rm14834 ao14837

The Object Detection Challenge – Report

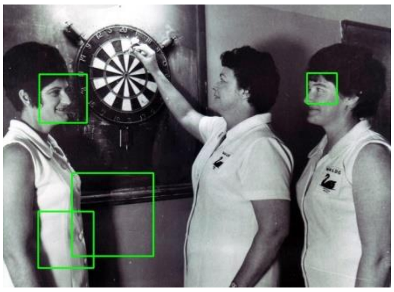

Figure 1: dart15.jpg

Figure 2: dart13.jpg

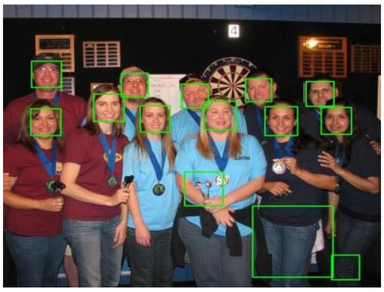

Figure 3: dart5.jpg

Figure 4: dart14.jpg

	dart5.jpg	dart15.jpg
True positives (TP)	11	2
False negatives (FN)	0	1
True positive rate (TPR)	1	0.66

Figure 5: Face detector performance on 2 test images

Figure 6: dart4.jpg

Practical difficulties in assessing the TPR:

- Need to preliminary determine an unbiased ground truth (a labelling of the actual faces in the image) which could potentially be very time consuming if the test set is large.
- The ambiguity of detection candidates is another problem that is likely to arise when trying to categorise one as a true positive or a false positive. Situations that might occur in this context are:
  - Target object is partially occluded by different objects or even by other target objects and detection captures only a small fraction of it (e.g. *Figure 3* some faces partially occlude others, but here detections are accurate enough)
  - Detection covers only part of an unoccluded object or considerably more than just the object (e.g. *Figure 4* contains a large detection that includes not only the second face, but also the trophy)
  - Multiple detections of the same object (e.g. this is also noticed in *Figure 4)*
  - Target object pose is no longer frontal, but rotated at different angles (e.g. in *Figure 1* the third face is rotated at 90 and detection captures only a portion of the profile face making the detection ambiguous)

You can trivially obtain a 100% TPR in a detection task by simply classifying every fixed size region of the image as a detection. The main disadvantage here is getting an enormous amount of false positives even when the object to be detected is missing from the image. Additionally, if you train the classifier by adding more and more features that are specific to your detection task, you can easily obtain a TPR of 100%. However this comes at the expense of overfitting and most likely losing precision for future detection tasks on different data.

$$F1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{2*TP}{2*TP+FP+FN} \quad recall = \frac{TP}{TP+FN} \quad precision = \frac{TP}{TP+FP}, \quad where$$

TP = number of true positives
FP = number of false positives
FN = number of false negatives

Lastly, we can use the F1 score method to evaluate the detector performance, but in order to compute the TP, FP and FN values we need to formalize a set of rules. Assuming our ground truth is a set of rectangle coordinates which correctly mark the true locations of the objects searched for in our detection task, we can give the definition of what a true positive, false positive and false negative is as follows:

1. We classify a detection candidate as a *true positive* if it:
   - covers at least 50% of the area represented by one of the rectangles mentioned previously
   - does not overlap another detection candidate which had been considered a true positive for the same object
   - Is at most twice the size of the rectangles in the ground truth
2. If any one of these three conditions is not met then we count a detection candidate as a *false positive*. If after checking all detection candidates there still remain unmatched elements from our ground truth set, that is elements which do not have a corresponding true positive detection, then we count these as *false negatives*.

