Improving the Accuracy of Object Based Supervised Image Classification using Cloud Basis Function Neural Network for High Resolution Satellite Images

Imdad Ali Rizvi

imdadrizvi@iitb.ac.in

Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India.

B.Krishna Mohan

Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India. bkmohan@iitb.ac.in

Abstract

A lot of research has been undertaken and is being carried out for developing an accurate classifier for extraction of objects with varying success rates. Most of the commonly used advanced classifiers are based on neural network or support vector machines, which uses radial basis functions, for defining the boundaries of the classes. The drawback of such classifiers is that the boundaries of the classes as taken according to radial basis function which are spherical while the same is not true for majority of the real data. The boundaries of the classes vary in shape, thus leading to poor accuracy. This paper deals with use of new basis functions, called cloud basis function (CBF) neural network which uses a different feature weighting, derived to emphasize features relevant to class discrimination, for improving classification accuracy. Multi layer feed forward and radial basis function (RBF) neural network are also implemented for accuracy comparison sake. It is found that the CBF NN has demonstrated superior performance compared to other activation functions and it gives approximately 3% more accuracy.

Keywords: Accuracy assessment, Image Segmentaion, Image classification, Object based image analysis, Radial basis functions neural network.

1. INTRODUCTION

Object based image classification methods are increasingly used for classification of land cover/use units from high resolution images, and often the final result is close to the way a human analyst would interpret the image. Object based image classification does not operate directly on single pixels, but image objects which refer to homogeneous, spatially contiguous regions. These are obtained by dividing an image, namely image segmentation, which is a challenging problem due to the fact that it is no longer meaningful to carry out this task on a pixel-by-pixel basis. The fine spatial resolution implies that each object is now an aggregation of a number of pixels in close spatial proximity, and accurate classification requires that this aspect be considered. To deal with the problem of complexity of high resolution images, the image is first segmented into homogeneous regions, and a set of features are computed for each region segment. These

segments are classified using one or more of the machine learning algorithms. In the present study, various activation functions for artificial neural network classification are considered. This method basically includes three steps. 1) Image segmentation to extract the regions from the pixel information based on homogeneity criteria. 2) Calculation of spectral parameters like mean vector, texture, NDVI and spatial/shape parameters like aspect ratio, convexity, solidity, roundness and orientation for each region. 3) Classification of image using the region feature vectors using suitable classifiers such as NN.

Supervised classification is one of the most commonly undertaken analyses of remotely sensed data. The output of a supervised classification is effectively a thematic map that provides a snapshot representation of the spatial distribution of a particular theme of interest such as land cover. The goal of a supervised image classification system is to group images into semantic categories giving thus the opportunity of fast and accurate image search. To achieve this goal, these applications should be able to group a wide variety of unlabelled images by using both the information provided by unlabelled query image as well as the learning databases containing different kind of images labelled by human observers. In practice, a supervised image classification solution requires three main steps: pre-processing, feature extraction and classification [1]. Based on this architecture, many image classification systems have been proposed, each one distinguished from others by the method used to compute the image signature and/or the decision method used in the classification step. Artificial Neural Network (ANN) and Support Vector Machine (SVM) are commonly used advanced methods for supervised classification of remotely sensed data [2]. The serious drawback of SVM is that the boundaries of the classes as taken according to radial basis function networks are spherical while the same is not true for majority of the real data. The boundaries of the classes vary in shape, thus leading to poor accuracy. This work is developed on the modified RBFs neural network based classifier for object based classification of high resolution satellite remotely sensed images. The new basis functions, called cloud basis functions use a different feature weighting, derived to emphasize features relevant to class discrimination as discussed in [3]. Further, these basis functions are designed to have multiple boundary segments, rather than a single boundary as for RBFs. This new enhancement to the basis function along with a suitable training algorithm allows the neural network to better learn the specific properties of the problem domain. The boundaries of classes considered are not spherical but a set of boundaries is considered for each class, which promises higher accuracy theoretically. This technique is emphasized specifically for multi-spectral satellite images. Thus it was aimed to propose a suitable classifier for the high resolution satellite remotely sensed images and to test the applicability of modified cloud basis functions for the field of remote sensing.

