An Ensemble Approach for the Detection And Classification Of Mixed Pixels of Remotely Sensed Images

Shabnam Choudhury
Electrical Engineering
National Institute of Technology
Rourkela, India
Email: 216ee1266@nitrkl.ac.in

Anik Chaudhuri
Electrical Engineering
National Institute of Technology
Rourkela, India
Email: 216ee1265@nitrkl.ac.in

Dr Dipti Patra
Electrical Engineering
National Institute of Technology
Rourkela, India
Email: dpatra@nitrkl.ac.in

***Abstract*—Accuracy has been necessity condition to render fine spatial resolution in the mapping of land patches. Every remotely sensed image can be characterized as objects whose accuracy varies as a function of the spatial resolution. The objects can be assigned a spectral signature which demonstrates the reflection factor of the pixels. This further prompts the confusion of pixels occupying more than one class. Such pixels can be termed as mixed pixels, whereas pixels of a homogeneous class designated as pure pixels. This paper sights the mixed pixels of the satellite images and focuses on their classification employing the texture feature. The primary reason of focus on texture is that it is based on the spatial arrangement of intensities of an image. The texture features such as colour intensity, energy and local binary pattern are used to analyze the mixed pixel problem. Supervised learning techniques such as artificial neural network and ensemble bagged learning has been reviewed and compared to handle the mixed pixel issue, thus imparting better resolution.**

Keywords*—*Ensemble bagged learning; Local binary pattern; Mixed pixel; Spatial resolution.

I. Introduction

Remotely sensed images are the multispectral category of images [1] which is meant to extract data within the specified wavelength, in particular, spectral patterns. Mixed pixel [2] is the pixel that contains the average energy reflected from the multiple bands of a multispectral image. Hence, it does not belong to a unique class i.e. possesses a heterogeneous nature. The main cause of mixed pixel is the low spatial resolution of the sensor as a result of which the adjacent consistent pixels occupy the same pixels. Hence the result is a mixture of the individual end members.

Spatial resolution can be defined as the number of independent pixel per inch. The clarity of the image is determined accurately by fine spatial resolution, but that can be unavailable and also cost-ineffective. Hence coarse resolution can be taken into account. Nevertheless such images contain large proportions of mixed pixels which further deteriorate the quality to a great extent. Hence the processing of coarse spatial resolution to fine spatial resolution is the main challenge. Due to the presence of mixed pixels, it is quite difficult to estimate the area coverages and position the boundaries. In order to tackle this issue, accurate classification algorithms and validation of training samples are adopted.

II. Literature Study

The generalised land cover classification techniques [4] are hard and soft classification. A hard classification technique maps the mixed pixel into a class of larger proportion irrespective of the estimation of the class probability as it directly predicts the class boundary. Consequently, this approach leads to information loss deteriorating the quality of land patches appearance. Mixed pixels can thus be treated as the error, noise in the allocation of class for hard classification methods. This paved the way for an approach which takes into consideration the conditional probabilities estimation and is based on sub pixel classification. Soft classification generates multiple fractions of images by estimating class proportion of each pixel. These multiple layers can be used in various learning algorithms for further classification. Un-mixing algorithms [5] were first developed using reflectance parameters for the chemical mixture. In the separation of the pixel spectrum each endmember in the pixel is estimated as its abundance value. It incorporates two kinds of models- linear and non-linear. Linear model [6] computes the pixel reflectance as the linear combinations of pure class reflectance weighted by their respective surface proportions. But the drawback is that it neglects the multiple scattering among end members. Alternatively, researchers went for the non-linear model [7] but complexity was the primary issue. Probabilistic approaches such as maximum likelihood [8] were also used in which the variances and covariance of the class are taken which determines the probability of a pixel to belong to a class. But this type of classifier requires an adequate data sample which has to be sampled for the estimation of the mean vector and variance-covariance matrix. The variance-covariance matrix stability fails during high correlation.Markov random field approach [3] was also adopted for spatial resolution of multi-spectral images by taking the class membership of each pixel and assigning them a particular class label and considering the spatial attraction model. Fuzzy based classification [5][6] such as fuzzy c-means clustering has also been employed for classifying pixels into class membership functions based on maximum similarity. But it fails to resolve the mixed pixel issue completely as it classifies with variable membership grades.

