SMAI Assignment 2

Akshay Gupta, Roll No 20171058

3 November 2019

1 Problem-1

1.1 Eigenfaces

Eigenfaces are the eigenvectors of the covariance matrix of face images, and are used in the field of computer vision in facial recognition, and in general pattern recognition.

1.2 Number of Eigenfaces required

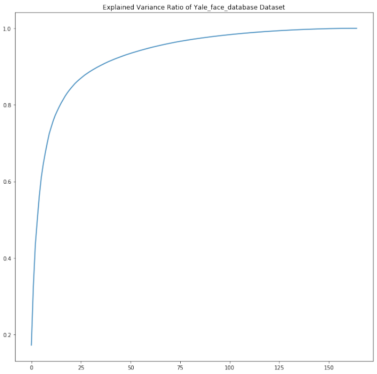
The number of eigenfaces required to satisfactorily represent the images in each of the datasets is as follows:-

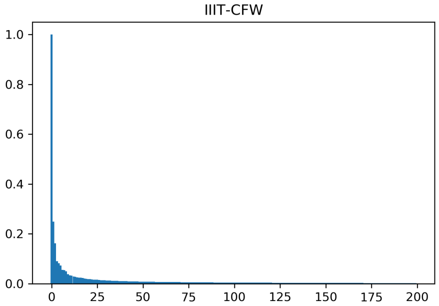
IMFDB : 60

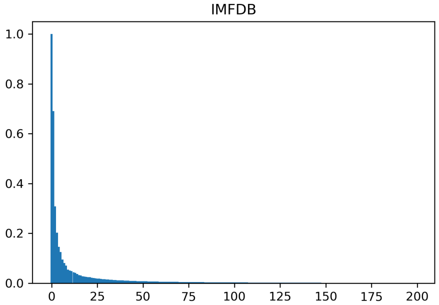
Yale Face Database : 75

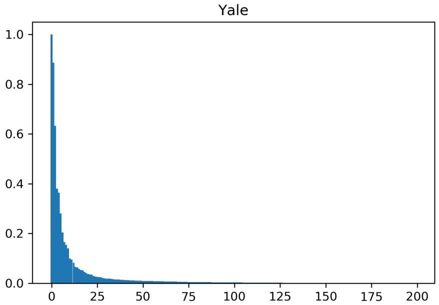
IIIT-CFW : 72

The numbers were determined by analysis of the eigenvalue spectrum of the datasets and keeping the eigenvalues such that their value is above a threshold.







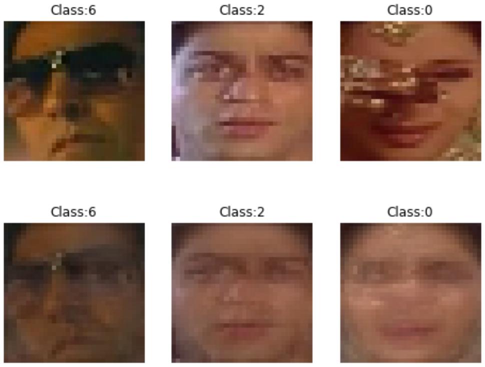


We keep those eigenvectors/faces whose sum of eigenvalues is 0.95 of the total sum.

1.3 Reconstructed Images and related problems

In IMFDB, the most difficult people to represent are Shilpa Shetty and Katrina Kaif. In IIIT-CFW it is Amitabh Bachan and Aishwarya Ray. This was determined empirically, by calculating reconstruction error for each class separately. The most difficult to represent classes were the ones which had error greater than 0.95 of the max reconstruction error.

Some reconstructed images are shown below.



2 Problem 2

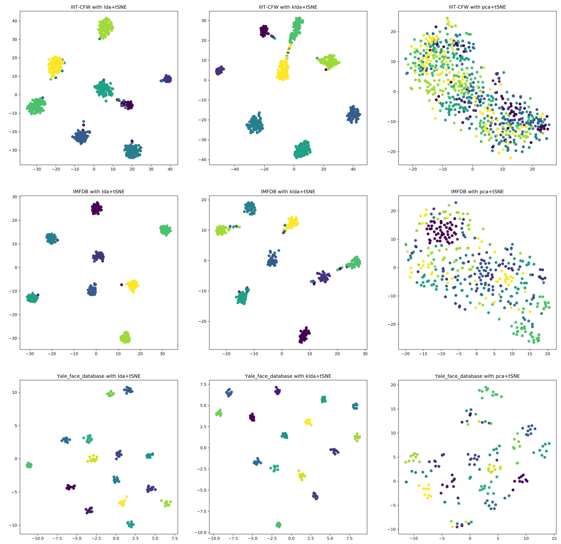
2.1 Classification Results

The classifer used to get the performance for variousfeatures is MLP. Since in MLP weights are initialized randomly features some set of features perform almost same. Concatenation of features performs good like concatenation of all original features or concatenation of LDA and KLDA.Analysing the dataframe and taking the original features (no concatenation).For every dataset ResNetfeature performs best and KPCA worst.

3 Problem 3

3.1 t-SNE Plots

The 2D t-SNE and PCA plots are shown below:-



3.2 Observations

It is observed that, with the exception of a few outliers, the data points with the same classes are closer and are clustered in the t-SNE.

