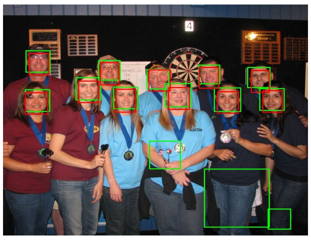
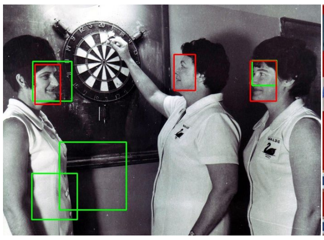
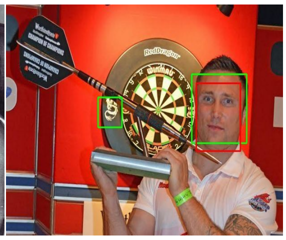
Dartboard Challenge Report

Subtask 1: Viola-Jones Object Detector

Below are the results of running the given frontal face classifier, trained using AdaBoost for Viola Jones, to detect human faces from the frontal view on images dart4, dart5, dart13, dart14 and dart15. The ground truths are shown by the red bounding boxes, while the faces detected by the classifier are shown by the green bounding boxes.

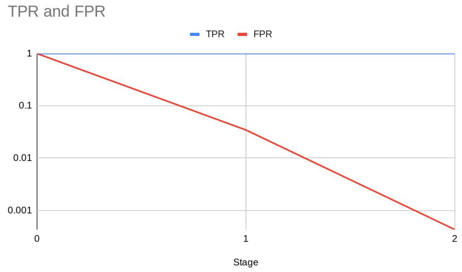

There are some difficulties with calculating the True Positive Rate. This is mainly due to what we perceive as being a face. For example, some may say that the forehead, chin and ears are part of the face, while others may say that these are not included as part of a human face. In order to allow for this difference in opinion, a smaller threshold is used to find the number of true positives when calculating the intersection over union between the detected and ground truth bounding boxes.

It is possible to achieve a 100% True Positive Rate on any detection task. As the True Positive Rate is calculated by the number of successful detections (true positives) divided by the number of ground truths (sum of true positives and false negatives). If the detector was weak and classified every output as a positive, then there will not be any false negatives and therefore achieve the 100% True Positive Rate for any detection. In the Viola-Jones detector, this is achieved by classifying all Haar features in every sliding window as a detection. Therefore, the True Positive Rate is used in conjunction with the precision, where the precision is the number of successful detections (true positives) divided by the number of detections (sum of true positives and false positives). The precision measures the correctness of the results while the True Positive Rate measures the completeness of the results. If the detector was weak and achieved a 100% True Positive Rate on any detection, the precision would be very low as there will be a very high number of detections. The F1 score calculated below for all test images measures the precision against the True Positive Rate.

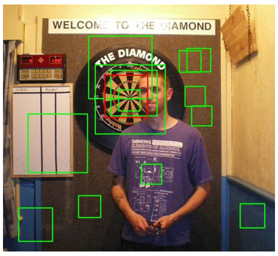
Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Av.
TPR	0.000	0.000	0.000	0.000	1.000	1.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	1.000	1.000	0.333	0.396
F1	0.000	0.000	0.000	0.000	1.000	0.880	0.000	0.000	0.000	0.400	0.000	1.000	0.000	0.667	0.500	0.286	0.296

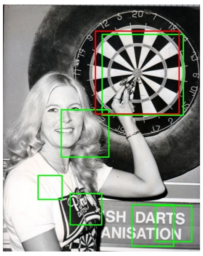
Subtask 2: Building and testing our own detector

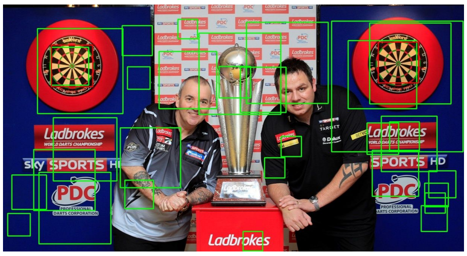
The Viola-Jones detector can be adapted and trained using AdaBoost, utilising an attentional cascade method, to identify images of dartboards for later parts of the coursework.



