A. Data Preparation

Face datasets from http://vis-www.cs.umass.edu/lfw/ ,which is one of the lists, which is guided from the project description. Prepared N =1000 training images for face and non-face respectively, N=100 test images for face and non-face respectively. Images resized and cropped into 20 x 20 resolution. The dataset is saved in each folder named as follow:

- */resized_face/:* N =1000 training images for face
- */resized_nonface/:* N=1000 training images for nonface
- */resized_face_test/:* N=100 test images for face
- */resized_nonface_test/:* N=100 test images for nonface

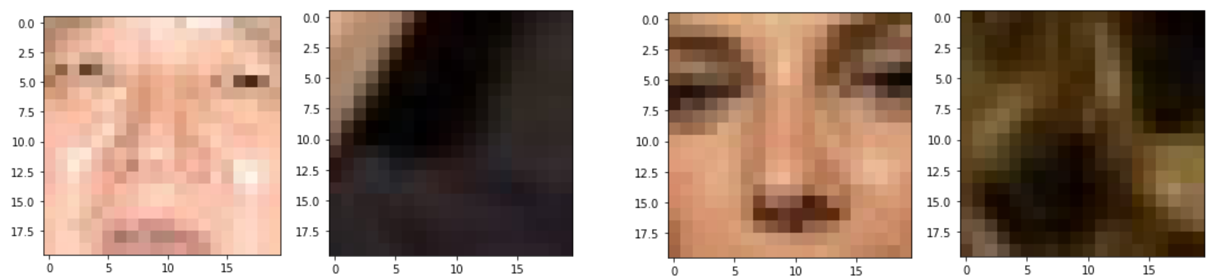

Figure 1. Data preparation for train set and test set

Figure 1 shows the example of the dataset. The original image data resolution is 250 x 250. The face region selection starting from the center of the image and offset 30 in each direction. This method provides face train/test set containing features, e.g., eyes, nose and mouth. The one thing that could not be challenge for test images is that most of their dataset has only front face data, not including side views. The variability of face would not be classified correctly based on this learning model. That’s what we could expected in this study. Nonface training data set is from the background of the face dataset, the nonface train/test dataset gained randomly from the corner of the images.

After images set two different categories with face and non-face then, the image resolution is resized into 20 x 20 x 3. The data is reshaped into 1200 dimensional data and build learning model. The objective of this project is that evaluating face classification performance using single gaussian model, mixture gaussian, mixture of t-distribution, and factor analysis. The study compares the performance between different distribution. The performance is evaluated with false positive rate, false negative rate and misclassification rate. Also, ROC (Receiver operating characteristic) curve plot gives usefulness of the test.

B. Result

B-1. Single Gaussian

(a) **Visualization estimated means and Covariance matrix for face and nonface**

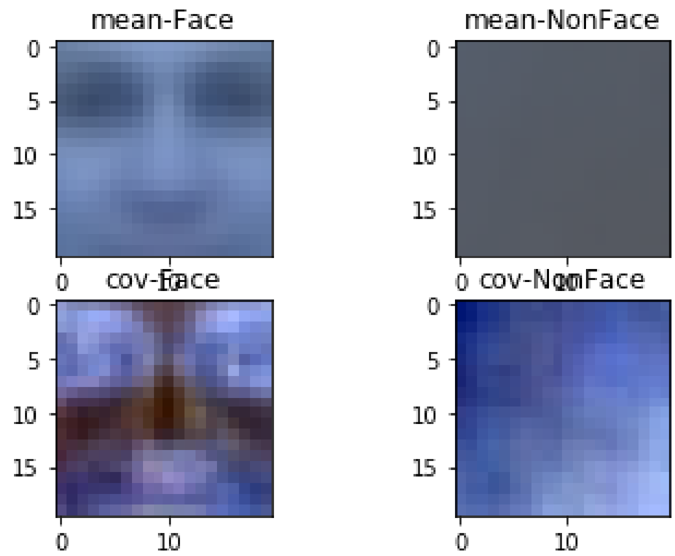

Figure 3. Single Gaussian mean and covariance matrix

(b) Evaluation the learned model on the testing images

False Positive: 0.99 False Negative: 0 Misclassification: 0.495

(c) **Plot ROC curve**

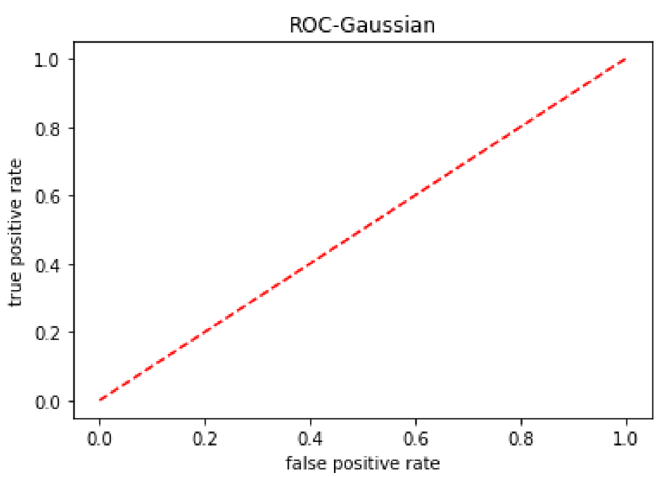

Figure 4. Mixture Gaussian ROC(k=3)

