SMAI ASSIGNMENT 2 - TECHNICAL REPORT

Author - Paryul Jain
Date of Submission : 3rd Nov 2019
paryul.jain@research.iiit.ac.in

1 Setup and How to Run

You need to have installed the following python packages : *numpy*, *pandas* , *sklearn* , *scipy* , *seaborn* and *jupyter-notebook* (or any other equivalent).

Download the **dataset** folder and the python notebook here.

Make sure to have the **dataset** folder in the same directory as the python notebook. (20171083.ipynb)

Open a jupyter-notebook and open the corresponding folder where the downloaded files are kept. Run the python notebook.

2 Datasets And Preprocessing

Three Data Sets are used. All the images have been reduced to 32*32 from the original sizes. Also the 3 channels have been kept intact.
1.IIIT-CFW Dataset : 672,32,32,3 (cfw_dict).
2.IMFDB Dataset : 400,32,32,3 (imfdb_dict).
3.YALE Dataset : 165,32,32,3. A dictionary for each dataset is made with the class name and class index which helps in identification of classes later. Each dataset is flattened to make further processing easy.

3 Feature Extraction Methods

We have used 6 feature extraction methods. PCA, LDA, Kernel PCA, Kernel LDA, VGG and RESNET. VGG and RESNET features were provided to us and are available in the corresponding dataset folders. Each function takes dataset, labels and number of dimensions as the arguments and returns extracted features in a new dimensional space.

For each dataset we then compute the extracted features for each method and show them in a 3D scatter plot.(Figure 1)

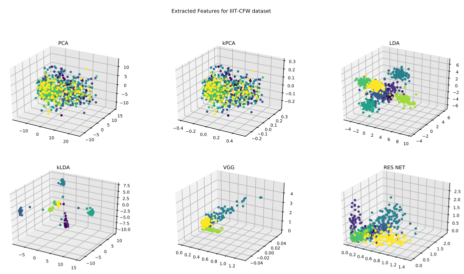

Figure 1: Extracted Features for IIIT-CFW

3.1 Eigen Faces and PCA

We use the concept of Eigen Faces - set of vectors which help to visualise images in a compressed form. They form a set of basis vectors to represent a set of images. We do Principal Component Analysis on all the three datasets and find the compressed form of the images. We then reconstruct back the images using this compressed set(Eigen Faces). Analysis of reconstruction error among the datasets and within the datasets has been done and some results are shown here.

IIIT-CFW data, being sparse requires 100 Principal Components to satisfactorily reconstruct the images back. And so does YALE (15 classes), requires around 75 PC's. IMFDB data requires 50 PC's. We plot the eign value spectrum of the datasets to infer this.(Figure2)

Class wise reconstruction also gave interesting results. As we know women are more expressive of thier emotions and also the IMFDB dataset has different color pics of some actors/actress , which make them more difficult to be reconstructed when less number of principal components are used. (Figure 3)

