Stat 231: Machine Learning and Pattern Recognition

Project II: Detecting Faces in Images by Boosting Technique

Mahati Kumar

004414435

Step 2.1: Construction of weak Classifiers

Feature Design and Feature Value Calculation:

As a part of the Adaboosting or Adaptive Boosting Procedure, we are to design some number of features. These features have two portions to them, the shaded and the unshaded areas, and a difference in the intensity between these portion is calculated for every images in the dataset. This is what taken as the feature value for a particular image, for a specific feature type. Here, I have designed 5 different types of Haar - features as shown below. Thus for every image the feature value is calculated for every feature type and for all possible valid scales and positions of these feature windows.

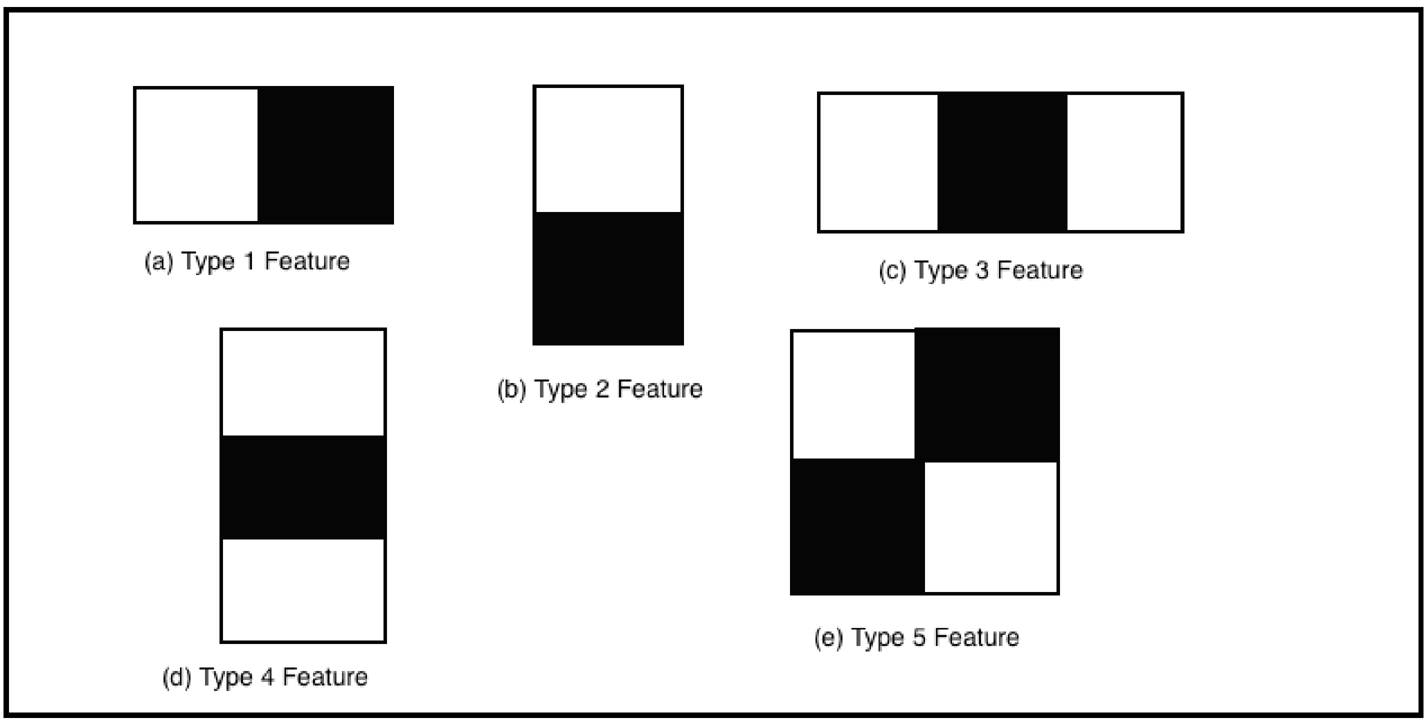


Figure (2.1)

For the above designs of the features in a 16x16 image I was able to obtain 32384 valid features, with split being 8704 of the Type 1 and Type 2 features each, 5440 of the Type 3 and Type 4 features each and the rest 4096 of the Type 5 features. I found the feature values i.e the intensity difference values of every feature on application on an image, for 10000 face and 10000 non-face 16x16 images. I have used only in 16x16 images throughout my computations due to the limitation of time.

Histogram plots:

The figures 2.1.a to 2.1.e are the histogram plots of 5 randomly selected features of each Type, for Faces and Non-faces. The red dashed-dotted line corresponds to non-faces, and the Blue solid line corresponds to the faces.

