Assignment-2 : Report

Meher Shashwat Nigam, 20171062

I. EIGENFACES

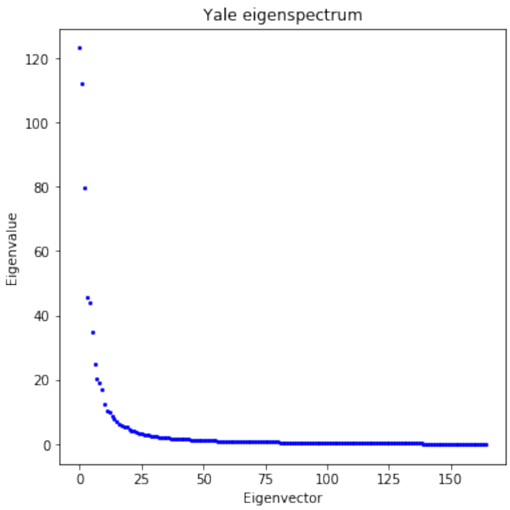
A. What are Eigenfaces?

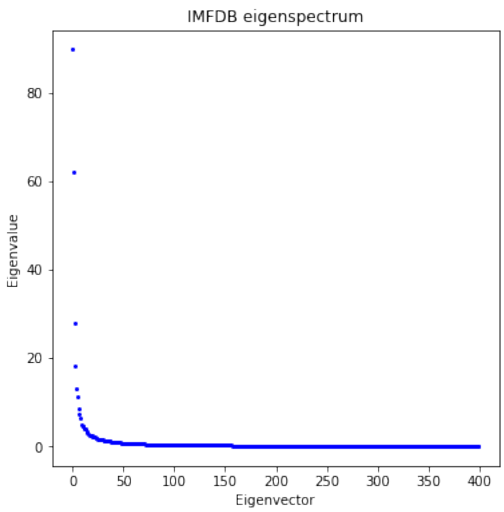
Eigenfaces is the name given to a set of eigenvectors obtained from principal component analysis (PCA) of a high-dimensional vector space of face images. The eigenfaces themselves form a basis set of all images used to construct the covariance matrix.

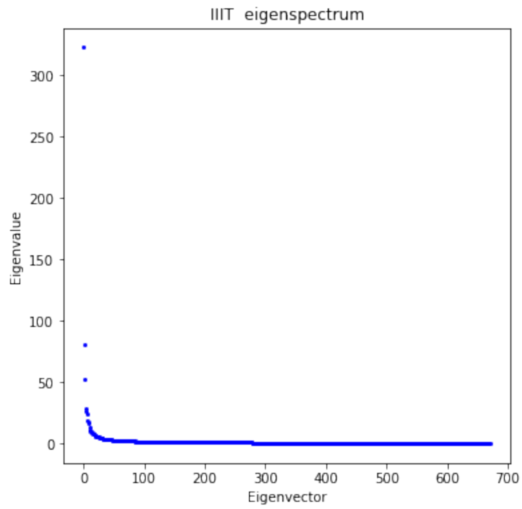
B. How many eigenvectors/faces are required to satisfactorily reconstruct a person in the three datasets?

For representing 95% of image data variance we take the top eigenvectors whose sum of corresponding eigenvalues is 95% of the sum of all eigenvalues.

Observing the eigenvalue spectrum of the three datasets in descending order:







Obtained number of eigenvectors for 95% variance representation:

Dataset	No. of eigenvectors
Yale	65
IMFDB	128
IIIT-CFM	360

C. Which person/identity is difficult to represent compactly with fewer eigen vectors?

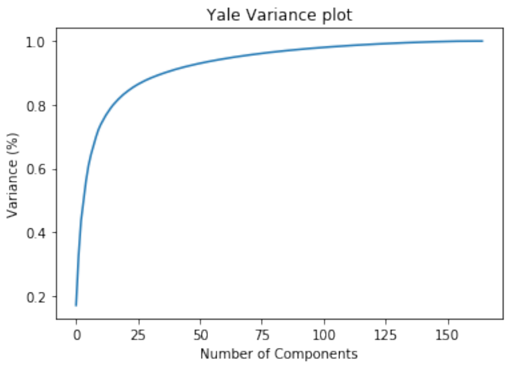
For a given dataset each class was reconstructed using the same number of eigenvectors (as obtained above) for 95% variance representation.

