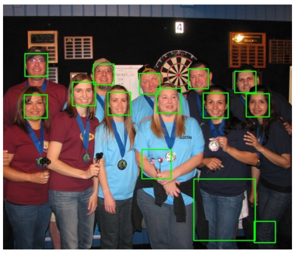
CW2: The object Detection Challenge

Group 80 Carina XU (zx16649) Boyang ZHAO (bz16563)

SUBTASK 1: THE VIOLA-JONES OBJECT DETECTOR

In this task, five pictures are detected using frontalface.xml. The result is shown as follow.

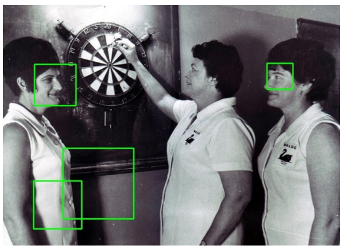
a) dart 4



b) dart 5

c) dart 13

d) dart 14



e) dart 15

Figure 1. Face detected images

The TPRs of dart5.jpg and dart15.jpg are calculated as follow.

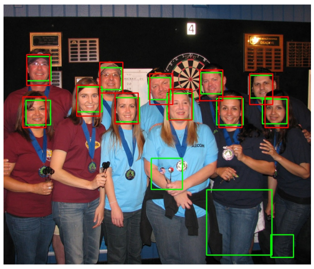
$$TPR = {}^{TP}/_{POS}$$

- Dart 5: TPR = 1
- Dart 15: TPR = 0.666667

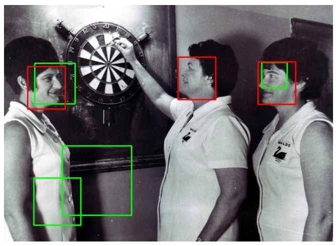
The main difficulty during calculating TPR is to set TP. It was hard to decide Whether it can be defined as a positive one when the bounding box draw around a part of the face.

It is always possible to achieve a TPR of 100% because the detector can always detect all faces even with low accuracy (i.e. other things are also marked as face). Thus, the ground truth is used to calculate the TPR and F1-score for dart5.jpg and dart15.jpg.

The ground truths of dart5.jpg and dart15.jpg are shown in figure 2. The red rectangles are the ground truths we set.



a) dart 5



b) dart 15

Figure 2. Ground truths of the detected images

The overlap area between the detected rectangle and ground truth is calculated to ensure that they are true positive (i.e. they are faces we want to detect).

Then the F1-score is calculated using following function. $F1Score = \frac{2\times({}^{TP}/_{DP}\times TPR)}{{}^{TP}/_{DP}+TPR}$ where DP is detected positive. The test on dart15.jpg is shown as follow.

```
FP2
TP2
TPR = 0.666667
F1 score = 0.571429
Program ended with exit code: 0
```

Figure 3. Test run output for calculating F1-score in dart 15

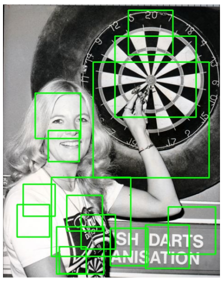
SUBTASK 2: BUILDING & TESTING YOUR OWN DETECTOR

There are three different stages during the training, the TPR and FPR of each stage are shown in the following table.

	0-Stage	1-Stage	2-Stage
TPR	1	1	1
FPR	1	0.038956	0.00047313

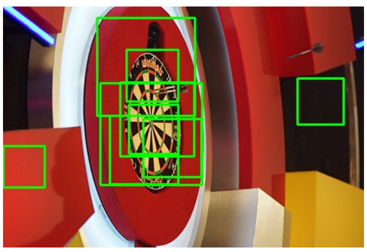
Table 1. TPR vs FPR for three stages

In the training process, we only using positive picture. As a result, all TPRs for three stages are 1. In the 0-Stage, the boosting procedure only classified all image as positive. Thus, FPR is also 1. When it comes to 1-Stage and 2-Stage, sampled patching from negative images, only few of negative images are classified as positive (i.e. find the dart in the image containing no dart boards). The accuracy is improved by stage during the training. After the training procedure, a classifier is created and store in the file cascade.xml. After using this to detect the dart in the image, some of the results are produced in Figure 4 as follow.