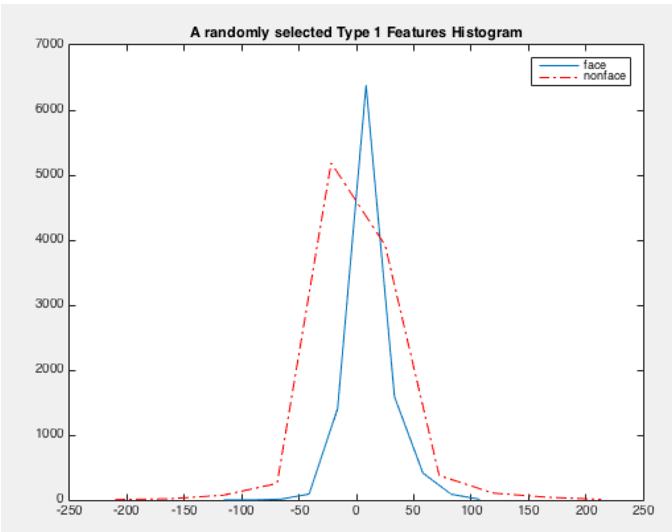


Figure (2.1.a)

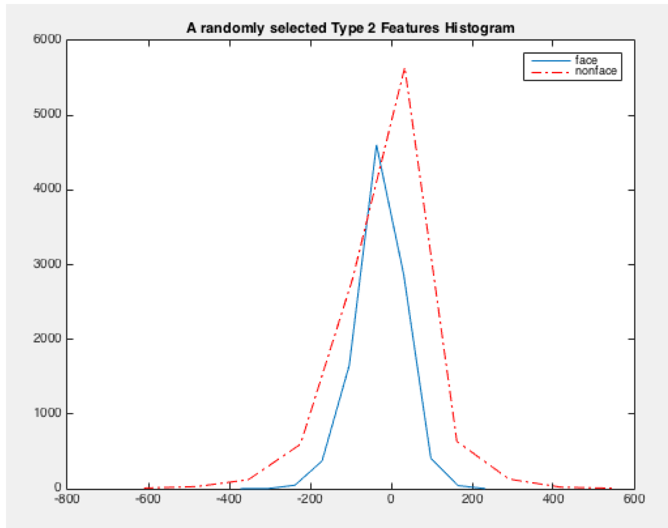


Figure (2.1.b)

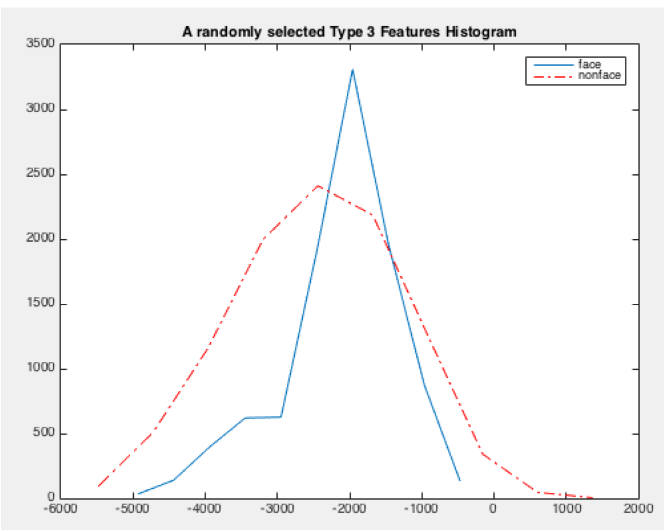


Figure (2.1.c)

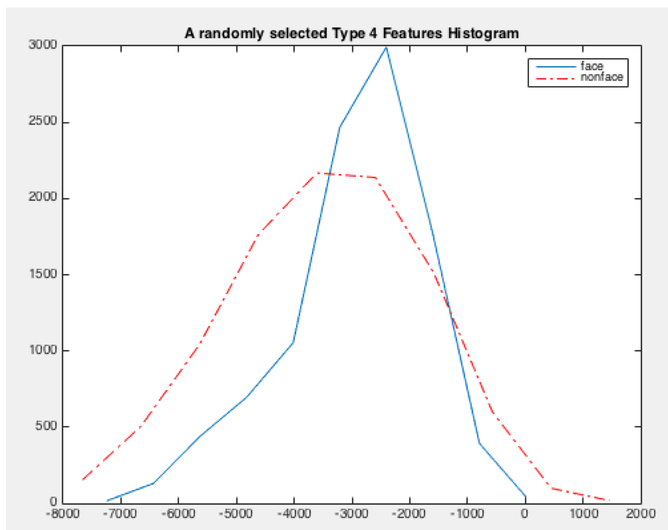


Figure (2.1.d)

