Pathak, Virendra and Dikshit, Onkar (2005) *Neuro-textural classification of Indian urban environment.* Geocarto International, 20(3). pp. 65-73.

Neuro-Textural Classification of Indian Urban Environment

VIRENDRA PATHAK†, ONKAR DIKSHIT‡
†Institute of Engineering & Technology, Lucknow, U. P., India
‡ Indian Institute of Technology, Kanpur, U. P., India.
*Email: vireniet@hotmail.com

Abstract

Experiments were conducted to see the effects of a set of factors on the Resilient backpropagation (*Rprop*) artificial neural network classification of an Indian urban environment using IRS-1C satellite data. Factors investigated were sample size, number of neurons in hidden layers and number of epochs. The effect of including texture information in the form of neighbourhood information and grey level co-occurance matrix (GLCM) features in the classification process has been explored. Statistically similar overall classification accuracy is achieved for *Rprop* and Gaussian maximum likelihood classification (GMLC). Investigations have revealed that a large sample size gave higher test accuracy; variation in number of neurons in hidden layer did not affect the overall classification accuracy significantly; lesser number of epochs resulted in higher overall test accuracy. Incorporation of texture information by both approaches improved classification accuracy in a statistically significant manner.

Key words: ANN, GMLC, GLCM, Texture, *Rprop*

1. INTRODUCTION

Attempts have been made in the recent past to use Artificial Neural Networks (ANNs) for alleviating some of the problems with conventional parametric pattern recognition techniques (Benediktsson *et al.* 1990; Bischof *et al.* 1992; Paola and Schowengerdt 1995a; Kanellopoulos and Wilkinson 1997). A few advantages of ANNs over conventional statistical techniques for pattern recognition are as follows. ANNs have an intrinsic ability to generalize; no assumption is made about the statistical distribution of the input data; these are capable of forming highly non-linear decision boundaries in the feature space; ANNs offer a more robust approach to land cover discrimination than that currently obtained using conventional supervised image classification techniques; ANNs have potential to provide higher classification accuracy with small training data compared to Gaussian maximum likelihood classification (GMLC) (Atkinson and Tatnall 1997). The accuracy of the classification using ANNs is a function of a wide range of factors. Paola and Schowengerdt (1995b) have presented a detailed review on backpropagation neural net

works and discussed various factors affecting classification accuracy. Investigations into these factors have focused especially on issues like type of network, its size and complexity, training set size, the learning algorithm, number of training iterations etc. (Foody *et al.* 1995a, 1995b, Foody and Arora 1997).

Most methods of classification use the grey scale values of a set of corresponding pixels taken from different spectral bands of the same scene to determine land use. However, a single ground cover usually occupies a region of neighbouring pixels and improved identification may be obtained by considering an entire region rather than a single pixel. The variability of grey values within the region can be taken into account together with the actual grey values. This variability constitutes the texture of the ground cover. Texture is a fundamental characteristic of image data and is often crucial to target discrimination (Woodcock and Strahler 1987). For spatially complex and spectrally mixed classes, the classification accuracy could improve if the spatial properties of classes were incorporated into the classification criterion (Lee and Philpot 1991). Texture methods are most appropriate under condition of high local variance like urban environments.

Most of the studies reported for land use classification have used standard back propagation methods, which are based on gradient descent algorithms with inherent limitations (Kanellopoulos *et al*. 1992; Paola and Schowengerdt 1995a; Shaban 1999). A better algorithm *Rprop* (Riedmiller and Braun, 1993) that overcomes the disadvantages of gradient descent has not attracted much attention of the remote sensing community despite its attractive feature. This paper, therefore, presents investigations of using *Rpop* artificial neural network along with textural information in image classification of an urban environment with IRS-1C satellite data.

2. OBJECTIVES, STUDY SITE AND DATA RESOURCE

The objective of this study was to evaluate the results of *Rprop* classification in comparison to GML classification (GMLC). Another objective was to investigate effects of a set of factors on the accuracy of *Rprop* classification. The factors investigated were sample size, number of neurons in hidden layers and number of epochs**.** Further, the effects of including texture properties in the form of neighbourhood information and grey level co-occurrence matrix (GLCM) based features on classification accuracy have also been investigated. The study area is a typical urban Indian city Lucknow, the state capital of the northern Indian state Uttar Pradesh. The geographical extent of Lucknow ranges between North latitudes 26^{o} $45^{'}$ and 27^{o} and the East longitudes 80^{o} $50^{'}$ and 81^{o} $5^{'}$. A central extract (512x512 pixels) from the area is chosen for this study. Figure 1 shows false colour composite (FCC) of the study area. Twelve classes covered the majority of urban land use features (Table 1). The satellite data for the study area included four multispectral bands of IRS-1C, LISS-III sensor with 23.5 m pixel size.