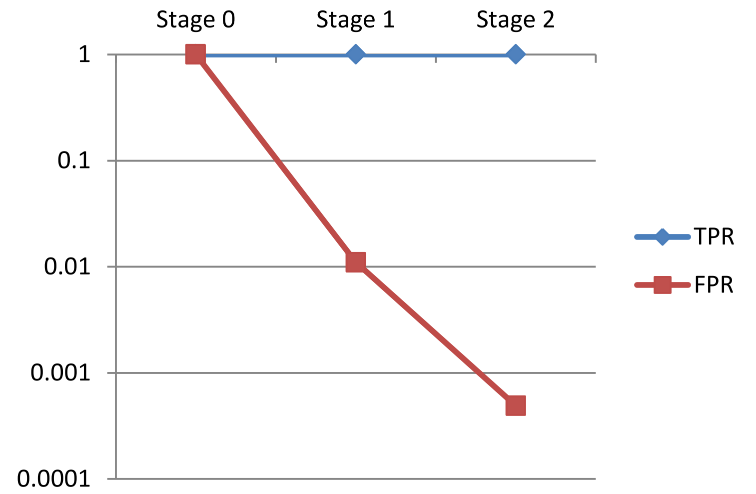

Figure 7: **TPR** vs **FPR** plot

- As expected the FPR drops logarithmically which is expected for classifiers using boosting due to more features being added at every stage, thus reducing the likelihood of an object to contain all the features and implicitly being classified as a detection.
- We can observe that the TPR stays 1 on every stage meaning that on the training set, the detector correctly classifies all dartboards. This means we should expect the TPR to be generally high on future tests.
- If adding more stages the FPR will continue to drop, but at an increasingly lower rate and after a certain number of stages the TPR will also start to drop

While the graph predicts the false positives will decrease with each training stage, it does not accurately estimate the number of such detections due to not offering any useful information about the true negatives. With regards to the true positives, it predicts that most of the dartboards will be classified correctly, but we should expect some misclassifications because of the potential variability of the test data.

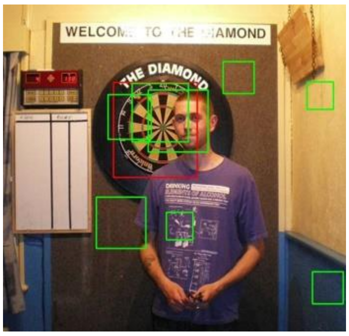

Figure 8: dart7.jpg

Figure 9: dart15.jpg

True positives (TP)	13
False positives (FP)	117
False negatives (FN)	7
True positive rate (TPR)	0.65
F1 score	0.17333

Figure 10: overall performance results of Viola Jones detector

Figure 11: dart3.jpg

Green detections = false positives
Blue detections = true positives
Red rectangles = ground truth information

Figure 12: dart4.jpg

When running the detector on our test set, both the TPR and the number of false detections have in fact much worse values than desired. We hoped to obtain a high TPR on the given test data and a relatively small number of false positives. Nevertheless, the resulted TPR turned out to be quite low compared to the expectations, many dartboards being missed. Moreover, the number of false positives is very large and needs to be refined further in order to obtain a meaningful classification. Also, after re-examining the detections produced by the algorithm, we found the following situations to be the most problematic:

- viewpoint changes (illustrated in *Figure 9* where the dartboard is missed because the viewpoint is shifted)
- lighting changes (situation captured in *Figure 12* where the upper half of the image is better illuminated resulting in poor classifications)
- partial occlusion (exhibited in *Figure 8* and as well as in *Figure 12)*

For this task we have based our detector solely on the evidence provided by the Hough-Transform. It is worth mentioning that, in order to handle the cases where dartboards have elliptical shapes, we run the Hough-Circle detection algorithm twice: once on the original version of the image and once on a version scaled in the *y* dimension in order to transform the ellipses in the images into circles enabling us to detect them using the same algorithm.

Initially, we considered using Viola Jones as a basis to be refined by the use of the Hough Space. However, we observed that some of the dartboards detected by the Hough Transform were not being found by Viola Jones at all, one such case being that *Figure 12* vs. *Figure 19*. Furthermore, by using our method of interpreting the Hough Space we managed to obtain a better TPR than that of the detector trained in part 2 whilst dramatically reducing the number of false positives.

Rationale for combining evidence:

- We start by running the Hough Transform on the non-scaled image as this correctly detects 15 out of the 20 dartboards in the test set whilst accounting for only 3 false positives
- We then run the circle detection algorithm on a scaled version of the image and we add new positives to the total set of detections only if they do not overlap any of the detections from the first step. This acceptance condition has the effect of limiting the slightly larger number of false positives produced by the second phase of our detection algorithm. However, it also has the drawback of eliminating a true positive generated in the second phase because it overlaps a misclassification produced during the first phase.