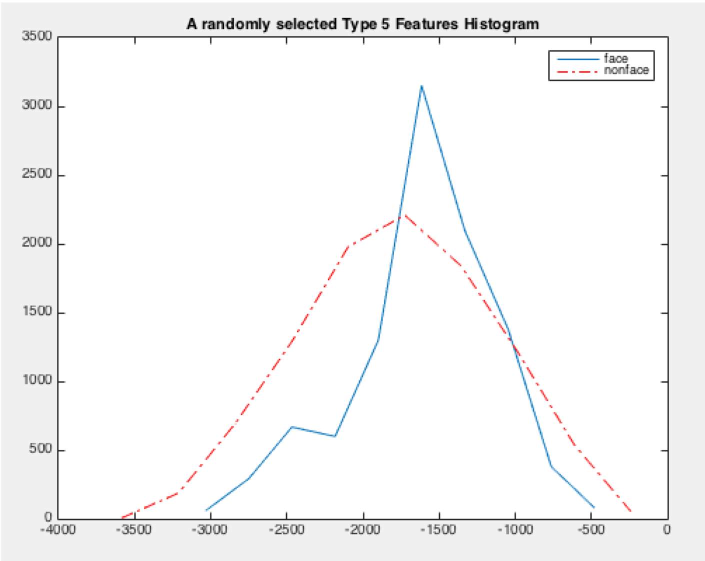
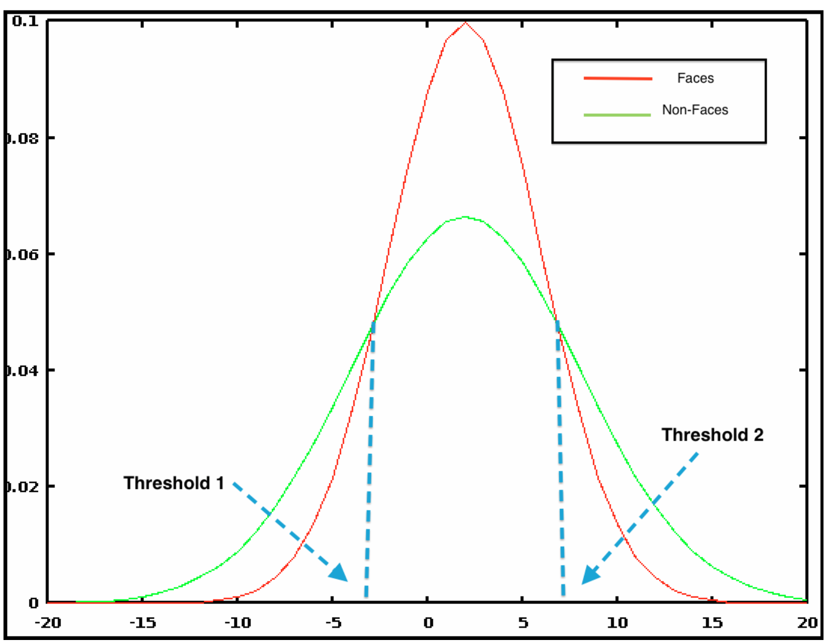


Figure (2.1.e)

Threshold Determination of Features:

As it can be seen from the plotting of the feature histograms above, we can evidently see that the Haar features follow an approximately gaussian distribution. Thus an ideal gaussian distribution of the face and non-face feature values can be looked upon as shown in the figure . Thus, taking advantage of this fact and with reference to the explanation in [1], parameterizing the two distributions with a Gaussian model, we can find the two intersection points that will represent the two thresholds. The formulae used for the same is provided and explained in [1].



[1]http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.145.643&rep=rep1&type=pdf

Step 2.2: Adaboosting

AdaBoost or Adaptive Boosting refers to a particular method of training a boosted classifier. As described Viola and Jones in [2], it is a Machine Learning approach that allows processing images extremely rapidly and achieving high detection rates. It involves the boosting or improvement over individual weak classifiers to a more stronger and super classifier which strings together multiple weak classifiers with updated weights in tandem. Here, I initially ran the Adaboost algorithm explained in [2] on a smaller dataset of faces and non-faces, ran the boosted classifier on the empty classroom images, added the false positives generated in the run to the non-faces dataset and then ran the same Adaboost algorithm on a larger dataset of 10000 faces and 10000 non-faces. The following Figure (2.2) is the set of top 10 best classifiers obtained at the end of the Adaboosting algorithm.