The following sections present the literature survey, theoretical background of *Rprop* and various factors considered for the study. The experimental design, results and conclusions for various investigations follow these.

3. LITERATURE SURVEY

The potential and capability of the neural network approaches over GMLC to deal with complex remotely sensed data has been demonstrated by many researches. Hepner *et al.* (1990) compared between BP ANN and GMLC in a land cover classification. They observed that the ANN could classify imagery better than GMLC using identical training sites although it was computationally very intensive. However, these conclusions were based on purely visual interpretation and without any mathematical comparison between the accuracy obtained in the ANN and GMLC techniques.

Similar observations were made by Foody *et al*. (1995a, 1995b) about the ability of ANN to have higher classification accuracy compared to the discriminant analysis (DA), especially when smaller training sets were available. The differences were only significant if the data were non-normally distributed. However, if the remotely sensed data satisfied the assumptions of conventional statistical classifiers (say normality in case of GMLC), no significant difference in classification accuracy was usually observed between conventional and neural networks classifications. When *a priori* information was available to the DA, the difference in accuracy between neural network and DA was found to be insignificant. It was also observed that the non-representative training data would lead, as expected, to significant difference between training data and testing classification accuracies and the effect was fairly similar for both the ANN and DA classifications.

Paola and Schowengerdt (1995a) compared a BP neural network with GMLC for classifying 12 urban land use classes using TM imagery for two cities. The neural network approach achieved higher test site accuracy than the GMLC, but required considerably higher computing time. The closer test and training site accuracies indicated that the neural network generalised better than the GMLC method. The classification time with ANN was about 15 times higher than with GMLC to produce the same overall test site accuracy. For the second city, the two maps were visually and numerically similar, although the neural network was superior in suppression of mixed pixel classification errors. These results indicated that the BP approach to neural network training was computationally intensive, taking at least an order of magnitude more time than the total classification time for GMLC. However, the classification time, once training was complete, was less for the neural network.

The accuracy of the classification using ANNs is a function of a wide range of factors. Paola and Schowengerdt (1995b) also presented a detailed review on backpropagation neural networks and discussed various factors affecting classification accuracy. Investigations into these factors have focused especially on issues such as type of network, its size and complexity, training set size, the learning algorithm, number of training iterations etc. (Foody *et al.* 1995a, 1995b, Foody and Arora 1997).

A large number of measures have been proposed to extract texture information from an image. For a comprehensive review of texture algorithms, one can refer to Haralick (1973), Gool *et al*. (1985), and Dikshit (1992). Conners *et al*. (1984) obtained higher classification accuracies by segmenting a high-resolution black and white image of urban area using GLCM-derived texture operators. Hlavka (1987) used edge-density texture measure with Thematic Mapper simulator (TMS) data and observed that urban and rural areas could be distinguished with texture alone. Use of two features of GLCM such as entropy and inverse difference moment derived from directional spatial co-occurrence matrices along with the spectral features improved the overall classification accuracy with SPOT data (Franklin and Peddle 1989). For an urban area, Lee and Philpot (1991) described a pattern-matching algorithm for classification which performed either as good as or superior to GLCM. Shaban (1999) obtained significant improvement in classification accuracy of Indian urban environment by a combination of texture and spectral features compared with pure spectral feature using SPOT images.

4. THEORATICAL BACKGROUND

The investigations in this paper have used *Rprop* artificial neural network and Gaussian maximum likelihood classification (GMLC) algorithms. The GMLC is a parametric classifier that relies on the second order statistics of a Gaussian probability density function (pdf) model for each class. It is often used as a reference for classifier comparison because, if the class pdf's are indeed Gaussian, it is the optimal classifier (Paola and Schowengerdt 1995a). Further details about GMLC are available with any of the standard text on remote sensing (Mather 1987, Richards 1993).

Many variants of neural network algorithms derive from the multilayer backpropagation neural network. For multispectral image classification, the most widely used input/output configuration is one input node for each input channel (typically each band of a multispectral image) and one output node for each desired class label. The number and size of the hidden layer is not determinate, though a few guidelines exist to help the user (Paola and Schowengerdt 1995a, Kanellopoulos and Wilkinson 1997). Every input and output node is connected to all of the hidden layers nodes. Each interconnection has an associated weight and as a whole contain (after training) the distributed, learned information about the classes. For complete details about ANNs one can refer to Zurada (1997), Rumelhart *et. al* (1986) and Dayhoff (1990).

A multilayer network typically uses sigmoid transfer functions in the hidden layers. These functions, are often called squashing functions since they compress an infinite input range into a finite output range. Sigmoid functions are characterized by the fact that their slope must approach zero, as the input gets large. This causes a problem when using steepest descent to train a multi layer network with sigmoid functions, since the gradient can have a very small magnitude, and therefore cause small changes in the weights, even though the weights are far from their optimal values.