The learned model performance is evaluated using ROC curve and covariance matrix. ROC shows that the detection of the face was not good enough with a distribution the potential reasons could be unimodal distribution, a single outlier could dynamically impact on the performance, and single dimension. The was not good enough to use in this case.

B-2. Mixture of Gaussian

- MOG model is tested with 3,5 components of MoG model (k=3, k=5). The overall performance is not change much. However, ROC -curve presents that the increased MoG model is more useful than smaller one.

(a) Visualization estimated means and Covariance matrix for face and nonface (k=3)

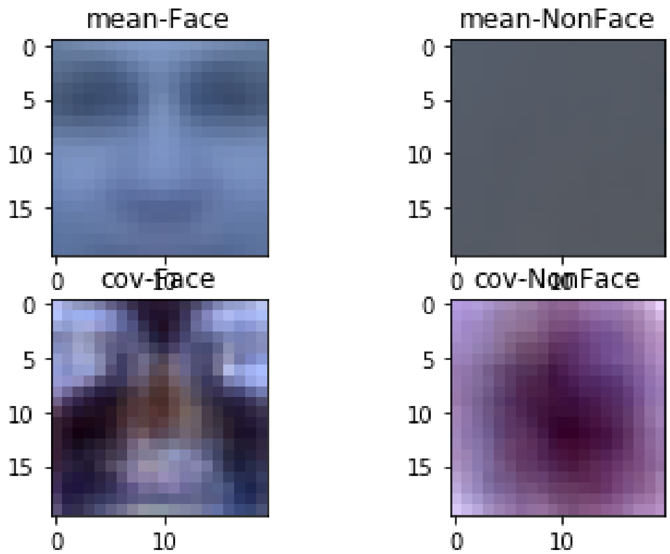

Figure 5. Mixture Gaussian mean and covariance matrix (k=3)

(b) Evaluation the learned model on the testing images (k=3)

False positive rate: 0.192 False negative rate: 0.093 Misclassification:0.1425

(c) Plot ROC curve(k=3)

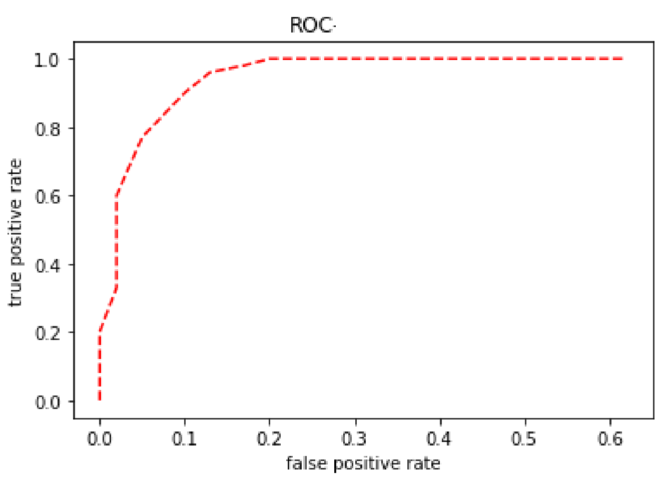

Figure 6. Mixture Gaussian ROC(k=3)

(a) Visualization estimated means and Covariance matrix for face and nonface (k=5)

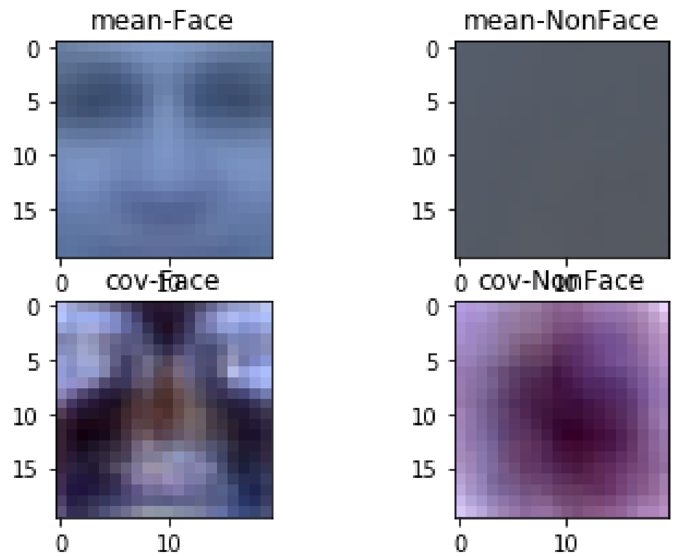

Figure 7. Mixture Gaussian mean and covariance matrix (k=5)

