CS 276A / STATS M231 Project 2

Yutong Lu, ID: 005027184

November 17, 2017

1 Implement AdaBoost

1.1 Haar filters

Display the top 20 Haar filters after boosting. Report the corresponding voting weights $\{\alpha_t : t = 1, ..., 20\}$. These top 20 Haar filters represent the top 20 explicit features of human faces.

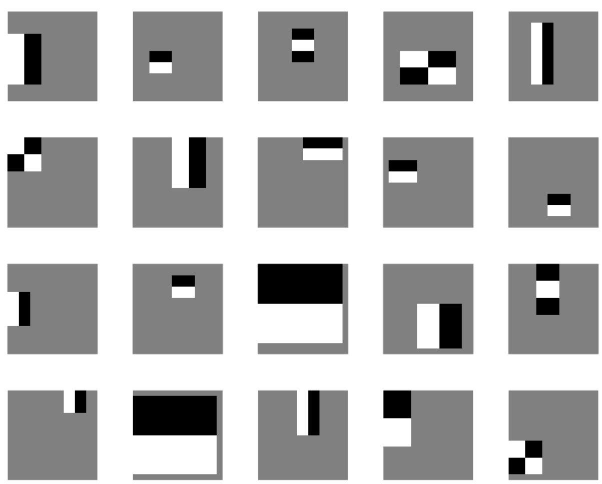

Figure 1: Top 20 haar filters

Knowing that $\alpha_t = \frac{1}{2} log \frac{1-e_t(h_t)}{e_t(h_t)}$, we can put α_t corresponding to the top 20 haar filters into a table.

1.1088	0.9643	0.8097	0.7355	0.6671
0.6634	0.5460	0.5741	0.5229	0.4757
0.4562	0.4507	0.4245	0.4215	0.4214
0.4172	0.3810	0.3820	0.3645	0.3464

1.2 Training error of strong classifier

Plot the training error of the strong classifier over the number of steps T.
It can be seen that with the increase of steps T, which simply means how many weak classifiers we have chosen, the training error of the strong classifier is decreasing. At the very first beginning, when we only have one weak classifier to build a strong classifier, the training error can be pretty large (over 0.3). When the step T is 100, we can see the performance of the strong classifier is pretty good that the error rate is below 0.1.

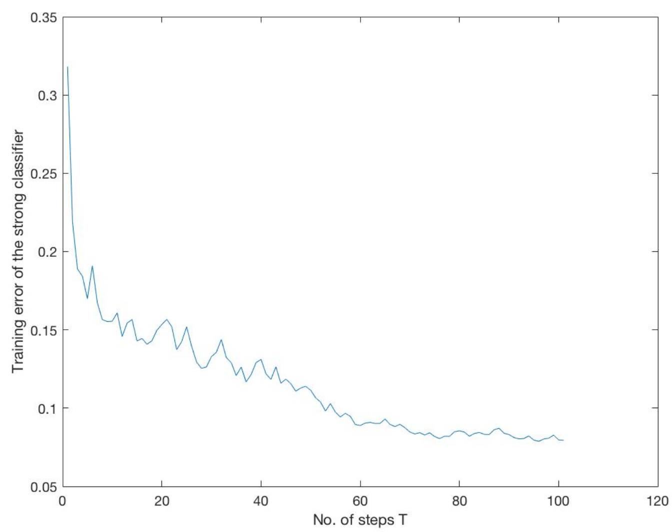

Figure 2: Training error of strong classifier

1.3 Training error of weak classifiers

At steps T = 0, 10, 50, 100, plot the curve for the training errors of the top 1, 000 weak classifiers among the pool of weak classifiers in increasing order. Compare these four curves. For each of these steps, as the no. of the weak classifier increasing, the training error is growing bigger. It is because the better weak classifier is selected with priority. Therefore it is sure that the error is larger and larger. When comparing among these four curves, it can be seen that a smaller step T has a smaller training error, and a very large step T (i.e. T= 100) could have a training error approaching 0.5.

