Applied Mechanics and Materials Vols. 738-739 (2015) pp 542-545 *Online: 2015-03-09*
© (2015) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.738-739.542

An Image Classification Algorithm based on SVM

Chunhua Qian [1,a] , Hequn Qiang [1,2,3,b] and Shengrong Gong [2,c]

[1] Department of Computer Science,Suzhou Polytechnic Institute of Agriculture,Suzhou, China

[2] Department of Computer Science & Technology, Soochow University, Suzhou, China

[3] Department of Computer Science, The University of Texas at Dallas, Richardson, Texas, USA

[a]qch_szai@sina.com, [b]borisqiang@gmail.com, [c]893917817@qq.com

Keywords: Image Classification; SVM; LIBSVM

Abstract. Image classification is a image processing method which to distinguish between different categories of objectives according to the different features of images. It is widely used in pattern recognition and computer vision. Support Vector Machine(SVM) is a new machine learning method base on statistical learning theory, it has a rigorous mathematical foundation, builts on the structural risk minimization criterion. We design an image classification algorithm based on SVM in this paper, use Gabor wavelet transformation to extract the image feature, use Principal Component Analysis(PCA) to reduce the dimension of feature matrix. We use orange images and LIBSVM software package in our experiments, select RBF as kernel function. The experimetal results demonstrate that the classification accuracy rate of our algorithm beyond 95%.

Introduction

Image classificaition is a hot topic of pattern recongnition and computer vision nowadays. Machine learning method is widely used in image classification.

There are two popular method in image classificaition, one is Artificial Neural Network(ANN), the other is Support Vector Machine(SVM). ANN is high dependence on the quantity and quality of samples, also easy to fall into local optimum problem with small quantity samples. SVM is based on statistical learning theory with the rigorous mathematical foundation. It trends to be the best method in image classification.

Section 2 contains the analysis of Support Vector Machine. In section 3, we propose an image classificaition algorithm based on SVM. In section 4, we use orange images and LIBSVM in our experiments, then give the experimental results and conclusion.

Analysis of Support Vector Machine(SVM)

SVM comes from optimal classification plane, shown as Fig.1.

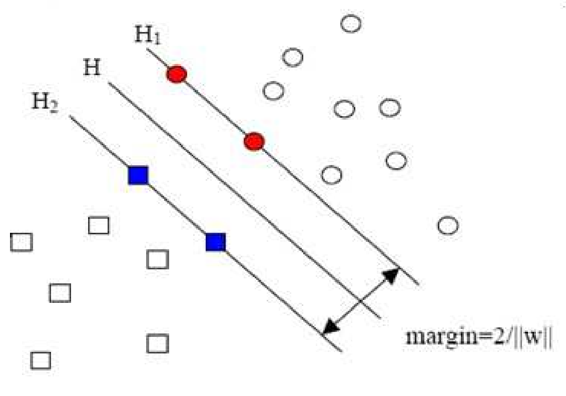

Fig.1 Optimal classification plane sketch-map

In Fig.1 the square point and circular points represent the two types of samples, H is the classification line, H1 and H2 are the two line that parallel to the classification line and cross the samples which be nearest to the classification line. The distance between H1 and H2 is called classification margin.

The optimal classification line can not only separate the two classes correctly, but also make the classification margin maximum. If extension optimal classification line to higher dimensional space, it becomes optimal classification plane.

Given a sample data set $\{(x_i, y_i) \mid i = 1,2,\cdots,k\}$,where x_i is input value and y_i is output value. In order to make sure the margin between two kinds of point maximum, use Lagrange method to transform the original problem to find maximum value of function through Eq.1

$$Q(\alpha) = \sum_{i=1}^{k} \alpha_i - \sum_{i,j=1}^{k} \alpha_i \alpha_j y_i y_j (x_i \cdot y_j) \tag{1}$$

Where, α_i is the corresponding multiplier of every sample. Then transform it into high dimensionality space through kernel funciton $K(x_i \cdot y_j)$. Shown as Eq.2

$$Q(\alpha) = \sum_{i=1}^{k} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{k} \alpha_i \alpha_j y_i y_j K(x_i \cdot y_j) \tag{2}$$

Solve the Eq.2 to get Eq.3

$$f(x) = \operatorname{sgn}\{\sum_{i,j=1}^{k} a_i^* y_j K(x_i \cdot y_j) + b^*\} \tag{3}$$

Where b^* is the threshold of classification.

