International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-3 Number-2 Issue-10 June-2013

Performance Analysis of Image Classification Algorithm Based on Feature Fusing Technique Model

Mukul Yadav[1], Gajendra Singh Chandel[2], Ravindra Gupta[3]

Abstract

Unclassified region deceases the efficiency and performance of PLSA and FLDA. The proper selection of feature sub set reduced the unclassified region of PLSA and FLDA. Now a day's binary classification are widely used in image classification. The mapping of data one space to another space creates diversity of outlier and noise and generate unclassified region for image classification. For the reduction of unclassified region we used radial basis function for sampling of feature and reduce the noise and outlier for feature space of data and increase the performance and efficiency of image classification. Our proposed method optimized the feature selection process and finally sends data to FLDA classifier for classification of data. Here we used fisher classifier. As a classifier FLDA suffering two problems (1) how to choose optimal feature sub set input and (2) how to set best kernel parameters. These problems influence the performance and accuracy of FLDA. Now the pre-sampling of feature reduced the feature selection process of FLDA for image classification.

Keywords

Image classification, feature reduction, FLDA, RBF

1. Introduction

The content based image classification reduces semantic gap between retrieval image and query image. The lower content of query image such as colour, texture and dimension play important role for feature selection process. The flow of feature selection process generates a negative result of query processing [4]. The feature selection, proper parameters adjustment can modify the FLDA classification performance. The parameters that should be optimal include define parameter C and the kernel function parameters such as the gamma (g) for the radial basis function (RBF) kernel[7]. To use a FLDA, one must choose a kernel function, set the kernel parameters and determine a soft margin constant C (penalty parameter). The classification algorithm is an alternative to finding the best C and gamma when using the RBF kernel function. RBF algorithms have the potential to generate both the optimal feature subset and FLDA parameters at the same time. Our dissertation aim is to optimize the parameters and feature subset selection process, without losing the FLDA classification performance [8]. The modified method performs feature selection and parameters setting in an new way. Based on whether or not feature selection is performed independently of the classification algorithm that constructs the classifier, feature subset selection algorithms can be classified into two section for provide feature vector of FLDA[10]. The rest of paper is organized as follows. In Section II discuss related work of image classification. The Section III discuss FLDA and ant colony. Section IV discusses proposed algorithm .section V discuss experimental result followed by a conclusion in Section VI.

Mukul Yadav, Department of Computer Science & Engineering S.S.S.I.S.T, Sehore.
Gajendra Singh Chandel, Department of Computer Science & Engineering S.S.S.I.S.T, Sehore.
Ravindra Gupta, Department of Computer Science & Engineering S.S.S.I.S.T, Sehore.

2. Related Work

In this section describe the related work in the field of image classification based on different technique. The classification part is crucial in concern of feature selection and mapping of data. Some author and research work used some optimization technique and other methods.

[2] Automatic Image Classification Using the Classification Ant-Colony Algorithm To enhance the versatility, robustness, and convergence rate of automatic classification of images, an ant-colony-based classification model is proposed in this paper. According to the characteristics of the image classification, this model adopts and improves the traditional Ant-Colony algorithm. It defines two types of ants that have different search strategies and refreshing mechanisms. The stochastic ants identify new categories, construct the category tables and

determine the clustering centre of each category. [1] 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.

[4] 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 centres of each intensity cluster.

[5] 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 BOWS 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.

[6] Wavelet based Multi Class image classification using Neural Network, A feature extraction and classification of multiclass images by using Haar wavelet transform and back propagation neural network. The wavelet features are extracted from original texture images and corresponding complementary images. The features are made up of different combinations of sub-band images, which offer better discriminating strategy for image classification and enhance the classification rate.

[3] SNMFCA: Supervised NMF-based Image Classification and Annotation A novel supervised nonnegative matrix factorization based framework for both image classification and annotation (SNMFCA). The framework consists of two phrases: training and prediction. In the training phrase, two supervised nonnegative matrix factorizations for image descriptors and annotation terms are combined together to identify the latent image bases, and represent the training images in the bases space. These latent bases can capture the representation of the images in terms of both descriptors and annotation terms. Based on the new representation of training images, classifiers can be learnt and built.