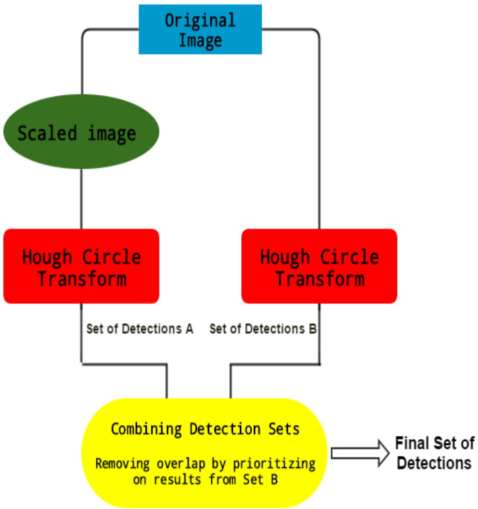

Figure 13: Flow diagram of the detector components

Figure 14: thresholded magnitude of dart11.jpg

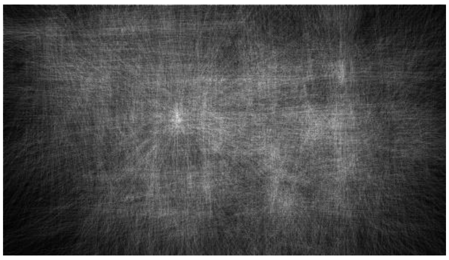

Figure 15: Hough Circle Space of dart11.jpg

Figure 16: dart11.jpg

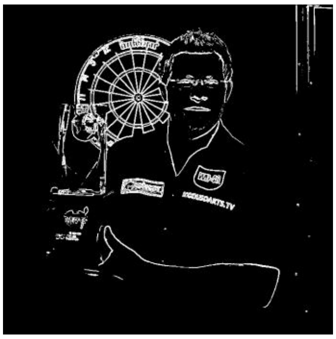

Figure 17: thresholded magnitude of dart4.jpg

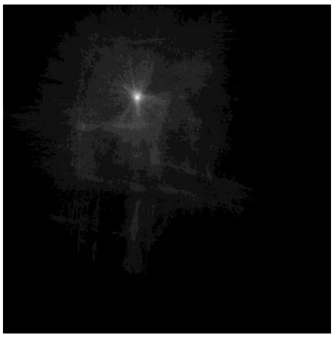

Figure 18: Hough Circle Space of dart4.jpg

Figure 19: dart4.jpg

Merits of this approach:

- Dartboards subject to viewpoint changes are detected
- The majority of the elliptical dartboards are detected
- A very limited number of false positives on the given test set, so an improved F1 score

Drawbacks:

- Slower than Viola Jones, particularly when the size of the image is large
- Does not work well if the image is blurred or the dartboards are poorly lit as in these cases edges may not be picked up
- Does not work well when half or more of a dartboard is occluded

TP	17
FP	3
FN	3
Precision	0.85
Recall	0.85
F1 score	0.85

Figure 20: overall performance results

Having evaluated our detector based on the Hough transform, we needed to find a way to reduce the number of false positives during the first stage of the algorithm and to improve detection performance in the problem cases identified in the previous part.

Rationale for enhancing the algorithm by detecting points where multiple lines intersect:

The effectiveness of this method stems from the fact that dartboards have many lines that intersect at the centre. Therefore, we modified the circle Hough Transform algorithm in order to detect the kind of intersections mentioned above.

The idea behind the circle detection algorithm presented in lectures is to use the information contained in the gradient to vote for the centres of possible circles. From the relationship between the gradient direction and the edge direction, which says that the edge direction is perpendicular to the gradient direction, we observed that it was possible to modify the Hough transform for circles such that each point in 2-d space would vote for a point along the edge direction at that location in the image rather than along the gradient direction. To do this, we add $\pi/2$ to the gradient direction used in the formula of the Hough transform. By making this change, each pixel in the image now votes for points along the edge direction located at distances between *minRadius* and *maxRadius* away. Therefore, we can observe strong peaks only in places where multiple lines intersect. This last fact is very useful in the detection process as the points on the strong diagonal edges of a dartboard contribute a large amount of votes towards the centre of an area occupied by it.

