

A Comparison Study on Back-Propagation Neural Network and Support Vector Machines for the Image Classification Problems

Kwang-Kyu Seo^{1*}

영상분류문제를 위한 역전파 신경망과 Support Vector Machines의 비교 연구

서광규^{1*}

Abstract This paper explores the classification performance of applying to support vector machines (SVMs) for the image classification problems. In this study, we extract the color, texture and shape features of natural images and compare the performance of image classification using each individual feature and integrated features. The experiment results show that classification accuracy on the basis of color feature is better than that based on texture and shape features and the results of the integrating features also provides a better and more robust performance than individual feature. In additions, we show that the proposed classifier of SVM based approach outperforms BPNN to incorporate the image classification problems.

Key Words : Image Classification, Color, Texture and Shape Features, Support Vector Machines (SVMs), Back Propagation Neural Network (BPNN)

요약 본 논문은 영상 분류 문제를 위한 support vector machines (SVMs)의 적용을 통한 분류의 성능을 다루고 있다. 본 연구에서는 영상 분류 문제에서 자연영상을 대상으로 색상, 질감, 형상 특징벡터를 추출하고, 각각의 특징벡터와 이들을 결합한 특징벡터를 사용하여 역전파 신경망과 SVM 기반의 방법을 적용하여 영상 분류의 정확성을 비교한다. 실험결과는 각각의 특징벡터중에는 색상 특징벡터값을 이용한 영상 분류가 그리고 각각의 특징벡터보다는 이들을 결합한 특징벡터를 이용한 영상 분류가 보다 우수함을 보여준다. 그리고 알고리즘간의 비교에서는 정확성과 일반화 성능 측면에서 역전파 신경망보다 SVMs이 우수함을 보였다.

1. Introduction

Image classification and retrieval has gained wide popularity from different communities during the last two decades and it is an interdisciplinary research activity these days. Many applications such as digital libraries, image search engines, and medical decision support systems require effective and efficient image classification and retrieval techniques to access the images based on their contents, commonly known as content-based image retrieval [1, 2]. Features like color, texture, shape, spatial

relationship among entities of an image and also their combination are generally being used for the computation of multidimensional feature vector. The features such as color, texture and shape are known as primitive features. Images have always been an essential and effective medium for presenting visual data. With advances in today's computer technologies, it is not surprising that in many applications, much of the data is images. There have been considerable researches done on pattern recognition area using artificial neural networks [3, 4, 5]. Although numerous theoretical and experimental studies

¹Dept. of Industrial Information and Systems Engineering, Sangmyung University, Professor

*Corresponding author :Kwang-Kyu Seo(kwangkyu@smu.ac.kr)

reported the usefulness of the neural network such as back-propagation algorithm in classification studies, there are several limitations in building the model such as finding an appropriate neural network structure and converging to a local minimum solution.

This study presents the effectiveness of support vector machine (SVM) based approach to detect the underlying data pattern for image classification problems using color, texture and shape features and integrating them. This work also compares with the performance of image classification using each individual feature and integrated feature of natural images by the proposed methods. In addition, bearing in mind that the optimal parameter search plays a crucial role to build an image classification model with high classification accuracy and stability, we employ cross-validation to find out the optimal parameter values of kernel function of SVMs. In order to evaluate the classification accuracy of SVMs, we compare its performance with that of back-propagation neural network (BPNN).

2. Image Features

The performance of the image classification depends on the image features to describe the image content sufficiently and adequate classification mechanism. In this paper, color, texture and shape information are used to represent image features.

2.1 Color

Color is a dominant component to human perception. Image representation scheme is summarized by three major factors. First, the representation must be closely related to human visual perception. Second, the representation must include the global and local information in the image. Third, the representation must be as compact as possible for efficient storage and fast search. As a result of these considerations, we use RGB color model [6] and HSV (Hue, Saturation, Value) color model [7]. For global image representation and fast search, RGB color histograms of image, which are quantized into each R, G, and B coordinate of the divided bins are extracted. For local information, image is divided

into rectangular regions. A set of HSV joint histograms in each rectangular region are extracted, and dominant hue, saturation, value are used as features in that region. This results in compact image representation for efficient storage and fast search. As the size of the region becomes smaller, the local variations of color information are captured by the histograms. The size of the region should be small enough to emphasize the local color and large enough to emphasize a statistically valid histogram.

