# The Object Detection Challenge

## The Viola-Jones Object Detector

The field of Image Processing always raised the problem of performance in the case of face recognition due to large formats of data it has to iterate through. The first real-time detection method was the Viola-Jones object recognition framework, that provided a scalable solution to the problem in question and introduced the concept of a sliding window. To detect the presence of the object, each sliding window was integrated and sampled for chosen Haar features. Moreover, the accuracy of this method was improved with the aid of a strong classifier trained using AdaBoost, a greedy algorithm whose purpose is to select the optimal weight features on the basis of the resulting errors.

Figure1: *The detected representation of Dart4, Dart5, Dart 13, Dart14, Dart 15 images*

To assess the performance of the classifier we tested out the given algorithm on a set of images and then calculated the true positive rate(TPR) for each of our images that takes into account the number of positive labelled elements and the number of true positives.

However, there are some practical difficulties in evaluating the true positive rate as being accurate due to the broad variation in which this objects might be found. This is why the models used to train the classifier should be clearly defined prior to their assessment. Therefore, firstly we need to decide upon a set of measures and rules that define a face. To make the task in hand more feasible, the Viola-Jones framework requires full view of frontal upright faces that can impose some constraints, but for the sake of our reasoning we considered to be positive all representations that contain at least 50% of the faces. Consequently, we obtained the TPR of value 100% (11/11) for the image Dart5 and 60% (2/3) for the image Dart15.

Secondly, the TPR formula does not take into account the number of false positives which gives us the possibility of achieving  a TPR of 100% on any detection task. This can be achieved by selecting a weaker threshold such that all elements that should be labeled as positives will be recognised. Although this method will maximise the TPR, it will also make our classifier prone to detecting more false positives by allowing more noise to affect the output.

Consequently, we choose to assess the performance of our classifier by calculating instead the F1 score where *F1 = 2 *(precision * recall)/(precision + recall)* given *precision = number of correct positives/all positives returned* and *recall = number of correct positives/number of positives that should have been returned.* So for the *Dart5* and *Dart15* images we obtained 0.87 respectively 0.(66) , where an $F_1$ score reaches its best value at 1 and worst at 0. It is worth observing that the new computed accuracy score of Dart5 image decreased from 100% with TPR to 87 % with F1 score which describes much better the results obtained with our classifier just by taking into account also the false positive rate.

**Building and testing the trained dart detector**

Figure 2: Error devepolment during training stages


Next we concentrate on understanding the AdaBoost algorithm behaviour given a set of positives and negatives samples of dart boards. The training positive set was generated by variating the viewing angle and the contrast of one main image allowing a better generalisations of the classifier.

Throughout the 3 stages of training within the boosting procedure, we were updated of the state of the classifier with the aid of the TPR and FPR of each stage as shown in *Figure2*. We can therefore observe that TPR is maintained at the value of 1 throughout all stages of training which implies that true positives are always detected. Moreover, the FPR decreases from 1 to 0.0016 after only 3 stages of training meaning the number of false positives detected at the end is very low which might lead to overfitting if the boosting procedure was to be extended further.

**Testing the classifier performance**

*(Figure3 : Dart2, Dart14, Dart10, Dart11 )*

In order to test the accuracy of the resulted TPR and FPR values during our boosting procedure we will compare them with the F1 score computed across all our 16 detected images. The F1 score of 0.165 (16.5 %) was obtained on the basis of a ground truth .

While observing the results of our classification we can notice some strong contradictions with the initial TPR and FPR values. Therefore, not all true positives are detected as it can be seen in the *Dart11* image, which can be due to the lack of variations within the positive training set.

Secondly, the number of false positives is higher than expected because of the lack of similarities between the two training sets provided. For instance, in the *Dart 2, 10* and *14* images we obtain the F1 scores of 9%, 15% respectively 10% due to an increasing number of false positive although the elements detected do not resemble the targeted ones. As stated before this will have a strong influence on the F1 score computed even if the TPR and the FPR don't quite include this in their calculation.

For the reasons above, we can strongly affirm that the TPR and the FPR do not even closely characterise the performance of our classifier. However, provided with more case specific training sets the performance of the classifier might get closer to our expectations.

### Integration of the Viola-Jones Object Detector with Circle Hough Transform

*(Figure4: Dart12, Dart5)*

In order to combine this new evidence with the output of our Viola-Jones detector we had to explore their advantages and disadvantages. While the Circle Hough Transform outputs no false positives, but fails to detect any ellipses, the Viola-Jones detector picks up every dartboard, but with an increasing number of false positives. For this reason, after merging the above two concepts we succeeded to label *Dart12* image correctly by comparing the values of the hough space described by the bounded areas even if the circle detection was firstly shifted. However, the limitations of our implementation relies on the fact that the hough values variate quite drastically between images making false positives hard to identify using a general hough threshold as it can be seen in the *Dart5* image.

### Evaluating the detector

Consequently to using a combined version of the two detection algorithms, we obtained a F1 score of value 0.476(47.6%) which considerably improved from 0.165 (16.5 %), value obtained when using only the Viola-Jones. This is due to reducing the number of false positives that however cannot be forward pursued because of the lack of generalisation. However, as later computed, the F1 score performed on the results produced by only the Circle Hough Transform is 0.864 (86.4%) which can be argued as being superior because of its confidence given by a not so permissive threshold. Yet we can decide between the two approaches described above depending on our ultimate purpose.

### Combining the detection algorithms

As pre-processing to our merge we choose to threshold the rectangles detected by the Viola-Jones algorithm taking into consideration the hough space in order to eliminate irrelevant information. Moreover, we also merged the circles produced by the Hough Transform to reduce the information processed with the assurance that they can intersect only when they describe the same dartboard and later converted them to rectangles to ease our computations. The next step was to test whether the separate evidence intersect and allow the converted circles to have higher influence due to their precision. The remaining squares were then merged together depending on their average hough space value and then added to the final output before being again thresholded for eliminating extra information that might have been derived during the process.

*Figure5 : flow diagram, depicting combined evidence from the Hough Transform and Viola-Jones detector*