Many learning algorithms [9] supervised and unsupervised have adopted for proper classification of mixed pixels. Unsupervised classification algorithms such as K-means clustering were employed to cluster the pixels of a given class thereby segregating the other classes ones. Supervised learning has been carried out in these stages of training, classification, and output. It develops the training set of pixels by identifying the particular areas and spectral characteristics. This is followed by classifying the pixels into their respective classes based on their estimated probabilities of belongingness. Finally, the categorised land cover pixel is generated in the output. Another category of learning is the semi-supervised type of learning in which both labelled and unlabeled data sets are used for improving the accuracy of algorithms. Earlier the backpropagation method for classification of IRS-ID satellite images has been adopted and the comparison was made against maximum livelihood classifier and the accuracy was calculated as 85.19%. Feed forward neural networks were also employed for classification. The statistical approaches are generally not useful as it uses the probability density functions which are based on the Gaussian distribution. In this paper, the mixed pixel issue has been identified and classified using the supervised learning approach using artificial neural network which is then further compared with the ensemble learning approach using classifications such as Linear Discriminant Analysis, linear support vector machines, KNN, commonly known as ensemble bagged tree. The feature used for classification is the texture of the image, taking into account the attributes such as colour intensity, energy content, and local binary pattern.

Artificial neural network is a basically a non-parametric approach, learns the class probabilities from the data itself. This paper is motivated to employ back propagation method to classify the mixed pixel present in the remotely sensed image based on texture features and the accuracy has been improved to 89%.But the underlying limitation is the it fails to show satisfactory performance for external datasets.

Section II explains the hierarchical study area of the paper. Section III represents the methodology proposed followed by section IV which contains the simulation results. Section V gives the conclusion statement.

III.THEORETICAL STUDY

Texture analysis of an image [10] refers to the quantification of an image based on their texture content and spatial arrangement of pixel position. Texture based classification is useful as it explains the local variations quantifying smoothness, roughness and regularity of the surface. There are basically two approaches of texture analysis- structural and statistical. The structural approach explains images as asset of repeated patterns (texels).The statistical approach emphasises on the intensity arrangement, hence is more popular. This type of analysis is helpful when the classification of the region of interest has to be performed based on texture value of the image rather than the threshold of the features.

A. Colour Intensity

Colour intensity [11] defines the RGB (Red, Green, Blue) intensity of an image. Landsat -8 images are taken for analysis which uses multispectral scanner mapped in Band 4-Red of wavelength 0.64-0.67m with resolution 30m. Hence, we have worked upon the red pixel value of the image.

B. Co-occurrence matrices

Co-occurrence matrices [12] are a very relevant method to represent the spatial correlation of gray scale values of images at a given offset such as the distance of variation and direction. Gray level co-occurrence matrices (GLCM) have the main advantage of sparsity, hence is a very useful tool for texture feature extraction. It calculates the number of occurrences a particular pixel with intensity p in the specified direction to adjacent pixel q, represented as $G(p,q)$ where G represents elements of GLCM matrix. The rotational invariance of the image is checked by varying the pixel in different directions such as 0, 45^o, 90^o, 135^o degrees. The elements of the GLCM matrix can be used to find the energy content. Energy is taken as the angular second moment (ASM) feature. Angular second moment describes the homogeneity of a distribution. It can be found from the GLCM matrix as the squared value of the individual elements. The equation for calculating energy is represented by equation (1)

$$\sum_x \sum_y G(x,y)^2 \tag{1}$$

G represents the elements of GLCM matrix. x and y are the row and column values.

C. Local Binary Pattern

The local binary pattern is the latest emerging statistical feature[13] used for texture classification based on its neighbourhood. In this technique, the centre pixel intensity is compared with the neighbourhood pixel value, thereby thresholding the neighbouring pixel value and the output is represented as a binary number. Mathematically it can be represented by function shown below.

$$f(k) = \begin{cases} 1 & k > 0 \\ 0 & otherwise \end{cases}$$

k is represented as the pixel intensity difference between two points. The resultant binary number represented in the form of histogram. Since they make no assumptions regarding probability distribution, LBP is invariant to transformations in gray-scale.

