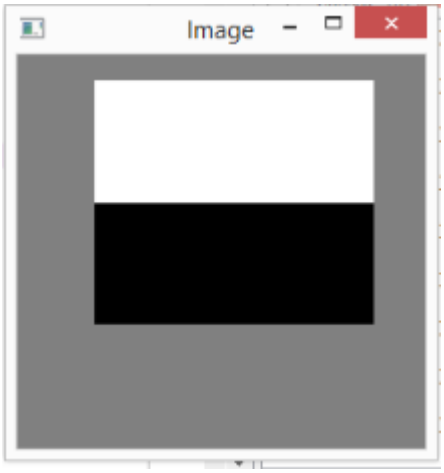
Name – Vivek Onkarnath Rathi
ECE 763 Computer Vision Project 02

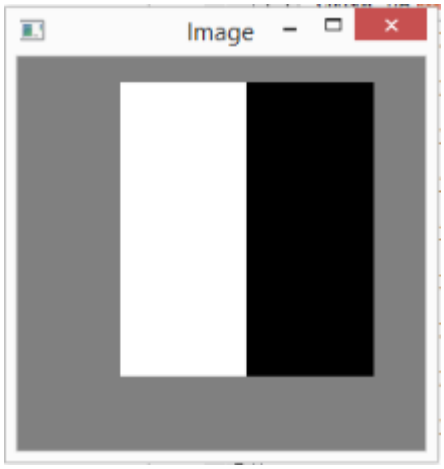
Adaboost Algorithm (Ensemble Learning) is implemented for face detection.
2000 training images (1000 Face & 1000 Non Face) & 400 test images (200 Face & 200 Non Face) were used.
All images were 16*16 grayscale images.
Haar features were input to the adaboost, i.e. weak learners; finally giving a strong classifier.

Haar Features-

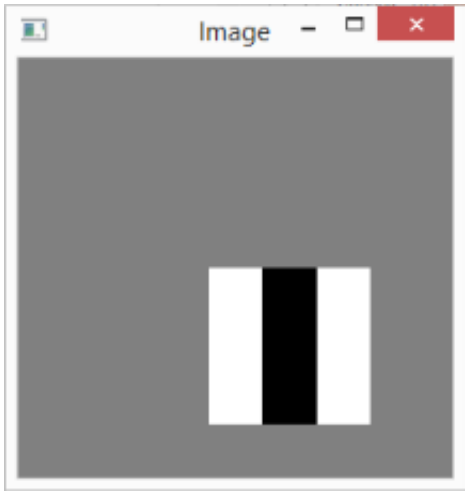
1. Haar features are used for face classification.
2. Each haar feature forms a weak classifier
3. Haar features can be visualized as rectangular features (positive & negative rectangles)
4. Haar features can be of various types/configuration such as type-2-x, type-2-y, type-3-x, type-3-y, type-4
5. Visualizing haar features-