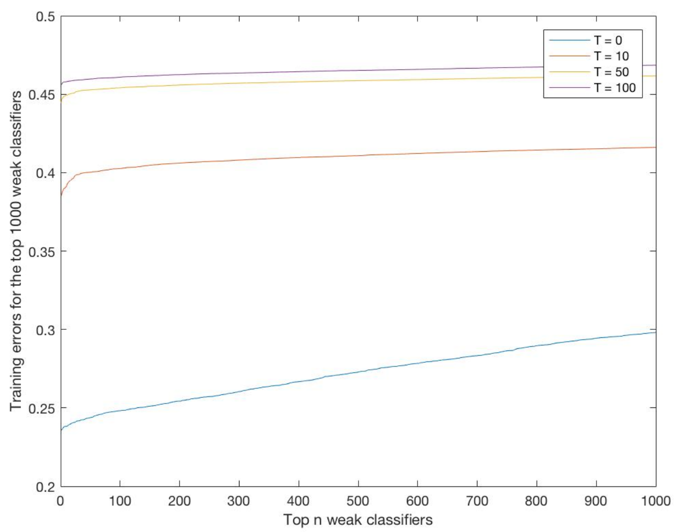

Figure 3: Training error of weak classifier

1.4 Histograms

Plot the histograms of the positive and negative populations over F(x), for T = 10, 50, 100, respectively.
When the step is getting bigger, it can be seen that two histograms representing positive faces and negative faces are more apart from each other, to be exactly, they will have less overlap.

1.5 ROC

Based on the histograms, plot their corresponding ROC curves for T = 10, 50, 100, respectively.