[7] Local Naive Bayes Nearest Neighbor for Image Classification An improvement to the NBNN image classification algorithm that increases classification accuracy and improves its ability to scale to large numbers of object classes. The key observation is that only the classes represented in the local neighbourhood of a descriptor contribute significantly and reliably to their posterior probability estimates.

[8] Research on Image Classification Based on a Combination of Text and Visual Features A text-image co-occurrence data become available on the web, mining on those data is playing an increasingly important role in web applications. Utilizing description information to help image classification.

3. FLDA and RBF Network

Classification plays a big role in image classification. In this section of paper describe FLDA classifier for classification purpose and radial bias network for feature optimisation of feature vector of input of classifier. The optimised feature generates a better result in compression of FLDA classifier. FLDAs are learning systems that use a hypothesis space of linear functions in a hyperspace [14], trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. The aim of Classification via FLDA is to find a computationally efficient way of learning good separating hyper planes in a hyperspace, where 'good' hyper planes mean ones optimizing the generalizing bounds and by 'computationally efficient' we mean algorithms[16,19] able to deal with sample sizes of very high order. The basic problem that a FLDA learns and solves is a two-category classification problem. Follow the method of Bennett's discussion (2000), suppose we have a set of l observations. Every observation can be represented by a pair $\{x_i, y_i\}$ where $x_i \in R^N$ and $y_i \in \{-1, 1\}$. That is, each observation contains an N-dimensional vector x and a class assignment y. Our aim is to find the optimal separating hyperplane; that is, the flat (N-1)-dimensional surface that best separates the data. For time being we assumed that a separating hyperplane exists, and is defined by normal vector w. On the either side of this plane we construct a pair of parallel planes such that:

$w \cdot x_i \geq b + 1$ for $y_i = 1$

$w \cdot x_i \leq b - 1$ for $y_i = -1$

where,b indicates the offset of the plane from the origin. This Often, a non-linear solution plane is required to separate data. To repeat the steps and maximize the separation between two non-linear functions can be computationally expensive [12,8]. Instead, the kernel trick is used: input data are mapped into a higher dimensional feature space via a specified kernel function. These data are linearly separable in the higher dimensional space. A method of accommodating errors and outliers in the input data was developed , and can be implemented simply by allowing an error of up to ξ in each dimension (resulting in a 'fuzzy margin') and adding a cost function C(i) to the optimization equation (Burges). We then want to minimize:

$\frac{1}{2}\|w\|^2 + C \cdot (\Sigma\, \xi_i)$

Subject to the constraint:

$y_i\,(w \cdot x_i - b) + \xi_i \geq 1$

This is substantially harder to solve than the separable case. In the Chang and Lin's LIB FLDA manual, the constraints, minimization conditions and resulting decision functions are defined for each type of classification, along with algorithms to solving the required quadratic programming problems. A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin. If a function 'h' satisfies the property $h(\boldsymbol{x})=h(\|\boldsymbol{x}\|)$, then it is a radial function. Their characteristic feature is that their response decreases (or increases) monotonically with distance from a central point. The centre, the distance scale, and the precise shape of the radial function are parameters of the model, all fixed if it is linear [5]. A typical radial function is the Gaussian which, in the case of a scalar input, is

$$h(x)=\exp((-(x-c)^2)/(r^2)) \quad \ldots (1)$$

Its parameters are its centre c and its radius r. A Gaussian RBF monotonically decreases with distance from the centre. In contrast, a multiquadric RBF which, in the case of scalar input monotonically increases with distance from the centre. Gaussian-like RBFs are local (give a significant response only in a neighborhood near the centre) and are more commonly used than multiquadric-type RBFs which have a global Response. Radial functions are simply a class of functions. In principle, they could be employed in any sort of model (linear or nonlinear) and any sort of network (single-layer or multi-layer). RBF networks have traditionally been associated with radial functions in a single-layer network. In the Figure 4.4, the input layer carries the outputs of FLD function. The distance between these values and centre values are found and summed to form linear combination before the neurons of the hidden layer. These neurons are said to contain the radial basis function with exponential form. The outputs of the RBF activation function is further processed according to specific Requirements.