This paper is discussed under five different headings. In Section 2, proposed methodology for object based image segmentation and classification are elaborated. Some of the feature vectors are also discuss in the same section. In Section 3, comprises of experimental results and discussions. Section 4 summarizes the research findings and points out avenues for possible future works.

2. METHODOLOGY

The proliferation of high-spatial resolution multispectral imagery from satellite and aerial sensors (e.g. IKONOS from GeoEye, Inc., QuickBird from DigitalGlobe,Inc., ADS40 from Leica Geosystems, Inc.) has significantly changed the level of sophistication required in digital image processing [4]. In this paper we propose an approach for improving the accuracy of object based supervised image classification using Cloud Basis Functions Neural Network for high resolution remotely sensed multi-spectral satellite images, such as IKONOS or QuickBird.

Proposed Methodology for Object Based Image Segmentation

Over the last decade the analysis of Earth observation data has evolved from what were predominantly per-pixel multispectral-based approaches, to the development and application of

multiscale object-based methods. To empower users with these emerging object-based approaches, methods need to be intuitive, easy to use, require little user intervention, and provide results closely matching those generated by human interpreters. In an attempt to facilitate this, we present object-specific segmentation as an integrative object-based approach for automatically delineating image-objects (i.e., segments) from a high-spatial resolution remotely sensed image [5]. Object-based image analysis subdivides the image into meaningful homogeneous regions not only based on spectral properties but also on shape, texture, size, and other topological features, and organizes them hierarchically as image objects (also referred to as image segments) [6]. The segmentation procedure (extraction of the image objects) is controlled by the user-specified scale (size) or resolution of the expected objects. Object-based approaches have been successful for land-use and land-cover classification [7][8]. Classification of highresolution satellite images using standard per-pixel approaches is difficult because of the high volume of data, as well as high spatial variability within the objects. One way to deal with this problem is to reduce the image complexity by dividing it into homogenous segments prior to classification. This has the added advantage that segments can not only be classified on basis of spectral information but on a host of other features such as neighborhood, size, texture and so forth. The proposed methodology for object based image segmentation is shown in Fig 1.



FIGURE 1: Proposed methodology for object based image segmentation

Segmentation of the images is carried out using the region based algorithms such as morphological marker based watershed transform by employing the advantages of multi-resolution framework and multi-scale gradient algorithms. The segmentation of the color images is obtained using watershed transform to get its homogenous regions. Classification technique is then applied into these homogenous regions taking the shape, texture and spectral properties of the regions. The proposed algorithm is given below

- Apply multi-resolution framework (here Daubech6 family of wavelet transform is used) to input image.
- Use multi-scale gradient algorithms to calculate color gradient.
 The morphological gradient of each band of the image is calculated using equation (1)

$$G(f) = (f \oplus B) - (f \Theta B) \tag{1}$$

where G(f) = Morphological color gradient,

f = Given image

B = Structuring element.

$$G(f) = \sqrt{\left(G_r(f)^2 + G_g(f)^2 + G_b(f)^2\right)}$$
(2)

 $G_r(f)$ = Gradient of the red band, $G_g(f)$ = Gradient of the green band and

 $G_b(f)$ = Gradient of the blue band.

The multi-scale morphological color gradient is dilated with a square structuring element of size 2x2.

 The markers can be extracted from white top-hat or black top-hat transform. But extracted markers from either white or black top-hat will miss some of the objects. So, to utilize the advantage of both top-hat, markers are extracted using morphological laplacian [9], which can be defined as:

$$L(f) = g^{+}(f) - g^{-}(f)$$
(3)

where $g_{+}(f)$ = White top hat transform and $g_{-}(f)$ = Black top hat transform

For utilizing the spectral property of the image, markers are extracted from morphological color laplacian of the image; and is calculated using equation (4)

$$L(f) = \sqrt{\left(L_r(f)^2 + L_g(f)^2 + L_b(f)^2\right)}$$
(4)

where L(f) = Morphological color gradient,

 $L_r(f)$ = Gradient of the red band,

- $L_q(f)$ = Gradient of the green band and
- $L_r(f)$ = Gradient of the blue band.
- Apply connected component labeling to connect various labels.
- Morphological marker based watershed transform algorithm is used for region segmentation [10].
- Region merging is done to avoid over-segmentation.
- Mosaic image is generated.
- Inverse wavelet transform is used to generate high resolution image.