4 Problem 4

4.1 Formulation

For the problem, we first extract the features(using PCA, LDA, or other variants), and train a KNN classifier. Then, given a data point X and the corresponding class ID, we find the predicted ID(using our KNN classifier), and return yes or no according to whether it matches the given class ID or not.

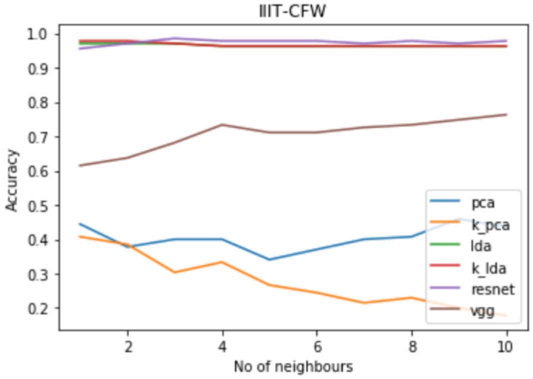
Metric used to compare performance is accuracy

$$Accuracy = \frac{No.of Correct Predicitions}{Total Samples} \qquad (1)$$

Another metric that used is precision

$$Precision = \sum_{C\epsilon Classes} \frac{TP(C)}{TP(C) + FN(C)} \qquad (2)$$

The accuracy v/s no. of neighbours graph was plotted for different features to check for the best value of k in KNN. The plot for IIIT-CFW is shown below



The plot for each dataset was analysed to select a value for no of nearest neighbors. Resnet performed good for all the datasets for k equal to 5.

5 Problem 5

5.1 Problem Statement

The problem I have chosen is that of emotion prediction. The problem is evidently not trivial - given a photo of a person, it is difficult to identify the features that may help us predict the emotion of the person. As such the problem may be intractable for non machine learning methods.

5.2 Applications

If we are able to build a system that can detect the emotion of a person with a single photo, it may be put to use in diverse ways. Some of the applications are :-

- For Blind People :- The emotion of narrator in a video can be described to a blind person so that he better interprets the context of the video.

- Robot emotion :- Applying this classifier on photos of robots, the emotions of robots can be predicted.

5.3 Model Pipeline

- We first take the Yale and IMFDB datasets and concatenate the two datasets to form a single dataset.

- Use ResNet features as our training set(additionally, if we have lower computational power, we may use PCA or LDA on this dataset to reduce dimensionality)

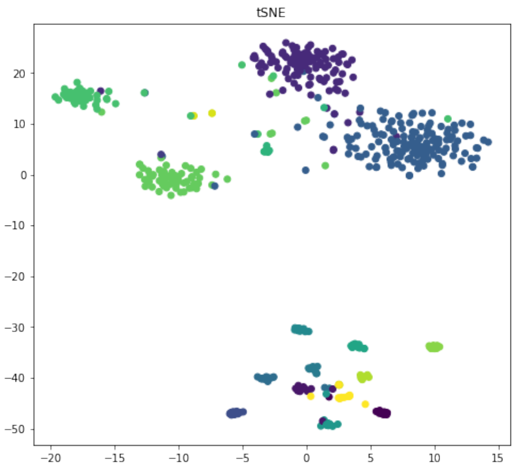
- Split the data into training and testing sets, then train our classifier.

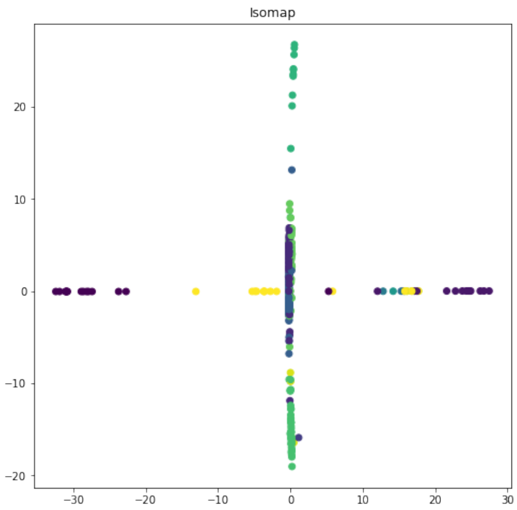
- Validate our classifier on the testing set.

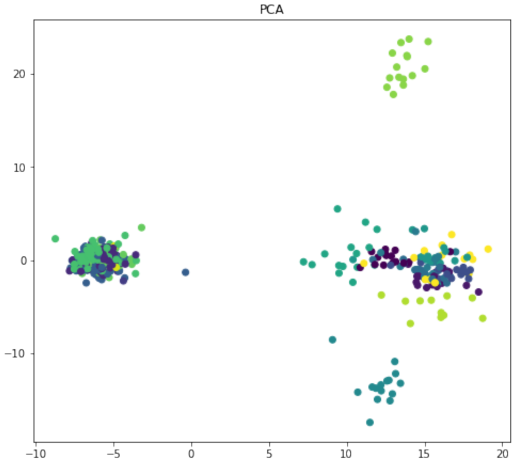
5.4 Results

The MLP classifier with KLDA features gives a cross validation score of 91.8 % and accuracy of 92.92 % after train test split.

The t-SNE, PCA, Isomap visualization of the data in 2D is as shown below.


tSNE


Isomap


PCA