# Pattern Recognition and Machine Learning – Project 3
Aravind Chandradoss

Q1)
Data set: 4 Categories of data set from **AR Database** namely **Normal (AR_1), Laugh(AR_4), Cooling Glass (AR_10) and Scarf(AR_13)** are used**.**

For Question 1a and 1b, A mathwork algorithm "mixGaussianEM" was used, which **generate** the needed **Gaussian model consisting of Mean and Sigma.** This sigma is later used for **visualizing** the Gaussian distribution. On the **generated EM model, PCA was applied to reduce the dimension to 2 and 3,** which was used for visualizing purpose.

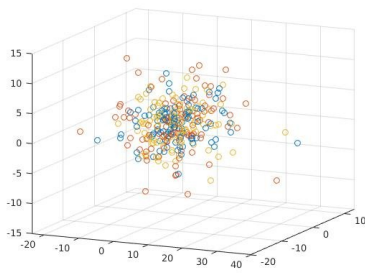The **dimension-reduced Gaussian distributions** are compared below,
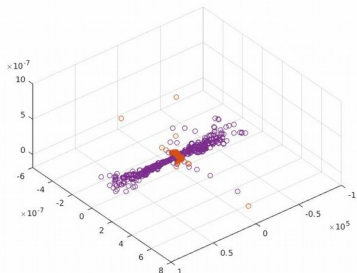
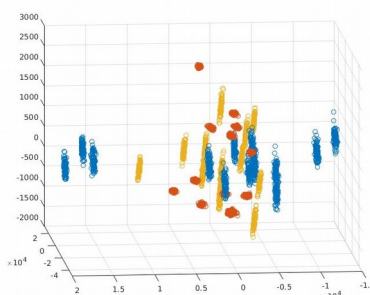Using all **P-dimensions**.

Reduced using PCA
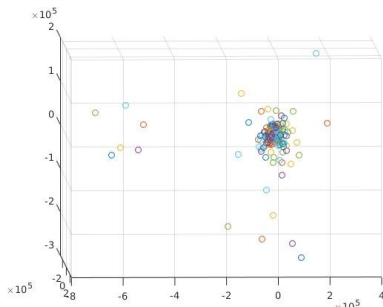
Using **PCA with 100 dimensions**

Reduced using PCA

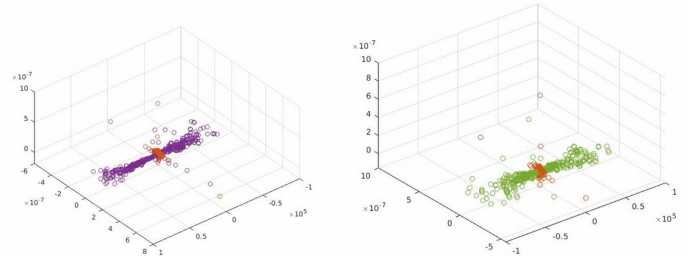Using **ICA with 100 dimensions**

Reduced using PCA

using **LDA**

Reduced using PCA

**Comparison between two classes using Gaussian distribution:**

**With all the P-classes**, the two classes have following distribution.

Inference: Although. Looked more or less similar in 3 dimension, both showed significant difference when mahalanobis distance was found (Discussed in Classification section) Thus, it can be used for Classification.