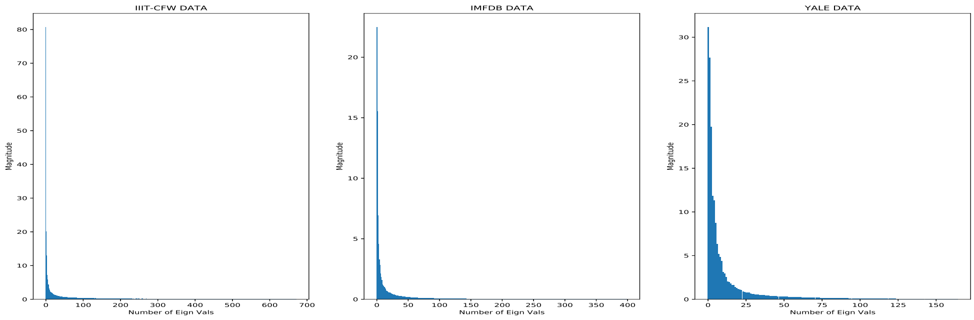

Figure 2: Eign Spectrums

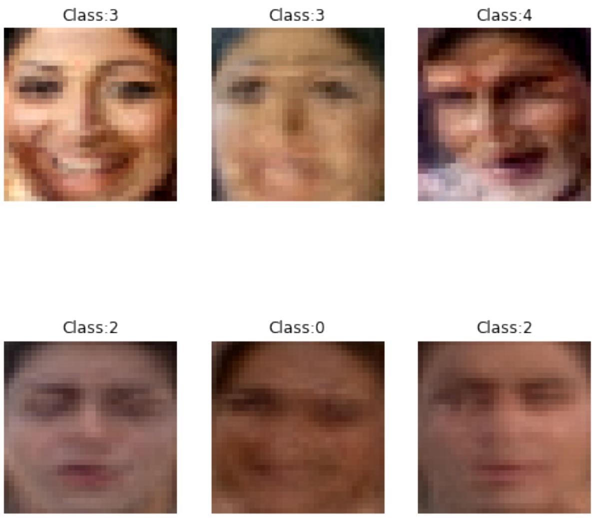

Figure 3: Reconstruction of IMDBF with PCA

4 Classification with Models

We user 2 models , MLP and KNN with different feature extraction techniques combined. We first The data is split into training and validation sets. From here we project the training data and validation data into different feature spaces using the above six mention feature extraction algorithms. We also analyse the data with multiple feature extraction techniques applied together to attain better accuracy.

4.1 MLP

We use a MLP (Multi Layer Perceptron) with 100 and 50 as sizes of linear hidden layers. Solver used is adam and the hidden layers have activation Relu.

4.2 KNN

A K-Nearest Neighbour algorithm is also used for classification of From here we project the training data and validation data into different feature spaces using the above six mention feature extraction algorithms. We also analyse the data with multiple feature extraction techniques applied together to attain better accuracy.

4.3 Results and Analysis

Results (Figure 4,5,6) show that RESNET features gives very high accuracy on all the datasets in case of MLP and KNN. Infact in case of RESNET + KNN the accuracy is 100% on YALE dataset. Since RESNET uses deep neural networks with skip connections and hence allows to train the model with many layers and thus prevents saturation and overfitting. These properties make the RESNET features very good for training and hence we see very high accuracy.

Simple PCA with KNN on YALE gives very low accuracy since all data is reduced in very low dimensions , can think of as points on a single line. Now here the chances of misclassifications are very high since we do not have enough dimensions to calculate distance. PCA is also not good since it does not take into the information of classes which LDA does and we can see from the KNN results that LDA performs better than PCA for all the datasets. (Figure 7,8,9). Here we also see that PCA is not good for classification. Which is infact true since it is generally used as noise removal , dimensionality reduction technique where PCA can throw away the discriminant dimensions.

We also show the confusion matrix between the classes of the dataset in case of MLP. Misclassifications can be explained, for instance in IMFDB dataset, there are some dark spots,

	Reduced Dimension	Error	Precision	Recall	F1-Score
PCA	75.0	50.370370	0.476960	0.487542	0.475749
KPCA	75.0	46.666667	0.518824	0.532512	0.522099
LDA	3.0	73.333333	0.252794	0.248400	0.243573
KLDA	7.0	82.962963	0.175948	0.178253	0.171330
VGG	4096.0	34.074074	0.627247	0.619664	0.613903
ResNet	2048.0	2.222222	0.975486	0.983820	0.979289
VGG + PCA	75.0	31.111111	0.661349	0.650267	0.652292
VGG + LDA	3.0	52.592593	0.468139	0.434944	0.434235
ResNet + PCA	75.0	2.222222	0.977083	0.983820	0.979945
ResNet + LDA	3.0	8.888889	0.903632	0.893531	0.895961

Figure 4: IIIT- CFW DATASET-MLP

	Reduced Dimension	Error	Precision	Recall	F1-Score
PCA	60.0	13.75	0.869769	0.884241	0.862957
KPCA	60.0	15.00	0.839078	0.860374	0.842577
LDA	3.0	42.50	0.567441	0.588427	0.551553
KLDA	7.0	26.25	0.745386	0.749429	0.724430
VGG	4096.0	8.75	0.922917	0.911298	0.914296
ResNet	2048.0	5.00	0.945933	0.948664	0.945486
VGG + PCA	60.0	11.25	0.904315	0.882077	0.890671
VGG + LDA	3.0	57.50	0.428233	0.435596	0.405710
ResNet + PCA	60.0	7.50	0.922771	0.929433	0.923553
ResNet + LDA	3.0	15.00	0.863636	0.873489	0.858733

Figure 5: IMFDB DATASET-MLP

	Reduced Dimension	Error	Precision	Recall	F1-Score
PCA	40.0	12.121212	0.923810	0.910714	0.895748
KPCA	40.0	12.121212	0.821429	0.886905	0.832653
LDA	3.0	27.272727	0.652381	0.690476	0.639683
KLDA	14.0	9.090909	0.886667	0.844444	0.845926
VGG	4096.0	45.454545	0.529762	0.494048	0.479762
ResNet	2048.0	3.030303	0.964286	0.964286	0.952381
VGG + PCA	40.0	48.484848	0.445556	0.450000	0.424762
VGG + LDA	3.0	69.696970	0.202222	0.238889	0.192540
ResNet + PCA	40.0	3.030303	0.964286	0.964286	0.952381
ResNet + LDA	3.0	18.181818	0.854762	0.821429	0.815873

