An Efficient Technique for Color Image Classification Based On Lower Feature Content

Jitendra Kumar, EC Department, TIEIT, Bhopal, India, Jitendra12kumar86@gmail.com
Neelesh Gupta, EC Department, TIEIT, Bhopal, India, neeleshgupta9826@gmail.com
Neetu Sharma, EC Department, TIEIT, Bhopal, India, neetusharma85@gmail.com
Paresh Rawat, EC Department, TCST, Bhopa, India, parrawat@gmail.com

Abstract
Image classification is backbone for image data available around us. It is necessary to use a technique for classified the data in a particular class. In multiclass classification used different Classifier technique for the classification of data, such as binary classifier and support vector machine .In this paper we used an efficient classification technique as radial basis function. A Radial Basis Function (RBF) neural network has an input layer, a hidden layer and an output layer. The neurons in the hidden layer contain Gaussian transfer functions whose outputs are inversely proportional to the distance from the center of the neuron. For classification of data support vector machine (SVM) is used as binary classifier. The some approaches commonly used are the One-Against-One (1A1), One-Against-All (1AA),and SVM as Ant Colony Optimization(ACO). SVM-ACO decrease unclassified data and also decrease noise with outer line of data. Here SVM-RBF reduce noise with outer line data and complexity more than SVM-ACO.
Keywords-- Image classification; feature sampling; support vector machine; ACO; RBF.

I .INTRODUCTION

A large part of this research work is devoted to finding suitable representations for the images due to large collection of image data available around us. From classification trees to neural networks, there are many possible choices for what classifier to use. Classification has delivered important meanings in our life. In general, the definition of classification simply means the grouping together of alike things according to common qualities or characteristics. Classification has essential part to play especially in assisting in the search process. By classifying things into different segments it enables us to retrieve things or information that we needed to look for, without the risk of too much time consuming in retrieving that particular things or information. The Support Vector Machine (SVM) approach is considered a good candidate because of its high generalization performance without the need to add a Prior knowledge, even when the dimension of the input space is very high.

II. Support Vector Machines

A Support Vector Machine (SVM) performs classification by constructing an *N*-dimensional hyper-plane that optimally separates the data into two categories. SVM models are closely related to neural networks. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron neural network In the parlance of SVM literature, a predictor variable is called an *attribute*, and a transformed attribute that is used to define the hyper-plane is called a *feature*. The task of choosing the most suitable representation is known as *feature selection*. A set of features that describes one case (i.e., a row of predictor values) is called a *vector*. So the goal of SVM modeling is to find the optimal hyper-plane that separates clusters of vector in such a way that cases with one category of the target variable are on one side of the plane and cases with the other category are on the other size of the plane. The vectors near the hyper-plane are the *support vectors*. The figure below presents an overview of the SVM process.

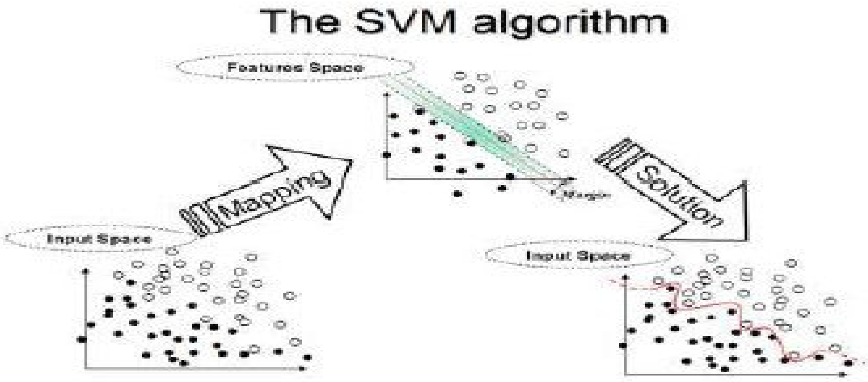

Fig1 Basic classification by SVM.