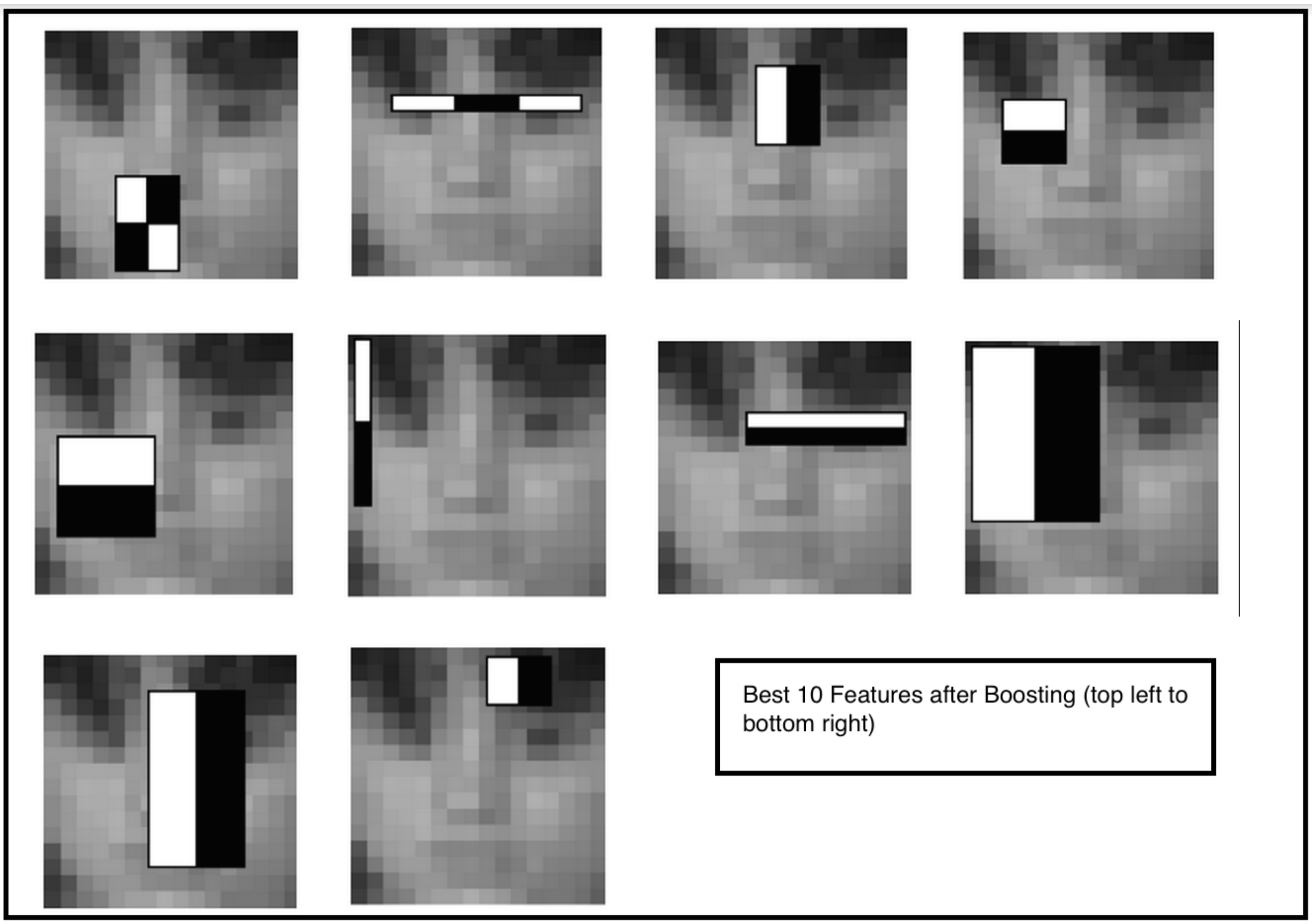


Figure (2.2)

[2] http://www.stat.ucla.edu/~sczhu/Courses/UCLA/Stat_231/Lect_note/viola01rapid.pdf

Curve for the errors of top 1000 weak classifiers: After Adaboosting

After Adaboosting, iterations T=0,T=10,T=50,T=100 each the top 1000 errors i.e the errors are sorted in the non-decreasing order and the top 1000 are taken and plotted as shown in Figure (2.2.a) below.
As it can be seen, with the increase of T, the error of weak classifiers increase. At T=100 only top 500 classifiers have their errors smaller than 0.5.