(b) Evaluation the learned model on the testing images (k=5)

False positive rate: 0.14 False negative rate: 0.03 Misclassification:0.085

(c) Plot ROC curve(k=5)

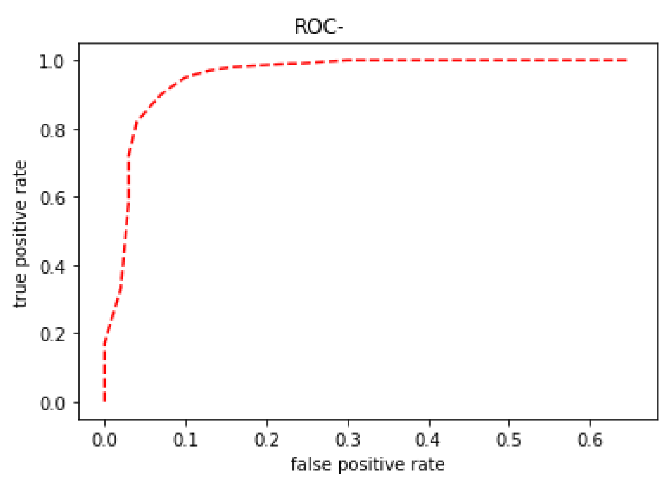

Figure 8. Mixture Gaussian ROC (k=5)

The mixture of gaussian is weighted sum of k normal distribution. The mixture of gaussian model express as a marginalization of a joint probability distribution between data and discrete hidden variables. Then, the expectation and maximization algorithm calculates the probability for each data points for each gaussian, and its probability of the datapoints belongs to the each of the gaussian. The estimated mean and covariance matrix visualization does not change that much with different gaussian clusters. However, ROC graphs show that the performance detecting the face become better when it raised to the k = 5. The MoG model based on ROC evaluation, it is much better performance then the single gaussian performance.

B-3. T-distribution

(a) Visualization estimated means and Covariance matrix for face and nonface

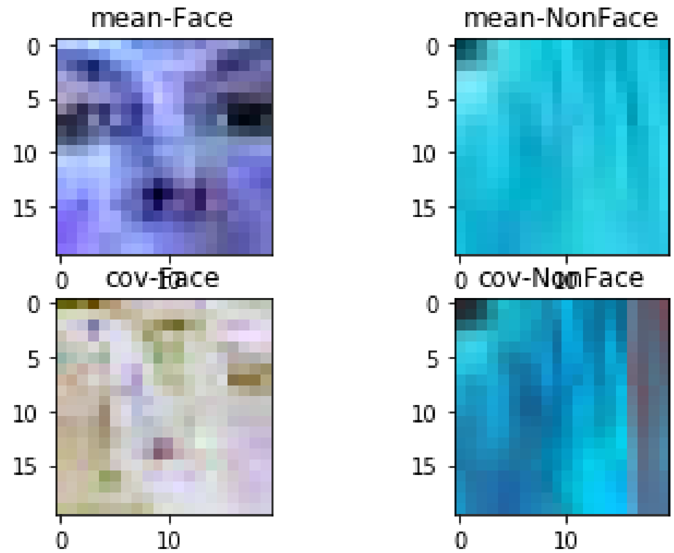

Figure 9. T-distribution mean and covariance matrix

(b) Evaluation the learned model on the testing images

False positive rate: 0.165 False negative rate: 0.045 Misclassification:0.16

(c) Plot ROC curve

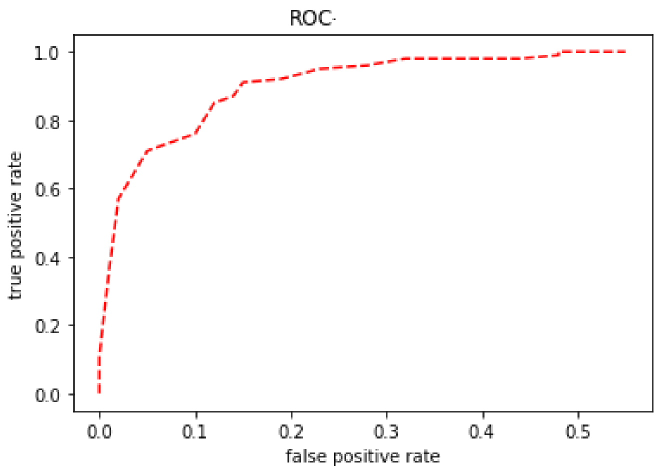

Figure 10. T-distribution ROC

The T-distribution model provides so much different mean and covariance results. T-distribution mean and covariance matrix give brighter result. The reason could be from the vector factor gives this much difference on the images. Interestingly the mean face has the clear on the features in nose and eyes. The ROC seems reasonable to use this distribution.