IV. PROPOSED METHODOLOGY

Classification of remotely sensed images is basically done in the major steps [14][15] including selection of a proper training samples or data, selection of proper application area, proper pre-processing and evaluation process. Selection proper satellite data depends broadly on the resolution of the sensor, feasibility of data and attribute of area of study. Since satellite images are nothing but the aerial photographs, proper sites should be considered to avoid coarser resolution. Data pre-processing which includes geometric, atmospheric and radiometric corrections is also an indispensable for image classification. Decision of suitable classification methods is also of utmost importance. Two methods have been proposed in this paper the first approach is to use the extracted features to train an artificial neural network which was further improved using ensemble classification. The extracted features are colour intensity, energy, local binary pattern.

A. Artificial Neural Network

Artificial Neural Network [16][17] is an interconnection of layers of neurons, multilayer preceptor(MLP) . It consists of three layers-input, hidden and the output layer. The layers consist of nodes which produce output based on the activation function. Weighted directed edges form the connection between the neurons which can be modified using learning algorithms. The nodes of the input layer receive the value and duplicate it to multiple layers, hence are passive in nature. While the nodes of other layers hidden and output are active in nature, they are capable of modifying the data. Proper selection of weights improves the accuracy of a neural network .The first step is the computation of number of endmembers from the input data followed by training of samples in the input layer. The spectral bands correspond to number of neurons in the input layer while the output layer gives information regarding estimated endmembers.

Backpropagation technique is mainly motivated to train the network for any random mapping input to output. It calculates the difference between actual output and the expected output and propagates the error back to the network. It mainly calculates the gradient of the objective function, which is further optimized and weights are updated to minimize the function.

The training of the network is shown in Fig. 1

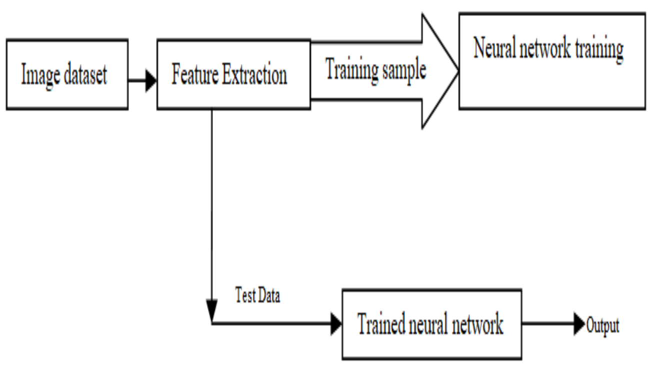

Fig. 1. Shows the neural network training and testing flow graph

1) The feature extraction block signifies the feature extracted here, i.e, colour intensity, energy and local binary pattern.The detection of mixed pixels is based on the observation that pixel values between 60-255 are classified as land; 10-43 as water; 44-57 as a mixed pixel.
2) The training data block contains 7200 data points of land, water, and mixed pixels.
3) The test data contains 2400 data points of mixed pixels.
4) The output of the network is finally categorised into class 1 (00)- land; class 2 (01)-water; class 3 (11)-mixed.

The complexity and time-consuming drawback of ANN paved the way to ensemble learning approach.

B. Ensemble Learning

Ensemble refers to the technique of averaging multiple classes to improve the performance of the hypothesis. Multiple models can be combined to form an optimum decision boundary [18] which is often a tedious process for the single classifier. Moreover, it allows reliable handling of few or large data, hence quite flexible. The effective ensemble methods are – Bagging, Boosting, Adaboost, and Stacking.

Bagging is a technique in which different data sets are sub-sampled with replacement, and the training dataset and sample size are the same. Ensemble decisions are taken based on majority voting.

Boosting also resamples data but they assign each consecutive classifier, the most informative data. Since it is applicable to only binary classification, the Adaboost algorithm was developed.

In AdaBoost algorithm, each iteration maintains a set of weights for the trained datasets followed by a learning algorithm to minimize the weighted error. Finally, the weights are updated after minimization. The key point in this algorithm is to put more weights on misclassified pixels rather than correctly classified.

In stacked generalisation, two level of the learning process is employed. In the first level, different learning algorithm employed to generate data followed by a combination of the individual learner, known as meta-classification.