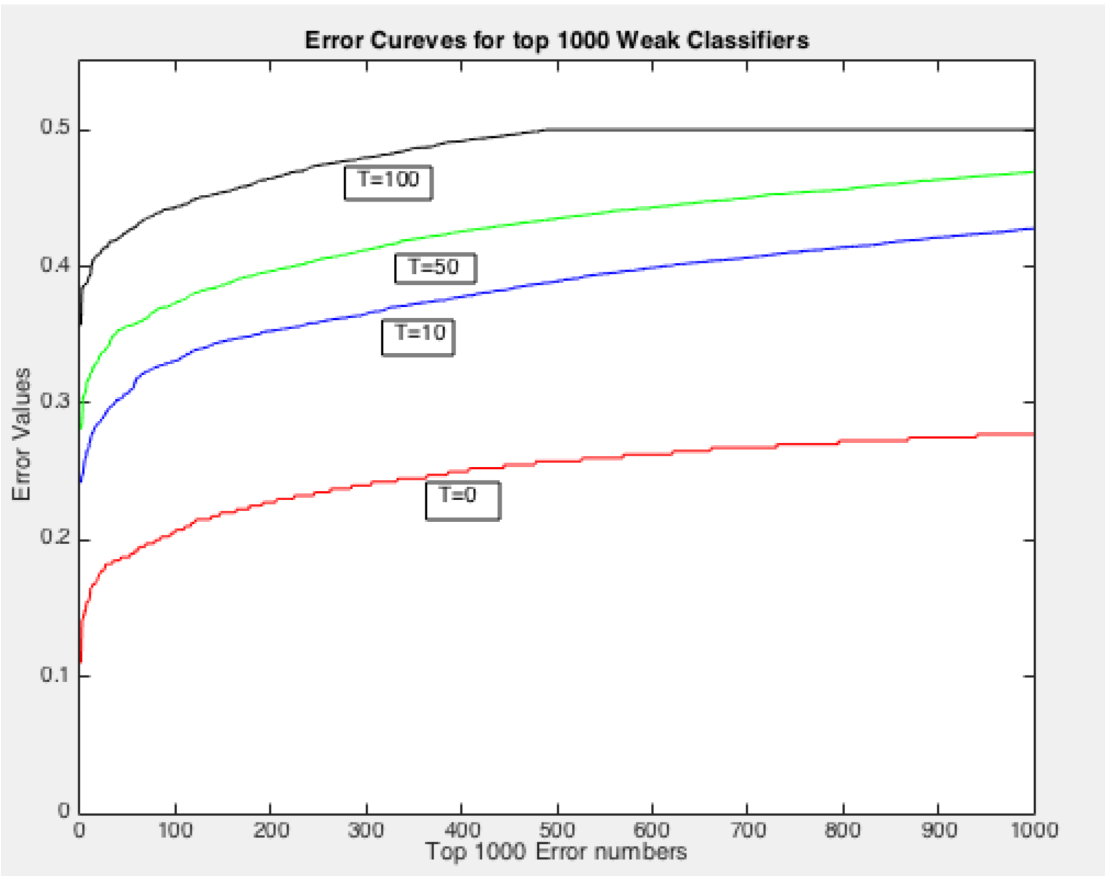


Figure (2.2.a)

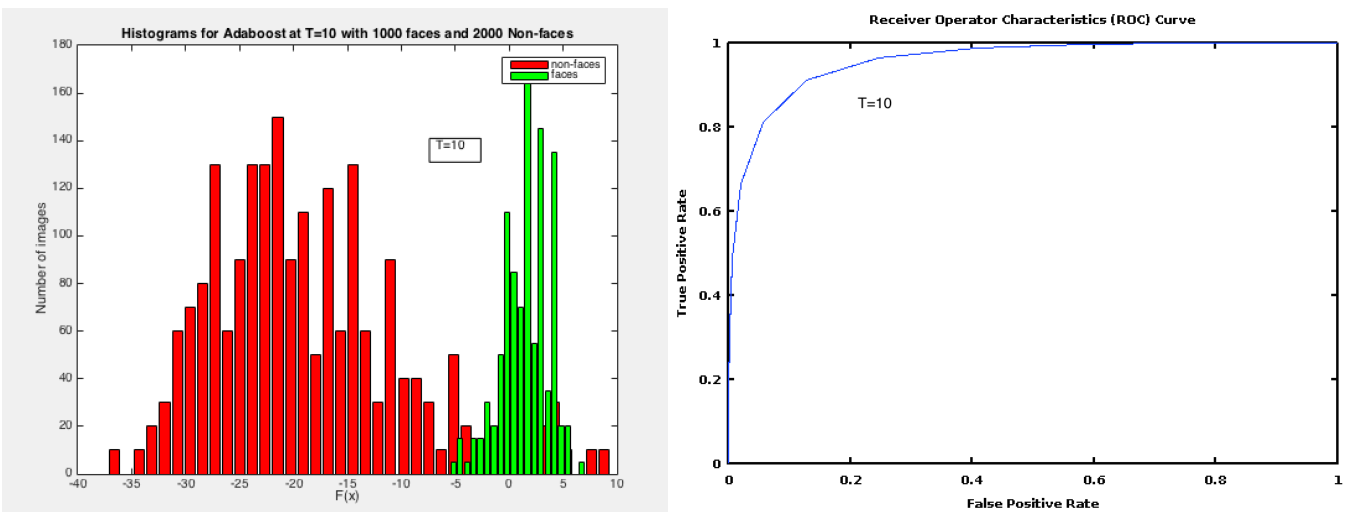
Histograms and ROC Curves

The histograms of the positive and the negative populations, i.e of the faces and the non-faces, are as shown in Figures 2.2.1 to 2.2.3. The Adaboost algorithm was run again on 1000 face and 2000 nan-face test images. The respective ROC (Receiver Operator Characteristics) Curve is also plotted to the right side of each figure. With the increase of T, the algorithm performs better, but the performances are already quite similar after T = 50. Figure 2.2.4 shows all the three ROC curves, for T=10,T=50 and T=100, on the same plot so that their movement with T can be easily seen.

