Face Recognition Using Bayes Classifier and KNN Classifier

Sahil Khanna
Department of Telecommunication
University of Maryland College Park
skhanna9@umd.edu

I. Introduction

Face Identification has been a major topic of discussion and has been widely used as a solution for many security problems. In this project, face recognition is implemented using Bayes rule and KNN rule. The purpose of this project is to gain the insights on these rules and their implementations. This project provided me with the opportunity to explore the conceptual and practical implementation of various machine learning concepts like Bayes Decision Rule, K-nearest neighbor rule, Principle Component Analysis and Fisher Linear Discriminant Analysis.

II. Difficulties Faced

Some of the challenges, I solved during the development of this projects are as follow:

1) Covariance matrix is singular because of fewer samples than dimensions. This problem has two solutions: First solution is to use Pseudo Inverse and Second solution is to add a small constant to the diagonal of the covariance matrix.
2) Memory out of bound error in case of calculating covariance matrix for each class. This problem was solved using dimension reduction tools such as PCA and LDA.

III. Algorithm

The algorithm used for both the models is written in Octave. Octave is a high-level language which has a same interface as MATLAB. The algorithm uses some library functions:

1) Cov(): It is used to calculate the covariance matrix.
2) Eig(): This function is used to calculate the eigen vectors and eigen values.
3) Mean(): This function is used to calculate the mean of the matrix.
4) Randperm(): This function is used to take random output in case of the tie situation in KNN model.
5) Unique() and Accumarray(): These function are used to calculate the k nearest neighbors and there majority labels.
6) Max(), Count(), Sqrt(), Sum(), Sort(), Nnz(): Some of general purpose functions which are utilized at various places throughout the algorithm.

IV. Results

The efficiency of the models is calculated based on the accuracy of the predicted labels. The observations are taken by varying the different parameters such as K values, dimensions of the reduced vectors, using different decision rules and varying percentage of retained variance while implementing PCA. I have considered data.m file for my training and testing. You can refer figure 1 for the visualization of the data set.

1) The first result of Bayes Classifier is taken by giving two faces for each class as a training sample and third face as a testing sample. It is shown in table 1 that the accuracy of the Bayes Classifier is less when dimension is reduced using linear discriminant analysis as compare to principle component analysis. In this case, I did not perform accuracy calculation without the reduction because the dimension of the covariance matrix was huge and octave was unable to handle this large memory.

Table 1 Accuracy of the Bayes Classifier.

Bayes Classifier		
	Principle Component Analysis (retains 95% variance)	Linear Discriminant Analysis
Accuracy	61.5	53.5

2) The second result of Bayes Classifier is taken by changing the retained percentage of variance while

implementing principal component analysis providing the same samples. It can be seen from the Table 2 that the accuracy of the classifier increase as the dimensions increases. The trend can be seen in figure 2.

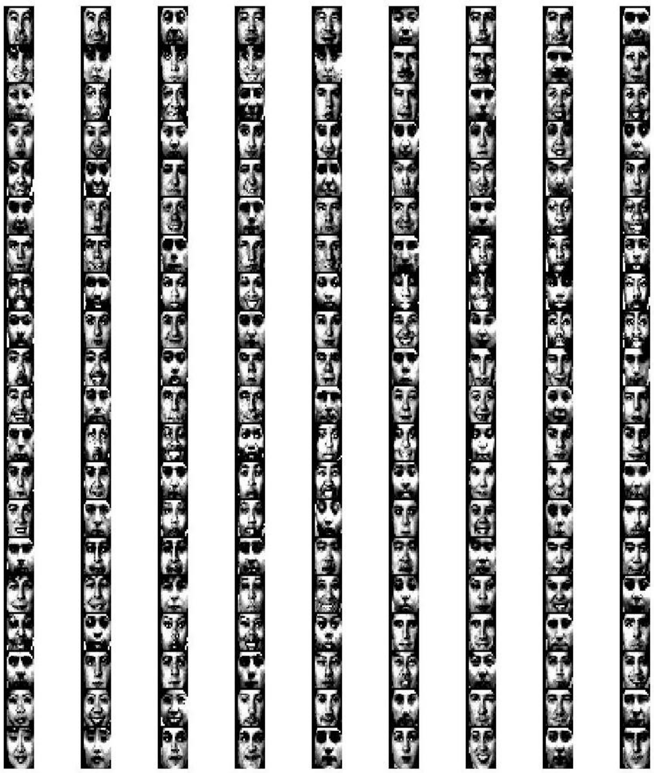

Figure 1 Face Data set with 600 sample faces

Table 2 Variation in accuracy with change in dimensions.

Bayes Classifier after PCA		
S.No.	Retains % of variance	Accuracy
1	50	20.5
2	55	25
3	60	28
4	65	34.5
5	70	40.5
6	75	46.5
7	80	49.5
8	85	52.5
9	90	59
10	95	61.5
11	98	62.5
12	99	63

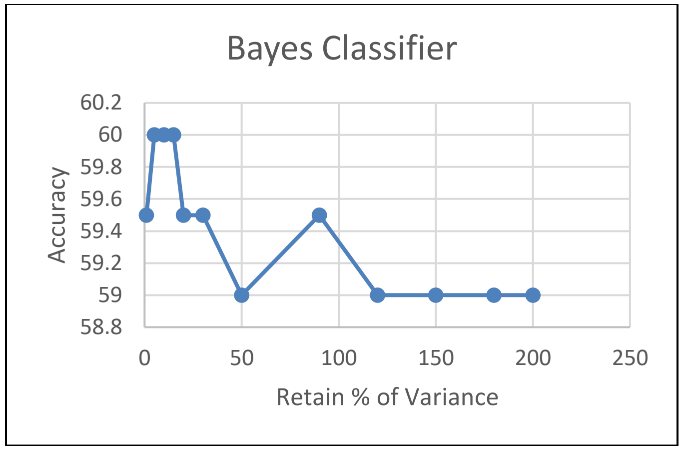

Figure 2 Trend for accuracy as increment of feature vector in PCA

3) The Third result of Bayes Classifier is taken by extracting the same number of features after PCA reduction as it is extracting from LDA i.e one less than total number of classes. It can be seen from the table 3 that the principal component analysis provides better results than linear discriminant analysis.