With PCA Subspace
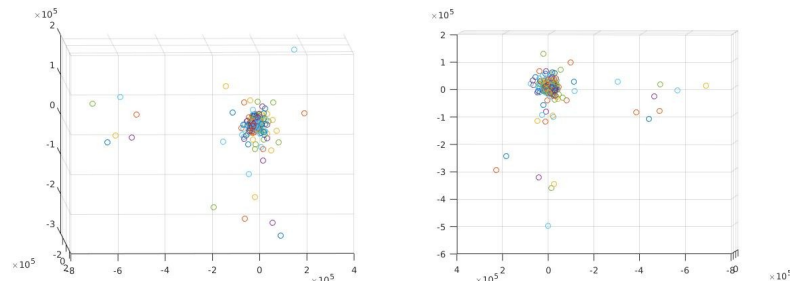
Inference: most of the data point (after dimension reduction) were found together and slightly distributed along two axis (principle axis). But in higher dimension, mahanobis distance are found to be significant than this visualization.
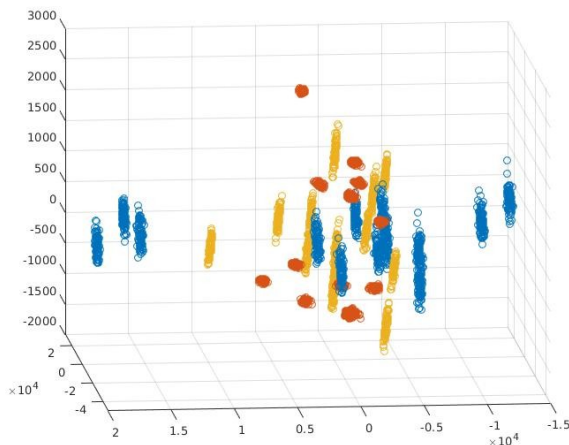
Using ICA Subspace:
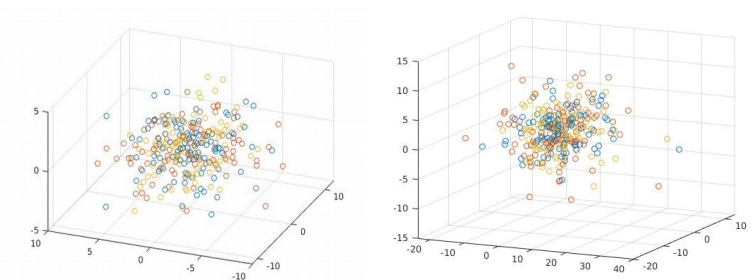
Inference: As expected, the distribution is found to be scattered and spherical in nature. Even after repeated projection data point, it was found to be spherical in nature

(Not like PCA, getting distribute along each principle axis). Thus, ICA might not be a good tool for classification.

Using LDA Subspace:

Inference: Interestingly, LDA gave impressive result. Even under reduced dimension, it showed well defined gaussian distribution. Thus, LDA can be a good tool for classification of the chosen data set.

Q1 Part C – Classification:

Classification was done, using **mahalanobis distance,**

Method: Leave-one-out approach was done.
Results: 10 unseen images from the each data set was used for testing. The results are as follows,

```
Class ID 1 :: correct # 10 out of 10, Incorrect # 0 :: percent correctly classified = 100.000000
Class ID 2 :: correct # 10 out of 10, Incorrect # 0 :: percent correctly classified = 100.000000
Class ID 3 :: correct # 9 out of 10, Incorrect # 1 :: percent correctly classified = 90.000000
Class ID 4 :: correct # 10 out of 10, Incorrect # 0 :: percent correctly classified = 100.000000
>> |
```

Of all test image, only one was failed. And the **average accuracy** was found to be **97.5%**
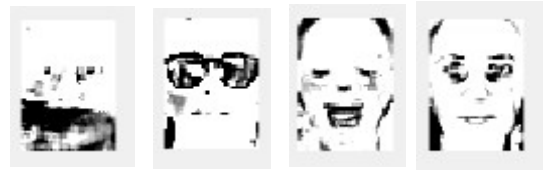
Accuracy
The following is the average accuracy:

Mixer Model	PCA
97.5%	89%

Q2 Reconstruction of images

The images that are reconstructed from mixer model subspace (i.e. using mu and sigma from the generated model) are as follows,

Using P-dimension

The above are some randomly chosen images from the generated images. One can clearly see that, the generated image has the vital feature (namely scarf, cooling glass, laugh and normal expression) in them. The above was using all the p-dimension

Using PCA subspace:

Interestingly, when the images generate using PCA subspace were check, some significant images were found, which could be used as a sample template for classification. These images had the exact template for the classes (Where white being considered as the main feature). Those images are as follows,

Scarf   cooling glass   laugh   normal
(Where the white pixels are considered as vital features)

Using ICA Subspace:

Scarf   cooling glass   laugh   normal

Although, the images generate from ICA seemed to have some vital features, it did not generate the expected images. It was like, one in every ten image has some significant feature in it. Thus, ICA did not really well in case of reconstruction. But, it can used for classification provided that we know the appropriate vectors which can be used for classification.

Using LDA Subspace:
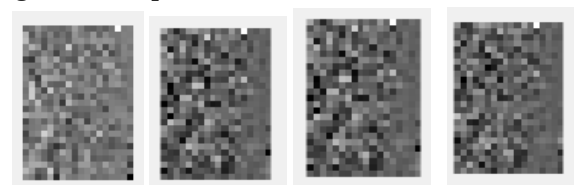
For, the generated image were found to be more pixated.