The output of the watershed transform may result in over-segmentation. To merge the adjacent region or the homogenous regions; region merging criterion is implemented. Each segmented object or region is assigned the average grayscale of each band to generate the

mosaic color image. To get the final segmentation at high resolution image, low frequency coefficient of the wavelet is replaced with mosaic image, while detailed coefficients of the wavelet are modified so as to avoid noise introduced back into the finer image. Inverse wavelet transform is then applied on these modified images to get the high resolution segmented image.

2.2. Feature Vector Extraction

In general, physical features have certain associations with spectral features, hence they can be identified by using multi-spectral information from the remotely sensed images. Features of objects can be further divided into three categories

- Geometric
- Spectral or thematic
- Textural

A feature vector of all the regions present in the image is calculated. For this work totally 8 features were calculated. The first three values correspond to the values of region's average color in multi-spectral space. The next three features are related to the shape of the region such as solidity, aspect ratio and eccentricity. The next features correspond to the texture features of each region like contrast ASM etc.[11].

2.3. Object Based Image Classification

Many classifiers are available for classification of multi-spectral satellite images. These include discriminate analysis, maximum likelihood classification scheme, etc. A major disadvantage of these classifiers is that they are not distribution free. This has prompted significant increase in use of ANN for classification of remotely sensed images [12]. Several other reasons can be sited in favor of Neural Network (NN) based classifiers as listed below [13].

- Each of the (region) parameters will be in a different numerical range, some in [0,1], some in [0, 255], etc. Rescaling all parameters to a single range can affect the inter-class and intra-class separation.
- NN classifiers can detect and use to their advantage non-linearity in data patterns.
- Ancillary data can be included in NN classifiers.
- NN architectures are flexible which can be easily optimized for performance.
- NN can handle multiple subcategories per class.

Multi Layer Feed Forward (MLFF) and Radial Basis Function (RBF) NN classification techniques are widely used remote sensing applications. In this study we consider one more type of NN classifier called Cloud Basis Function (CBF) NN.

2.3.1. Multi Layer Feed Forward Neural Networks

Typically an MLFF NN consists of a set of sensory units (source nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes. The input signal propagates through the network in a forward direction on a layer-by-layer basis. Learning in MLFF NN consists of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, an input pattern is applied to the sensory nodes of the network and its effect propagates through the network layer by layer. Forward pass is followed by a backward pass. During backward pass the error signal (difference between actual output of the network and the desired output) is propagated backward through the network, against the direction of synaptic connections. Hence it is named as back-propagation [14]. The synaptic weights are adjusted to make the actual response of the network move closer to the desired response. Following parameters are considered while implementing RBF ANN,

- Number of input nodes = 7
- Number of output nodes = 9
- Number of hidden layers = 1
- Number of nodes in each hidden layer = 8
- Learning rate = 0.79
- Momentum = 0.5
- Normalization factor for patterns = 255

2.3.2. Radial Basis Function Neural Networks

Design of RBF NN can be viewed as a curve fitting approximation in a high dimensional space. Learning is equivalent to finding a surface in a multidimensional space that provides a best fit to the training data. Correspondingly generalization is equivalent to the use of this multidimensional surface to interpolate the test data. The construction of a RBF NN, in its most basic form, involves three layers with entirely different roles. The input layer is made up of source nodes (sensory units) that connect the network to its environment. The second layer, the only hidden layer in the network, applies a nonlinear transformation from the input space to the hidden space; where hidden space is in general of high dimensionality. The output layer is linear, supplying the response of the network to the activation pattern applied to the input layer. A mathematical justification for the rationale of a nonlinear transformation followed by a linear transformation may be traced back to an early paper by [15]. According to this paper, a pattern-classification problem cast in a high dimensional space is more likely to be linearly separable than in a low-dimensional space. Following parameters are considered while implementing RBF ANN,

- Number of input nodes = 7
- Number of output nodes = 9
- Number of hidden layers = 1
- Number of nodes in each hidden layer = 8
- Learning rate = 0.85
- Momentum = 0.5
- Normalization factor for patterns = 255

