

# Human face detection by Boosting technique

Chufeng Hu | UID: 504687715 | huchufeng@g.ucla.edu

The object of this project is detecting human face by AdaBoost and RealBoost. The training set contains 11838 human face images and 45356 non face images. All the training images are  $16 \times 16$ . The testing set contains four pictures taken in class. In this project, I first convert training and testing images into integrate matrix and save them as txt files in order to speed up the computation.

## AdaBoost

### 1. Construction of weak classifiers.

Generate classifiers with different size and location. Each classifier calculates a specific area on the data image and return a feature score. Classifiers are generated begin with (width = 2 & height = 2) to (width = 15 & height = 15), classifiers may be vertical or horizontal concatenated rectangles.

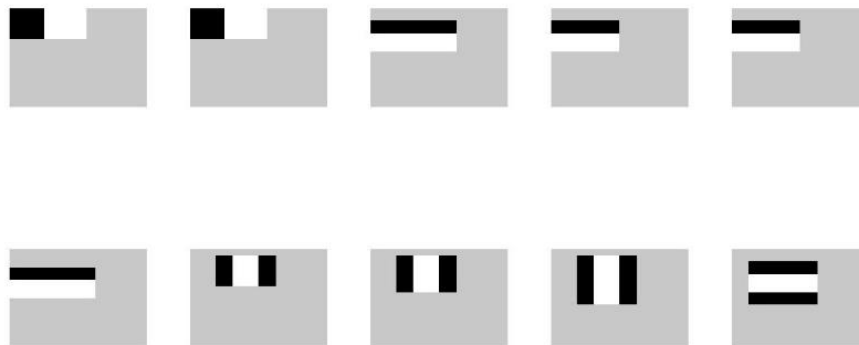


Figure 1

Figure 1 shows the top ten weak classifiers chosen by AdaBoost. Each classifier computes the difference of the sum of black and white blocks.

### 2. Training error of weak classifiers

At iteration  $T = 0, 10, 50, 100$ , plot the top 1000 training error of the weak classifier candidates.

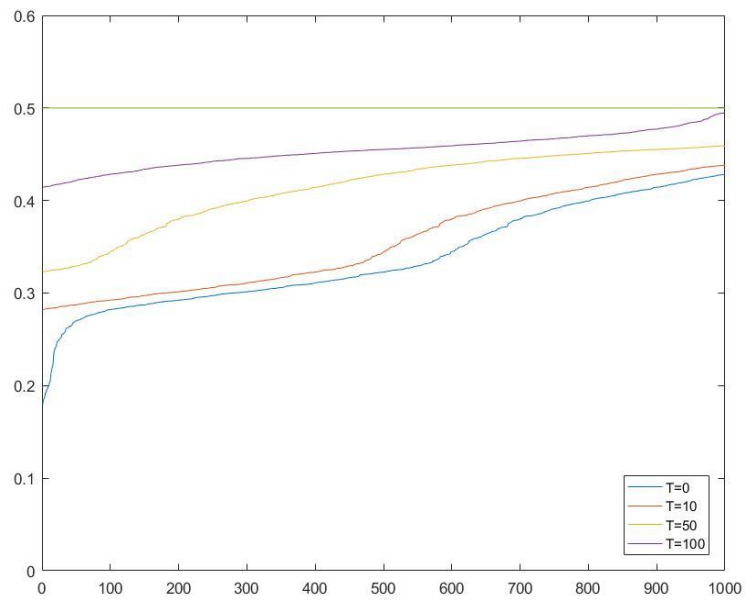
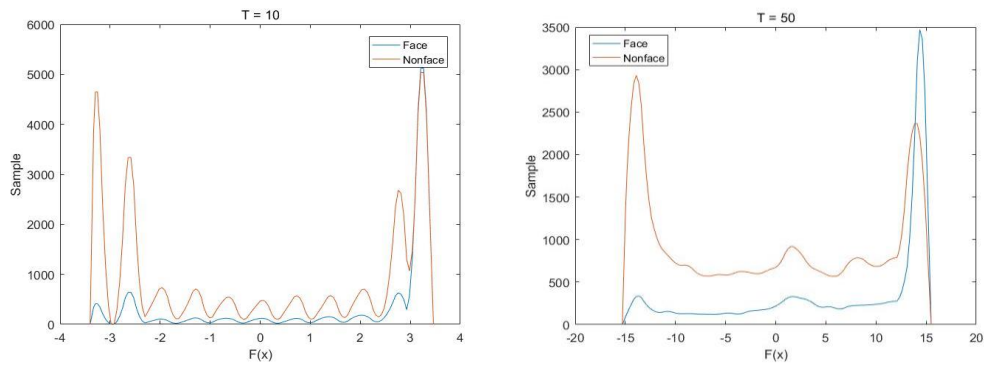


Figure 2

Figure 2 shows the training error of weak classifier candidates in different iterations. The error of remaining classifiers increases up to 0.5 as AdaBoost keep selecting minimum error classifiers. Algorithm stops if all the candidates have error close to 0.5.