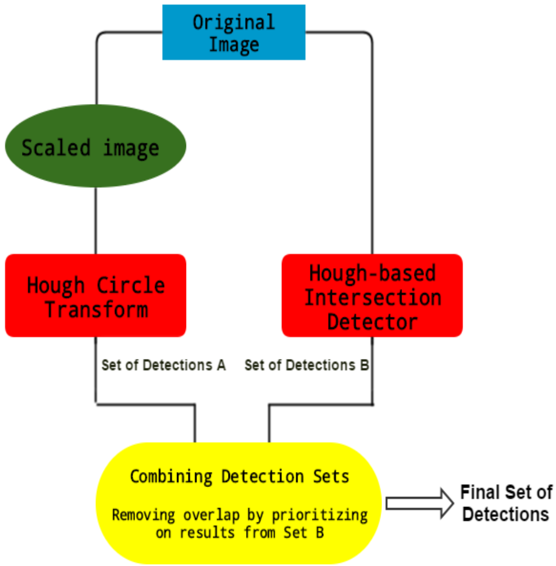

Figure 21: Flow diagram of final detector

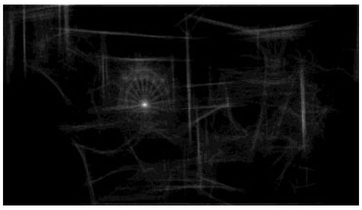

Figure 22: Hough Intersection Space of dart11.jpg

Figure 23: dart11.jpg

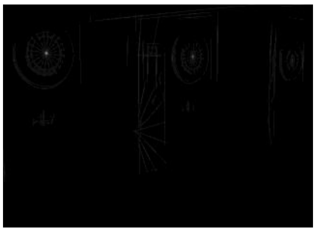

Figure 24: Hough Intersection Space of dart10.jpg

Figure 25: dart10.jpg

It is immediately obvious from the Hough spaces of the two images above that by using this technique the centres of each dartboard are well differentiated from the rest of the space whereas the space generated by the unaltered Hough transform contains many more peaks to filter through. Therefore, it is much easier to pinpoint the exact locations of the dartboards, and the likelihood of generating false positives is drastically reduced. This is especially noticed in cases with occlusion (e.g. *Figure 22*).

Benefits of this approach:

- Faster than the previous algorithm as it does not need to consider as many radiuses due to the intersecting lines being much stronger than the circles in the test images, thus enabling more relevant points to contribute with votes
- It detects all partially occluded dartboards in the given test set
- It performs much better than the circle detector in cases where ellipses are present because the diagonal edges inside the dartboards are still strong enough even if the object shape is elliptical
- Does not produce any false positives on the test set given and reduces the number of false negatives to only one, so F1 score is further improved

TP	19
FP	0
FN	1
Precision	1
Recall	0.95
F1 score	0.97

Figure 20: overall performance results

Drawbacks:

- It still has problems to deal with blurred dartboards
- It has difficulties dealing with very large dartboards because, in that case, the diagonal lines do not intersect at the centre anymore meaning the classification will result in false negatives
- It will not detect dartboards which have weak intersecting edges
- It does not perform well on dartboards which are almost fully occluded

Further improvements:

In order to increase performance further, we considered how it was possible to deal with blurring in the images. One possible way was to sharpen the edges by subtracting a blurred version of the image from the original image. This has the effect of enhancing edges and, indeed, resulted in the correct classification of the leftmost dartboard in dart8.jpg, the last false negative, but it also resulted in a number of false positives on the images with many intersecting lines such as dart14 or, where the dartboard is very strongly illuminated initially, as in dart10.jpg. We ultimately decided against using the sharpening idea in the final submission however, this is one method of pre-processing the images which is worth exploring if the test set for the detection task contains blurry images.