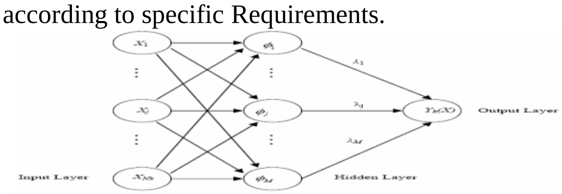

Figure 1: Structure of Radial Basis Function Neural Network.

In order to specify the middle layer of an RBF we have to decide the number of neurons of the layer and their kernel functions which are usually Gaussian functions. In this paper we use a Gaussian function as a kernel function. A Gaussian function is specified by its center 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[8,9]. 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

International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-3 Number-2 Issue-10 June-2013

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.

4. Proposed Model for Classification

In this section describe the proposed model of image classification based on FLDA and feature optimization using radial bias network. Image feature selection process decides the performance of image classifier. We put the optimized feature sub set selection using radial bias network. The output of RBF network proceeds input for FLDA classifier. The basic idea of the proposed technique is to carry out the training process of the hidden layer of RBF neural classifiers by taking into account the class-memberships of the training samples. In particular, clusters are generated by grouping training samples belonging to the same class in order to avoid the creation of mixed clusters. Moreover, the widths of the kernel functions are selected by using a supervised procedure aimed at limiting the widths of kernels located in boundary regions between classes while maintaining, at the same time, a certain overlapping inside the internal regions of each class. In the following, a detailed description of the proposed technique is provided. The complete description of model shown in figure.

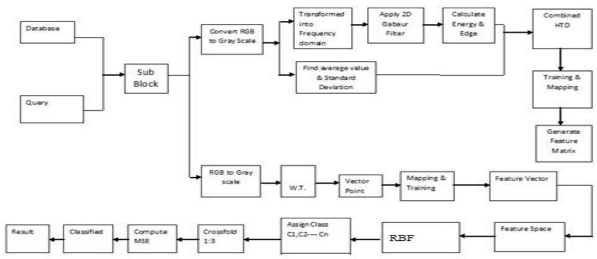

Figure 2: shows proposed model for image classification based on RBF network

5. Implementation Details

In this section we implement three methods for image classification using FLDA with DAG and FLDA with RBF network. We evaluated performance of our algorithm using a general-purpose image database containing 500 JPEG images with size of 256*256 or 256*384 pixels from COREL photo gallery. These images are divided into4 categories, and there are 100 images in each semantic category. We test the performance of; the retrieval performance is measured by precision and recall, they are defined below.

$$\text{Precision}=\frac{\text{number of relevant images retrieved}}{\text{number of images retrieved}} \quad \ldots\ldots\ldots (2)$$

$$\text{Recall}=\frac{\text{number of relevant images retrieved}}{\text{number of relevant images in database}} \quad \ldots\ldots\ldots (3)$$

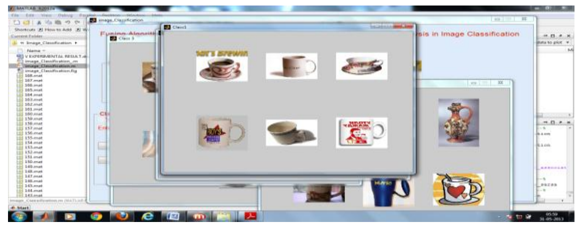

Figure 3: shows that the classified result of FLDA method for 100 cup image of 1000 databases and accurate classified image is 30

Data set	Method	Precision (%)	Recall (%)
Data set 1	FLDA DAG	86.66	80.21
	FLDA RBF	91.33	83.60
Data set 2	FLDA DAG	90	91
	FLDA RBF	93	96.06
Data set 3	FLDA DAG	83.33	79.81
	FLDA RBF	80	79
Data set 4	FLDA DAG	83.33	76.83
	FLDA RBF	93.33	83.33
Data set 5	FLDA DAG	86.66	78.66
	FLDA RBF	90	78.6

Figure4: shows the comparative result of FLDA-DAG and FLDA-RBF

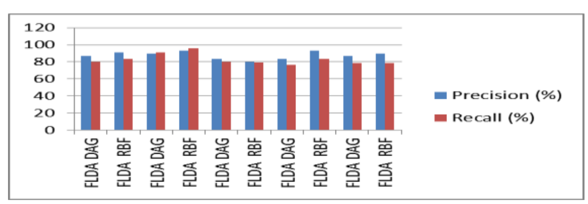

Figure 5: shows the comparative result analysis of FLDA with DAG and RBF network. The classification performance is better than FLDA-DAG

