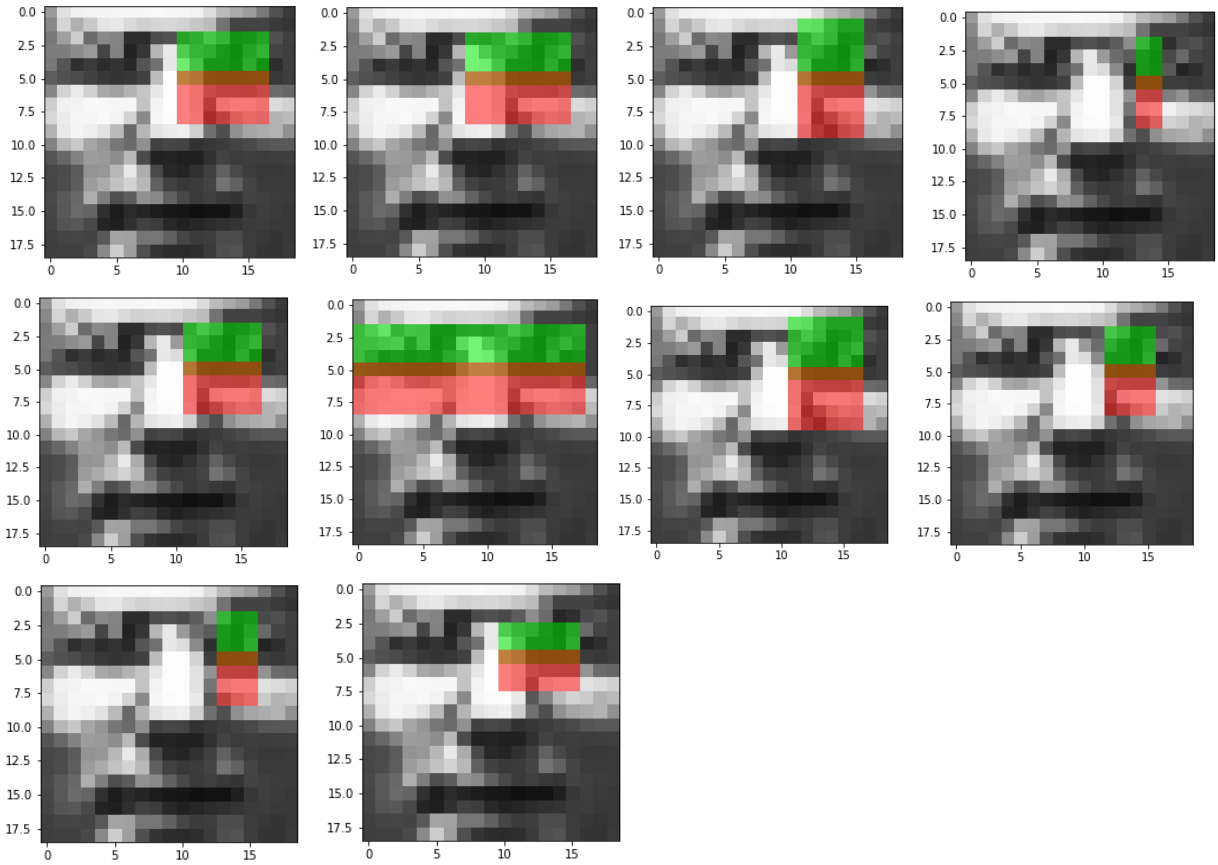
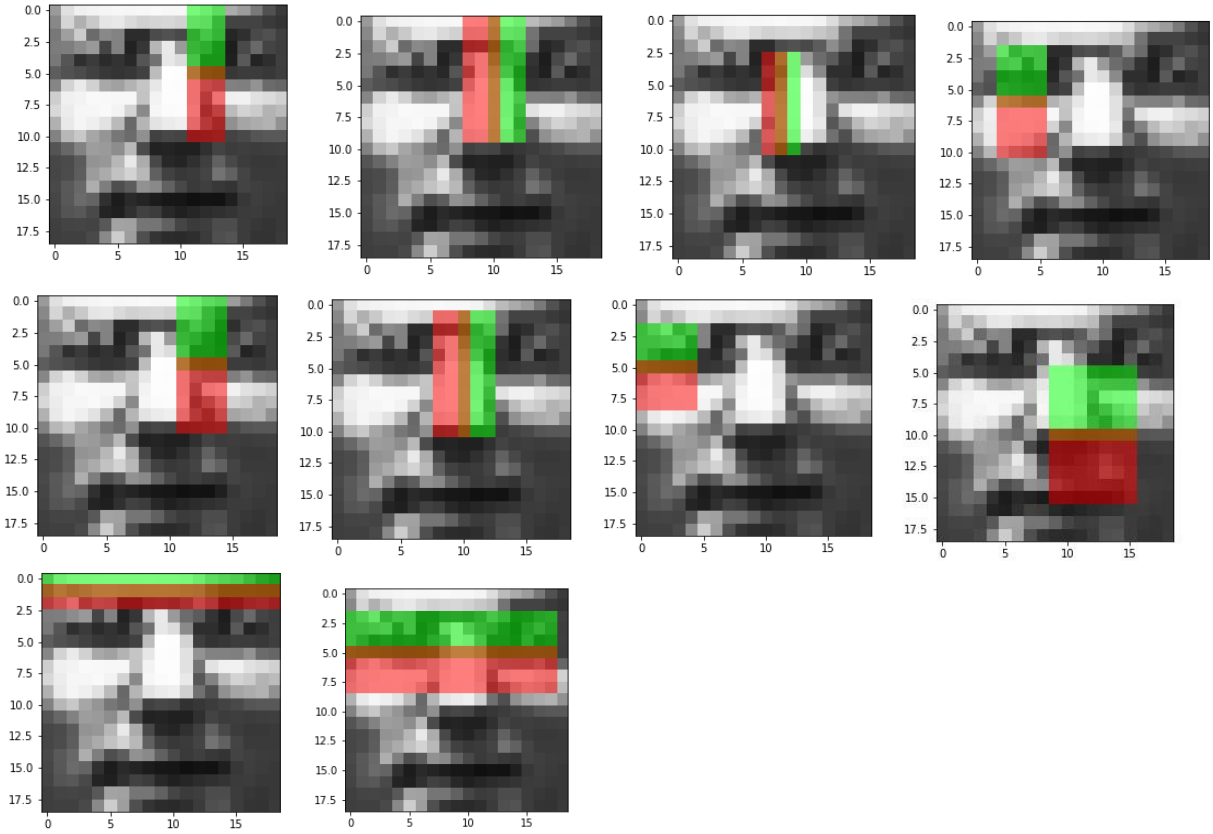
Process flow:

1) The training and testing data were stored in pickle files for easy loading and usage
2) The environment used was google colab.
3) Input image size of 19,19 was used. Total 500 face images and 1000 non-face images were used for training
4) All 5 haar features: type2-x, type2-y, type3-x,type 3-y, type 4 were used.
5) The total number of features calculated was 51705. Each of these features is considered as a weak classifier
6) First the weights for the training examples are initialized. As the face and non-face examples have different values, the weights initialized for these examples are different.
7) The features calculated are applied to the training examples and stored before starting the iteration for finding the optimal weak classifiers.
8) The weak classifiers are trained to find the optimal ones. Based on reference [2], an optimal method is used for finding the error of each possible threshold in constant time and the error of all thresholds in linear time.
9) The threshold is set to the value of the feature at which the error is a minimum
10) Total 10 weak classifiers are found.

Before boosting top 10 features:



After boosting top 10 features:



Total images in test data = 2000

Total face images = 1528

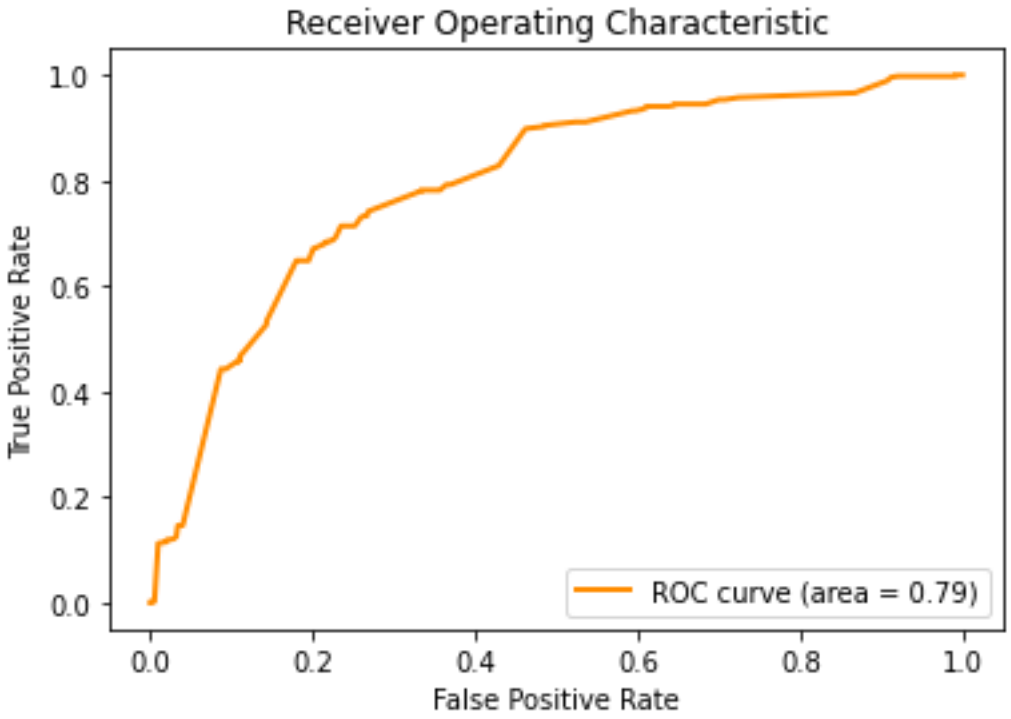
Total non-face images = 472

False Positive Rate: 164/1528 (0.107330)

False Negative Rate: 257/472 (0.544492)

Accuracy: 1579/2000 (0.789500)

ROC Curve:



Reference

1) https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf
2) https://github.com/aparande/FaceDetection
3) https://github.com/jasonleaster/FaceDetection/blob/master/FaceDetection/adaboost.py