The purpose of *Rprop* training algorithm is to eliminate these harmful effects of the magnitudes of the partial derivatives. Only the sign of the derivative is used to determine the direction of the weight update, the magnitude of the derivative has no effect on the weight update. The size of the weight change is determined by a separate update value. The update value for each weight is increased by a suitable factor, whenever the derivative of the performance function with respect to that weight has the same sign for two successive iteration. The update value is decreased by another factor whenever the derivative with respect to that weight change sign from the previous iteration. If the derivative is zero, then the update value remains the same. Whenever the weights are oscillating the weight changes sign from the previous iteration. If the weights continue to change in the same direction for several iterations, then the magnitude of the weight change will be increased. *Rprop* generally converges much faster than the other algorithms (Riedmiller and Braun, 1993).

Following paragraphs presents the theoretical background about factors such as sample size, number of neurons in the hidden layers, number of epochs and texture considered for the study.

The training data must be representative of the class with which it is associated. In addition, these classes must have some seperability in the feature space for the classifier to be able to discriminate them. According to a general guideline given by Mather (1987), in the case of a single variable and the estimation of a single property (such as mean or the variance) a sample size of 30 is usually held to be sufficient. For the multivariate case the size should be at least 30*p* pixels per class, where *p* is the number of features (spectral bands), and preferably more. Researchers have used different training sets where size of the data varied considerably. Civco used one training sample, the mean vector, per class in one of the study (1991) and 10 samples per class in another (1993). Hepner *et.al.* (1990) used what they termed 'the minimal training set' consisting of a 10 by 10 training site for each class. A few others used similar training set sizes (Xiao and Liu 1991, Benediktsson et al. 1990). The largest training set used consisted of 22000 Patterns. (Heermann and Khazenie 1992). Foody *et al.* (1995a) reported that classification accuracy was significantly increased as a result of increasing the number of training cases for abundant classes in the image.

Toratora (1978) suggested that for the multinomial distribution, the sample size is given by the following equation

$$n = BP_i(1 - P_i)/b_i^2$$

Where B is the upper (α/k) x 100th percentile of the χ^2 distribution with 1 degree of freedom. P_i, $i = 1,\ldots..k$, is the proportion of the population in the ith category and b_i is the absolute precision desired. In the majority of the cases, an absolute precision is set for the entire classification and not for the each category. Therefore, $b_i = b$ and the only sample size calculation is required for the P_i close to 1/2. In the worst-case scenario, the sample size can be obtained from the simple equation

$$n = B / 4b^2$$

B can be determined from the χ^2 table with 1 degree of freedom and $1-\alpha/k$, where $\alpha = 1-R$, R is the desired reliability (confidence level).

Generally, for classification of multispectral imagery, a three layer (single hidden layer) fully interconnected network is sufficient and is most commonly used strategy (Benediktsson *et al.* 1990; Civco 1991; Paola and Schowengerdt 1995a). The number of nodes in a hidden layer required for a particular classification problem is not easy to deduce. The neural network architecture, which gives the best results for particular problem, can only be determined experimentally. This can be a lengthy process especially for large classification task. To define a network size which is appropriate for a given classification problem, it is necessary to examine the total number of input features and the number of output classes. For a three layer network, a general guideline on this issue has been given by Paola and Schowengerdt (1995a). Kanellopoulos and Wilkinson (1997) stated that ideally the first hidden layer of a network with two hidden layers should contain two to three times the number of inputs such that a sufficient number of hyper-planes can be formed to define hyper-regions. The second hidden layer effectively combines the hyper-planes or hyper-regions from the previous layer to form sub-regions defining each class. To allow two or three regions per class, as often employed in statistical classification of remotely sensed data, they found it useful to make the number of nodes in second hidden layer roughly equal to two to three times the total number of classes. However they caution that it is not possible to rely on such heuristics and each classification problem needs to be carefully examined in its own right.

Bischof *et al.* (1992) minimized the 'sum squared error' in 50 training cycles, Civco (1991) obtained a root mean square error of 0.18 after 250000 iterations of the class mean vectors. Kanellopoulos *et al.* (1992) found that it required 900 training cycles to achieve 81 percent classification accuracy on test data. Paola and Schowengerdt (1995a) achieved 93.4 percent test site accuracy after 50000 iterations.

Texture can be defined as a repeated variation in tone (spectral response) over relatively small areas. Texture features provide information about the spatial distribution of spectral variations, Haralick *et al*. (1973) suggested that textural and spectral properties are present in an image simultaneously but under a given condition, one property can dominate the other. They presented one of the most widely used approaches to texture analysis, the grey level co-occurrence matrix (GLCM) approach. The intermediate result of the approach is in the form of so-called co-occurrence matrices. From these matrices, a large number of texture features can be computed and used for classification. For details about GLCM texture features, the reader can refer Haralick *et al*. (1973).