Please note: Here, F(x) of the histograms is the formula for H(x) from Prof. Zhu's slides,

$$H(x) = \mathrm{sign}\{\textstyle\sum_{t=1}^{T} \alpha_t g_t(x) - \frac{1}{2}\sum_{t=1}^{T} \alpha_t\}$$

I have used this formula, as I used +1 and 0 to represent the classification of face and non-face respectively.
And - In each of the images the 0 point on the x axis is at different locations and do not sync with each image below.



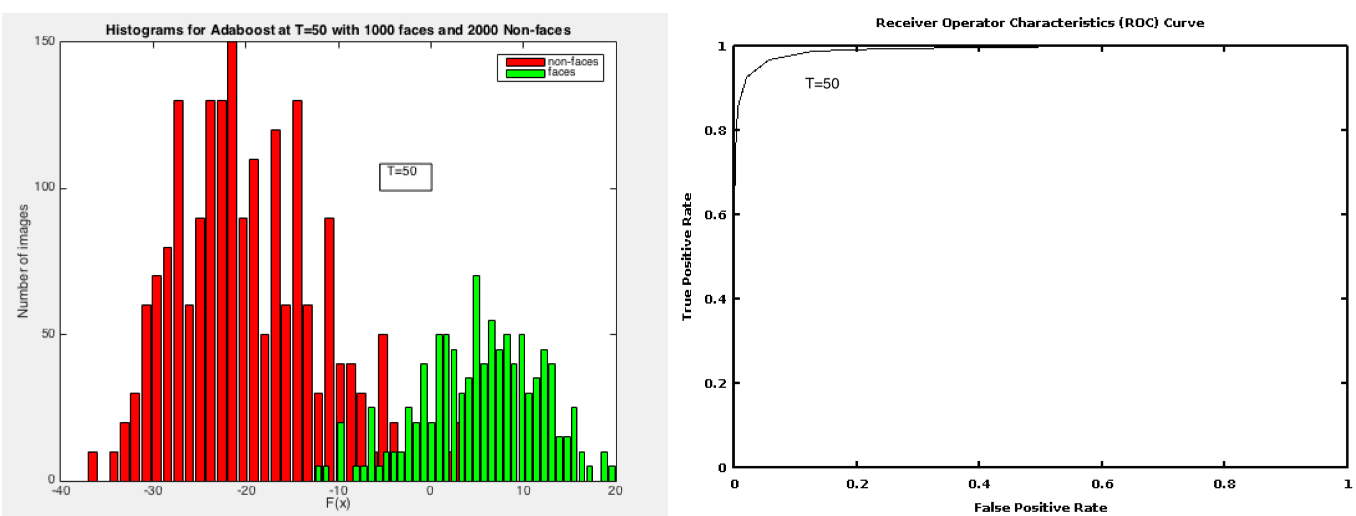


Figure (2.2.1) - top two images

Figure (2.2.2) - bottom two images

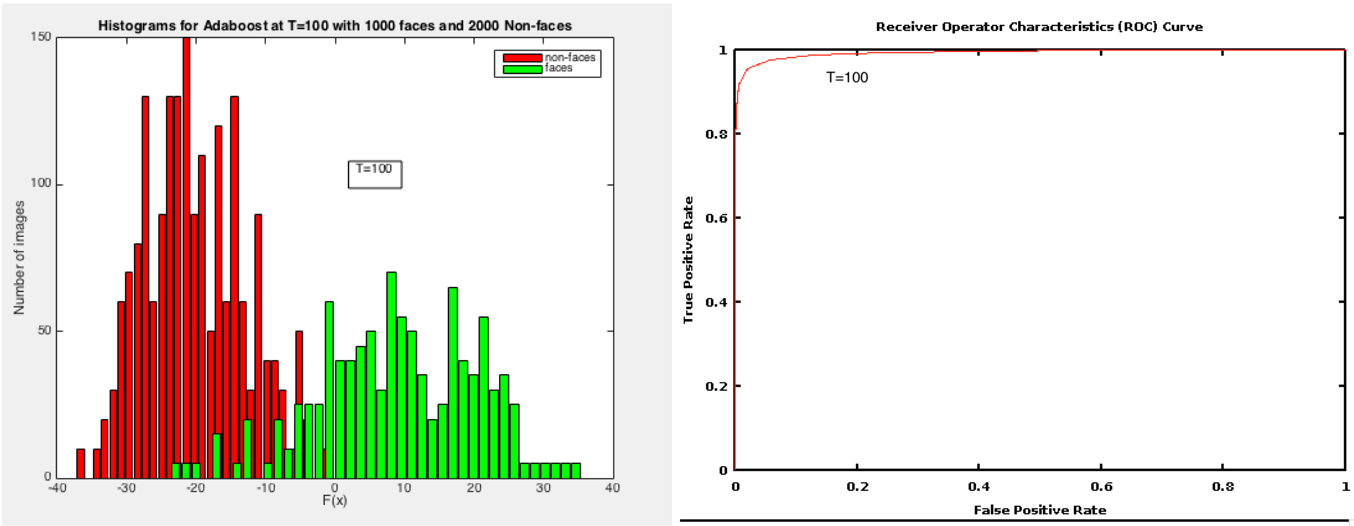


Figure (2.2.3)

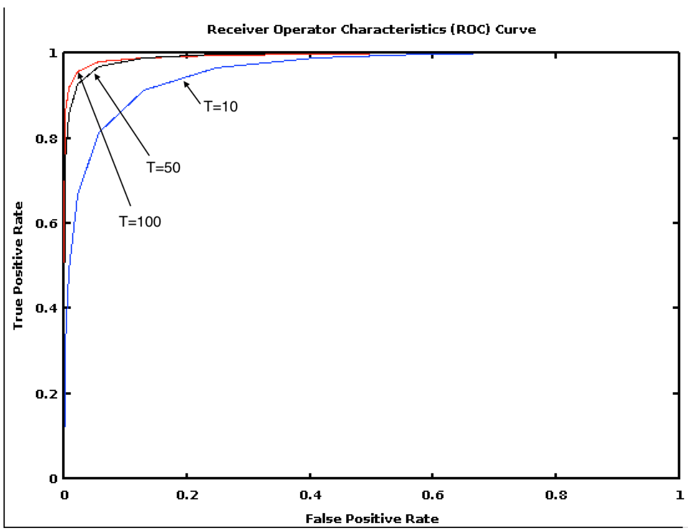


Figure (2.2.4)

Step 2.3: Real-boosting