The three most important advantages of SVM are as follows:

(1) SVM aiming at the limited samples. The goal of SVM is to get the optimal solution with limited information. It can avoid the excessive learning phenomenon.

(2) SVM transform into an optimization problem. It can get the global optimal solution in order to solve the local extremum problem which can't avoid in Artificial Neural Network (ANN).

(3) The introduction of nuclear technology in SVM can transform the problem into high dimensionality feature space through the nonlinear transform. It designs the linear discriminant function in high dimensionality feature space to instead of the nonlinear discriminant function in problem space, and solve the dimension problem ingeniously. So the complexity of SVM is independent to the dimension of samples.

Image classification algorithm based on SVM

The main process of image classification algorithm based on SVM: suppose L sample images, the size of every image is $M \times N$. The Gabor features after Gabor wavelet transformation are $\{W_{u,v,i}\}M \times N, v = 0,1,...,4, u = 0,1,...,7$. Reconstruct the same scale and direction Gabor feature of different images, we can get forty feature matrices $\{O_{u,v}\}L \times (M \times N)$, the dimension of matrix $\{O_{u,v}\}$ is $L \times (M \times N)$, where the i-th row correspond to transform the Gabor feature matrix $W_{u,v,i}$ of the image I_i with direction u and scale v to the row vector. Then use PCA method to reduce the dimension of $\{O_{u,v}\}L \times (M \times N)$, the forty matrices $M_j, j = 0,1,...,39$. Take M_j as training samples send to the SVM classifier SVM_j, calculate the Optimal decision function f_j of every SVM classifier. The process is described as Fig.2.

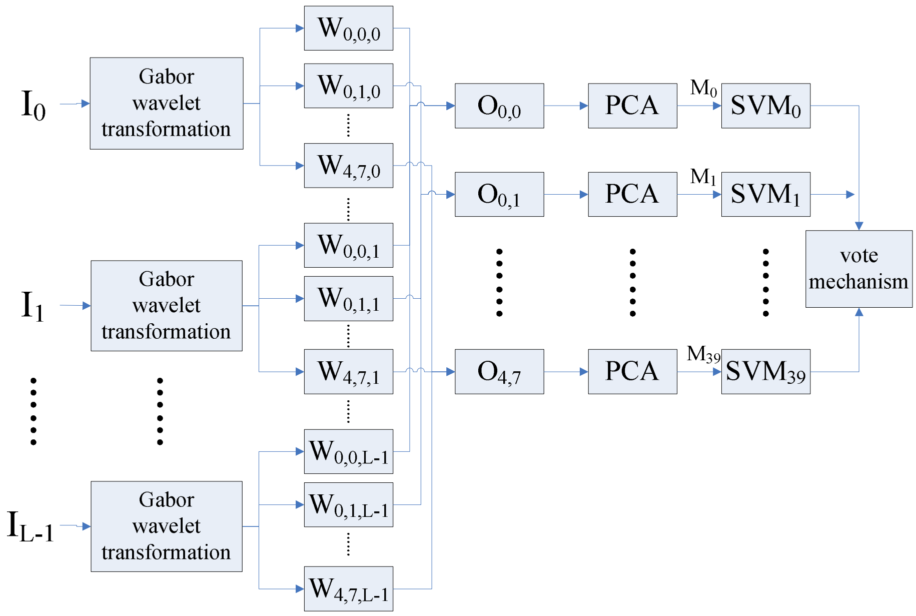

Fig.2 The process of image classification algorithm based on SVM

In this paper, we use orange images as experimental objects. Do the same transformation to every image, then get the forty decision results, each decision result recorded as one vote, according to the vote amout to decide which class the image belongs to.

The algorithm shown as follows:

Step 1: Extract the feature. Calculate the Gabor basis function as Eq.4 and calculate the Gabor feature matrix as Eq.5

$$g_{u,v}(x,y)=\frac{\left\|k_{u,v}\right\|^2}{\sigma^2}\exp\left[\frac{k_{u,v}{}^2\left(x^2+y^2\right)}{2\sigma^2}\right]\bullet\left\{\exp\left[ik_{u,v}\begin{bmatrix}x\\ y\end{bmatrix}\right]-\exp\left[-\frac{\sigma^2}{2}\right]\right\} \tag{4}$$

$$X=\begin{vmatrix}X_{11} & X_{12} & \cdots & X_{1J}\\ X_{21} & X_{22} & \cdots & X_{2J}\\ \vdots & \vdots & \vdots & \vdots\\ X_{N1} & X_{N2} & \cdots & X_{NJ}\end{vmatrix} \tag{5}$$

The feature matrices are described as $\{W_{u,v,i}\}M\times N, v=0,1,\ldots,4, u=0,1,\ldots,7, i=0,1,\ldots,L-1$.

