IJCSC 0973-7391 Volume 7 • Number 2 March 2016 - Sept 2016 pp. 136-139 available online at www.csjournalss.com

Comparison of SVM based Face Recognition using Global Approach Technique

Anitha Ashishdeep, Priyanka Sharma
Computer Science and Engineering Department, Institute of Technology
Nirma University, India
{14mcecs1, priyanka.sharma}@nirmauni.ac.in

Abstract: Face recognition is an important application in thedomain of image processing based applications. Various Machine learning algorithms has been proposed for face recognition. Here we applied Support Vector Machine (SVM) algorithm and compared the efficiency of various SVM kernels(Linear, Polynomial, RBF) for face recognition. We have used Global approach technique for face recognition. Global approach technique converts entire face to a single feature vector containing grey values of the entire face. Here we have used Principal Component analysis (PCA) method for feature reduction.When PCA is applied on images of faces it returns Eigen faces. Eigen faces are set of Eigen vectors. The simulation results gives optimal parameters for each kernel. It has been proved that RBF kernel outperforms other kernels.

Keywords: SVM, PCA, Grid Search, Gaussian(RBF) kernel, Polynomial kernel, Linear kernel.

1. Introduction

Face recognition is important in various areas like surveillance, banking systems, residential security. Face recognition techniquecompares given face to all the other faces in the database and gives a ranked list of matches. Face recognition can be done by two methods a) global method and b) component based method.

a) Global approach technique: In this method, the input to the classifier is a single feature vector. This vector contains grey values of the entire face image. It is good in classifying frontal static view of the face.This approach does not perform well for pose changes or any emotions.
b) Component based approach technique: In Component based approach, entire face is divided intocomponents like left eyes, right eyes, area between the eyes, right lip, left lips, space between the bridge of the nose and centre of the mouth, left cheek and right cheek. These components are used as an input for classification. This method is robust against pose changes and emotions. [1]

Outline of this paper is as follows: Section 2 gives description of SVM and its various kernels. Section 3 describes PCA which is used for feature reduction. Section4 describes parameter selection for SVM kernel. Section 5 contains experimental results which shows performance of various SVM kernels with optimal parameters.

2. Support vector machine

Support vector machine is an efficientsupervisedmachine learning algorithm. SVM produces best hyper plane forclassifying any new set of examples.

Here classification is done by finding a hyper plane which has maximum distance to the closest data points in the training set. Data points which are closest to the decision hyper plane are called support vectors.

SVM is widely used because itsstrength lies in classifying anynon-linear decision boundary.If data points are not linearly separable,it uses a kernel function which maps the input space into a high dimensional feature space. The data in the feature space is then separated using optimal hyper plane.

In SVM, a decision function*f* is calculated to find a hyperplane for the given training data points (xi,yi),where i=1 to Nand yi∈{-1,1} is the class label.The general form of SVM decision function is:

$$f(x)= \sum^{N}_{i=1} \alpha_i y_i K(xi, xj)$$

WhereK(xi, xj) is the kernel function and α_i is the coefficient weights and xi and xj are feature vectors.

SVM Kernels:

There are various kernels functions like linear, polynomial and Gaussian kernel.

Linear Kernel: Linear kernel is the simplest kernel function which is best suited for linearly separable data points. Herekernel is defined as:

Linear Kernel: $K(xi, xj)= xi \cdot xj+b$ *[2]*

Polynomial Kernel: Polynomial kernel are well suited for the training data set where all dataare normalized. Kernel function is defined as:

Polynomial Kernel: $K(xi,xj)=(\gamma.xi.xj+1)^d$ *where dis the degree of the polynomial.*

Gaussian Kernel (RBF kernel): The main strength of this method is its applicability for any dimensional data. Here kernel function is defined as:

RBF kernel: $K(xi,xj)=exp(-\gamma.||xi-xj||^2)$ *where* $\gamma=1/2\sigma^2$. *[2]*

3. Principal Component Analysis

Feature reduction is a very important step in any application. Reducing the feature helps in reducing computational cost. It is also important in reducing the complexity of the system which in return avoids over fitting. Here we have used Principal Component Analysis (PCA) for feature reduction. PCA represents variance in training data with as few dimensions as possible,here instead of dropping certain features for reducing the features, PCA compresses features together and select the most important feature combination. When PCAis applied on images of faces it returns Eigen faces. Eigen faces are set of Eigen vectors. They represent variance in collection of images and use this information to compare the images of distinct faces.

The function we have used here is RandomizedPCA. We have selected top 150 Eigen faces, and we have also used whitening feature of Randomized PCA. The goal of whitening is to make the input less redundant.

4. Selecting Parameters for various Kernels

Parameterselection is a very important factor for improving efficiency of the model. Here we have used the grid search approach. Grid search is an efficient way to find the best values of parameters(C,γ). In grid search, best value for the pair(C,γ) is chosen by performing cross fold validation.

The value of C defines cost ofclassification. It controls the influence of individual support vectors.A small value of C gives high bias and low variance.A large value of C gives low bias and high variance as shown in Figure 1.Small C makes the cost of misclassification low.