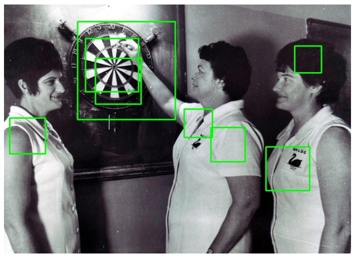


a) dart 9

b) dart 10



c) dart 12



d) dart 15

Figure 4. Dart detected images

The F1-score of all result images are calculated as follow.

$$F1Score = \frac{2\times\left({TP}/{DP}\times TPR\right)}{{TP}/{DP} + TPR}$$

where DP is detected positive

- Dart 0: F1-score = 0.11
- Dart 1: F1-score = 0.25
- Dart 2: F1-score = 0.33
- Dart 3: F1-score = 0.29
- Dart 4: F1-score = 0.18
- Dart 5: F1-score = 0.08
- Dart 6: F1-score = 0.11
- Dart 7: F1-score = 0.05
- Dart 8: F1-score = 0.11
- Dart 9: F1-score = 0.13
- Dart 10: F1-score = 0.13
- Dart 11: F1-score = 0.75
- Dart 12: F1-score = 0.2
- Dart 13: F1-score = 0.55
- Dart 14: F1-score = 0.05
- Dart 15: F1-score = 0.67

The TPR in the testing data is lower than the one in training set. On the other hand, the FPR is higher. Although the TPR and FPR in table 1 show high accuracy of the classifier, it has low F1 score on testing data. High accuracy in training data cannot prove high efficiency in testing data.

SUBTASK 3: INTEGRATION WITH SHAPE DETECTORS

Figure 5 illustrates the two of the examples detected by the detector.

a) dart 13

b) dart 15

Figure 5. 2D representation of the Hough Space, the thresholded gradient magnitude image, and result images

The overall F1-score for all text images is

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} = 2\times\frac{{}^{14}/_{31} \times {}^{14}/_{20}}{{}^{14}/_{31} + {}^{14}/_{20}} = 0.60$$

The key merits and shortcomings of the implementation are listed as follow.

KEY MERITS

- Lines on the dart can be easily detected. Easy to find the intersected point of the lines. The accuracy of the detector is improved.
- The logic of the programme is simple, and have clear structure.

SHORTCOMINGS

- It relies too much on canny calculator so that it can only detect edge in clear images.
- If other part in the images has more lines intersect in one point than the dart. It is detected as the dart as well.

In this subtask, Hough line detect is used to find the centre of the dart. The process is shown in the flow chart.

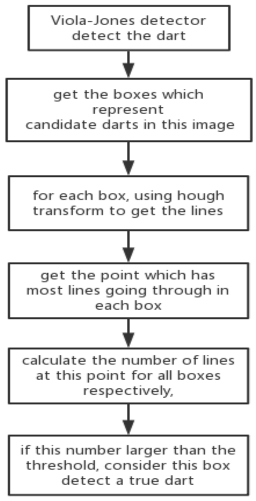


Figure 6. Processes combined Hough transform and Viola-Jones detector

- The box contains the dart has typical concentric lines, so this feature can be used to detect dart.
- Use Hough transform to locate each line in the box
- The point which has most lines going through is probably the centre of dart
- Compare the number of lines at this

probable point in each box and make the threshold equal to the maximum among them minus two. This won't miss the true dart if there are other points have more lines than the dart.

SUBTASK 4: IMPROVING YOUR DETECTOR

In this subtask, the detector is improved to detect the blurry dart. The rationales are listed as follow.

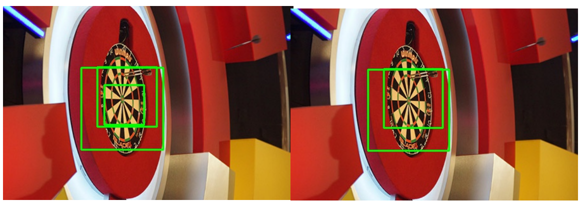
- Reinforce the light intensity and the contrast ratio to make the image more sharpen and light.
- Calculator for the detector is

$$g(x) = \alpha f(x) + \beta$$

- This approach can make the edges more sharp and clear, so the line detector can achieve a better performance. Which means more lines will go through the centre of the dart than others.
- Light process has an effect to erase and cover the noisy pixel in the picture.

In this task, the detection is improved compare to subtask three.

a) dart 3



b) dart 12

Figure 7. Detected images compare to subtask 3

The overall F1-score for all text images is

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$= 2\times\frac{{}^{16}/_{26} \times {}^{16}/_{20}}{{}^{16}/_{26} + {}^{16}/_{20}} = 0.70$$

Comparing to subtask 3, the overall F1-score has increased by 0.1.

The key merits and shortcomings of the implementation are listed as follow.

KEY MERITS

- Almost perfectly detect the dart area by the line detector.
- With increase the light intensity and contrast ratio, the line becomes more easy to recognize and locate.

SHORTCOMINGS

- Rely too much on the lines detector which is unable to recognize a precise shape area.
- Some picture with more edges can be tough noises for the detection.