Step 2: Reconstruct the features.Reconstruct the Gabor feature matrix of every image with same scale and direction, the new feature matriecs are $\{O_{u,v}\}L\times(M\times N)$.

Stpe 3: reduce the dimension and denoising. Suppose the cumulative variance contribution rate $\varphi(p)$, reduce the dimension and denoise $\{O_{u,v}\}L\times(M\times N)$ based on PCA method, we can get the new matrices $M_j, j=0,1,\ldots,39$.

Step 4: Parameter selection. Select the type of kernel function and parameters, take M_j as training samples, train the SVM classifier $SVM_j, j=0,1,\ldots,39$, obtain the optimal decision function f_j.

Experimental results and conclusion

In this paper, we use the LIBSVM softerware package in our orange images classification experiments. LIBSVM is a simple, easy to use and fast effective SVM softerware package developed by Lin Chih-Jen[5]. It not only provides the compiled executable file in Windows series system, but also provides the source code to improve, modify and apply on the other operating system conveniently. But the LIBSVM software provides many default parameters and can not be adjusted. So we design the new parameters in our experiments to improve the accuracy rate of image classification.

The experiment process based on LIBSVM shown as follows:

Step 1: prepare the data set according to the requirements format of LIBSVM software package

Step 2: execute a simple scaling operation on the data. The range of original data may too large or too small, SVMscale can zoom the data into appropriate scope in order to accelerate the training speed.

Step 3: select suitable kernel function.

Step 4: use cross – validation method to select the optimal parameters C and g.

Step 5: obtain the SVM model based on the data set training with the parameters C and g

Step 6: Test and classification based on the obtained model in step 5.

The parameter selection and classification result are shown as Table 1.

Table 1 Parameter selection and classification result based on LIBSVM

SVM kernel function	RBF
kernel parameter	Γ=0.1
the number of training samples	60
the number of support vectors	3
the number of test samples	40
fault classification samples	2
the accuracy rate	95%

The experimental results demonstrate that the classification algorithm based on SVM has good performance in orange images classification. Subsequent experiments show that the accuracy rate beyond 95%.

Acknowledgement

This paper was financially supported by the Scientific research fund for young teachers of Suzhou Polytechnic Institute of Agriculture and Jiangsu Overseas Research & Training Program for University Prominent Young & Middle-aged Teachers and Presidents.

References

[1] Yang Hong-Ying, Wang Xiang-Yang, Wang Qin-Yan, Zhang Xian-Jin. LS-SVM based image segmentation using color and texture information[J]. Journal of Visual Communication and Image Representation, 2012, Vol.23(7), pp.1095-1112

[2] Hosseini S. A, Ghassemian H.A. new fast algorithm for multiclass hyperspectral image classification with SVM[J]. International Journal of Remote Sensing, 2011, Vol.32(23), p.8657-8683

[3] Izquierdo-Verdiguier Emma, Laparra Valero, Gomez-Chova Luis, Camps-Valls Gustavo. Encoding Invariances in Remote Sensing Image Classification With SVM[J]. IEEE Geoscience and Remote Sensing Letters, Sept. 2013, Vol.10(5), pp.981-985

[4] Wu Jianxin. Efficient HIK SVM Learning for Image Classification[J]. IEEE Transactions on Image Processing, Oct. 2012, Vol.21(10), pp.4442-4453

[5] Lin Chih-Jen. http://www.csie.ntu.edu.tw/~cjlin/.

[6] Agrawal Saurabh, Verma Nishchal K., Tamrakar Prateek, Sircar Pradip. Content Based Color Image Classification using SVM[C]. 2011 Eighth International Conference on Information Technology: New Generations, April 2011, pp.1090-1094

Sensors, Measurement, Intelligent Materials and Technologies III
10.4028/www.scientific.net/AMM.738-739

An Image Classification Algorithm Based on SVM
10.4028/www.scientific.net/AMM.738-739.542