
dart9	1.00	1.00	1.00
dart10	1.00	1.00	1.00
dart11	1.00	1.00	1.00
dart12	1.00	1.00	1.00
dart13	1.00	1.00	1.00
dart14	1.00	1.00	1.00
dart15	1.00	1.00	1.00
Average	1.00	1.00	1.00