I ran the Real Boosting algorithm, which is simple mathematical modification over the Adaboost algorithm, over on a smaller dataset of faces and non-faces, ran the boosted classifier on the empty classroom images, added the false positives generated in the run to the non-faces dataset and then ran the same algorithm again on a larger dataset of 10000 faces and 10000 non-faces. (just like adaboost)

The Histograms of the positive and negative populations over the F(x) axis for T = 10, 50, 100 are shown in Figures (2.3.a) to (2.3.c). And finally the corresponding ROC Curves are also shown the Figure (2.3.d).

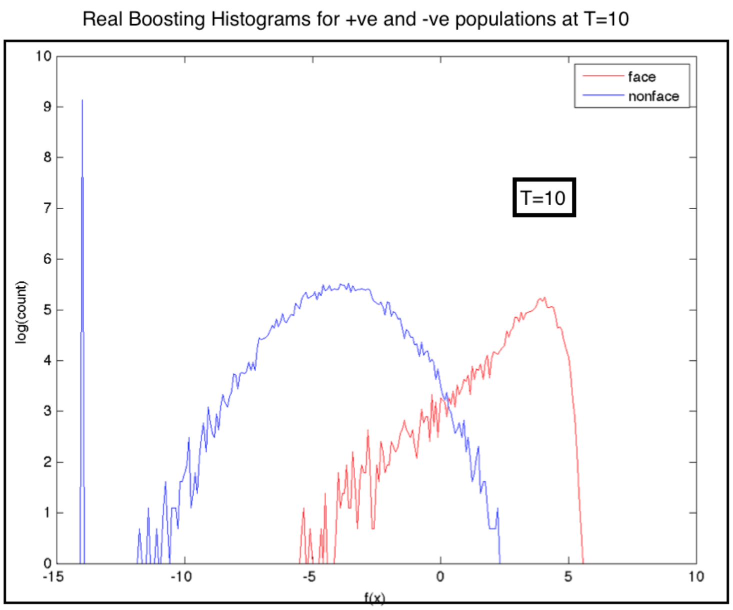


Figure (2.3.a)