### 3. Plot histogram of data

At iteration  $T = 10, 50, 100, 150$ , compute AdaBoost score for each image, and plot the histogram, the histogram and threshold of each classifier will change when the data weight change.



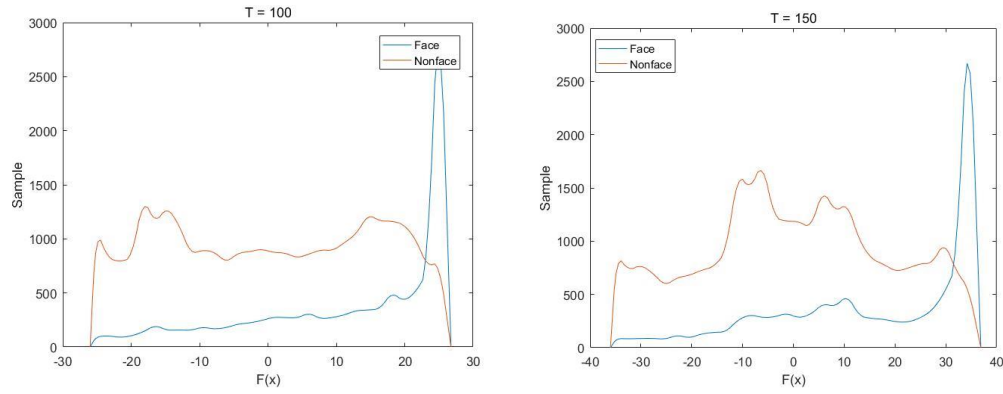


Figure 3

The distribution of face and non-face data become more separated as the iteration number increasing.

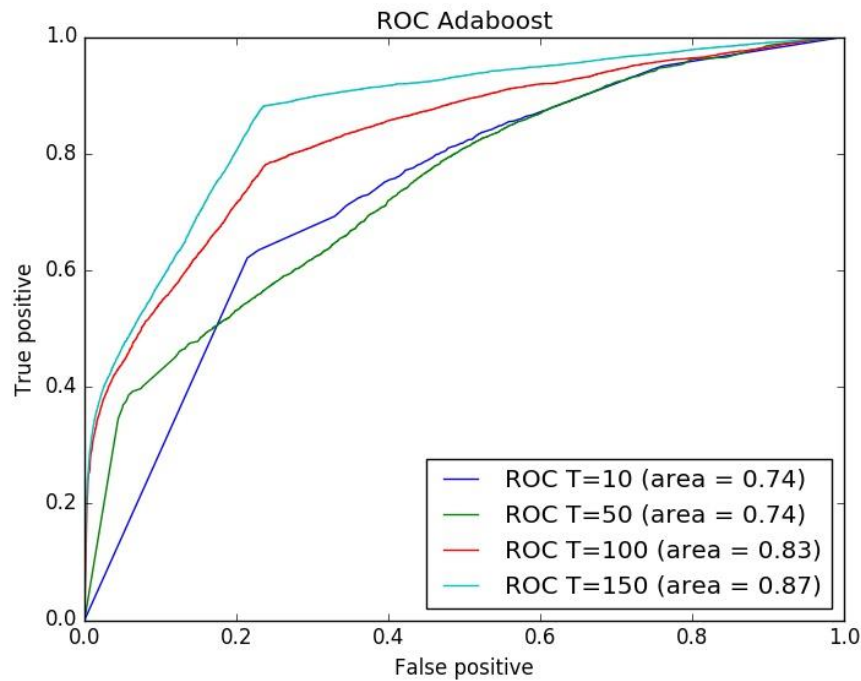


Figure 4

Figure 4 shows the ROC curve based on the histogram of data in each iteration  $T$ . Which gives us a presentation about the trade-off between “true positive” and “false positive” prediction. Based on the AUC, the performance of AdaBoost increases as we adding more classifiers.

## RealBoost

### 4. Histogram of RealBoost algorithm

RealBoost uses the classifiers, which are selected by AdaBoost in the previous steps. In this part, I choose 30 bins for one classifier, and the bin size is  $\frac{1}{30}$  of the range of the classifier's

score on training set. The data set could be more separated if a larger number of bins are used to simulate the projected space distribution.

The training data was projected on a one-dimension space, and I smoothed the shape of histogram in order to let them be more close to distributions.