5. EXPERIMENTAL METHODOLOGY

Samples (S1 to S5) were determined at five different reliability values with desired precision of 5% (Toratora 1978) (Table 2). Test samples sets of the similar size as of training samples were used to test accuracy of classifications. A three layer (single hidden layer) fully interconnected network was used for *Rprop* classification. The experiments were carried out in three stages. In the first stage, classification was done with spectral features using *Rprop* and then with GMLC. For *Rprop* classification of first stage a network with 4-12-12 configuration was used (four input bands, twelve hidden nodes and twelve output classes). Training of network was carried out for 1000 epochs.

In the second stage, a study was conducted to evaluate factors affecting classification accuracy using *Rprop*. The factors studied were sample size, number of neurons in hidden layer and number of epochs. For studying effect of sample size, classifications with different sample set (S1 to S5) and varying number of neurons in hidden layer were carried out for 1000 epochs. To understand effects of variation in number of nodes in hidden layer and number of epochs, classifications were carried out with sample size S5 while varying number of nodes in hidden layer (8 to 20) at different number of epochs ranging from 1000 to 50000 (1k to 50k).

In the third stage, effect of adding texture information in the form neighbourhood information and GLCM texture feature was studied. In the first approach, texture information from band 1 was included by capturing information from 3x3 window and taking central pixel from rest of the bands making a total of 12 input nodes (Bischof *et al*. 1992). In the second approach, texture information from all bands were used from 3x3 window making a total of 36 input nodes for a four band data (Hepner *et al*. 1990).

To study the effect of adding texture features derived from GLCM method, *mean* (mean), *variance* (var), *homogeneity* (hom), *contrast* (con) *and dissimilarity* (dis) were taken for the study. Shaban and Dikshit (1998) reported that window sizes 7 and 9 give best result for texture extraction from Indian urban areas. Based on the results of their study, GLCM texture features at window size 7 and 9 were considered. For experimentation using GLCM texture feature in conjunction with spectral features, first GML classification was carried out and then *Rprop* classification was done using information derived from second stage of experimentation for number of neurons in hidden layers and number of epochs. Training was carried out till 1000 epochs or MSE became less than a threshold value of 0.001. The input value of every pixel feature vector was normalised between 0-1. The number of neurons in hidden layer was kept equal to three times number of input bands for all experimentations related with GLCM texture and number of neurones in output layer was equal to the number of classes.

The overall classification accuracy and the accuracy of the individual classes were assessed by computing kappa coefficients (κ) and associated asymptotic variances (Bishop *et al*. 1975). Pair-wise statistical tests were performed to assess the significance of any differences observed between two classifications using a Z statistic as given by the following equation (Congalton *et al.* 1983). In this equation Z_{ab} is the Z statistic for comparison of classification a and b; κ_a and κ_b are the kappa coefficient of classification a and b; and σ_a^2 and σ_b^2 are the asymptotic variances of κ_a and κ_b respectively. The difference between two classifications was considered to be statistically significant at the 95% confidence level if the absolute value of the *Z* statistic exceeded 1.96.

$$Z_{ab} = \frac{\kappa_a - \kappa_b}{\sqrt{\sigma_a^2 + \sigma_b^2}}$$

6. RESULTS

(1) Comparison of classification results by *Rprop* and GMLC are presented in Table 3, from which the following observations can be made:

(a) Classification using *Rprop* provides similar accuracy to that of GMLC. Though training accuracy appears to be higher in *Rprop* but that is statistically insignificantly different from GMLC.

(b) Accuracies for most of the classes were similar with no significant difference.

(2) Investigations using five different sample sets have shown (Table 4) that with the increase in sample size, test accuracy increase for all networks having different number of neurons in hidden layer. Maximum overall accuracy was achieved for sample size S5, which has highest number of samples per class.

(3) Table 5 presents results showing the effect of number of neurons in hidden layer on classification accuracy. The following observations emerge from this table:

(a) Variation of neurons in hidden layer did not have any significant effect on the classification accuracy for the test set.

(b) With increase in number of neurons in hidden layer, it was observed that overall training accuracy increased with the increase in number of neurons at higher epochs.

(4) Results of experiments with different number of epochs have been presented for sample set S5 (Table 5). The following points emerge from this table:

(a) With an increase in number of epochs test accuracy decrease, higher test accuracy was achieved with fewer epochs. Maximum accuracy was achieved with 1000 epochs, which was similar to accuracy achieved at 5000 epochs.