2.3.3. Proposed Object Based Supervised Image Classification using Cloud Basis Function Neural Networks

Rather than treating image as set of pixels if we treat it as a set of objects more information can be extracted, as with pixels only intensity values can be used. And with the construction of regions, knowledge is given to the system to classify. This is similar to the way human brain analyzes an image by breaking it down into various objects and uses features such as shape, texture, color and context along with the its cognizance powers to interpret the image. Therefore, dividing the image into regions and then opt for classification is better than per pixel classification. Hence cloud basis function neural network is used which is essentially a form of neural network with modification in radial basis function neural network, the algorithm is as follows:

Creating the modified radial basis function neural network

- Define the input nodes, which take in as input the data from the images.
- Define the intermediate nodes for basis function mapping, which map the inputs to the basis space through the Gaussian functions.
- Define the output nodes, which form the classes in the image.

Programming the training algorithm for the neural network

- Apply k-means clustering for initial data to find the possible basis function centers, μ.
- Form the basis function mappings.
- Calculate the scale factors, for each of the basis function centers with respect to each of the other basis function centers.

$$\omega_{p,j} = \sqrt{\frac{1}{2} \sum_{i=1}^{d} \left(\mu_{ip} - \mu_{ij}\right)^2}$$
(5)

where, $\mathcal{O}_{p,j}$ = Scale factor of the boundary segment between pth mean and jth mean

 $\left(\mu_{ip}-\mu_{ij}
ight)$ = Euclidian distance between pth mean and jth mean

d = Number of features in each object

And the default scale factor as the mean of all the scale factors as

$$\omega_{0,j} = \frac{1}{k} \sum_{p=1}^{k} \omega_{p,j}$$
(6)

where, $\mathcal{O}_{0, i}$ = Default scale factor for jth mean

Compute the output matrix of the basis function mapping, φ, for the input samples.

$$\phi_{j}(x \mid \mu_{j}, \{\omega\}_{j}, U_{j}) = \exp\left(-\frac{\sum_{i=1}^{d} u_{ii,j}(x_{i} - x_{j})^{2}}{(Sel(\{\omega\}_{j} \mid x))^{2}}\right)$$
(7)

where , $\varphi_j(x \mid \mu_j, \{\omega\}_j)$ = Basis function output for sample x belonging jth to cluster

$$Sel(\{\omega\}_j \mid x)$$
 = Boundary segment selected for sample x
 $(x_i - u_{ij})^2$ = Euclidian distance between sample x and jth mean

Compute the post basis function weight matrix, W.

$$W = (\varphi)^{\neg} T$$
(8)
where $(\varphi)^{\neg} =$ Pseudo inverse of the output of the basis function matrix
 $T =$ Target Vector

- Compute the output of the network for the input samples and the error in the output with respect to the target vector T as the Euclidean distance from the target vector.
- Update the scale factors and the basis function centers based on the error in the output
 of the network using the supervised iterative gradient descent algorithm.

$$\{\omega\}_{j}^{m} = \{\omega\}_{j}^{m-1} - \frac{\partial E}{\partial \{\omega\}_{i}^{m-1}}$$

$$\tag{9}$$

 After iterative gradient descent is complete for the training iteration, the network output for all the training samples is calculated.

$$Network _Output = \Phi * W$$
(10)

where, Φ = Basis function output matrix and W = Post basis function weight matrix

- According to the network output, classify the pixels and partition the training set into two sets of classified {X^c} and misclassified samples {X^M}.
- If the number of misclassified samples is less than a set threshold, or if the number of misclassified samples doesn't change in successive cycles, stop training.
- For all the classes for which the number of misclassified samples is greater than the set threshold, add a basis function to improve the representation of the class.
- Repeat the training algorithm till a maximum number of epochs are completed or till the number of misclassified samples do not change with the increasing basis functions

Classifying the test images using the network

- Input the test images for classification
- Obtain the output matrix for the classification details of the image

