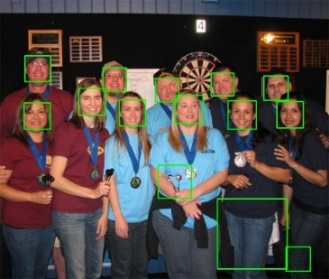
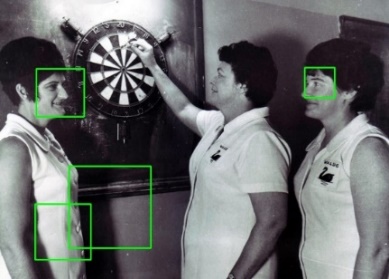
Image Processing and Computer Vision – The Dartboard Challenge

Subtask 1

The first subtask was to use the Viola-Jones object detector to find frontal human faces and to analyse its performance. To do this we first annotated the images by drawing rectangles around the faces present in order to generate our ground truth, then ran the provided Viola-Jones face detection algorithm on the images, which generated its own rectangles around what it detected as faces, these images are shown below.



Figures 1-5: The bounding boxes generated by the face detector.

We then edited the face.cpp file to compare the coordinates of accurate rectangles which we drew and the code generated ones, by providing our coordinates in a csv file, this gave us the true positive rate for the algorithm which allowed us to calculate the F1 score. Our results are shown below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | dart4.jpg | dart5.jpg | dart13.jpg | dart14.jpg | dart15.jpg |
| TPR | 1 | 1 | 1 | 1 | 1 |
| F1 Score | 1 | 0.88 | 0.667 | 0.5 | 0.667 |

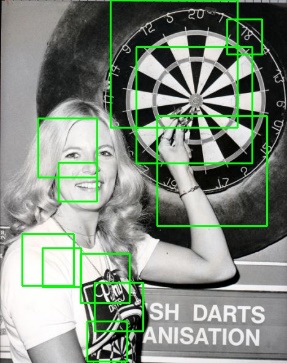
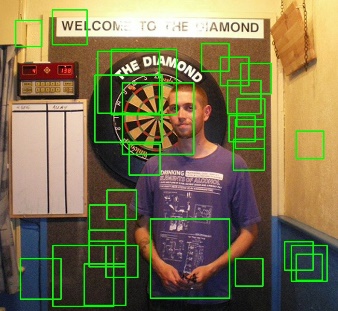
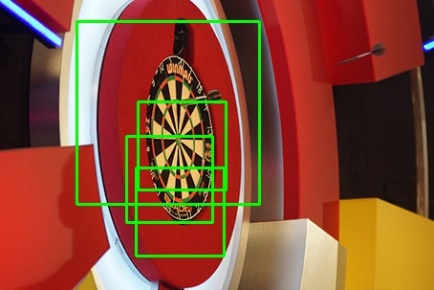
The TPR results show that the face detection algorithm accurately identified all the faces which are in all of the images. However the F1 scores show that when the algorithm is run, it falsely detected a number of objects as faces in all the images we tested it on other than dart4.jpg.

It can be difficult to accurately calculate the TPR as it can be hard to define what is a true positive, for example in image 15, none of the three faces are truly front facing, so it is hard to decide which are to be counted as faces and which are not. It can also be hard to decide if an object has actually been detected, for example of there is a large box and a face which is far smaller than the box, happens to be in it, does this count as a detection? How much bigger would the box need to be than the face before it stops counting as a detection?

Why is it always possible to achieve a TPR of 100%?

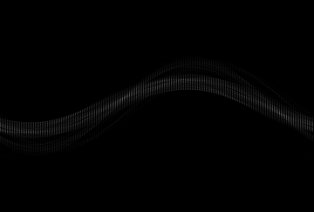
This is possible as you can write an algorithm which detects every part of an image to be a face, therefore it will correctly detect all of the faces as faces achieving a TPR of 100%. However it will also detect everything that is not a face, as a face.