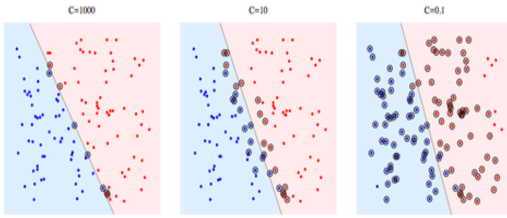

Figure 1. Large value of parameter C gives low bias and high variance (Source: https://www.quora.com/What-are-C-and-gamma-with-regards-to-a-support-vector-machine)

Value of γis a very important parameter for RBF kernel.

For smaller value of γ, there is high variance and thus support vector can influence in deciding the class of other input values.For larger value of γ, the radius of the area of influence only includes the support vector itself. There is low variance and high bias and thus support vectors don't influence in deciding the class

Degree of the polynomial is an important factor for polynomial kernel.

5. Experimental results

Here we have used pre-processed images from Labeled Faces in the Wild(LFW). This data set contains approximately 13,000 images of faces.

Here we have considered folks that have a minimum of 40 pictures in the data set each having a 0.4 aspect ratio. Table 1 shows the optimal parameters obtained by Grid search of each kernel.

Table 1. Optimal Parameters (C and γ) for SVM kernels(linear, Polynomial,RBF)

Kernel	C	γ	degree	Accuracy
Linear	0.11	0.0	-NA-	79%
Polynomial	0.5	0.005	1	84%
RBF	10	0.001	-NA-	87%

Here, Figure 2 shows classifications results obtained using Polynomial kernel with C=0.5, γ=0.005 and degree=1.

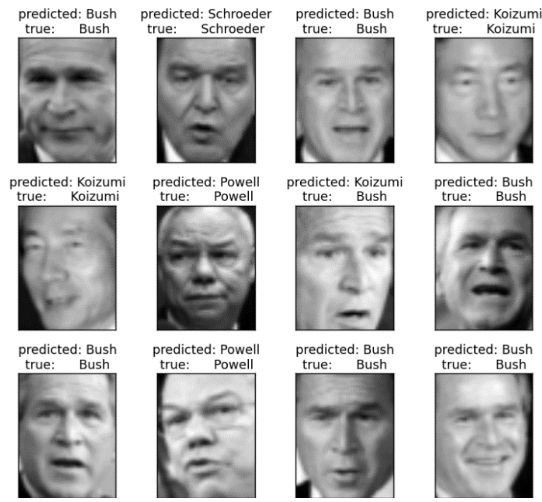

Figure 2. Results using Polynomial kernel

Here, Figure 3 shows classification results obtained using RBF kernel with C=10 and γ=0.001.

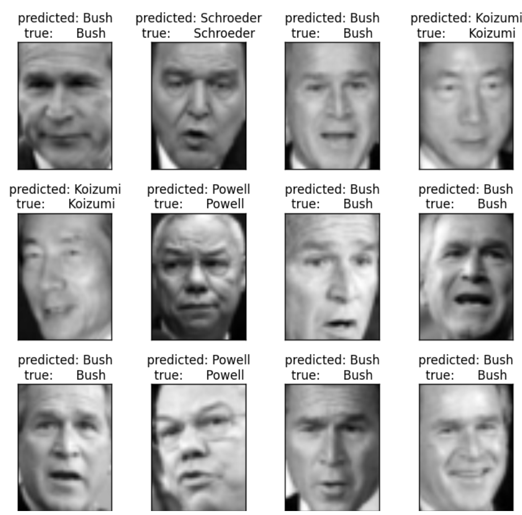

Figure 3. Results using RBF kernel

For our LFW image data set, the efficiency of each SVM kernel changes in accordance with the value of parameters(C,γ, and degree). We obtained the best parameter values for each kernel using Grid search and the study shows that RBF kernel outperforms other two kernels.

IJCSC 0973-7391 Volume 7 • Number 2 March 2016 - Sept 2016 pp. 136-139 available online at www.csjournalss.com

RBF kernel is powerful in classifying nonlinear boundary compared to linear kernel. Also RBF kernel is better than polynomial kernel because it has less parameters so it is simpler to tune.

Conclusion

The global approach with PCA and RBF SVM kernel gives an accuracy of 87% which is a good result with modest amount of training data. However, Global approach with PCA is not good for handling expressions and pose changes. It is also sensitive to lightning conditions.

References

1. Heisele, Bernd, Purdy Ho, and TomasoPoggio. "Face recognition with support vector machines: Global versus component-based approach." Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 2. IEEE, 2001.
2. Artificial Intelligence Research and Development: Recent Advances and Applications edited by LledóMuseros, OriolPujol, NúriaAgell, ISBN 978-1-61499-451-0 (print) | 978-1-61499-452-7 (online)
3. Scikit- learn Machine learning in python. URL:http://scikit-learn.org/stable/

DOI: 10.090592/IJCSC.2016.122