The ensemble learning is definitely a better approach than single learning as it measures the disadvantage of a weak solution of a single classifier. Also, ensemble draws perfect hypothesis with good approximation.

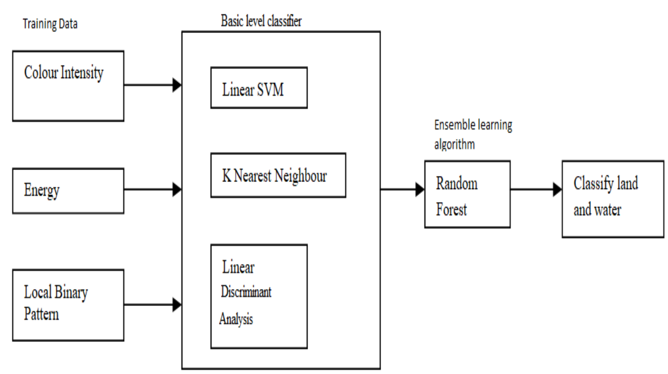

Fig. 2. Ensemble learning approach to classify land and water

The flow graph in Fig. 2 takes 8,457 data sets for each feature. The features used are colour intensity, energy, and local binary pattern. It is then fed into the ensemble tree classifier block which is followed by random forest algorithm. In the random forest algorithm, three input features are separated at each node. Then the linear combination of input features is created at each node. Finally, the best split is selected randomly. Random forest algorithm improves accuracy by reducing the correlation which increases the error bounds. It is simple and can be easily parallelized.

Ensemble bagged trees [19][20] make the average predictions of the random data sets. They are best suited for verifying

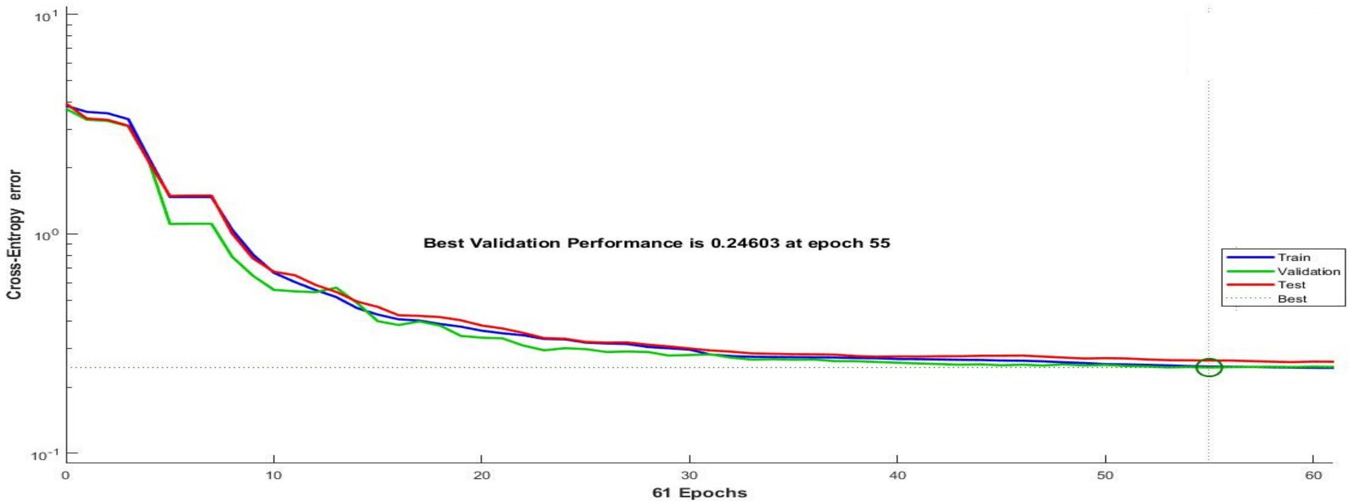

Fig. 3. Performance graph of the trained neural network

method, hence can draw good results in complex models. Ensemble classifiers can be used for both large and small volume of data. The classifiers combined here are Linear SVM, Linear Discriminant Analysis, Fine KNN, Ensemble bagged tree.