In this study, the RGB features are the normalized color histograms of an image which are quantized into 18 bins each R, G, and B coordinate. We uniformly quantized HSV space into 18 bins for hue (each bin consisting of a range of 20 degree), 3 bins for saturation and 3 bins for value for lower resolution. In order to represent the local color histogram, we divided image into equal-sized rectangular regions and extract HSV joint histogram that has quantized 162 bins for each region.

2.2 Texture

Texture analysis is an important and useful area of study in computer vision. Most natural images include textures. Scenes containing pictures of wood, grass, etc. can be easily classified based on the texture rather than color or shape. Therefore, it may be useful to extract texture features for image clustering. Like as color feature, we include a texture feature extracted from localized image region.

The co-occurrence matrix [8] is a two-dimensional histogram which estimates the pair-wise statistics of gray level. The $(i, j)^{th}$ element of the co-occurrence matrix represents the estimated probability that gray level i co-occurs with gray level j at a specified displacement d and angle θ . By choosing the values of d and θ , a separate co-occurrence matrix is obtained. From each co-occurrence matrix a number of textural features can be extracted. For image clustering, we used entropy which is mostly used in many applications.

2.3 Shape

Image shape is one of most important and primitive features for the representation and indexing of image databases. Applications with gray scale or binary images have to use shape features for retrievals. Although,

humans can effectively use color or texture to differentiate among natural scenes, many artificial images cannot be distinguished on the basis of color or texture alone. In this paper, we use shape features which can represent the global shape structure of the image.

We describe the shape information contained in an image on the basis of its significant edges. A histogram of the edge directions is used to represent the shape attributes. The edge information contained in the database images is generated in the preprocessing stage using the Canny edge detection [9] (with $\sigma=1$, Gaussian masks of size=9, low threshold=1, high threshold =255). The corresponding edge directions are quantized into 72 bins of 5 degree each. Scale invariance is achieved by normalizing these histograms with respect to the number of edge points in the image. Edge histogram has some advantages for image retrieval applications [9]. First, use of edge directions captures the general shape information. Second, histogram of the edge directions is invariant to translations in the image. Third, normalization scheme can achieve scale invariance.

3. Support Vector Machines

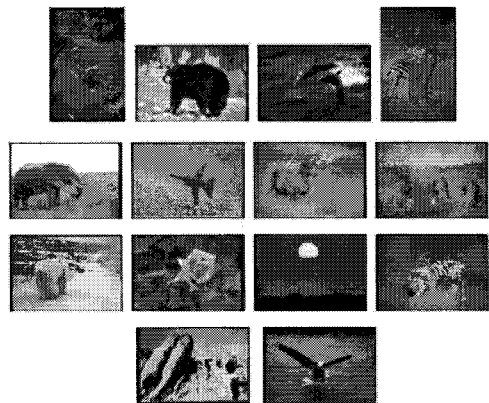
SVMs use a linear model to implement nonlinear class boundaries through some nonlinear mapping input vectors into a high-dimensional feature space. The linear model constructed in the new space can represent a nonlinear decision boundary in the original space. In the new space, an optimal separating hyperplane is constructed. Thus, SVMs are known as the algorithm that finds a special kind of linear model, the maximum margin hyperplane. The maximum margin hyperplane gives the maximum separation between decision classes.

The training examples that are closest to the maximum margin hyperplane are called support vectors. All other training examples are irrelevant for defining the binary class boundaries [10]. SVMs are simple enough to be analyzed mathematically since it can be shown to correspond to a linear method in a high dimensional feature space non-linearly related to input space. In this sense, SVMs may serve as a promising alternative combining the strengths of conventional statistical methods that are more theory-driven and easy to analyze,

and more data-driven, distribution-free and robust machine learning methods. The more detailed description of SVMs is omitted in this paper.

