CW2 : Object detection task

Minzhi Fan & Sonny Peng

Subtask1 : The Viola-Jones Object Detector

Figure 1 : face detectors of picture 4,5,13,14,15

These five pictures are detected using frontalface.xml.The TPRS of picture 5 and 15 are calculated by following:

TPR = TP/TP+FN ⟶ Dart5 : TPR = 1; Dart15:TPR = 0.6667

The difficulty during calculating TPRs is it is hard to determine which is true positive. It is hard to decide if it is the positive one when a bound box is drawn around a part of face. And the reason why it is challenging for computer is that the colours of skin, facial expressions, light .etc, which have great impact on the accuracy of detection.

It is always possible to achieve a 100% TPR because the Viola-Jones face detector always can identify every part of the image such that every object is detected. But the accuracy may be low , eg, some non-facial parts also marked as face, which leads to high false positive rate(FPR). A better detector should have 100% TPR and low FPR.

There are several rules created in order to calculate F1 score:
1.A face only considered valid only if the mouth and eyes are visible and it must be a human face.
2.The bound box must circle out the eyes and the mouth fully and the size of box cannot be larger than 1.5x dimensions.
F1 score can be calculated using formula:
F1 = 2TP/(2TP+FN+FP) -> dart15 : F1 = 0.57; dart 5 : F1 = 0.88
We can also use ground truth to ensure true positive by calculating the overlap area between detected faces and setting ground truth.

Subtask2 : Building & Testing your own Detector

There are three different stages during training. TPR and FPR is shown in the following table. In training process, only positive images are used so that all TPRs are 1 which means all objects are detected. At stage 0, all pictures are classified as positive so that FPR is 1. At stage 1 and 2, only few of negative images are classified as positive, thus the FPR is low. The accuracy is improved stage by stage during training and finally a classifier is generated, cascade.xml. And following images use this classifier to detect.