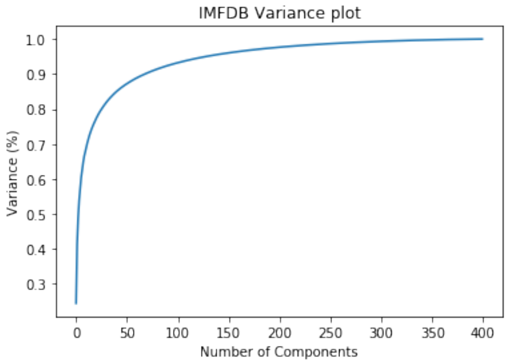
The class having the maximum reconstruction error in each dataset is as shown:

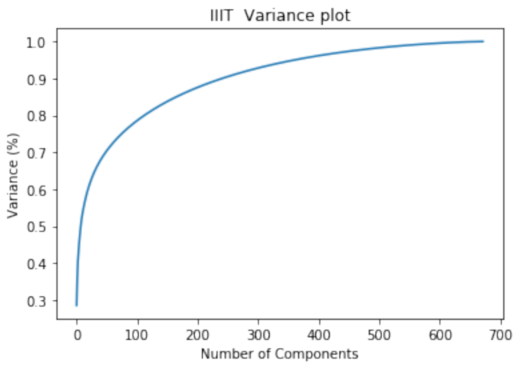
Dataset	Highest reconstruction error
Yale	Class 0
IMFDB	Class 2
IIIT-CFM	Class 2

D. Which dataset is difficult to represent compactly with fewer eigen vectors?

As explained above from the eigenvalue spectrum, we calculate how many eigenvectors (principal components) would be required to represent 95% variance of the whole dataset. The following plots show the normalized cumulative sum of eigenvalues for representing the datasets:







It can be seen easily that the IIIT-CFM cartoon dataset has a less steep curve and requires more eigenvectors for representation. This is due to the fact that cartoons don't have similar facial features and have very unique features per sample.

II. MLP Classification

A. Making the MLP classifier

Made a Multilayer perceptron classifier with the following architecture and properties:

Hidden Layers	2 layers, each of size 1000
Solver	Adam
Activation	ReLU
Maximum Iterations	350

B. Feature extraction

For each dataset the features were extracted using the following methods:

- PCA
- Kernel PCA
- LDA
- Kernel LDA
- ResNet features
- VGGFace
- LDA of VGGFace features
- PCA of ResNet features

C. Training and Validation

Each dataset was split into train and validation data in a 3:1 ratio. The classifier was trained on features extracted from the train data. The validation data was projected onto the obtained features and tested using the trained classifier obtained.

D. Results obtained

Yale Dataset

Features/Combination of features	Reduced Dimensional Space	Classification error	Accuracy	F1 score
PCA	65	16.666667	83.333333	0.818878
Kernel PCA(poly=3)	65	7.142857	92.857143	0.875556
LDA	14	0.000000	100.000000	1.000000
Kernel LDA(poly=3)	14	0.000000	100.000000	1.000000
VGG	14	39.393939	60.606061	0.565629
Resnet	14	0.000000	100.000000	1.000000
LDA of VGG	14	35.714286	64.285714	0.615952
PCA of resnet	65	0.000000	100.000000	1.000000

IMFDB Dataset

Features/Combination of features	Reduced Dimensional Space	Classification error	Accuracy	F1 score
PCA	128	16.00	84.00	0.844542
Kernel PCA(poly=3)	128	17.00	83.00	0.835527
LDA	7	25.00	75.00	0.751871
Kernel LDA(poly=3)	7	25.00	75.00	0.751871
VGG	7	10.00	90.00	0.908798
Resnet	7	8.75	91.25	0.917343
LDA of VGG	7	44.00	56.00	0.555811
PCA of resnet	65	7.00	93.00	0.931552

IIIT-CFM Dataset

Features/Combination of features	Reduced Dimensional Space	Classification error	Accuracy	F1 score
PCA	360	36.904762	63.095238	0.623251
Kernel PCA(poly=3)	360	35.714286	64.285714	0.625336
LDA	7	63.095238	36.904762	0.362442
Kernel LDA(poly=3)	7	63.095238	36.904762	0.362442
VGG	7	29.629630	70.370370	0.667560
Resnet	7	1.481481	98.518519	0.985303
LDA of VGG	7	41.666667	58.333333	0.552684
PCA of resnet	65	1.785714	98.214286	0.982361

E. Best methods for datasets

Yale dataset gave 100% accuracy with LDA and ResNet methods.