4. Research Data and Experiments

To show the classification effectiveness of the proposed methods, we checked the classification accuracy. Classification results using color, texture and shape features and integrated individual features such as color+texture, color+shape, texture+shape and color+texture+ shape features of real world images will be shown. All experiments were performed on a Pentium IV with 1Gbytes of main memory and 100 Gbytes of storage. We experimented on 2,800 images where most of them have dimensions of 192x128 pixels. The 2,800 images can be divided into 14 categories with 200 images each such as such as flower, bear, dolphin, zebra, elephant, airplane, horse, lion, polar bear, rose, sunset, tiger, valley and eagle as shown in Figure 1.



[Figure 1] Some images of the selected 14 categories used in the experiments

4.1 Configuration of BPNN and SVMs

As mentioned before, BPNN is used for a comparative algorithm to evaluate the performance of the proposed SVMs. BPNN selected for a comparative algorithm in this work consists of one hidden layer and an output layer. The network has a number of inputs determined by the number of color, texture and shape features of images used for classification. Each input is scaled by a constant

factor that is chosen to be the largest estimate of that color, texture and shape features of images in the training data set. The hidden layer employs a logistic activation function and the size of the hidden layer is determined iteratively. The output layer is set at fourteen neurons. The target output for the network is chosen to be one for the correct class and zeros for all the other classes indicating.

SVMs have a lot of parameters to be set such as selection of kernel and the representation of the data. Moreover, it turns out to be sensitive to specific choices of representation and kernel. In order to produce robust results, it is very important to set these parameters. In order to implement SVMs we must decide which kernels to select and then the appropriate kernel parameters are chosen. In this paper, we adopt RBF kernel function. For the kernel parameters of RBF kernel, we used cross-validation to find the best values of parameters for SVMs.

4.2 Training BPNN and SVMs

For training BPNN and SVMs, the data set is divided into the following subsets; a training set of 80% (2,240 images) and a validation set of 20% (560 images) of the total data set (2,800 images).

First, we present the training method for BPNN. BPNN is trained using a back-propagation algorithm with adaptive learning and momentum. There are three aspects related to practical training of BPNN: the size of data, the number of hidden layers, and the stopping criteria. The best-hidden layers number can be obtained by trial-and-error approach. Therefore, BPNN is trained with varying sizes of hidden layer, and the best one is selected as a final configuration. When it comes to stopping criteria, we watch both training and validation errors. Typically, training error is going down and validation error first also goes down but then may begin to rise. This is the moment to stop training. In additions, we adjust learning rate η , and momentum term α . Finally, the BPNN structure for validation set is determined as one hidden layer with 32 neurons and η and α are 0.6 and 0.25 respectively,

Second, we explain the training method for SVMs. In order to train the SVMs, we conduct linearly scaling input

data to the range [0, 1]. As mentioned above, the appropriate parameter values of RBF kernel are found using cross-validation. All the pairs of (ν, γ) for RBF kernel are tried and the one with the best cross-validation accuracy is selected. We realize that trying exponentially growing sequences of ν and γ is a practical method to identify optimal parameters. Pairs of (ν, γ) are tried and the one with the best cross-validation accuracy is picked. We found that the optimal (ν, γ) was $(2^{-1}, 2^{-5})$ with the highest cross-validation.

4.3 Experimental Results

After finding the optimal parameters of the proposed algorithms, the whole training data was trained to generate the final classifier. As mentioned before, we evaluated the classification performance of natural images with the individual and integrated features for SVMs with RBF kernel and compared it with that of BPNN as shown in Table 1. Table 1 shows both training and validation success rates that were achieved under individual and coupled features and the performance results of the proposed approaches.

As shown in Table 1, classification accuracy on the basis of color feature is better than that based on texture and shape features. Integrating the results of the color+texture+shape feature based classification provides a better and more robust performance than the other coupled features. The color+texture features based classification is closed to the performance of all integrated features.

In additions, SVMs with RBF kernel have consistently given better performance than that of BPNN. SVMs provide 98.35% success on the training set and 96.96% on the validation set with color+texture+shape features. Examining the results, we can see that BPNN manages to achieve 92.37% success on the training set and 89.82% on the validation set. Compared to the performance of SVMs with the other kernels such as linear, polynomial, and sigmoid as not shown in this paper, the BPNN gave relatively good results. This means that the neural network method is robust in both the performance and less sensitivity to parameters.