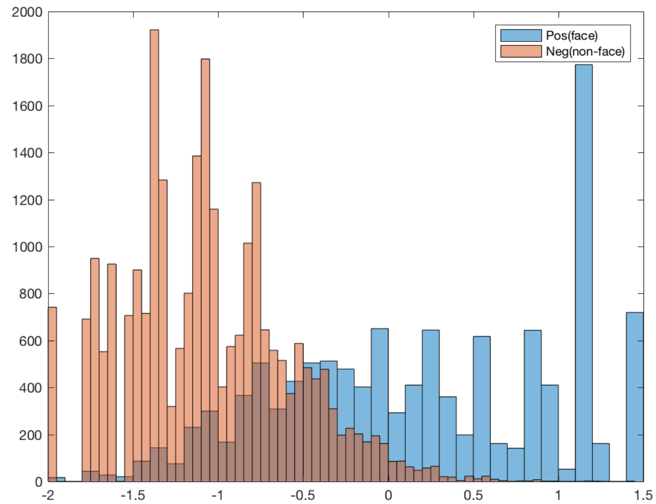

Figure 4: Histogram for T = 10

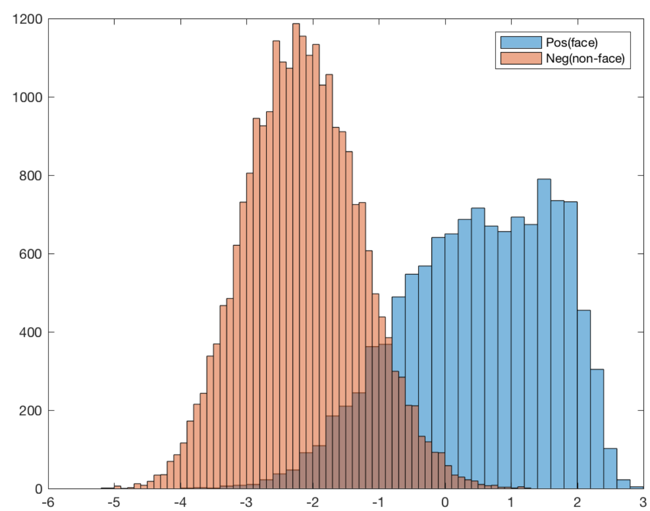

Figure 5: Histogram for T = 50

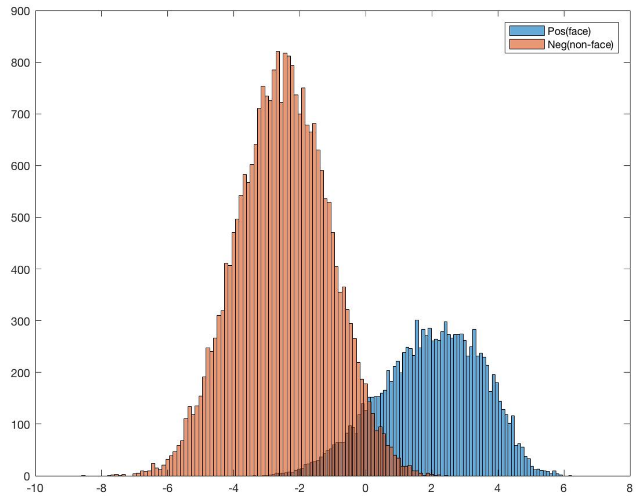

Figure 6: Histogram for T = 100

By showing three ROC curves for T = 10, 50, 100 on one figure, we can see clearly that a larger step number can get to a better ROC curve. A better ROC curve simply means the curve is closer to the left upper corner.

1.6 Face detection

Display the detected faces in both of the provided images without hard negative mining. Without negative mining, we just show the frame around an object if it is recognised as a face and the results for all three pictures are shown below.

1.7 Face detection with hard negative mining

Display the detected faces in both of the provided images with hard negative mining. After hard negative mining, the result of face detection should be better since hard negative mining is kind of cheating since we add some non-face data to our original features. The recognization results are shown below.

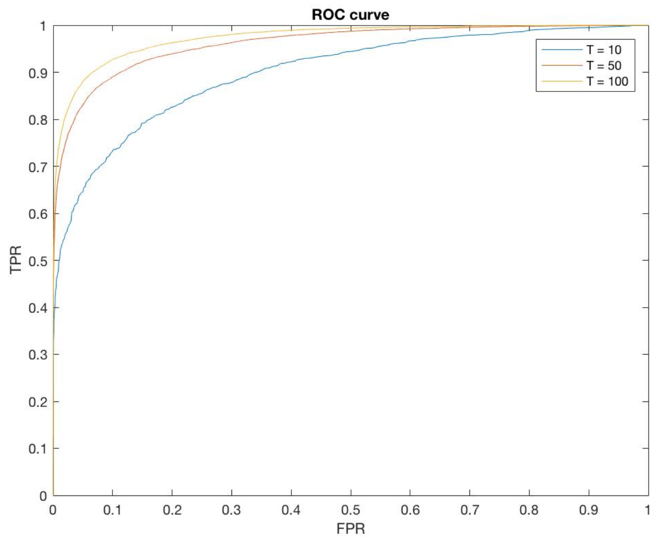

Figure 7: ROC for T = 10, 50, 100

Figure 8: Face detection for the first picture

Figure 9: Face detection for the second picture

Figure 10: Face detection for the third picture

Figure 11: Face detection with hard negative mining for the first picture

Figure 12: Face detection with hard negative mining for the first picture

Figure 13: Face detection with hard negative mining for the first picture

2 Implement RealBoost

2.1 Histograms

Plot the histograms of the positive and negative populations over F (x), for T = 10, 50, 100, respectively.
Generally, RealBoosting will have better results than AdaBoost since AdaBoost only fits the classifier using weights on the training data while RealBoost fits the classifier to obtain a class probability estimate. And this procedure makes RealBoost more accurate. From the histograms shown below and comparing them with histograms gained from AdaBoost, it can be seen that faces and non-faces are even more apart and have even less overlap.

2.2 ROC

Based on the histograms, plot their corresponding ROC curves. Compare them with the ROC curves in (e).
The ROC curve gained by RealBoost outperforms the curve gained by AdaBoost that all these three curves are more approaching to the left upper corner than curves of AdaBoost.