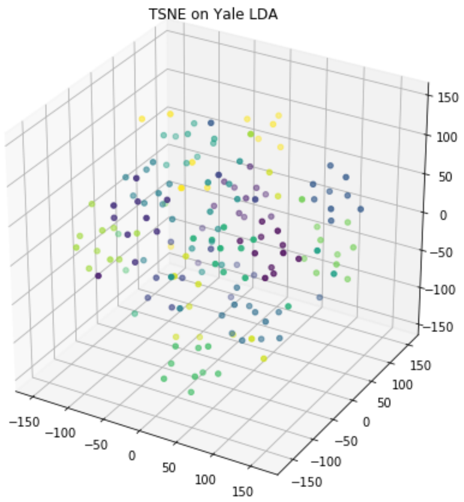
IMFDB dataset gave 93% accuracy with features obtained from PCA of ResNet features

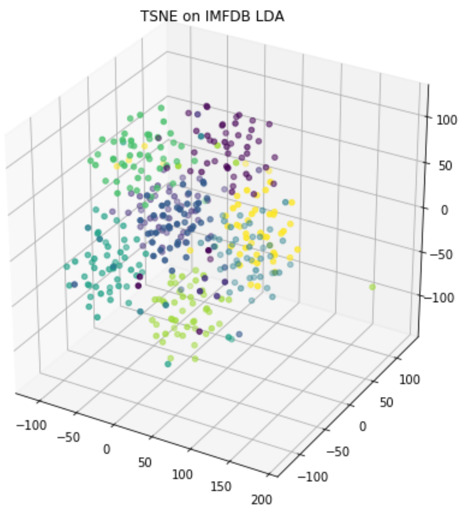
IIIT-CFM dataset gave 98.5% accuracy with features obtained from ResNet.

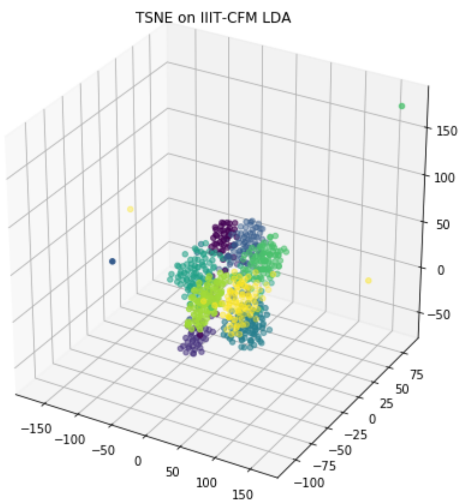
III. t-SNE based Visualization

t-Distributed Stochastic Neighbor Embedding (t-SNE) is an unsupervised, non-linear technique primarily used for data exploration and visualizing high-dimensional data.

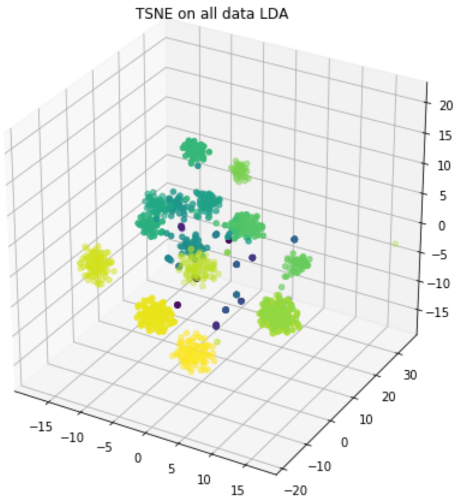
A. Visualization of t-SNE on datasets after applying LDA







B. Visualization of t-SNE on combined dataset after LDA



In the 3D t-SNE scatter plot for all the datasets combined, it can be seen that samples with the same label are closer to each other than others with a few exceptions like the purple scatters, which denotes large intra-class variance.