Chapelle et al. [1] use SVM to realize histogram-based image classification. They select several classes (include 386 airplanes, 501 birds, 200 boats, 625 buildings, 300 fish, 358 people, 300 vehicle) of the Corel database as the image database, distinguish different kinds of object through the SVM classifier. Hong bao Cao, Hong-Wen Deng, and Yu-Ping Wang et al. [2] Segmentation of M-FISH Images for Improved Classification of Chromosomes With an Adaptive Fuzzy C-means Clustering Algorithm An adaptive fuzzy c-means algorithm was developed and applied to the segmentation and classification of multicolour fluorescence in situ hybridization (M-FISH) images, which can be used to detect chromosomal abnormalities for cancer and genetic disease diagnosis. The algorithm improves the classical fuzzy c-means algorithm (FCM) by the use of a gain field, which models and corrects intensity in homogeneities caused by a microscope imaging system, flairs of targets (chromosomes), and uneven hybridization of DNA. Other than directly simulating the in homogeneously distributed intensities over the image, the gain field regulates centers of each intensity cluster. Sai Yang and Chunxia Zhao etld[3] A Fusing Algorithm of Bag-Of-Features Model and Fisher Linear Discriminative Analysis in Image Classification A fusing image classification algorithm is presented, which uses Bag-Of-Features model (BOF) as images' initial semantic features, and subsequently employs Fisher linear discriminative analysis (FLDA) algorithm to get its distribution in a linear optimal subspace as images' final features. Lastly images are classified by K nearest neighbour algorithm. The experimental results indicate that the image classification algorithm combining BOW and FLDA has more powerful classification performances. In order to further improve the middle-level semantic describing performance, we propose compressing the BOF distribution of images distributing loosely in high-dimensional space to a low-dimensional space by using FLDA, the images are classified in this space by KNN algorithm. SooBeom Park, Jae Won Lee, Sang Kyoon Kim etld[4] Content-based image classification using a neural network A method of content-based image classification using a neural network. The images for classification are object images that can be divided into foreground and background. To deal with the object images efficiently, in the pre-processing step we extract the object region using a region segmentation technique. Features for the classification are shape-based texture features extracted from wavelet-transformed images. The neural network classifier is constructed for the features using the back-propagation learning algorithm. Among the various texture features, the diagonal moment was the most effective.

III. Radial Basis Function

A Radial Basis Function (RBF) neural network has an input layer, a hidden layer and an output layer. The neurons in the hidden layer contain Gaussian transfer functions whose outputs are inversely proportional to the distance from the center of the neuron. An RBF network positions one or more RBF neurons in the space described by the predictor variables (x,y in this example). This space has as many dimensions as there are predictor variables. The Euclidean distance is computed from the point being evaluated (e.g., the triangle in this figure) to the center of each neuron, and a radial basis function (RBF) (also called a kernel function) is applied to the distance to compute the weight (influence) for each neuron. The radial basis function is so named because the radius distance is the argument to the function

$$Weight = \mathrm{RBF}(distance) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

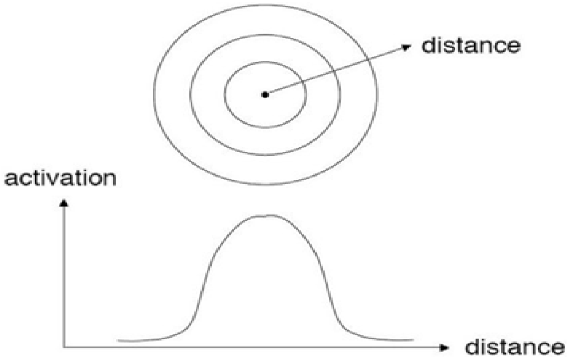

Fig2. RBF distance vector and activation

In Gaussian function many input node are used as input layer and one hidden layer which combine linearly at the output layer.