6. Conclusion And Future Work

RBF- FLDA reduces the semantic gap and enhances the performance of image classification. However, directly using FLDA 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 classification ratio are good classifier. In this dissertation, we have explored unclassified region data on multi-class classification. We have designed RBF- FLDA to alleviate the two drawbacks in the traditional FLDA. Here RBF play a role of feature sampling technique. 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. Performance of result evaluation shows that our RBF- FLDA is better classifier in compression of DVM-Dag.

References

[1] Soo Beom Park, Jae Won Lee, Sang Kyoon Kim "Content-based image classification uses a neural network" in 2003 Elsevier.
[2] Wei-jiu Zhang, Li Mao, Wen-bo Xu " Automatic Image Classification Using the Classification Ant-Colony Algorithm" in International Conference on Environmental Science and Information Application Technology 2009.
[3] Liping Jing, Chao Zhang and Michael K. Ng "SNMFCA: Supervised NMF-based Image Classification and Annotation" in IEEE 2011.
[4] 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 transactions on fuzzy systems, vol. 20, no. 1, February 2012.
[5] 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.
[6] Ajay Kumar Singh, Shamik Tiwari and V.P. Shukla "Wavelet based Multi Class image classification using Neural Network" in International Journal of Computer Applications (0975 – 8887) Volume 37– No.4, January 2012.
[7] Sancho McCann and David G. Lowe "Local Naive Bayes nearest Neighbor for Image Classification" in IEEE 2012.
[8] Lexiao Tian, Dequan Zheng and Conghui Zhu "Research on Image Classification Based on a Combination of Text and Visual Features" in Eighth International Conference on Fuzzy Systems and Knowledge Discovery 2011.
[9] Shaohua Wan"Image Annotation Using the SimpleDecisionTree" in International Conference on Management of e-Commerce and e-Government 2011.
[10] M. Kim, M. Madden, and T. A. Warner, "Forest type mapping using object-specific texture measures from multispectral IKONOS imagery: Segmentation quality and image classification issues," Photogramm. Eng. Remote Sens., vol. 75, no. 7, pp. 819–829, Jul. 2009.
[11] Y. O. Ouma, J. Tetuko, and R. Tateishi, "Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and nonforest vegetation in very-high-resolution optical-sensor imagery," Int. J. Remote Sens., vol. 29, no. 12, pp. 3417–3456, Jun. 2008.
[12] M. Simard, S. S. Saatchi, and G. De Grandi, "The use of decision tree and multiscale texture for classification of JERS-1 SAR data over tropical forest," IEEE Trans. Geosci. Remote Sens., vol. 38, no. 5, pp. 2310–2321, Sep. 2000.
[13] M. Pesaresiand and J. A. Benediktsonn, "A new approach for the morphological segmentation of high-resolution satellite imagery," IEEE Trans. Geosci. Remote Sens., vol. 39, no. 2, pp. 309–320, Feb. 2001.
[14] J. A. Benedktsonn, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," IEEE Trans. Geosci. Remote Sens., vol. 43, no. 3, pp. 480–491, Mar. 2005.
[15] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," IEEE Geosci. Remote Sens. Lett., vol. 8, no. 3, pp. 542–546, May 2011.
[16] M. Dalla Mura, J. AtliBenediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 10, pp. 3747–3762, Oct. 2010.
[17] D. Tao, X. Tang, X. Li, and X.Wu, "Asymmetric bagging and random subspace for FLDAs-based relevance feedback in image retrieval," IEEE Trans. Pattern Anal. Mach.Intell., vol. 28, no. 7, pp. 1088–1099, Jul. 2007.
[18] G. Guo, A. K. Jain, W. Ma, and H. Zhang, "Learning similarity measure for natural image retrieval with relevance feedback," IEEE Trans. Neural Netw., vol. 13, no. 4, pp. 811–820, Jul. 2002.
[19] V. Vapnik, The Nature of Statistical Learning Theory. NewYork: Springer-Verlag, 1995.