The graph above shows that for each stage of the attentional cascade the TPR is 1, but with each additional stage the FPR is reduced (1 in the first stage, to 0.000430125 in the final stage). This implies that the cascade initially classifies everything as positive, and with each iteration it refines its parameters to reduce the number of false positives and hence the FPR. The graph implies that the classifier can be further improved on the test data by adding further stages, although this carries the risk of overfitting.







The example images above show the result of images 7,9, and 14 after being processed by the classifier. It is obvious here that although it can accurately find the ground truths, there is an abundance of false positives that also occur. The table below shows the performance of the classifier on all of the 16 input images.

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Av.
TPR	1.000	1.000	0.000	1.000	1.000	0.000	0.000	1.000	0.500	1.000	0.000	1.000	0.000	1.000	0.500	1.000	0.625
F1	0.286	0.333	0.000	0.200	0.200	0.000	0.000	0.133	0.105	0.286	0.000	0.667	0.000	0.222	0.063	0.500	0.187

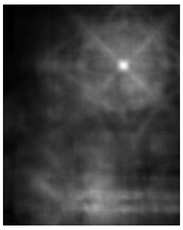
The classifier was trained on a single bitmap that was contorted in various ways within a specific range of angles and contrasts. This is one of the primary factors in explaining why the TPR was reduced from 1 to 0.625 when moving from training data to the test images. The classifier had a relatively small sample size to train on so certain angles and contrasts that occur in the test images would not have been experienced before by the classifier so images with dartboards at certain angles and contrasts may be missed by the Viola-Jones detector. This also meant that the classifier was not trained on images that were partially obscured, so the detection for images such as dart7 is affected due to an object being in front of the board. Due to the nature of classifiers, the classifier is likely to perform better on training data where it can iteratively optimise its parameters for the data, but is unlikely to perform as well on unseen data due to overfitting.
As we can see from the example images, the classifier produces a large number of false positives that have a negative effect on the F1 score. There is potential to reduce the number of false positives detected by the classifier by training it on a larger set of negative images, in order to give it a larger sample on which to evaluate its parameters.

Subtask 3: Integration with Shape Detectors

The Viola-Jones detector is not very accurate and can be optimised by combining it with a shape detector. As a major feature of dartboards are concentric circles, a circle detector was chosen to be used in conjunction with the Viola-Jones detector to create an accurate dartboard detector. By using the Hough transform for circles to create a circle Hough space, this space can be thresholded to find the circles within an image. The detector can be improved by finding circles within bounding boxes from the Viola-Jones detector and then thresholding the Intersection Over Union to find the most fitting bounding box.