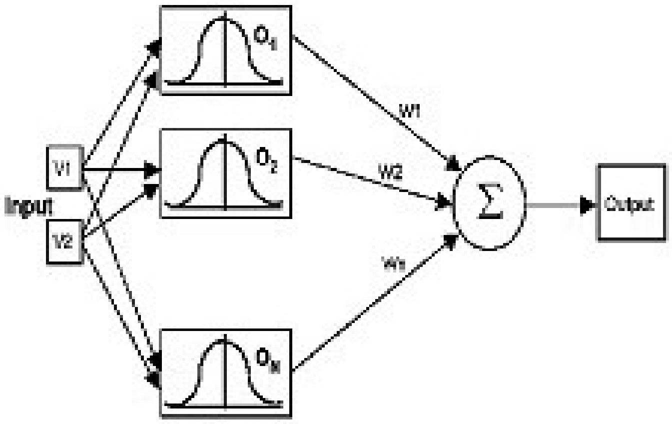

Fig 3 Basic structure of RBF network

In this paper we use a Gaussian function as a kernel function. A Gaussian function is specified by its centre and width. The simplest and most general method to decide the middle layer neurons is to create a neuron for each training pattern. However the method is usually not practical since in most applications there are a large number of training patterns and the dimension of the input space is fairly large. Therefore it is usual and practical to first cluster the training patterns to a reasonable number of groups by using a clustering algorithm such as K-means or SOFM and then to assign a neuron to each cluster. A simple way, though not always effective, is to choose a relatively small number of patterns randomly among the training patterns and create only that many neurons. A clustering algorithm is a kind of an unsupervised learning algorithm and is used when the class of each training pattern is not known. But an RBFN is a supervised learning network. And we know at least the class of each training pattern. So we'd better take advantage of the information of these class memberships when we cluster the training patterns. Various methods have been used to train RBF networks.

$$K_{RBF}(x_1, x_2) = \exp(-p \| x_1 - x_2 \|^2) \quad \ldots\ldots(2)$$

One approach first uses K-means clustering to find cluster centers which are then used as the centers for the RBF functions. We have use one pass clustering algorithm as.

Output: centers of clusters
Variable
C: number of clusters
c*j* : center of the *j*-th cluster
nj : number of patterns in the *j*-th cluster
di j : distance between x*i* and the *j*-th cluster
begin
C =1; c1 x1;*n*1 :=1;
for *i* :=2 to *P* do /* for each pattern */
for *j* :=1 to *C* do /* for each cluster */
compute *di j*;
if *di j* _*R*0 then
/* include x*i* into the *j*-th cluster */

c*j* (c*j**n**j* +x*i*)=(*n**i*+1);
ni :=*ni*+1;
exit from the loop;
end if
end for
if x*i* is not included in any clusters then
/* create a new cluster */
C :=*C*+1;
c*C* x*i*;
nC :=1;
end if
end for
end

It is quite efficient to construct the middle layer of an RBF since we can finish clustering by going through the entire training patterns only once. However, K-means clustering is a computationally intensive procedure, and it often does not generate the optimal number of centers. Another approach is to use a random subset of the training points as the centers.

IV. Experimental Result

In our experiment we have taken 600 famous images of data set. For research we use MATLAB 7.8.0 for different images of data set. Hera we take three different class of input data image as given below.

Now we have performed efficient technique method SVM-RBF for above three input images. Here only one output result will be shown for every three input data. Calculate precession and recall value on behave of above input data.

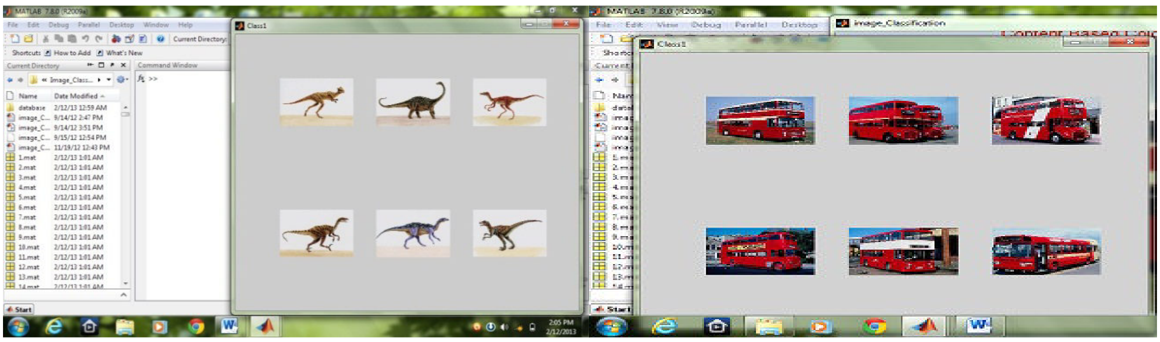

Fig 4 Classification for 1^{st} & 2^{nd} input data set. Precession and recall value will be 92.3% and 84.8% & 94% and 91%.