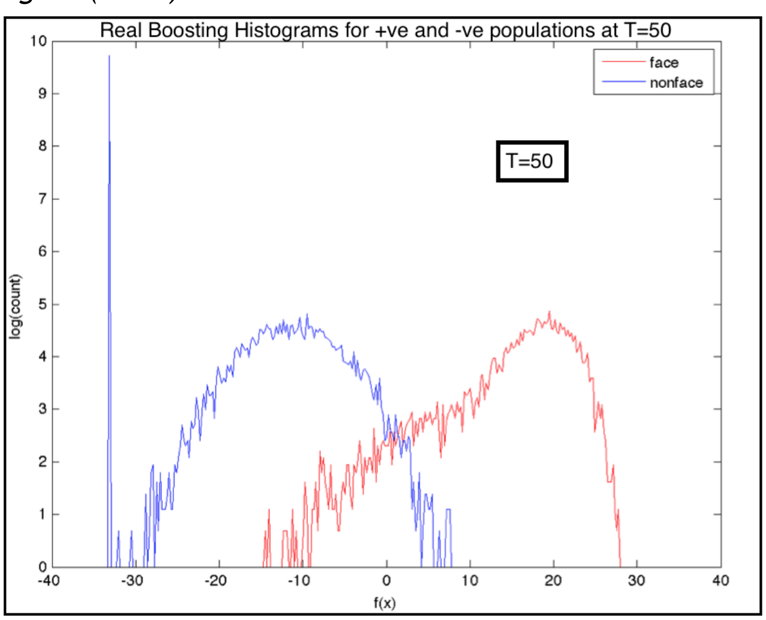


Figure (2.3.b)

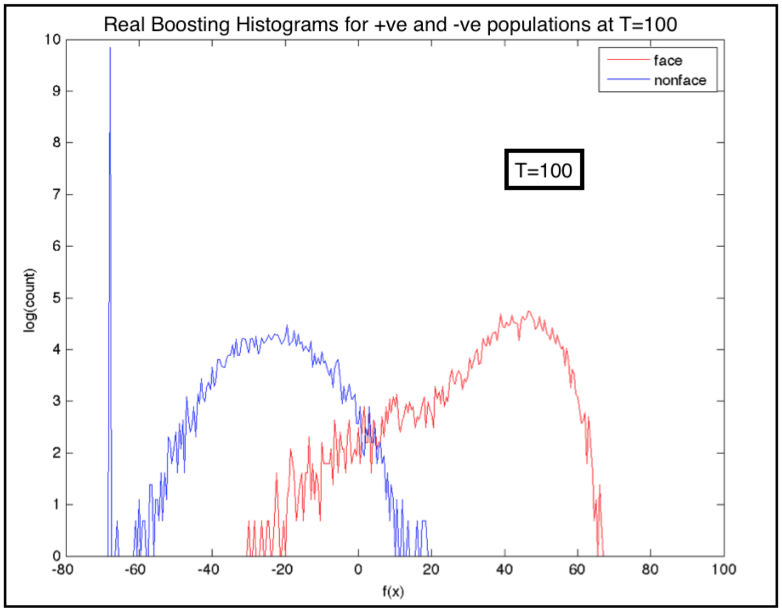


Figure (2.3.c)

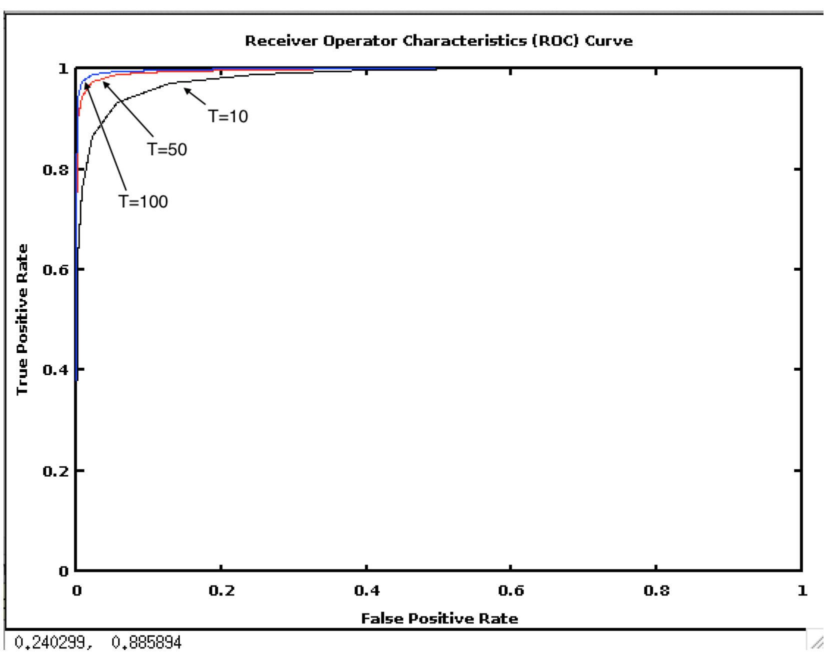


Figure (2.3.d)

Step 2.2 and 2.3: Testing on Classroom Images

The classroom image was rescaled multipe to include as many faces as possible. In the comparison below it can be clearly seen that the Real Boost results are much better than the adaboost ones, as there are comparatively more non-faces also being detected as faces (false-positives) in adaboost over realboost.

Ada Boosted Classifiers testing on Classroom image:

Real Boosted Classifiers testing on Classroom image:

Thank you for reading.