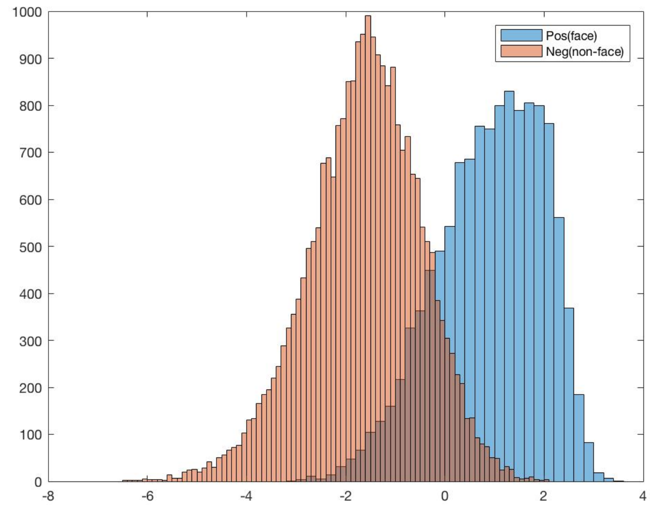

Figure 14: Histogram for RealBoost at T = 10

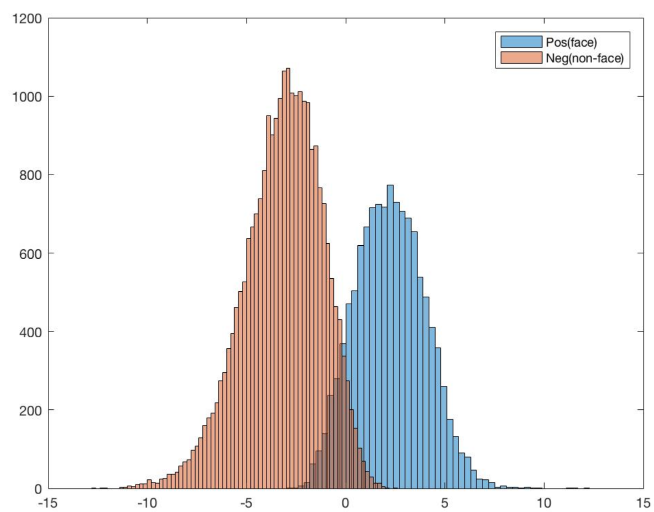

Figure 15: Histogram for RealBoost at T = 50

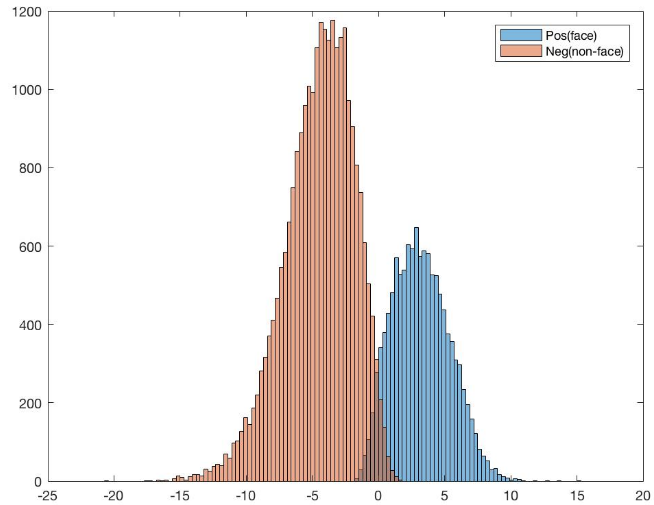

Figure 16: Histogram for RealBoost at T = 100

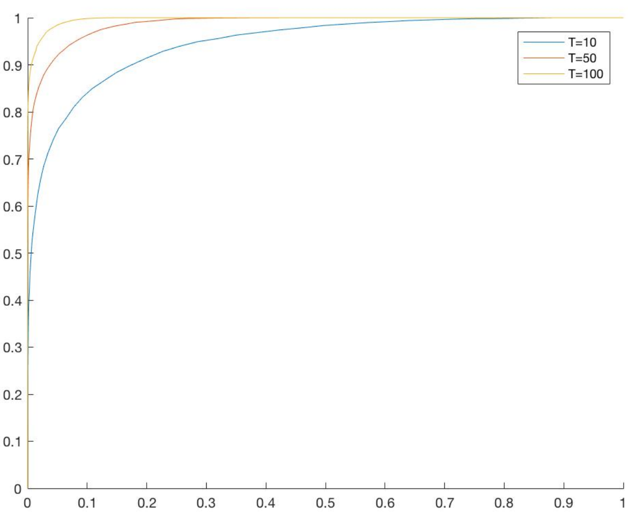

Figure 17: ROC curve for RealBoost