(b) Increase in number of epochs increases training accuracy in a significant manner with highest at 50000 epochs.

(5) Texture in the form of neighbourhood information was added using two approaches. Table 7 shows result of these investigations from which the following points emerge:

(a) First approach of incorporating texture property using neighbourhood information improves test accuracy that is statistically insignificant.

(b) Second approach using neighbourhood information from all the bands simultaneously improves test accuracy significantly.

(6) Tables 6, 7 and 8 show results of adding GLCM texture features with spectral bands in GML and *Rprop* classification. The following point comes out from experiments:

(a) Accuracy using window size 9 is slightly higher than 7 for both the classification methods.

(b) Texture features except *homogeneity*, when used in conjunction with spectral features improve classification accuracy in a statistically significant manner with GML as well as *Rprop* classification.

7. CONCLUSIONS

From the results of experimentation following conclusions can be drawn:

1. Use of ANN in the form of *Rprop* provides statistically similar accuracy as with GMLC for both, spectral and textural classification.
2. With the increase in size of sample set, test accuracy increases in *Rprop* classification.
3. Variation in number of neurons in hidden layer does not affect test accuracy in a significant manner.
4. Increase in number of epochs increases training accuracy but decreases test accuracy. However, higher test accuracy is achieved with smaller number of epochs. This could be due to over-fitting of data at higher epochs.
5. Inclusion of texture in the form of 3x3 window neighbourhood information simultaneously from all bands increases test accuracy in a significant manner.

6. GLCM texture features except *homogeneity*, when used in conjunction with spectral features, increase accuracy in a significant manner for GML as well as *Rprop* classification.

REFERENCES

Atkinson PM, Tatnall ARL (1997) Neural networks in remote sensing. *International Journal of Remote Sensing* 18:699-709

Benediktsson JA, Swain PH, Ersoy OK (1990) Neural network approaches versus statistical methods in classification of multisource remote sensing data. *I.E.E.E. Transactions on Geoscience and Remote Sensing* 28:540-551

Bischof H, Schneider W, Pinz AJ (1992) Multispectral classification of Landsat images using neural networks. *IEEE Trans. on Geoscience and Remote Sensing* 30:482-490

Civco DL (1991) Landsat TM image classification with an artificial neural network. *Proceedings, ASPRS-ACMS Annual meeting, Baltimore, MD,* Vol. 3 (Falls Church, VA: American Society for Photogrammetry and Remote Sensing), pp.67-77

Congalton RG, Oderwald RG, Mead RA (1983) Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques, *Photogrammetric Engineering and Remote Sensing* 49:1671-1678

Conners RW, Trivedi MM, Harlow CA (1984) Segmentation of a high-resolution scene using texture operators. *Computer Vision, Graphics and Image Processing* 25:273-310

Dayhoff JE (1990) *Neural Network Architectures: An Introduction*, Van Nostrand Reinhold, New York.

Dikshit O (1992) *The classification of texture in remotely-sensed environmental imagery.* Ph.D. Thesis, University of Cambridge, U.K.

Franklin SE, Peddle DR (1989) Spectral texture for improved class discrimination in complex terrain. *International Journal of Remote Sensing* 10:1437-1443

Foody GM, McCulloch MB, Yates WB (1995a) The effect of training set size and composition on artificial neural network classification. *International Journal of Remote Sensing* 16:1707-1723

Foody GM, McCulloch MB, Yates WB (1995b) Classification of remotely sensed data by an artificial neural network: Issues related to training data characteristics. *Photogrammetric Engineering and Remote Sensing* 61:391-401

Foody GM, Arora MK (1997) An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing*18:799-810

Gool LV, Dewaele P., Oosterlinck A (1985) Texture analysis, anno. 1983. Computer Vision Graphics, and Image Processing 29:336-357

Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics* SMC-3:610-621

Hepner GF, Logan T, Ritter N, Bryant N (1990) Artificial neural network classification using a minimal training set: comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing* 56:469-473

Heermann PD, Khazenie N (1992) Classification of multispectral remote sensing data using a back – propagation neural network. *I.E.E.E. Transactions on Geoscience and Remote Sensing* 30:81-88

Hlavka CA (1987) Land-use mapping using edge density texture measures on Thematic Mapper simulator data. *IEEE Trans. on Geoscience and Remote Sensing*, GE-25, 104-107

Kanellopoulos I, Varfis A, Wilkinson GG, Megier J (1992) Land cover discrimination in SPOT HRV imagery using an artificial neural network-a 20-class experiment. *International Journal of Remote Sensing* 13:917-924

Kanellopoulos I, Wilkinson GG (1997) Strategies and best practice for neural network image classification. *International Journal of Remote Sensing* 18:711-725.