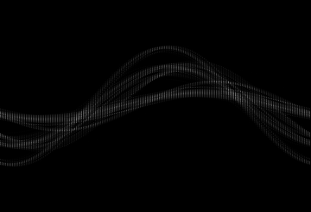
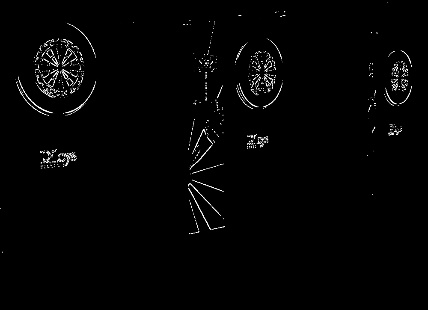
Subtask 2

 The graph above shows how as the classifier was trained it always maintained a true positive rate of 1, meaning it always managed to correctly identify all of the dartboards present in the images, however on the first stage, the false positive rate was also very high at 1, meaning that it classified everything that wasn’t a dartboard as a dartboard. This means on this stage of training the classifier is classifying everything as a dartboard, so is essentially useless. After the second stage of training the false positive rate was dropped to 0.01274 meaning only a small number of the inputted negative images were wrongly classified as containing a dartboard image. On the third and final stage of training the classifier, the false positive rate was dropped to 0.00071 meaning that there was even less false positive classifications within the training set.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dart | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| TPR | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| F1 Score | 0.154 | 0.5 | 0.2 | 0.25 | 0.222 | 0.182 | 0.154 | 0.0690 |
| Dart | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| TPR | 1 | 1 | 0.667 | 1 | 1 | 1 | 1 | 1 |
| F1 Score | 0.103 | 0.167 | 0.1 | 0.333 | 0.4 | 0.167 | 0.074 | 0.333 |

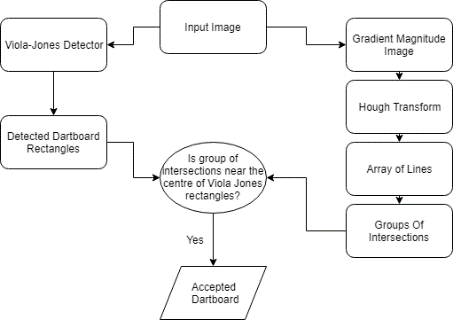
Comparing how the classifier preformed identifying dartboards with 3 rounds of training against the face detection algorithm of the previous task, it is clear to see from both the images and the data results that the face detection algorithm was far more accurate. Although the true positive rate for all except one of the dartboard results was 1, the F1 scores were far lower which indicated that the false positive rates were far higher. This is confirmed by looking at the outputted images where it is obvious that there are far more bounding rectangles than there are dartboard images.

Subtask 3

Our detector managed to produce a very good result on dart4.jpg, recognising the one present dartboard and not miss identifying any other object. Therefore it had a TPR of 1 and and F1 score of 1, meaning perfect detection.

On dart10.jpg only one of the three dartboards in the image was detected. From inspecting the threshold gradient magnitude image, you can see that this is likely as the only the nearest dartboard has clearly define lines, where as the two further away dartboard have lines which are either broken up or merged together, because of this our detector would not have accepted them as lines and therefore missed them.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dart | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| TPR | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| F1 Score | 1 | 1 | 1 | 0.4 | 1 | 1 | 0.333 | 0.667 |
| Dart | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| TPR | 0.5 | 1 | 0.3333 | 1 | 1 | 1 | 0 | 1 |
| F1 Score | 0.667 | 0.4 | 0.5 | 0.4 | 0.667 | 0.333 | 0 | 0.667 |

Overall our detector was definitely an improvement over just using the Viola-Jones detector alone. Although three true positive rates dropped, there was significant increase in the F1 scores, including five of our images having F1 scores of 1, meaning all dartboards and only those dartboards were detected. Its main shortcomings were in image 14 were it did not detect either dartboard, even though to a human eye they seem fairly clear, as their orientation is front on and there is nothing blocking any part of them. It also did not manage to remove all false positives, as the line detector often found groups of intersections which were not dartboards.

The rationale behind why we combined our detector with the Viola-Jones one, in this way, is that the Viola-Jones already correctly identified all the dartboards, our detector simply tried to remove the false positive results, which is did relatively well.