Dart9

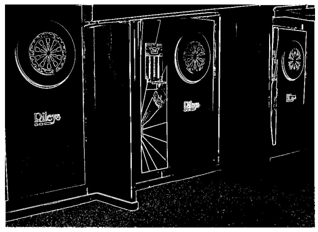
Dart9 Magnitude



Dart9 Circle Hough Space

Dart10

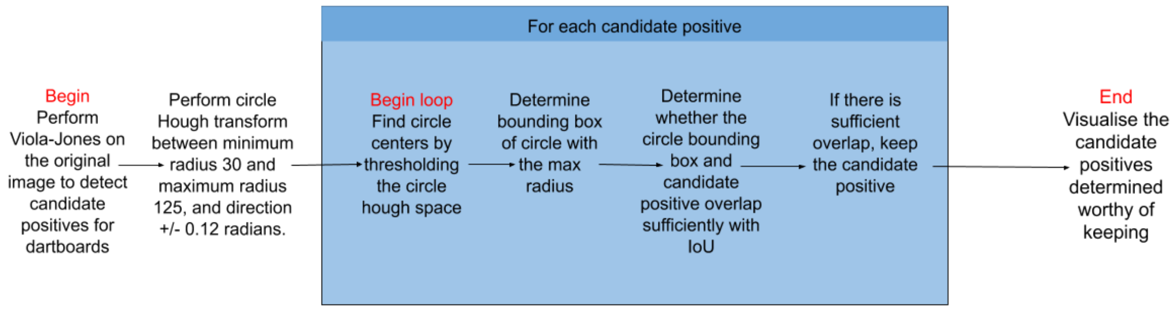
Dart10 Circle Hough Space



Dart10 Magnitude

Image	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Av.
TPR	1.000	1.000	0.000	1.000	1.000	0.000	0.000	1.000	0.500	1.000	0.000	0.000	0.000	1.000	0.500	1.000	0.563
diff.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-1.00	0.000	0.000	0.000	0.000	-0.062
F1	1.000	1.000	0.000	1.000	1.000	0.000	0.000	0.667	0.667	1.000	0.000	0.000	0.000	0.667	0.286	1.000	0.518
diff.	+0.714	+0.667	0.000	+0.800	+0.800	0.000	0.000	+0.534	+0.562	+0.714	0.000	-0.667	0.000	+0.445	+0.223	+0.500	+0.331

- The Viola-Jones detector combined with the circular hough transform is able to drastically reduce the number of false positives by removing those that do not contain a circle, thus improving the f1 score.
- The combined detector is still limited by the original accuracy of the Viola-Jones classifier, as it only refines the positives produced by the classifier and does not consider any area of the image that was originally missed.
- The combined classifier does performs well on images where the dartboard is front on but not so well when there is a significant angle. This is because the angle turns the image of the board into more of an ellipsis, which our classifier is unable to accurately detect.



- We recognise that the Viola-Jones detector was able to get most of the actual positives, but was producing too many false positives.
- We decided that we would take the output of the Viola-Jones and attempt to improve its F1 score by reducing the number of false positives, by comparing each candidate positive to the circle hough transform output.
- We decided to use IoU when comparing the circle and candidate positive to ensure that the overlap was not by chance, and that they were actually detecting roughly the same region.

Subtask 4: Improving the detector

There are a number of ways that the detector could be improved. For example, we only use the circle hough transform during our image processing but there are other distinctive features to a dartboard that could be identified using similar techniques. In order to improve the detector a line detector and an ellipses detector, using hough transforms, could be integrated to firstly recognise the distinctive pattern of the board, and secondly to have better detection of boards that are angled relative to the camera (which is one limitation of using the circle detector). We would begin by finding the circles and ellipses within the image using the respective detectors to narrow down the regions within the image in which a dartboard may be contained. We would then introduce the line detector to find lines within those regions identified. By finding the line intersections, we could match these with the centre of the circles and ellipses to accurately detect dartboards in the image.

It would also be possible to utilise the colours in an image to better recognise dartboards. The accuracy of the detector could be improved significantly if it were able to spot regions of red or green, black and white and use segmentation to better determine whether they were adjacent. This method may be limited if the image is grayscale, however, segmentation could still be used for these images by using segmentation for darker and lighter areas adjoined in a pattern.

Lastly, altering the training parameters for the Viola-Jones would increase the number of possible regions a dartboards may be contained in. This is due to there being more positive samples and these samples would contain more distorted dartboards with elliptical shapes. This would improve our Viola-Jones detector for images such as dart10 as these images are taken from an angle resulting in elliptical circles. This would allow our shape detectors to help find dartboards in more regions as they rely on the output bounding boxes of the Viola-Jones and therefore improve the accuracy of our dartboard detector.

Sign off:
Ben Fozard (bf17813): 1

Isaac Lipszyc (il17557): 1

Isaac Lipszyc