

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 achieved an F1 score of 1 when tested against the sixteen 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, the Viola-Jones object detector has been used 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 sixteen test 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.

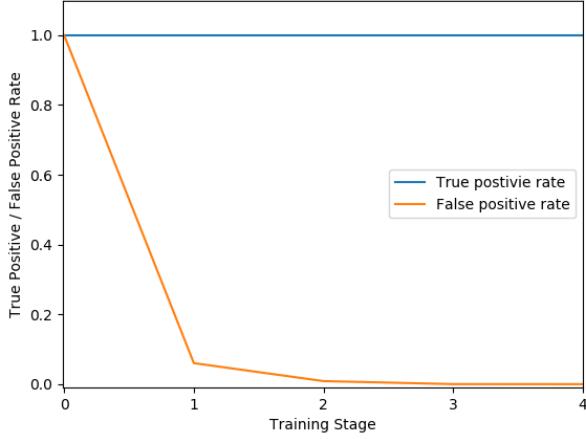


Fig. 2: True positive rate plotted against false positive rate when training the cascade classifier on 500 images of dartboards and 500 negative images over four stages (0-4).

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 only.

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

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

To measure how varying the number of training images affects performance, two identical Viola-Jones classifiers were trained on sets of 500, and 1000 true images and negatives. Training on 1000 images improved the F1 score by 0.06. 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. Because of this, the false positive rate is much higher for the test images.

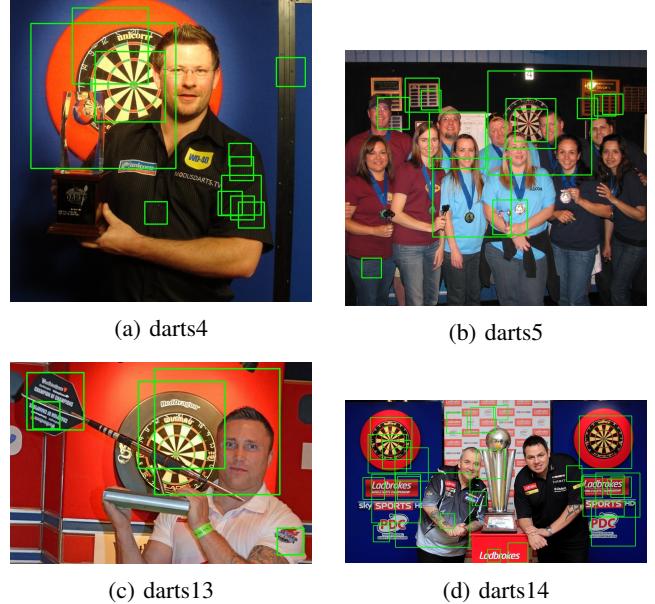


Fig. 3: Four images from the test data set. Green rectangles show where the Viola-Jones classifier has detected a dartboard.

The performance of the dartboard detector at this stage is remarkably worse than the face detector from task one. One reason for this being that the face detection classifier has been trained on a larger number of stages, which means that there is more opportunity for false positives to be filtered out from the result. This is supported by a recall value of 1 for the dartboard classifier, meaning that all dartboards have been successfully detected. However, the precision of 0.21 brings the F1 score down.

Additionally, the face detector has been trained on a much larger data set than the dartboard detector. We propose that further increasing the size and variety of the training data could increase the precision of the model.

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.

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

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.

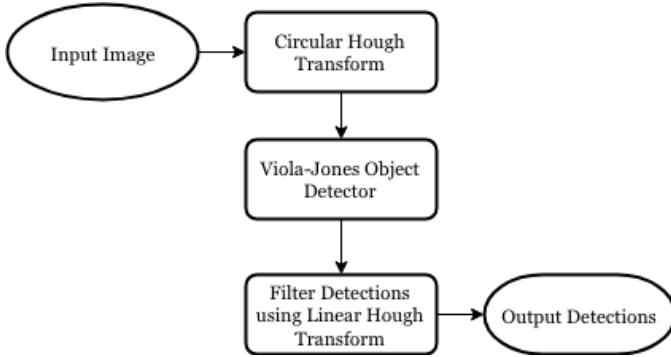


Fig. 4: Flow diagram of image detection algorithm.

Reasons for building the classifier in this way include:

- Although the Viola-Jones classifier can filter out background information, applying the hough transform first allows it to focus on areas of interest;
- Each area is checked by the linear hough transform to ensure that there is a dartboard present. A large number of intersecting lines indicates truth.

