Assignment 1

Swathi Guptha
International Institute of Information Technology, Hyderabad
Hyderabad, India, 500032
mittalswathi08@gmail.com

This report deals with implementing various feature extraction techniques on IMFDB dataset, CFW dataset and YALE dataset. Various combinations of feature extractions are fed to classifiers and performance is analyzed for the same.

1. Eigen Faces

The dataset is loaded from the respective directories. Few pre-processing techniques like resizing (32 X 32), greyscale scale conversion are performed. In order to learn features of classes efficiently the dataset is shuffled and is split into train-test dataset (3:1).

1.1. What are eigen faces?

Eigen faces are an orthogonal basis set from which most faces can be constructed. They are blurry depictions of faces that each highlight a certain type of feature. The idea is that anyone's face can be reconstructed from a suitable linear combination of eigen faces. They are used in facial characterization and recognition applications.

1.2. How many eigen vectors/faces are required to \satisfactorily" reconstruct a person in these three datasets?

To determine the number of eigen face satisfactory to reconstruct a person's face, recognition rate is taken as performance metric. Here recognition rate is defined as ratio of successfully recognized test images to the total number of test images. E.g. there are 100 test images out of which it recognizes 60 test images successfully i.e. whether they are known faces or unknown faces, and then recognition rate is 60 %. for all the three dataset. From fig (1) it can be inferred that eigen vectors with variance around 90 % are sufficient to get the maximum recognition rate possible for the given data set.

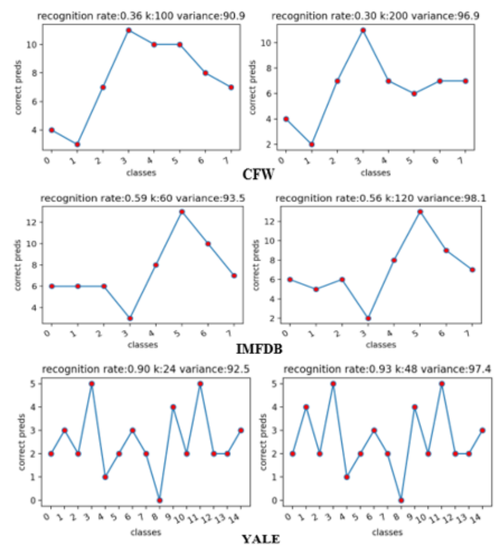

Fig (1) Shows recognition rate for different eigen values(variance)

1.3. Reconstruction of eigen faces

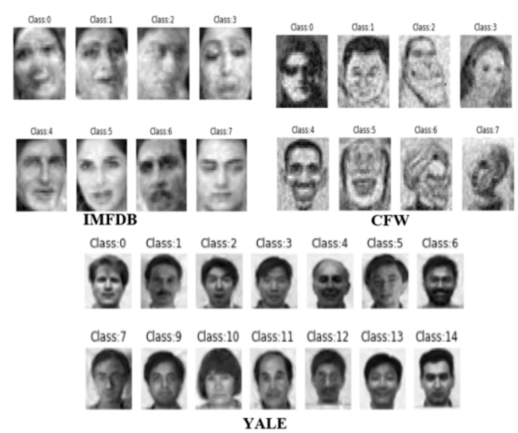

Fig (2) Reconstructed Images

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

From fig (1) and fig (2) it can be inferred that the CFW data has minimum recognition rate and maximum distorted (noisy) images respectively. The reason behind CFW having such low performance rate maybe because they are not natural faces. The features present in natural face might be more distinguishable from class to class than the features present in a cartooned face.

2. Result on various combinations of feature extractors and classifiers

F1 score for IMFDB data

	logistic	SVM	MLP	decision tree
PCA	0.77	0.79	0.84	0.19
kPCA	0.1	0.1	0.11	0.1
LDA	0.58	0.37	0.58	0.31
kLDA	0.1	0.1	0.1	0.1
PCA+LDA	0.1	0.1	0.12	0.1
VGG	0.88	0.4	0.89	0.53
RESNET	0.95	0.45	0.95	0.56

F1 score for CFW data

	logistic	SVM	MLP	decision tree
PCA	0.45	0.29	0.48	0.18
kPCA	0.12	0.12	0.14	0.14
LDA	0.27	0.24	0.29	0.17
kLDA	0.12	0.12	0.12	0.12
PCA+LDA	0.15	0.15	0.14	0.15
VGG	0.62	0.57	0.61	0.54
RESNET	0.99	0.34	0.98	0.51

F1 score for yale data

	logistic	SVM	MLP	decision tree
PCA	0.98	0.71	0.93	0.21
kPCA	0.02	0.02	0.1	0.07
LDA	0.9	0.26	0.93	0.07
kLDA	0.02	0.02	0.02	0.02
PCA+LDA	0.05	0.07	0.07	0.02
VGG	0.71	0.02	0.69	0.14
RESNET	1.0	0.21	0.98	0.19

Fig (3) Performance table

Various combinations of feature extractors (PCA, kernel-PCA, LDA, kernel-LDA, VGG, Resnet) and classifiers (Logistic regression, SVM, MLP, decision tree) are implemented on all the dataset for face recognition. From fig (3) it is pretty evident that feature extractor like PCA, LDA, VGG, Resnet are performing efficiently with classifiers like Logistic regression and MLP.

3. T – SNE

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.

3.1. Use t-SNE based visualization of faces

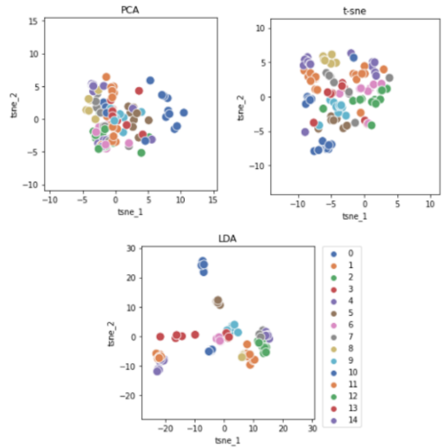

Fig (4) Visualization of dimensionality reduction algorithms

Like pca and LDA, t-sne is also a dimensionality reduction algorithm. However, pca retains only global variance unlike t-sne which retains local-variance. From fig (4), PCA has tried to separate the different points and form clustered groups of similar points. When compared to PCA, in t-sne the classes are clustered in their very little group. And LDA visualization maximizes the class separation.

4. KNN

KNN is a non-parametric and lazy learning algorithm. Non-parametric means there is no assumption for underlying data distribution. In other words, the model structure determined from the dataset.Lazy algorithm means it does not need any training data points for model generation. All training data used in the testing phase.

4.1. formulate the problem using KNN

	IMFDB	CFW	YALE
Accuracy	0.550	0.381	0.548

k-NN is one of the foremost basic classification algorithms in machine learning. It belongs to the supervised learning class of machine learning. k-NN is usually employed in search applications wherever you're looking for "similar" things. KNN can be applied for face verification. So, whenever a face need to be verified, we can measure how similar the new face is from the existing verified faces. The similarity can be measured by distance metric. The accuracy metric can be used for performance evaluation.

Link: Report