type-2-y haar feature

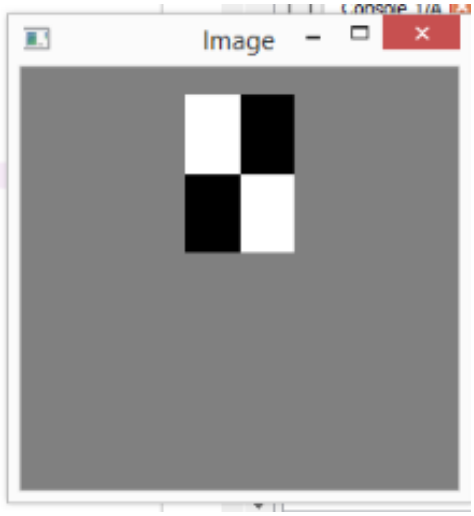


type-2-x haar feature



type-3-x haar feature

type-3-y haar feature



type-4 haar feature

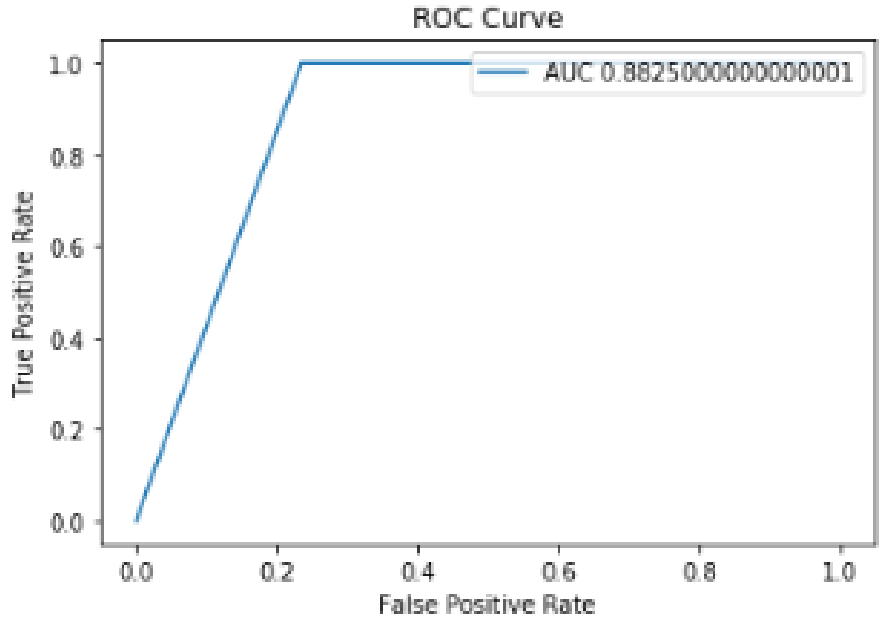
6. Haar features were obtained using open source code from scikit-learn package.
7. Haar feature is a scalar value, associated with rectangle co-ordinates and feature type as per above images.
8. Input image of 16*16 size can have almost 32176 (even more).

Adaboost Algorithm-

1. One needs to find threshold for a classifier. I used linear line search where each haar feature for given set of training images were considered as threshold & best possible threshold was found. It takes lot of time for 1000 training images. Due to time complexity, I decided to take mean of possible haar features specific to location for all face & non face training images. Averaging the mean for both face & non face gives the threshold for each haar feature i.e. the classifier
2. Initialize weights as (1/no. of samples).
3. Get errors for each classifier for given set of input images.
4. Find the classifier with the least error.
5. Find alpha.
6. Update the new weights.
7. Find error/accuracy for final classifier.
8. Repeat step 2 to 7 until algorithm converges. (training error<=10% stop)

**program takes about 12-15 minutes to run.

ROC Curve for top 10 haar features from adaboost (top 10/first 10 classifiers)



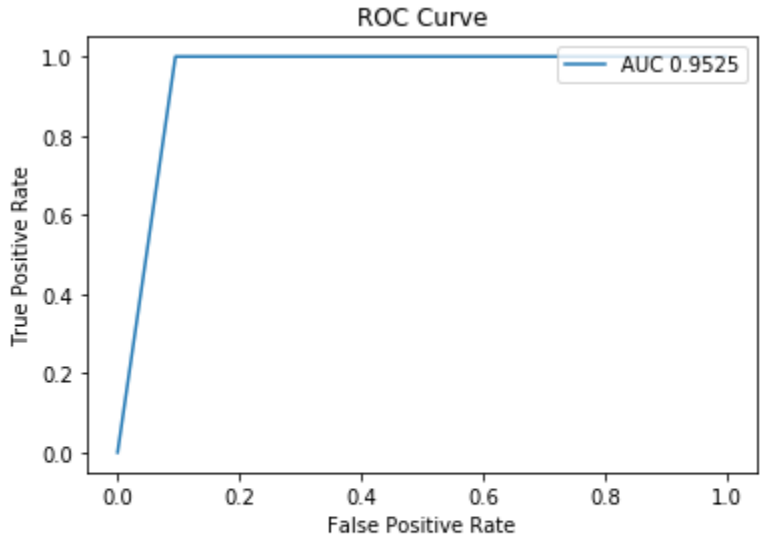
Confusion Matrix-

	Face	Non Face
Face	TP= 200	FP= 0
Non Face	FN= 47	TN= 153

```
True Positives: 200/200,
False Positives: 0/200,
True Negatives: 153/200,
False Negatives: 47/200,
Misclassification Error Rate: 11.75
```

Accuracy – 88.25%

ROC Curve for haar features until adaboost converges (here 47 weak classifiers / haar features)




```
True Positives: 200/200,
False Positives: 0/200,
True Negatives: 181/200,
False Negatives: 19/200,
Misclassification Error Rate: 4.75
```

Confusion Matrix-

	Face	Non Face
Face	TP= 200	FP= 0
Non Face	FN= 19	TN= 181

Accuraccy – 95.25%

Adaboost features-

Conclusion-

1. Adaboost combines weak learners to form a strong classifier
2. The above mentioned results are from a cascade i.e. weighted sum of only 10 classifiers.
3. We get overall accuracy of 88.25%
4. Overall accuracy can be increased if one considers more than 10 weak classifiers (say 100 i.e. >10)