Figure 6: YALE DATASET-MLP

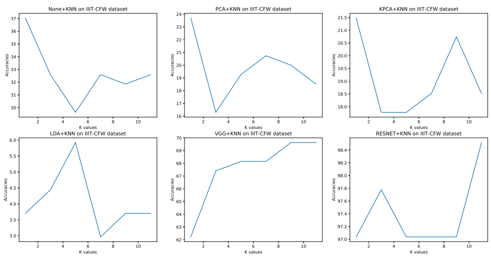

Figure 7: IIIT- CFW DATASET-KNN

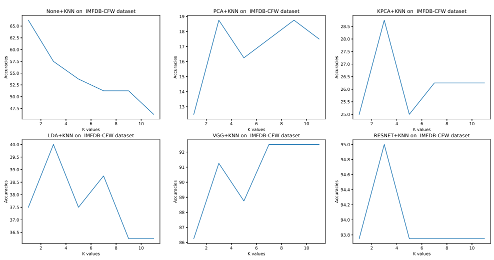

Figure 8: IMFDB DATASET-KNN

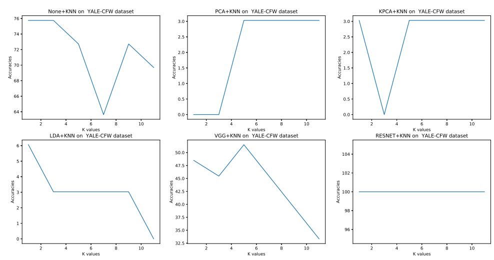

Figure 9: YALE DATASET-KNN

these are the classes of 2 similar looking actresses.

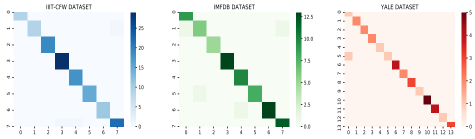

Figure 10: Confusion Matrices for MLP

5 TSNE

t-Distributed Stochastic Neighbor Embedding (t-SNE) is a non-linear technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional datasets.

t-Distributed stochastic neighbor embedding (t-SNE) minimizes the divergence between two distributions: a distribution that measures pairwise similarities of the input objects and a distribution that measures pairwise similarities of the corresponding low-dimensional points in the embedding.

In this way, t-SNE maps the multi-dimensional data to a lower dimensional space and attempts to find patterns in the data by identifying observed clusters based on similarity of data points with multiple features. However, after this process, the input features are no longer identifiable, and you cannot make any inference based only on the output of t-SNE. Hence it is mainly a data exploration and visualization technique.

6 Application

6.1 Gender Classification Problem

Given a dataset of images with multiple genders present. We need to predict the gender of any image when it is feeded onto the system.

6.2 Preprocessing

The dataset just needs to be changed at the labels , which is done using the load_data() function. Males are assignned Class 0, while females 1. We split the Datasets in 8:2 ratio for training and validation respectively. We then apply feature extraction techniques to the training and validation sets and pass on these to train the model.

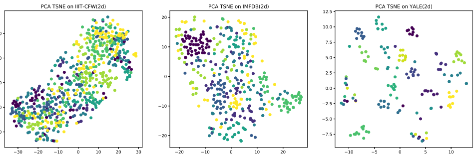

Figure 11: PCA+TNSE 2D

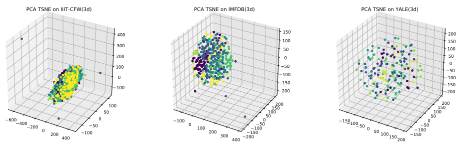

Figure 12: PCA+TNSE 3D

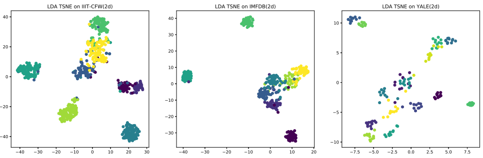

Figure 13: LDA+TNSE 2D

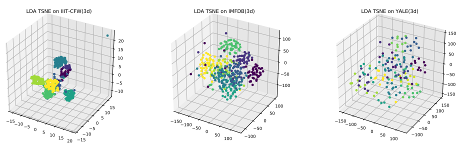

Figure 14: LDA+TNSE 3D

6.3 Classifier

Approach used is same as discussed in section 4 introduction and 4.1. The model uses the training data and learns the weights for the layers. The validation set is then passed to predict the labels.