IV. Face Verification using KNN

A. Formulation of the problem using KNN

We split the data into train and validation and do feature extraction using various methods on the train data. Then, we train the KNN classifier with this data. The validation data is projected onto the features initially obtained from the train data and the model is tested against the ground truth labels.

Given a sample, we first apply the same feature extraction to the sample and then with the KNN, find the K-nearest neighbours to the sample and assign the label through majority voting.

B. Accuracy and precision for all datasets for various features

Yale Dataset

Features/Combination of features	Reduced Dimensional Space	Accuracy	Classification error	Precision
PCA	65	80.952381	19.047619	0.861111
Kernel PCA(poly=3)	65	80.952381	19.047619	0.861111
LDA	14	92.857143	7.142857	0.940476
Kernel LDA(poly=3)	14	92.857143	7.142857	0.940476
VGG	14	69.696970	30.303030	0.666667
Resnet	14	100.000000	0.000000	1.000000
LDA of VGG	14	69.047619	30.952381	0.596296
PCA of resnet	65	100.000000	0.000000	1.000000

IMFDB Dataset

Features/Combination of features	Reduced Dimensional Space	Accuracy	Classification error	Precision
PCA	128	54.00	46.00	0.561711
Kernel PCA(poly=3)	128	54.00	46.00	0.563839
LDA	7	82.00	18.00	0.828730
Kernel LDA(poly=3)	7	82.00	18.00	0.828730
VGG	7	91.25	8.75	0.906534
Resnet	7	92.50	7.50	0.934325
LDA of VGG	7	63.00	37.00	0.659220
PCA of resnet	65	93.00	7.00	0.941487

IIIT-CFM Dataset

Features/Combination of features	Reduced Dimensional Space	Accuracy	Classification error	Precision
PCA	360	22.023810	77.976190	0.489140
Kernel PCA(poly=3)	360	25.000000	75.000000	0.521742
LDA	7	45.238095	54.761905	0.473022
Kernel LDA(poly=3)	7	45.238095	54.761905	0.473022
VGG	7	68.148148	31.851852	0.650242
Resnet	7	97.777778	2.222222	0.980176
LDA of VGG	7	60.119048	39.880952	0.549572
PCA of resnet	65	96.428571	3.571429	0.967142

C. *Accuracy and precision for all datasets by varying k*

Yale Dataset

K	Accuracy	Precision
3	95.238095	0.958333
5	92.857143	0.940476
7	95.238095	0.964286
9	95.238095	0.964286
12	92.857143	0.940476

IMFDB Dataset

K	Accuracy	Precision
3	83.0	0.833820
5	82.0	0.828730
7	82.0	0.825753
9	82.0	0.832089
12	81.0	0.817054

IIIT-CFM Dataset

K	Accuracy	Precision
3	41.666667	0.444979
5	45.238095	0.473022
7	45.238095	0.460913
9	44.047619	0.453854
12	44.642857	0.459166

V. Extension: Emotion Classification

A. *Problem Definition*

Create an emotion predicter/classifier. Given an image, identify the emotion being expressed. Both the Yale dataset and IMFDB dataset contain an 'emotion.txt. file, which holds ground truth emotion labellings for all images.

B. *Use cases and applications*

- Using facial emotion detection smart cars can alert the driver when he is feeling drowsy.
- Measure candidate's facial expressions during interviews to capture their moods and further assess their personality traits.
- Behavioral methods in market research use video feeds of a user interacting with a product, which is analyzed to gauge user's reactions and emotions.

C. *Experimental pipeline*

1) First load the images and the emotion.txt file. Map the images to their ground truth emotions to get the initial dataset and labellings.
2) Split the datasets into train and validation data in the ratio 4:1.
3) Use various feature extraction techniques (same of those used in the MLP classifier earlier.
4) Define a logistic regression classifier with C=0.01 for regularization.
5) Train the classifier for various features and evaluate metrics on the train data.