Lee J, Philpot W (1991) Spectral texture pattern matching: a classifier for digital imagery. *I.E.E.E. Transactions on Geoscience and Remote Sensing* 29:545-554

Mather PM (1987) *Computer processing of remotely-sensed images: An introduction*, John wiley & sons.

Paola JD, Schowengerdt RA (1995a) A detailed Comparison of back-propagation neural network and maximum-likelihood classifiers for Urban Land use classification. *I.E.E.E. Transactions on Geoscience and Remote Sensing* 33:981-996

Paola JD, Schowengerdt RA (1995b) A review and analysis of back-propagation neural networks for classification of remotely-sensed multispectral imagery. *International Journal of Remote Sensing* 16:3033-3058

Richards J.A. (1993) *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag.

Riedmiller M, Braun H (1993) A direct adaptive method for faster backpropagation learning: The RPROP algorithm. *Proceeding of the IEEE International conference on Neural Networks*

Rumelhart DE, Mcclelland JL (1986) *Parallel Distributed Processing*, Vols. 1 & 2 Cambridge, Mass.: MIT Press

Shaban MA, Dikshit O (1998) Textural Classification of High Resolution Digital Satellite Imagery. *Proceedings of the international Geoscience and Remote Sensing Symposium (IGARSS'98), Seattle Washington, July 1998* (Seattle)

Shaban MA (1999) *Digital classification of urban environments using satellite imagery.* Ph.D. Thesis, Indian Institute of Technology, Kanpur, India

Toratora RD (1978) A note on sample size estimation for multinomial populations. *The American Statistician.* 32 (3): 100-102

Woodcook CE, Strahler AH (1987) The factor of scale in remote sensing. *Remote Sensing of Environment* 21:311-332

Xiao JY, Liu ZK (1991) Classification of remotely–sensed image data using artificial neural networks. *International Journal of Remote Sensing* 12:2433-2438

Zurada JM (1997) *Introducation to artificial neural systems*. Jaico Publishing house, Mumbai.

Table 1 Classes in Lucknow study area and their brief description

S. no.	*Name*	*Description*
1	Agriculture-1	Agriculture area having crops at middle stage of growth
2	Agriculture-2	Agriculture area having crops at early stage of growth
3	Commercial	Central business area of the city
4	Educational institutes	Various educational Institutions
5	Government establishment	Different Government establishments
6	Grassy land	Big patches of lands having grass only
7	High residential	Residential areas with more than 600 persons/hectare
8	Medium residential	Residential areas with 400 persons/hectare
9	Park	Parks for recreational activities
10	Reserve forest	A big portion of land reserved for forest
11	River	River Gomti flowing from left to right
12	Water body	Various small water bodies in the study area

Table 2 Sample size

Samples (SS)	S1	S2	S3	S4	S5
Reliability in %	50	75	85	95	99
Sample size/class	35	46	53	71	90

Table 3 Comparison of classification accuracy (κ) between GMLC and *Rprop*

Class No.	Training			Test		
	a	*b*	Z_{ba}	*a*	*b*	Z_{ba}
1	0.75	0.82	1.11	0.78	0.79	0.19
2	0.82	0.90	1.57	0.77	0.83	0.96
3	0.57	0.53	-0.45	0.71	0.64	-0.97
4	0.51	0.44	-0.88	0.53	0.45	-1.01
5	0.19	0.32	1.99	0.11	0.22	1.75
6	0.92	0.90	-0.27	0.89	0.90	0.25
7	0.85	0.80	-0.84	0.81	0.83	0.36
8	0.91	0.91	0.00	0.85	0.83	-0.41
9	0.68	0.78	1.42	0.73	0.74	0.20
10	0.93	0.92	-0.25	0.90	0.87	-0.70
11	0.95	0.95	0.01	0.92	0.90	-0.27
12	0.87	0.96	2.27	0.93	0.98	1.46
Overall	0.74	0.77	1.35	0.74	0.75	0.26

a Using GMLC
b Using *Rprop* Classification
Z_{ba} Z-statistic

Table 4 Accuracy (κ) with different test sets (S1 to S5)

Neurons	Sample sets				
	S1	S2	S3	S4	S5
8	0.72	0.73	0.72	0.73	0.74
12	0.71	0.74	0.72	0.74	0.75
16	0.72	0.75	0.72	0.73	0.76
20	0.72	0.73	0.73	0.74	0.75