6.4 Results And Analysis

We compute the accuracy, error, precision, recall and f1-score for both the datasets. Tables(Fig.15,16) with accuracy shows that RESNET is the best feature extraction technique. MLP classifier is used and gives a very high accuracy. ISOMAPS of both the datasets are shown (Fig. 17,18) in both 2D and 3D after applying PCA. Also some of the miclassified images of IIIT-CFW and IMFDB are shown here.(FIgure 19,20)

The model is very useful and has many applications and forms the base for further classifications.

	Reduced Dimension	Error	Precision	Recall	F1-Score
PCA	75.0	11.851852	0.846284	0.792857	0.814815
KPCA	75.0	14.074074	0.805455	0.766667	0.783087
LDA	1.0	25.925926	0.627947	0.630952	0.629383
KLDA	1.0	29.629630	0.552929	0.547619	0.549399
VGG	4096.0	2.222222	0.973325	0.961905	0.967467
ResNet	2048.0	0.740741	0.995283	0.983333	0.989156
VGG + PCA	75.0	5.185185	0.968750	0.883333	0.917833
VGG + LDA	1.0	11.111111	0.841574	0.833333	0.837336
ResNet + PCA	75.0	0.740741	0.995283	0.983333	0.989156
ResNet + LDA	1.0	1.481481	0.978571	0.978571	0.978571

Gender classification for IIIT-CFW Dataset

Figure 15: IIIT-CFW-GENDER-DATASET

	Reduced Dimension	Error	Precision	Recall	F1-Score
PCA	60.0	10.00	0.901515	0.898496	0.899434
KPCA	60.0	5.00	0.949875	0.949875	0.949875
LDA	1.0	12.50	0.875000	0.875940	0.874922
KLDA	1.0	12.50	0.876263	0.873434	0.874293
VGG	4096.0	1.25	0.988372	0.986842	0.987451
ResNet	2048.0	0.00	1.000000	1.000000	1.000000
VGG + PCA	60.0	3.75	0.962164	0.963033	0.962447
VGG + LDA	1.0	16.25	0.839683	0.835213	0.836246
ResNet + PCA	60.0	1.25	0.987179	0.988095	0.987482
ResNet + LDA	1.0	1.25	0.988372	0.986842	0.987451

Gender classification for IMFDB Dataset

Figure 16: IMFDB GENDER DATASET

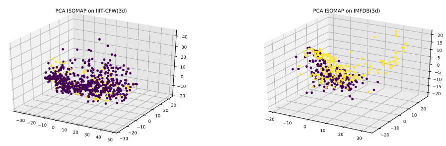

Figure 17: ISOMAP for GENDER Dataset using PCA(3D)

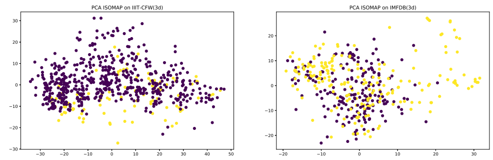

Figure 18: ISOMAP for GENDER Dataset using PCA(2D)

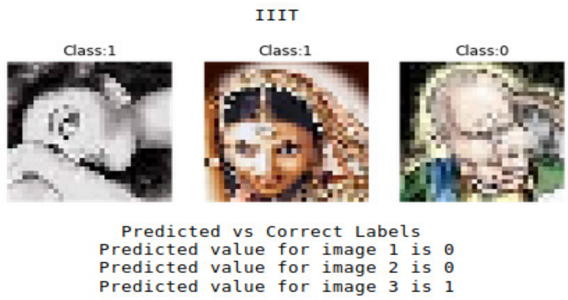

Figure 19: Misclassified Images IIIT-CFW Dataset

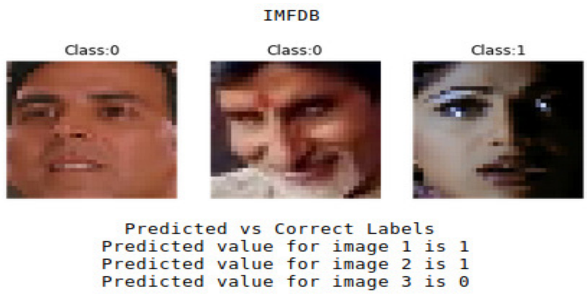

Figure 20: Misclassified Images IMFDB Dataset