Jae Yoon; Mahshad Farnoush; Jagadeesh Meesala

Facial Recognition with PCA and Classifiers as Bayes and KNN

Priniciples Of Machine Learning

2021

Abstract

We apply facial recognition techniques on the DATA and POSE data sets. "DATA.mat" consists of cropped 24x21 pixel images of 200 subjects, 3 images each. We first try dimensional reduction techniques on the data-set using principal component analysis (PCA). Have tried different dimensional reduction values and derived optimal value with high accuracy. Then applied classifiers k-NN and Bayes to classify neutral vs facial expression and similarly applied PCA followed by k-NN and Bayes classifiers on "POSE.mat" data set and computed the accuracy of classification.

1 Introduction

Implementation of Facial Recognition using k-NN and Bayes classifiers with principal component analysis(PCA) as optimization technique.

2 Results

2.1 Experiments with data.mat

PCA followed by Bayes

The dimensionality reduction offered by PCA is a very useful step for visualizing and processing high-dimensional datasets, while still retaining as much of the variance in the dataset as possible. Because PCA is translation variant, the faces must be frontal and well aligned on facial features such as the eyes, nose and mouth. PCA is useful for dimensionality reduction if the size of the training set is too small for the number of dimensions of the data. But if you are using all of the principal components, PCA won't improve the results of your linear classifier -if your classes weren't linearly separable in the original data space, then rotating your coordinates via PCA won't change that. The problem with image representation is its high dimensionality. Two-dimensional p x q gray scale images span am = pq-dimensional vector space, so an image with 24x 21 pixels lies in a 504-dimensional image space. The PCA method finds the directions with the greatest variance in the data, called principal components. By varying the number of principal components used, the dimensionality can be brought down as required.The following table gives the classification accuracy for PCA followed by the Bayes Rule for different number of principal components. It is seen that the accuracy does not vary by a large amount by increasing the number of principal components which means that most of the energy of the signal is contained within the first 20 principal components.

neutral	***smiling***
41	3
1	35

Table 1: confusion matrix

PCA followed by k-NN

The following table gives the classification results after PCA is applied to the image vector and by using the kNN classifier. Results with k=1,2 and 3 neighbors are reported with varying number of principal components. Here we see an increase in the classification accuracy as the number of principal components used in the data representation increase.

$accuracy$	$correct$	$misclassified$
0.6875	55	25
0.7375	59	21
0.8625	69	11
0.8125	65	15
0.8875	71	10
0.875	70	10

Table 2: k-NN accuracy

2.2 Experiments with pose.mat

(First) 9 images per subject are used for training and (last) 4 for testing. In general, we can see that the classification accuracy for every method are more than the previous data set as the number of training/testing samples are more.

PCA followed by Bayes

write up goes here.

PCA followed by k-NN

$accuracy$	$correct$	$misclassified$
0.7475	305	103
0.4975	203	205
0.5073	207	201
0.4975	203	205

Table 3: k-NN accuracy

3 Code Files

The following notebook files are used for pre-processing and classification application:

- Jagadeesh Meesala
  - github link:

- Jae Yoon
  - codelabs link:
- Mahshad Farnoush
  - codelabs link:

4 Conclusion

Optional. Results, consequences, future work.

5 References

PCA application: towardsdatascience PCA implmentation