Calculate the classification accuracy of the network

Following parameters are considered while implementing CBF NN,

- Total number of training samples taken = 135
- Maximum number of training iterations =10
- Maximum number of iterations for the iterative gradient descent for updating the scale factors calculated during each training iteration = 2
- Learning rate for the iterative gradient descent for updating the scale factors calculated during each training iteration = 0.2
- Maximum number of neighboring functions, for each basis function = 7 to 11
- Maximum number of misclassified samples, which when exceeded, a new basis function is to be added to the network = 10
- •

3. RESULTS AND DISCUSSION

We have implemented object based Multi Layer Feed Forward, Radial Basis Functions and Cloud Basis Function Neural Network in order to compare the accuracies with different activation functions. All these algorithms are implemented using C/C++ on Windows platform with Pentium 4 processor machine. The methodology is tested on a QuickBird window (2000 x 2000 pixels) of an urban fringe area comprising a few buildings, a quarry site, ponds, road, vegetation and foot paths. This image was retrieved on August 2001. On ground, it covers the Powai Area of Mumbai City. The study area is located between latitude (19 07' 14.69"N - 19 06' 39.98" N) and longitude (72 53' 43.07"E - 72 54' 29.08" E) as shown in Fig.2.



FIGURE 2: High resolution satellite image used as Study Area

The image was classified into 9 prominent classes covering a majority of the land cover features, Lake, Pool, Vegetation, Field, Road, Shadow, Bright Roof, and Dark Roof and Mountain. Accuracy and error statistics were computed for each activation function. Fig. 3 and Fig 4 depict output of object based classification using MLFF and RBF NN respectively which clearly indicates that object based classification is not a universal remedy, it is evident that regions are misclassified. For example, the roads and buildings or the grass and tree are spectrally similar and have a significant amount of spectral overlap. This is the primary reason for the large number of misclassifications between these classes. Similarly a part of the lake is being classified as a pool; an entire lake is classified as shadow, etc. This happens due to spectral

closeness of these regions. In order to reduce misclassification we need to take into consideration ancillary data (contextual data) available about the image.



FIGURE 3: Classified image using MLFF



FIGURE 4: Classified image using RBF

For example, if a region is classified as a shadow then there has to be tall structure in the vicinity of the shadow. Hence some improvement can be observed in Fig 5 which is classified image by CBF. There are plenty of geometrical information such as object feature, shape feature, texture, and contextual relation feature and so on. In this paper we add other feature information into feature space, which is area, entropy, shape index and contextual relation feature. Accuracy assessment is a general term for comparing predicted (i.e., classification) results to geographical reference data that are assumed to be true. This comparison is typically achieved by a basic subjective assessment of the observed difference in accuracy but should be undertaken in a statistically rigorous fashion. A set of reference objects representing geographic points on the classified image is required for the accuracy assessment. Randomly selected reference objects lessen or eliminate the possibility of bias.



FIGURE 5: Classified image using CBF

A random stratified sampling method was used to prepare the ground reference data. This sampling method allocates the sample size for each land use based on its spatial extent. A summary of accuracy and error statistics of all mentioned kernels can be found in Table 1.

Classes	ANN (MLFF)		ANN (RBF)		ANN (CBF)	
	CA	ΡA	CA	ΡA	CA	ΡA
Lake	0.8434	0.9357	0.8974	0.9127	0.8105	0.8045
Pool	0.9820	0.9658	1.0000	1.0000	1.0000	0.8750
Vegetation	0.8456	0.8834	0.8473	0.8494	0.9000	0.8750
Field	0.9248	0.8452	0.9308	0.8242	0.8954	1.0000
Road	0.5414	0.6143	0.5064	0.6873	0.9000	1.0000
Shadow	0.9632	0.9658	1.0000	0.9711	0.7988	0.8873
Bright Roof	0.9873	0.9921	1.0000	1.0000	0.9059	0.7660
Dart Roof	0.5360	0.6898	0.5760	0.6388	0.8514	0.8000
Mountains	0.7255	0.6585	0.5395	0.6445	0.9808	0.8750
Accuracy	0.8272		0.8542		0.8962	
Kappa Coefficient	0.8319		0.8491		0.8735	
CA = Consumer's Accuracy			PA = Producer's Accuracy			

Table 1. Accuracy and error statistics of object based supervised image classifier

The object based classification using CBF outruns the other kernel based NN classifiers in overall accuracy. The kappa coefficient which is 0.8319 and 0.8491 are low indicating the MLFF and RBF method are still an unsatisfactory one to classify remotely sensed images, where as for CBF it is recorded as 0.8735.