D. *Qualitative results*

Accuracy and other metrics for various features:

Yale Dataset

Features/Combination of features	Reduced Dimensional Space	Classification error	Accuracy	F1 score
PCA	65	60.606061	39.393939	0.300078
Kernel PCA(poly=3)	65	93.939394	6.060606	0.072727
LDA	10	45.454545	54.545455	0.453247
Kernel LDA(poly=3)	10	45.454545	54.545455	0.453247
VGG	10	93.939394	6.060606	0.050505
Resnet	10	100.000000	0.000000	0.000000
LDA of VGG	10	96.969697	3.030303	0.036364
PCA of resnet	65	100.000000	0.000000	0.000000

IMFDB Dataset

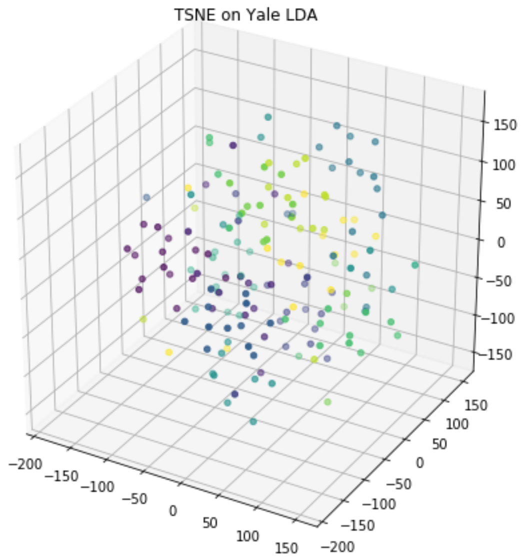
Features/Combination of features	Reduced Dimensional Space	Classification error	Accuracy	F1 score
PCA	128	37.50	62.50	0.413417
Kernel PCA(poly=3)	128	65.00	35.00	0.086420
LDA	6	46.25	53.75	0.345039
Kernel LDA(poly=3)	6	46.25	53.75	0.345039
VGG	6	51.25	48.75	0.316835
Resnet	6	47.50	52.50	0.349534
LDA of VGG	6	73.75	26.25	0.182424
PCA of resnet	65	37.50	62.50	0.483748

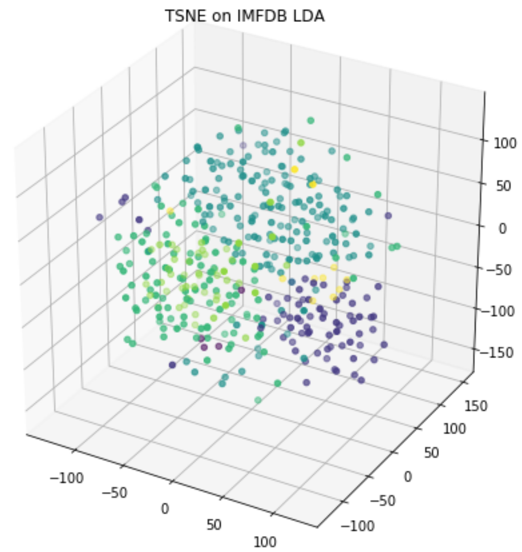
E. *Analyis of Qualitative results*

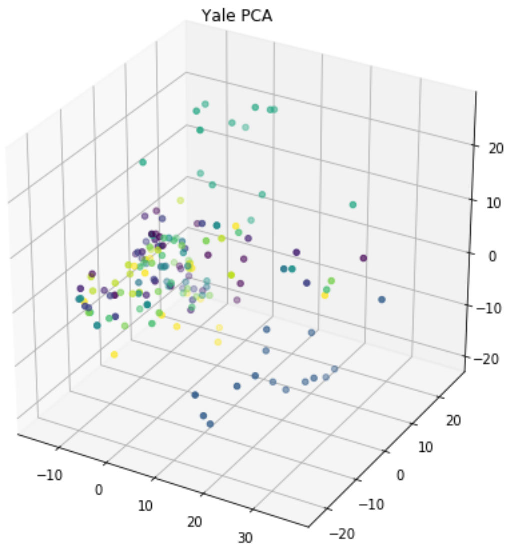
The best accuracy for Yale dataset was obtained from LDA feature extraction(54.5%) and on the IMFDB dataset, best accuracy was 62.5% obtained from PCA of ResNet features.