	TPR	FPR
Stage 0	1	1
Stage 1	1	0.0187
Stage 2	1	0.00151

Table 1 : TPR and FPR in training

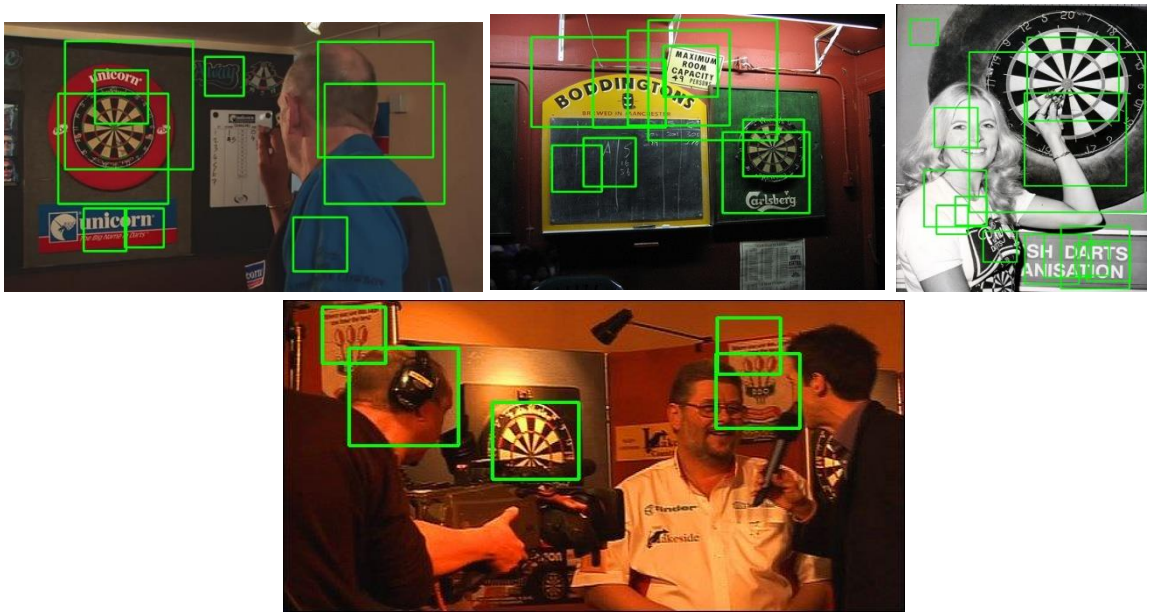


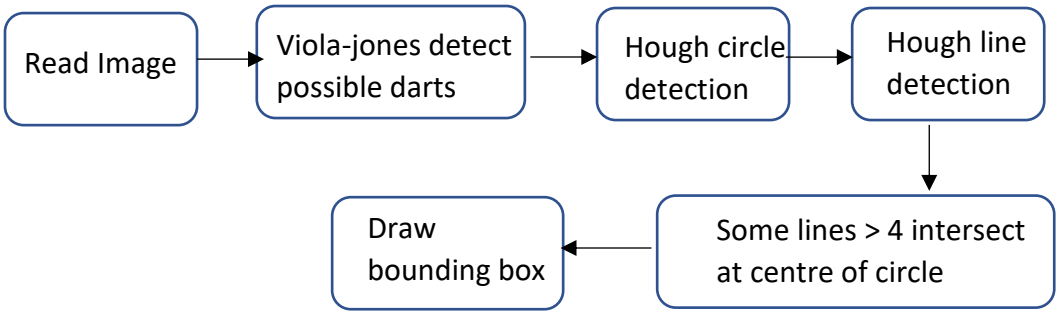
Figure 2: detectors using cascade.xml

The F1 score of all testing images are calculated below. F1 score looks at more factors ,thus it is more important in measuring the performance.

As shown in the table, F1 score is low, because many non-darts areas are detected as dart as well. And the TPR in testing data is lower than it in training data(high FPR).Even though there is high accuracy of classifier in training data due to table 1,but it cannot ensure high performance in testing data

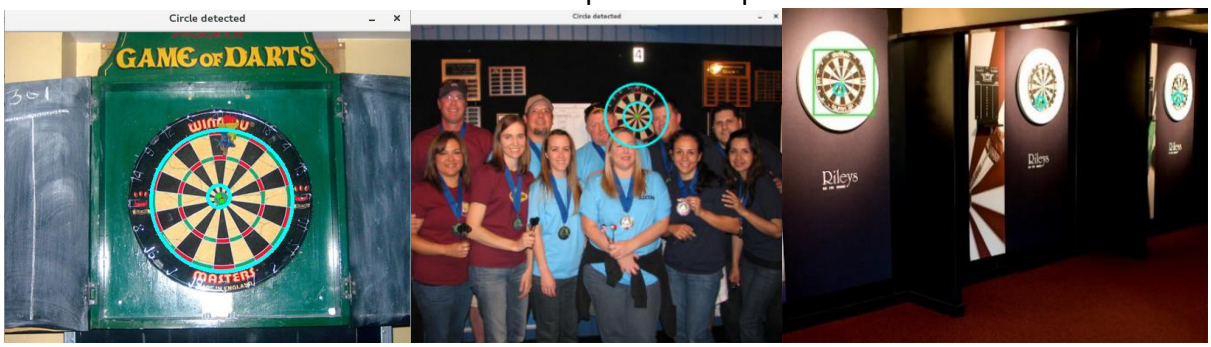
	F1 Score		F1 score
Dart1	0.18	Dart9	0.142
Dart2	0.2	Dart10	0.113
Dart3	0.2	Dart11	0.33
Dart4	0.095	Dart12	0.2
Dart5	0.11	Dart13	0.18
Dart6	0	Dart14	0.068
Dart7	0.045	Dart15	0.25
Dart8	0.041		

Subtask3 : Integration of shape detectors



A dart contains a circle and 20 lines, so we decided to combine Hough circle and line transform to establish the shape detector in order to improve the performance of subtask 2. Flowing is the flow charts.

- We use Viola-Jones to detect possible darts and generate boxes around them, but it produces lots of boxes which results in a high FPR.
- Hough circle is used to remove the boxes which do not contain a centre and a circle.
- We combine Hough line detection which detects many lines, only boxes contain more than four lines intersect at one point are kept. That is the final dart.



Picture	F1 Score	Overall F1 Score
1	1	
2	1	
3	0.286	0.32

In overall, our detector can detect most of the dart from the picture, even if it is obscured by other objects, this is because we combined high and low tolerance Hough circle detections on the picture, which ensure that both Hough circle detection is detecting on the same object which is centred at the same position. However, we failed when we tried to combine Hough line detection and Hough circle detection, which is shown from the third image. The overall F1 score is not good enough but still a good improvement from subtask2.

However, it is not so efficient when dealing with rotated targets, the reason is that our detector is restricted on pre-defined shape (line and circle) with only limited toleration, although we can increase the range of pass rate, but the result will often be under-fitting. Thus, if the target is too tilted, the circle detector will not allow it to pass, hence the low performance on the third picture.