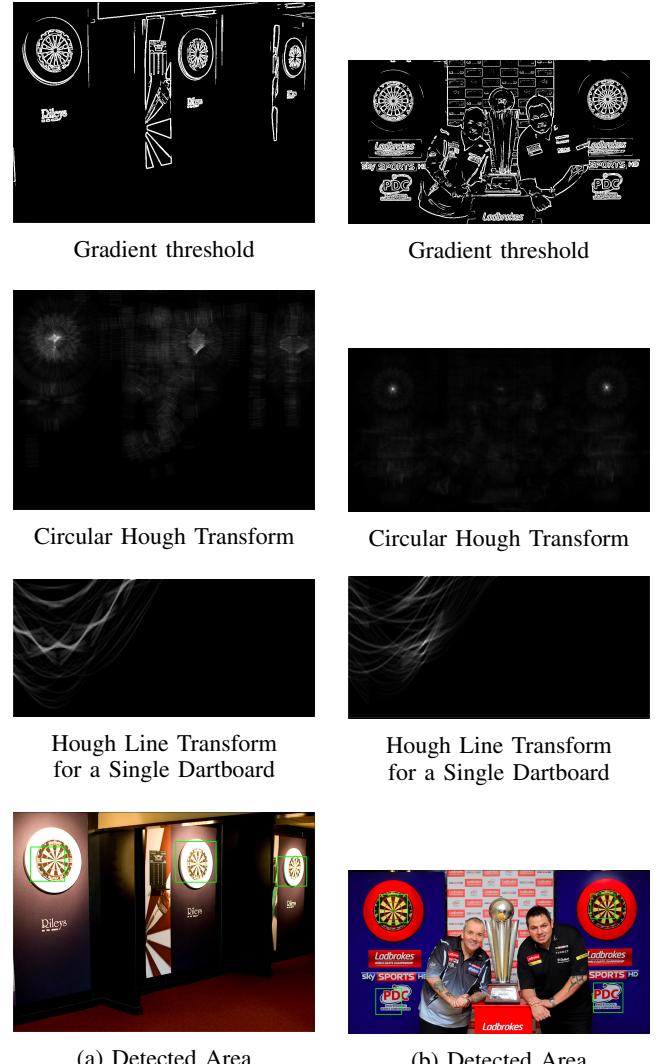


Fig. 5: Best and worst cases for the dartboard detector. Column a shows the best case *dart8*, which has a F1 score of 1. Column b shows the worst case *dart14*, which has an F1 score of 0.67

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 perfect precision, improvement can only be made to the recall. Two filters, using SURF and template matching have been used as final layers in the dartboard detector to reduce the number of false positives, and achieve a perfect F1 score for the test data set.

A. Speeded up robust features (SURF)

SURF aims to map the relationship between a descriptor and detector to find similarities in images. The SURF algorithm splits object detection into three stages: interest point detection, local neighbourhood description and matching. Interest point detection performs blob detection to detect characteristic regions in the object image. Local neighbourhood description uses the points of interest to define the features of the image. By comparing descriptors from two images, matching pairs of points of interest can be found.



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

By applying SURF as a filter, only areas with more than two significant points of interest have been kept in the list of detected dartboards.

B. Template Matching

Template matching is used to match parts of an image to a template. The dartboard used from part two only has 3 layers, therefore it lacks the later layers that remove falsely detected features. By having more layers, the false positive rate can be reduced.

Through experimentation, a template has been found that optimally rejects false positives, whilst keeping the true positives.

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

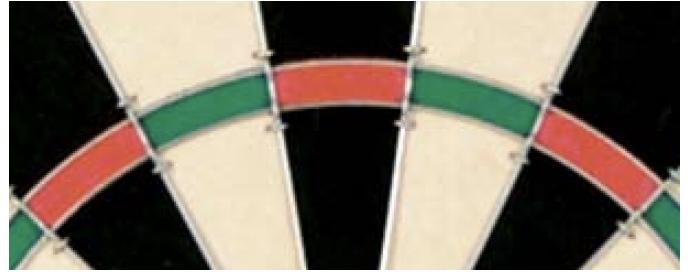


Fig. 7: The Template used in the Template Matching Extension.



Fig. 8: Template Matching Rejects the Two Remaining False Positives from *dart15*.

C. Results

Applying the SURF filter removes two false positives. One from *dart9*, and one from *dart14*. This leaves one false positive in *dart14*, which is removed by template matching.

We have decided to exclude a dartboard from *dart8* and *dart11* due to too much noise in the image.

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