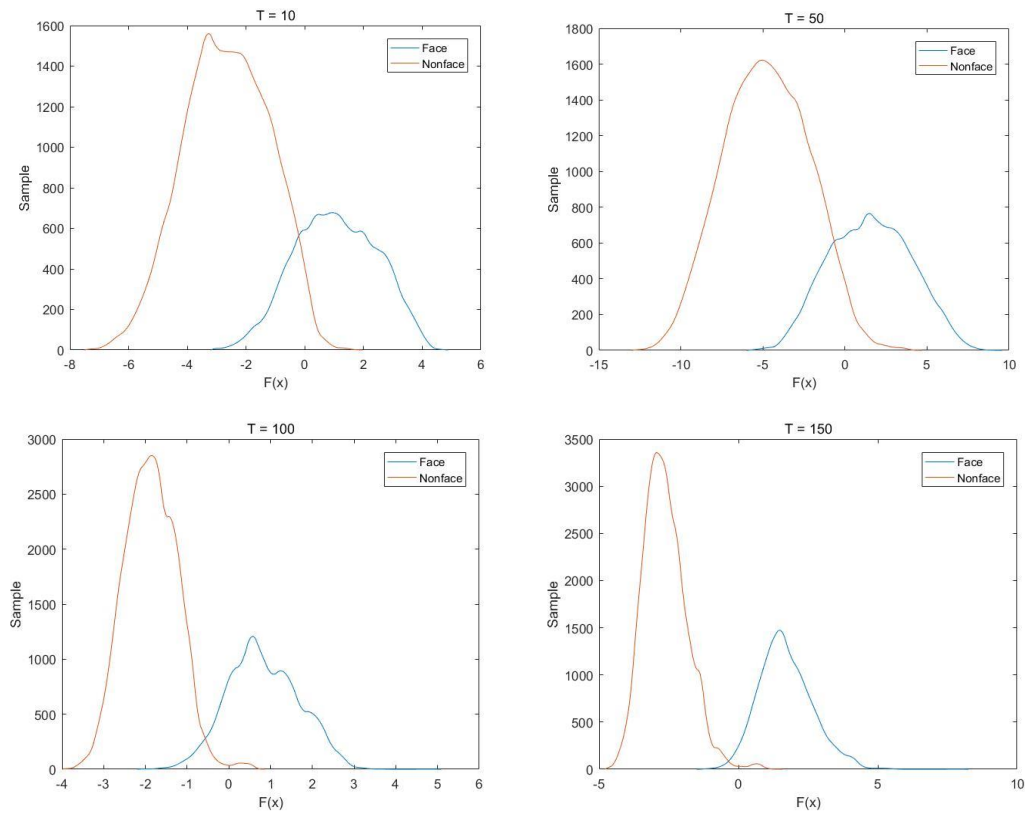


Figure 5

Compared with the histogram from AdaBoost, RealBoost separate face and non-face data better, and provides a wider margin.

## 5. ROC of RealBoost

Figure 5 shows the histogram from RealBoost, the two distributions become well separated as iteration number increasing. As RealBoost selecting more classifiers into the model, the accuracy of its prediction increases.

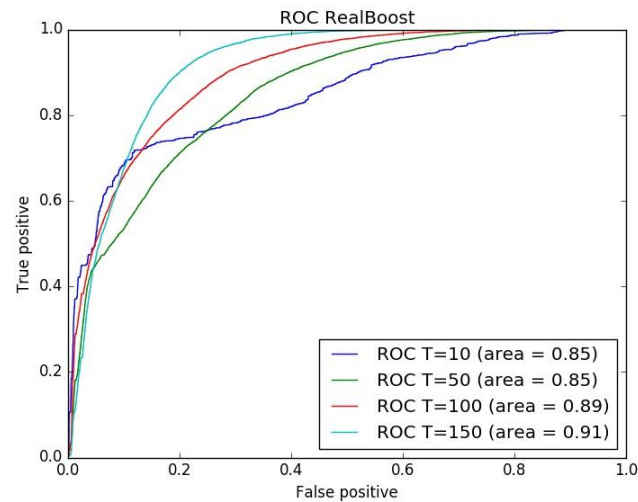


Figure 6

Figure 6 presents the ROC curve for different iterations in RealBoost. According to the AUC, RealBoost performs better than AdaBoost.

## 6. Face detection

To detect the test image, I first convert all test images and background images into grey scale map, then split the background image and add them into the training set.

When scanning the test image, I choose resolution from 64 to 128 by 16, which means there're 4 layers in the pyramid structure. The final face detection is made by AdaBoost and RealBoost with iteration = 150.

If rectangles are overlapped, compared these block and choose the one with max score. More details about the code is written in the "read\_me.txt" file.



Figure 7

Figure 7 shows the human face detection by AdaBoost. About 30 human faces are detected correctly and 10 faces are not recognized by this model.



Figure 8

Figure 8 shows human face detection by RealBoost, 25 face are recognized, and 6 faces are not recognized. Therefore, RealBoost performs better in the testing set. Both AdaBoost and RealBoost can detect human face in the test picture. Since RealBoost is more flexible than AdaBoost, it performs better in the final presentation.