Table 5 Accuracy (κ) for different number of epochs

Neurons	Epochs									
	1k		5k		10k		25k		50k	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
8	0.77	0.74	0.79	0.74	0.79	0.74	0.79	0.73	0.79	0.73
12	0.77	0.75	0.79	0.75	0.80	0.74	0.80	0.75	0.80	0.74
16	0.78	0.76	0.81	0.75	0.81	0.74	0.82	0.74	0.83	0.73
20	0.78	0.75	0.81	0.75	0.82	0.74	0.83	0.75	0.83	0.73

a Training accuracy
b Test accuracy

Table 6 Accuracy (κ) after adding GLCM texture features using GMLC and *Rprop*

Features	Training		Testing	
	GMLC	*Rprop*	GMLC	*Rprop*
*Mean7**	0.84	0.85	0.83	0.84
Mean9	0.85	0.87	0.84	0.84
Var7	0.79	0.83	0.78	0.81
Var9	0.80	0.83	0.79	0.81
Hom7	0.78	0.78	0.76	0.75
Hom9	0.79	0.81	0.77	0.77
Con7	0.79	0.82	0.78	0.80
Con9	0.80	0.82	0.79	0.80
Dis7	0.80	0.82	0.79	0.79
Dis9	0.81	0.81	0.79	0.80

**xw* where *x* is texture feature and number *w* is window size

Table 7 Accuracy (κ) using spectral + textural features (widow size 9) for the test set using *Rprop* classification

Class	*Features*								*Z-Statistic*						
	a	*b*	*c*	*d*	*e*	*f*	*g*	*H*	Z_{ba}	Z_{ca}	Z_{da}	Z_{ea}	Z_{fa}	Z_{ga}	Z_{ha}
1	0.79	0.83	0.90	0.94	0.93	0.98	0.99	0.96	0.60	1.98	2.82	2.51	3.83	4.22	3.45
2	0.83	0.88	0.90	0.93	0.94	0.87	0.93	0.94	0.88	1.39	1.94	2.22	0.65	1.92	2.23
3	0.64	0.49	0.69	0.76	0.79	0.69	0.78	0.74	-1.95	0.73	1.68	2.23	0.76	2.02	1.44
4	0.45	0.50	0.66	0.53	0.54	0.43	0.46	0.42	0.63	2.69	0.99	1.10	-0.30	0.03	-0.42
5	0.22	0.43	0.51	0.59	0.35	0.27	0.41	0.40	2.97	4.12	5.36	2.02	0.76	2.81	2.68
6	0.90	0.87	0.93	0.92	0.90	0.92	0.77	0.86	-0.74	0.56	0.27	0.01	0.29	-2.31	-0.93
7	0.83	0.85	0.88	0.87	0.90	0.79	0.88	0.88	0.42	0.89	0.70	1.42	-0.58	0.92	0.93
8	0.83	0.91	0.93	0.94	0.86	0.83	0.88	0.89	1.68	1.96	2.26	0.48	0.00	0.92	1.17
9	0.74	0.76	0.82	0.79	0.77	0.84	0.76	0.74	0.19	1.15	0.76	0.38	1.56	0.20	0.03
10	0.87	0.90	0.90	0.99	0.92	0.89	0.90	0.93	0.73	0.75	3.09	1.01	0.49	0.75	1.30
11	0.90	0.92	0.92	0.88	0.90	0.88	0.93	0.87	0.27	0.27	-0.50	0.00	-0.50	0.55	-0.73
12	0.98	0.96	0.88	0.95	0.98	0.94	0.95	0.94	-0.45	-2.44	-0.84	0.00	-1.17	-0.83	-1.17
Overall	0.75	0.77	0.83	0.84	0.81	0.77	0.80	0.80	1.41	4.25	5.09	3.56	1.46	2.88	2.66

a Spectral features only
b Textural information from neighbourhood (Input nodes-12)
c Textural information from neighbourhood (Input nodes-36)
d Spectral + *Mean9*
e Spectral + *Var9*
f Spectral + *Hom9*
g Spectral + *Con9*
h Spectral + *Dis9*
Z_{xa} Z-statistic with respect to *a* (*x* varies from *b* to *h*)

Table 8 Accuracy (κ) using spectral + textural features (widow size 9) for the test set using GML classification