Eventually, the image classification approach which is coupled proposed features is used for obtaining better

classification accuracy and SVMs give good results in case of providing and guaranteeing the selection of the appropriate kernel parameters.

[Table 1] Performance comparison between SVMs and BPNN with different features

	BPNN		SVM	
	Train. (%)	Val. (%)	Train. (%)	Val. (%)
Color(C)	78.35	76.96	82.41	80.18
Texture(T)	66.79	65.54	74.91	72.86
Shape(S)	63.17	62.32	69.64	70.36
C+T	91.07	89.11	97.14	96.07
C+S	88.04	85.18	95.04	92.14
T+S	81.96	79.82	88.13	86.07
C+T+S	92.37	89.82	98.35	96.96

5. Conclusion

In this paper, we presented that the classifier of the SVMs outperforms BPNN in the image classification problems. Color, texture and shape information of natural images are used as input elements to the proposed algorithms. Our experimentation results demonstrated that SVMs have the highest level of accuracies and better generalization performance than BPNN. The proposed method, SVMs served to exemplify that kernel-based learning algorithms are indeed highly competitive on image classification problems. In additions, we tested the accuracy performance with each individual feature and integrating features of images using BPNN and SVMs with RBF kernel. As a result, the classification accuracy on the basis of color is better than that texture and shape. The results of the integrating color, texture and shape features gave better and more robust performance than the other features. The integrating color and texture features are also provided the satisfactory classification performance. Eventually, the classification results using the integrating features provided the better classification accuracy.

Reference

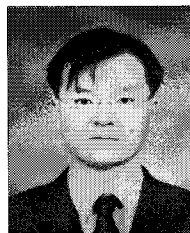
[1] Eakins, P. J., "Towards Intelligent image retrieval.", Pattern Recognition, vol. 35, pp. 3-14, 2002.
 [2] Jain, A. K., Murty, M. N. and Flynn, P. J., "Data

clustering: a review", ACM Computing Surveys, vol. 31(3), pp.264-323, 1999.

[3] Fournier, J., Cord, M., Philipp-Foliguet, S., "Back-propagation Algorithm for Relevance Feedback in Image retrieval", IEEE International Conference in Image Processing, Vol. 1, pp. 686-689, 2001.
 [4] Koskela, M., Laaksonen, J., and Oja, E., "Use of Image Subset Features in Image Retrieval with Self-Organizing Maps", LNCS, Vol. 3115, pp.508-516, 2004.
 [5] Park, S.-S., Seo, K.-K., Jang, D.-S., "Expert system based on artificial neural networks for content-based image retrieval", Expert Systems with Applications, Vol.29(3), pp. 589-597, 2005.
 [6] Berman, A. L., Shapiro, G., "Efficient image retrieval with multiple distance measures", Proceedings of SPIE, Storage and Retrieval for Image and Video Databases V, Vol. 3022, pp. 12-21, 1997
 [7] Jain, A. K., Vailaya, A., "Image retrieval using color and shape", Pattern Recognition, Vol. 29, pp. 1233-1244, 1996.
 [8] Haralick, R. M., Shanmugam, K., Dinstein, I., "Texture features for image classification", IEEE Trans on Sys, Man, and Cyb, SMC-3(6), pp. 610-621, 1973.
 [9] Jain, A.K, Vailaya, A., "Shape-based retrieval: a case study with trademark image databases", Pattern Recognition, Vol. 31, pp. 1369-1390, 1998.
 [10] Vapnik, V., Statistical learning theory. New York: Springer, 1998.

Kwang-Kyu Seo

[Regular Member]



- Aug. 2002 : Ph.D, Dept. of Industrial Engineering, Korea University
- Sep. 1997 ~ Feb. 2003 : Senior Research Scientist, Korea Institute of Science and Technology
- Mar. 2003 ~ Present : Assistant Professor, Dept of Industrial Information & Systems Engineering, Sangmyung University

<Major>

Production & Operation Management, Data Mining & CRM, Information System, Artificial Intelligence