Table 3 Variation of accuracy when feature extraction is same using PCA and LDA.

Bayes Classifier		
	Principle Component Analysis (features are 199)	Linear Discriminant Analysis (features are 199)
Accuracy	63.5	53.5

4) The first result of K nearest neighbor Classifier is taken by giving two faces for each class as a training sample and third face as a testing sample. It is shown in table 1 that the accuracy of the KNN is less when dimension is reduced using linear discriminant analysis as compare to principle component analysis.

Table 4 KNN accuracy with different number of features of training

K Nearest Neighbor (k = 1)			
	PCA (0.95% variance)	LDA	Without dimension reduction
Accuracy	58	56.5	59.5

5) The second result of KNN is taken by changing the retained percentage of variance while implementing principal component analysis

providing the same samples. It can be seen from the Table 5 that the accuracy of the classifier increase as the dimensions increases. The trend can be seen in figure 3.

Table 5 Variation in accuracy with change in dimension

KNN Classifier after PCA (k = 1)		
S.No.	Retains % of variance	Accuracy
1	50	22
2	55	24
3	60	27.5
4	65	33
5	70	38
6	75	46
7	80	48.5
8	85	53
9	90	58
10	95	58
11	98	58
12	99	58.5

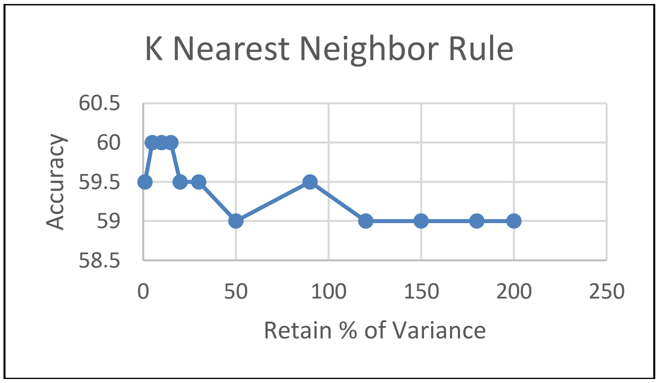

Figure 3 Trend for accuracy as increment of feature vector in PCA

6) The Third result of KNN Classifier is taken by extracting the same number of features after PCA reduction as it is extracting from LDA i.e one less than total number of classes. It can be seen from the table 3 that the principal component analysis provides better results than linear discriminant analysis.

Table 6 Variation of accuracy when feature extraction is same using PCA and LDA.

Bayes Classifier		
	PCA (features are 199)	LDA (features are 199)
Accuracy	58.5	56.5

7) The fourth result of the KNN is taken by changing the k values. It can be seen from the figure 4 that the accuracy increases as k increase up to one point and then become constant. This trend is because of only 2 sample for each class. Because of less number of classes the tie situation occurs frequently thus result in low accuracy.

Table 7 Accuracy variation with the value of k

KNN Classifier without dimension reduction		
S.No.	k	Accuracy
1	1	59.5
2	5	60
3	10	60
4	15	60
5	20	59.5
6	30	59.5
7	50	59
8	90	59.5
9	120	59
10	150	59
11	180	59
12	200	59

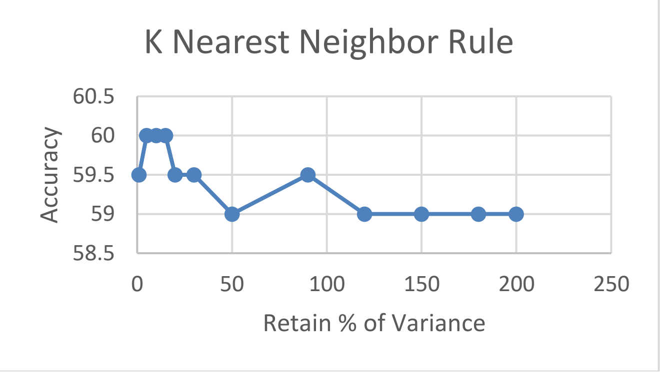

Figure 4 Trend for accuracy while changing k

V. CONCLUSION

It can be seen from the results that the accuracy of prediction depends upon various factors such as value of k, number of features, number of samples, feature extraction methods and the decision rule. There relations can be seen through these results. In my opinion Bayes decision rule is better from KNN rule to classify faces which can be derived from above results.

ACKNOWLEDGMENT

This project required in-depth research, a huge amount of work and dedication. Implementation would not have been possible if I did not have the support of many individuals.

I would like to extend mine sincere gratitude to all of them. First of all, I am thankful to professor Ramalingam Chellappa and professor Behtash Babadi for the provision of expertise and technical support in the completion of this project. Without their superior knowledge and experience, it would have been impossible for me to complete the project in the given timeline.

I would also like to express mine sincere thanks to Huai Jen Liang who devoted his time and knowledge in the implementation of this project and constant encouragement which helped me in completion of this project.