Class	Features						Z-Statistic				
	a	*b*	*c*	*d*	*e*	*f*	Z_{ba}	Z_{ca}	Z_{da}	Z_{ea}	Z_{fa}
1	0.78	0.92	0.81	0.93	0.81	0.88	2.44	0.38	2.69	0.38	1.67
2	0.77	0.93	0.88	0.83	0.90	0.93	2.89	1.85	0.97	2.34	2.88
3	0.71	0.83	0.76	0.78	0.76	0.76	1.87	0.72	1.06	0.70	0.70
4	0.53	0.52	0.63	0.44	0.56	0.51	-0.14	1.22	-1.20	0.33	-0.34
5	0.11	0.52	0.25	0.13	0.27	0.25	6.14	2.26	0.34	2.57	2.31
6	0.89	0.90	0.87	0.89	0.88	0.88	0.25	-0.47	0.00	-0.24	-0.24
7	0.81	0.89	0.89	0.84	0.84	0.83	1.55	1.54	0.62	0.63	0.42
8	0.85	0.93	0.87	0.86	0.89	0.89	1.55	0.27	0.24	0.74	0.75
9	0.73	0.83	0.79	0.80	0.79	0.79	1.51	0.90	1.12	0.89	0.91
10	0.90	0.95	0.95	0.94	0.93	0.93	1.24	1.23	0.90	0.58	0.60
11	0.92	0.90	0.90	0.93	0.92	0.92	-0.27	-0.27	0.29	0.00	0.00
12	0.93	0.93	0.96	0.93	0.92	0.93	0.00	1.03	0.00	-0.29	0.00
Overall	0.74	0.84	0.79	0.77	0.79	0.79	5.17	2.75	1.61	2.31	2.47

a Spectral features only
b Spectral + *Mean9*
c Spectral + *Var9*
d Spectral + *Hom9*
e Spectral + *Con9*
f Spectral + *Dis9*
Z_{xa} Z-statistic with respect to *a* (*x* varies from *b* to *h*)

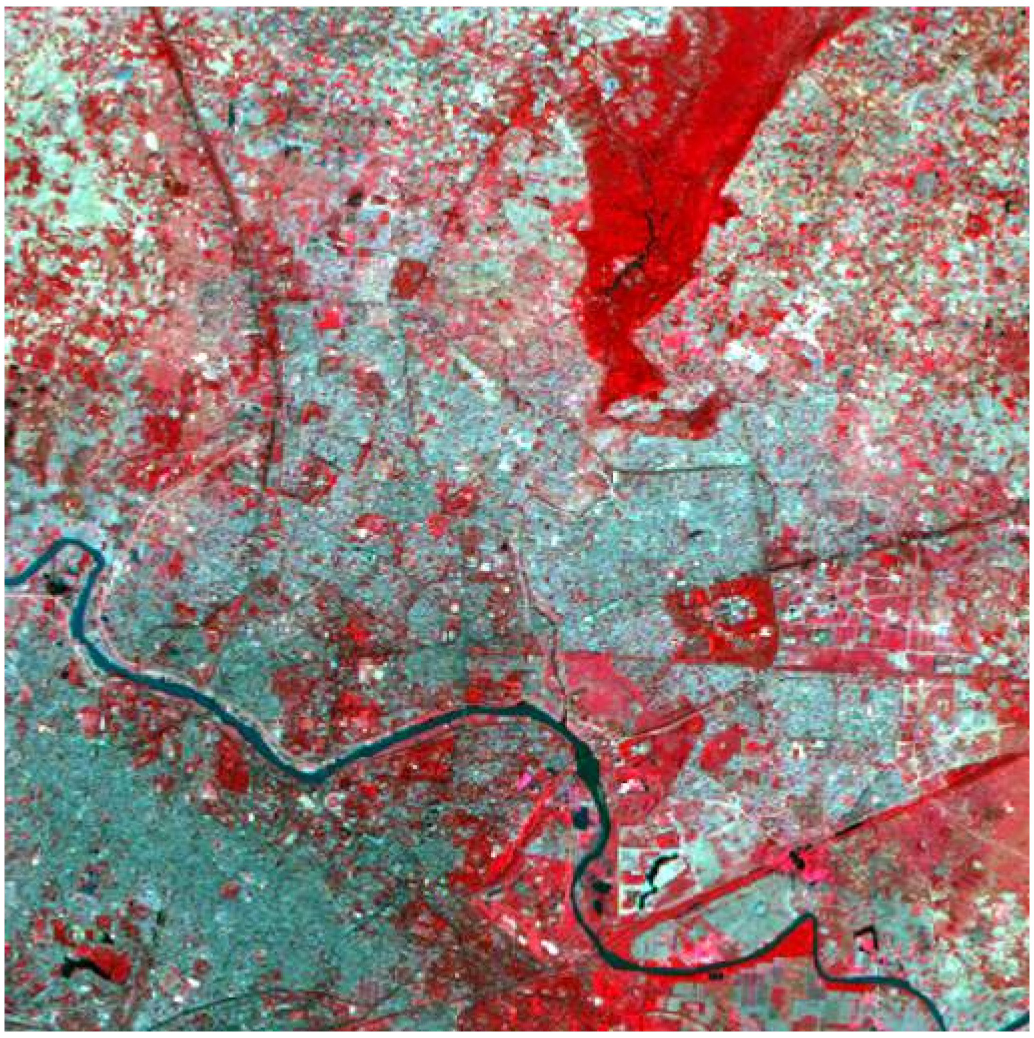

Figure 1.   FCC of study area