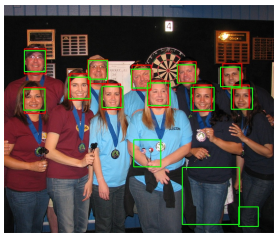
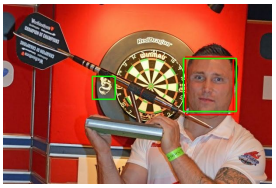
(a) dart4

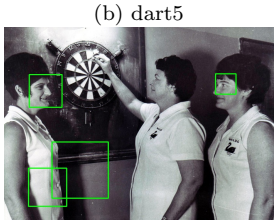


(b) dart5



(c) dart13

(d) dart14



(e) dart15

Figure 1: Frontal face detection with ground truth

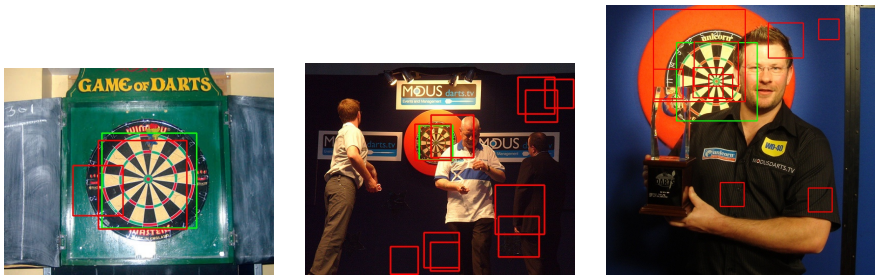
1 Face detection

TOGO Limitation of TPR. Difficult to define "true" values. The bouding region of a face is discrete and we may choose not to detect side on faces. Example below shows 2 possible "true" values for a detection. One give a IOU of under 0.7 but the other gives over 0.9.

Frontal face detection results		
Image name	TPR	F1-SCORE
dart4	1	1
dart5	1	0.88
dart13	1	0.666667
dart14	1	0.5
dart15	1	0

The results show that the true positivity rate, TPR, for all images is 1. Despite this we often see far lower F1 scores. This reveals the limitation of the TPR as a detector can always achieve a high TPR by having a low tolerance to classify something as a face. This would result in a very low F1 score as many false positives would arise yet the true positivity rate would remain high.

The image dart-15 reveals another limitation of these metrics. It can often to difficult to classify the ground truth. This classifier was created to detect frontal faces and therefore in dart15 I have claimed no frontal face present. Despite this the classifier has detected on of the faces brining its F1 score down. Similarly there is no clear boundry for a face and therefore the exact area that



(a) dart1 (b) dart6 (c) dart4

Figure 2: Dart board detection with ground truth

should be detected can vary resulting in the need for a larger tolerance for the intersection over area, IOU.

2 Dart board detection

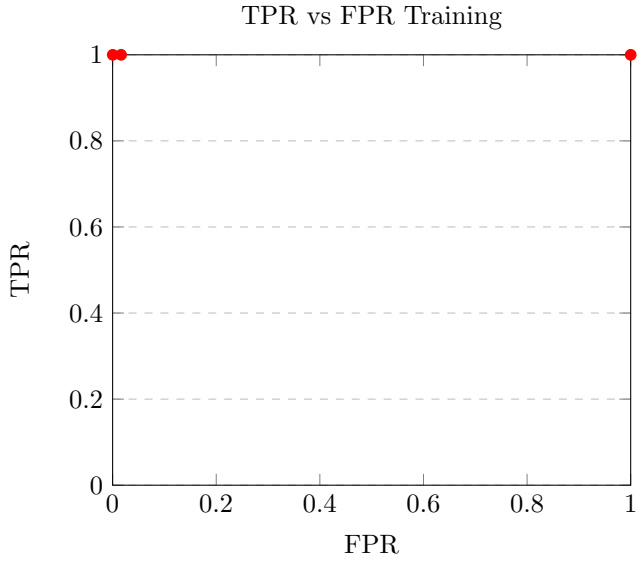


Figure 2 shows the change in FPR and TPR throughout the training process. At the beginning of the training all images are accepted throughout the stages more layers are added to the cascade. Each layer will remove of the regions classified as a dart board. This results in a drop in FPR as the training takes place. Throughout the training process the TPR remains and 1. However if more layers were added (and therefore the classifier removed more classified regions) then the TPR could also be reduced. This could still result in the classifiers F1 score increasing as the FPR would also be decreasing.

_____	Frontal face detection results	
Image name	TPR	F1-SCORE
dart1	1	0.666667
dart2	1	0.25
dart3	1	0.4
dart4	0	0
dart5	0	0
dart6	1	0.181818
dart7	0	0
dart8	0	0
dart9	0	0
dart10	0	0
dart11	0	0
dart12	0	0
dart13	0	0
dart14	0	0
dart15	1	0.5
Average	0.333333	0.133232

The average values in Table 2, show the unreliable results of the cascade for detecting dart boards. It not only has a very low TPR and also a very low F1-score suggesting there are many missed dart boards as well as regions falsely classified. The images in Figure **??** show the wide range of results from the classifier. a and b show successful classifications yet b and c show the high number of false positives. It is also noteworthy that the TPR values achieved during training were far higher than the average achieved during testing. This is as a result of the method used to generate positive images. These took one image of a dart board and generated variations (via movements/rotations) on that image. This means when classifying different types of dartboards of those with objects in the way (such as 2 c) the classifier performs far worse.