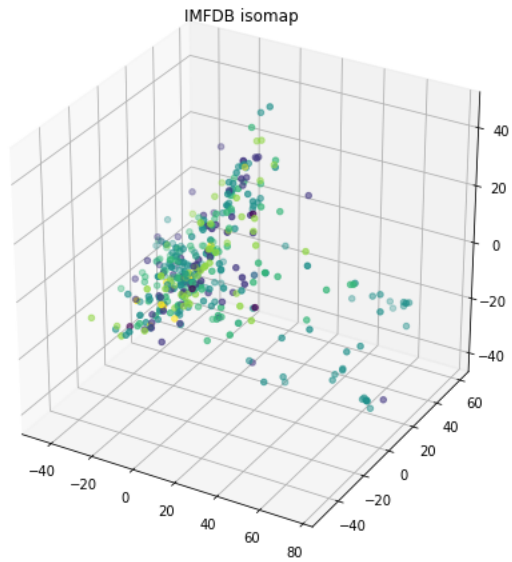
F. *Representation of data using different methods*

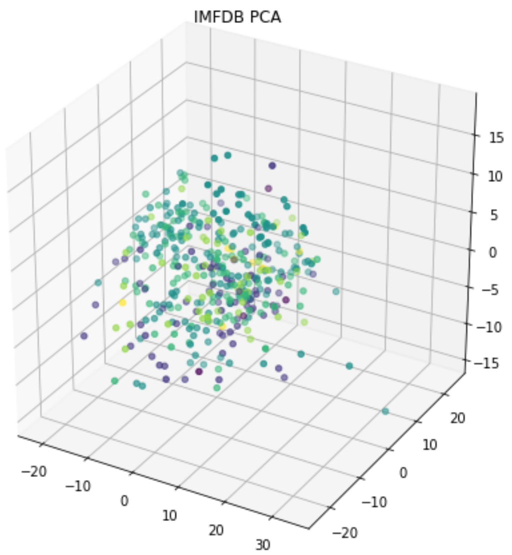
Here are the plots of TSNE (on LDA features), PCA and Isomaps for both datasets:

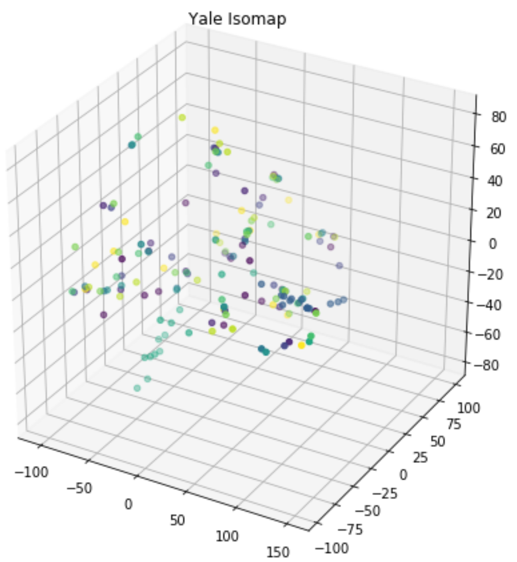












G. K-fold Validation

K fold validation for various values of of K applied on LDA features. Reported are the average accuracy and precision for each K.

For yale:

K splits	Accuracy	Precision
3.0	32.121212	0.407492
6.0	45.921517	0.432973
9.0	47.823262	0.460101
12.0	48.489011	0.415485
15.0	46.666667	0.410051
18.0	50.925926	0.393408
21.0	49.659864	0.396967

For IMFDB:

K splits	Accuracy	Precision
3.0	53.996933	0.388076
6.0	54.247701	0.387895
9.0	52.042649	0.354595
12.0	54.196376	0.445679
15.0	54.264008	0.452732
18.0	53.820817	0.462985
21.0	54.498747	0.444747

H. Quantitative results

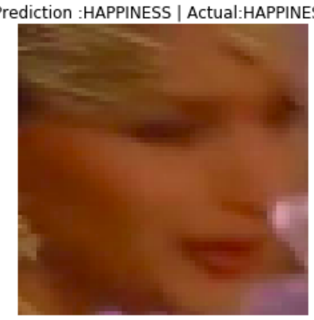
Examples of correct and wrong predictions from both datasets: Yale Dataset:

Prediction :centerlight | Actual:centerlight

Prediction :sad | Actual:happy

IMFDB Dataset:

Prediction :HAPPINESS | Actual:HAPPINESS



Prediction :NEUTRAL | Actual:HAPPINESS