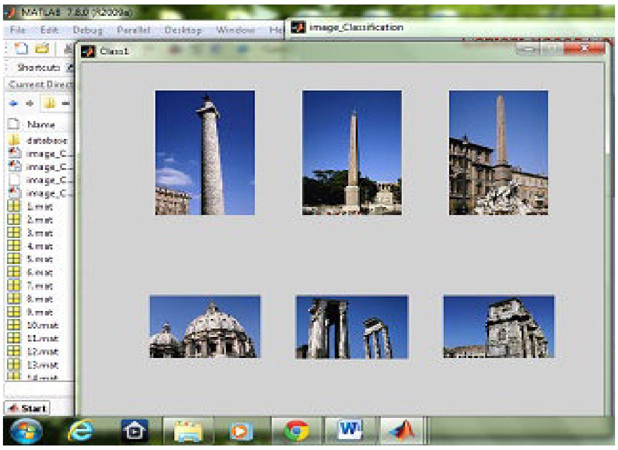

Fig 6 Classification for third input data set. Precession and recall value will be 93.3% and 89.4%.

Table 1 Comparative result

Data set no.	Input image	Classifier name	% Precession	% Recall
1	dinosaur	SVM-RBF	92.30	84.80
		SVM-ACO	91.33	83.60
		SVM-DAG	86.66	83.60
2	bus	SVM-RBF	94	91
		SVM-ACO	93	96.6
		SVM-DAG	90	91
3	tower	SVM-RBF	93.3	89.4
		SVM-ACO	80	79
		SVM-DAG	83.33	79.81

Conclusions

SVM-DAG based support Vector machine perform a better classification in compression of another binary multi-class classification. DAG applied a graph portion technique for the mapping of feature data. The mapping space of feature data mapped correctly automatically improved the voting process of classification. But DAG suffered a little bit problems with mapping of space data into feature selection process.SVM-ACO reduce the problem of feature selection from mapping of data feature. Performance of result evaluation shows that our RBF-SVM is better classifier in compression of SVM-ACO. SVM-RBF reduces the semantic gap and enhances the performance of image classification. However, directly using SVM scheme has two main drawbacks. First, it treats the core point and outlier equally, although this assumption is not appropriate since all outlier share a common concept, while each core point differs in diverse concepts. Second, it does not take into account the unlabeled samples, although they are very helpful in constructing a good classifier. In this dissertation, we have explored unclassified region data on multi-class classification. We have designed SVM-RBF to alleviate the two drawbacks in the traditional SVM.

Refrence

[1] O. Chapelle, P. Haffner, and V. Vapnik. “SVMs for Histogram-Based Image Classification,” IEEE Trans. on Neural Networks, 10(5): pp.1055-1065, Sep. 1999.

[2]Hongbao Cao, Hong-Wen Deng, and Yu-Ping Wang Segmentation of M-FISH Images for Improved Classification of Chromosomes With an Adaptive Fuzzy C-means Clustering Algorithm‖ in IEEE transaction on fuzzy system, vol. 20, NO. 1, February 2012

[3] Sai Yang and Chunxia Zhao “A Fusing Algorithm of Bag-Of-Features Model and Fisher Linear Discriminative Analysis in Image Classification” in IEEE International Conference on Information Science and Technology 2012

[4] SooBeom Park, Jae Won Lee, Sang Kyoon Kim “Content-based image classification uses a neural network” in 2003 Elsevier.

[5] J. Tenenbaum, V. Silva, and J. Langford, “A global geometric framework for nonlinear dimensionality

reduction," Science, vol. 290, no. 22, pp. 2319–2323, Dec. 2000.
[6] D. Tao, X. Tang, and X. Li, "Which components are important for interactive image searching?," IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 1, pp. 3–11, Jan. 2008.
[7] M. Belkin and P. Niyogi, "LaplacianEigenmaps and spectral techniques for embedding and clustering," in Proc. Adv. NIPS, 2001, vol. 14, pp. 585–591.
[8] Saurabh Agrawal, Nishchal K Verma , Prateek Tamrakar and Pradip Sircar "Content Based Color Image Classification using SVM" in Eighth International Conference on Information Technology: New Generations 2011.
[9] Priyanka Dhasal, Shiv Shakti Shrivastava, Hitesh Gupta, Parmalik Kumar" An Optimized Feature Selection for Image Classification Based on SVMACO" *International Journal of Advanced Computer Research, sept.2012.*