4. CONCLUSION

This paper attempts to study and compare the accuracy of object based image classifiers. The object based image analysis greatly reduced the salt-and-pepper classification effect in the classified image without adversely affecting the classified image accuracy. This greatly improves the visual effect of the classified image.

ANN has the advantages mainly of more tolerance to noise inputs and representation of boolean function apart from others, but too many attributes may result in over fitting. It was found that the neural network classifier trained using the standard back-propagation algorithm produced marginally better results compared to the other methods. The study shows that CBF NN improves the classification accuracy, though the CBF, being a relatively new technique in the remote sensing arena requires further study. A combined approach to classification using object based methods and contextual information available about the image, seems promising and needs further exploration.

5. REFERENCES

- [1] R.O. Duda, P.E. Hart P.E. and D.G. Stork, "Pattern Classification". 2nd edition, John Wiley & Sons (2001).
- [2] B. Tso and P.M. Mather, *"Classification methods for remotely rensed data"* London: Taylor and Francis (2001).
- [3] C.Ř. De Silva, S. Ranganath and L.C. De Silva, "Cloud Basis Function Neural Network: A modified RBF network architecture for holistic for holistic Facial Expression Recognition". Elsevier Pattern Recognition 41:1241-1253, 2008.
- [4] J. Im, J.R. Jensen and J.A.Tullis, "Object-based change detection using correlation image analysis and image segmentation". International Journal of Remote Sensing, 29(2), 399–423 (2008).
- [5] G.J Hay, G. Castilla, M. Wulder, and J.R Ruiz, "An automated object-based approach for the multiscale image segmentation of forest scenes" International Journal of Applied Earth Observation and Geoinformation 7, 339–359, 2005.
- [6] T. Blaschke, "A framework for change detection based on image objects", In Go"ttinger Geographische Abhandlungen, S. Erasmi, B. Cyffka and M. Kappas (Eds), 113, 1–9, Go" ttingen, 2005. Available online at http://www.definieris.com/pdf/publications/GGRS2004_Blaschke_G001.pdf (accessed 18 June 2009).
- [7] A.S. Laliberte, A. Rango, K.M. Havstad, J.F. Paris, R.F. Beck and Y.R. McNeely, "*Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico*". Remote Sensing of Environment, 93, 198–210, 2004.
- [8] R.C. Frohn, K.M. Hinkel, and W.R. Eisner, "Satellite remote sensing classification of thaw lakes and drained thaw lake basins on the North Slope of Alaska". Remote Sensing of Environment, 97, 116–126, 2005.
- [9] J. Eo and H. Kim, "A Detail Extraction Technique For Image Coding Using morphological Laplacian Operator". In Proceedings of the IEEE TENCON, pp.140-147, Korea, 1997.
- [10] F. Meyer and S. Beucher, "Morphological Segmentation". Journal of Visual Communication and Image Representation, 11: 21–46, 1990.
- [11] L. Costa and R. Cesar, "Shape Classification and Analysis", 2nd edition CRC Press, 259-310, (2009).
- [12] P. M. Mather, "Computer Processing of Remotely Sensed Images," John Wiley, NY, (2004).

- [13] J. Han, S. Lee, K. Chi, K. Ryu, "Comparison of Neuro-Fuzzy, Neural Network, and Maximum Likelihood Classifiers for Land Cover Classification using IKONOS Multispectral Data". In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, vol 6, pp. 3471-3473, Toronto, Canada, 2002.
- [14] S. Haykin, "Neural Networks: A Comprehensive Foundation" 2nd edition, Prentice Hall pp. 156-256, (1999)
- [15] T. Cover, "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition" IEEE Transactions on Electronic Computers, vol. EC-14(3),326-334, June 1965. Reprinted in Artificial Neural Networks: Concepts and Theory, IEEE Computer Society Press, Los Alamitos, Calif., 1992, eds. P. Mehra and B. Wah.