B-4. Factor Analysis

(a) Visualization estimated means and Covariance matrix for face and nonface

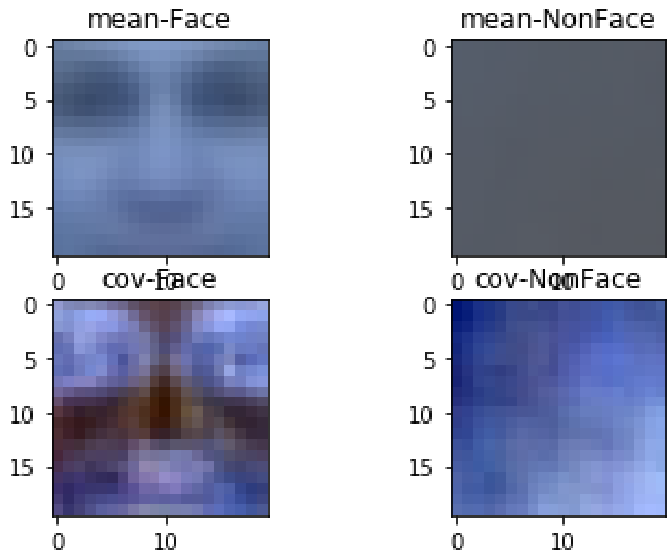

Figure 11. Factor analaysis mean and covariance matrix

(b) Evaluation the learned model on the testing images

False positive rate: 0.19 False negative rate: 0.08 Misclassification:0.135

(c) Plot ROC curve

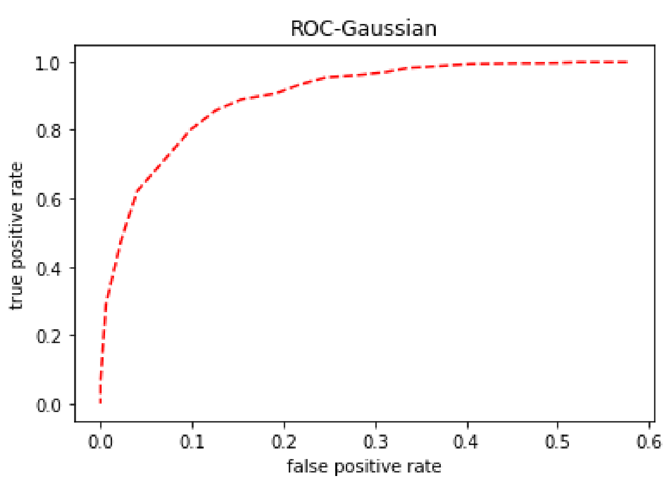

Figure 12. Factor analaysis ROC

The factor analysis contains vectors with diagonal covariance. The factor analysis model iterates 3 of the EM algorithms. The factor analyze associated model performance does not change that much on the mean and covariance matrix visualization. The false positive and misclassification rate is low enough, and it gives good enough ROC graph.

Conclusion

The face classification using different distributions gives different results of mean and covariance matrix, as shown in the previous result section. Differences between distributions are developed from single-gaussian approach. Mixed gaussian randomly assigned the mean and covariance matrix cluster and evaluate with weight item that trying to meet to fit into the model. The number of clusters could give more accurate result, but it is computationally more expensive as usual, and the performance was not that improved compared to the time increased to get performed. T distribution has additional variable called degree of freedom, and gamma function. Factor analysis added the vector with diagonal covariance for the fit. Interestingly, the result of mean and covariance matrix does not give enough distinct differences in this study case in single distribution and mixture gaussian and factory analysis, but only t-distribution gives distinctively different color of the mean and covariance images. Overall, the ROC performance is good to detect face except the single gaussian model. Increasing the number of iterations gives more reliable performance in this study. In the class, we discuss about these distributions, and the important thing to remember is there is no free launch. The computationally expensive model gives better performance (in my case mixture gaussian with k=5). However, it takes more time to get mean and covariance matrix. Thus, the right model could be determined based on where we implement or the purpose of the usage. Also, the dataset could be important factor to train and test.