Linear SVM linearly scales the training data, thus improves the accuracy. The hyperplane for a set of points p can be represented by equation (2)

$$F(p) = w^T p + b \tag{2}$$

where p is the pixel data, w is the weight vector(not necessarily normalized), b is the bias.The objective function is meant to maximize the margin between the hyperplanes.If the training set is linearly separable, the two classes of data can be separated by selecting two hyperplanes, so that the margin between them is maximized.

Linear Discriminant Analysis is mainly used for dimensionality reduction. It represents the direction in which maximum separation of classes occurs, as it maximizes the ratio of between-class variance to within-class variance.

Fine KNN classifiers are mainly used for obtaining finely detailed separations between classes. The number of neighbours is set to 1. It uses the representation based distances for classification, hence contains more information which improves the performance with respect to ordinary KNN.

V. RESULTS AND DISCUSSION

The simulation results of the adopted methods can be categorized into two sections:

A. Neural Network Approach

Performance of the neural network can be represented in the form of

1) Performance Graph.

The performance graph is shown in Fig. 3 plots the cross-entropy error against the number of epochs. Cross-entropy error explains better classification. Here the best result was achieved at 55 epochs with a validation error of 0.24603.

II)Receiving operating characteristic.

Fig. 4 plots for true positive rate against false positive rate for each output class. For a particular threshold, the ratio of sensitivity (true positive rate) to specificity (false positive rate) can be found from each point on the ROC curve.

B. Ensemble learning approach

Several individual classifier models are combined to enhance the machine learning approach. The single ensemble model decreases variance (bagging) bias (boosting), or improve predictions (stacking).

Ensemble methods can be divided into two groups:

- Sequential ensemble methods- where base learners are generated sequentially i.e to highlight the difference between the base learners. This boosts the overall performance.
- Parallel ensemble methods- where the base learners are generated in parallel i.e to highlight the independence nature of the base learners. This reduces the error by averaging.

It can be represented in the form of accuracy of different classifiers enlisted in the table below.The accuracy represents the faithfulness of the particular classifier to classify land, water and mixed pixels.From the accuracy value of the classifiers,we can make out that the ensemble classifier gives the best output relative to the other classifiers.

Classifiers Used	Accuracy
Fine KNN	94.2 %
Linear SVM	97%
LDA	97.5%
Ensemble bagged tree	98.7 %

TABLE I. ACCURACY OF CLASSIFIERS

Fig. 5 represents the classification of the Landsat 8 image of the Norwegian sea (taken from Landsat Image Gallery) of size 272 x 330. Using the proposed method the classification of land, water and mixed pixels in the image has been performed.

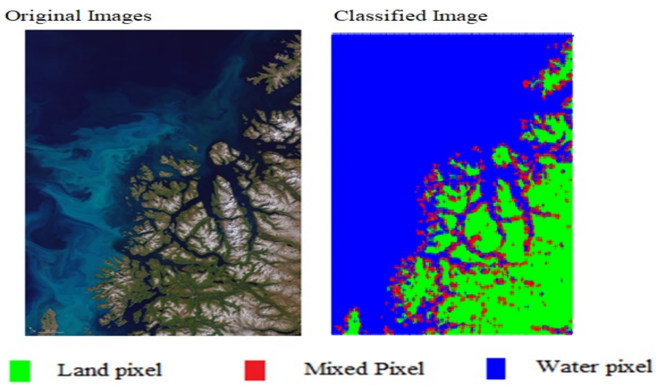

Fig. 4. LANDSAT image (272 x 330) of Norwegian sea

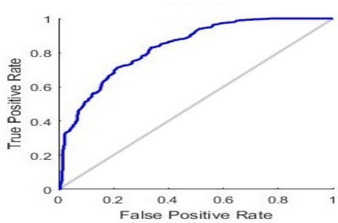

Fig. 5. Receiving operating characteristic of the trained neural network

VI.CONCLUSION AND FUTURE SCOPE

Mixed pixel problem in the satellite images is the primary reason behind the coarse resolution. Low-resolution images are of no use for useful processing of data in various fields. Moreover greater is the resolution , more complex and costlier is the design. Hence, to abate the issue is the key interest of this paper. Artificial neural network was successful to classify to a certain extent as it is independent of Gaussian distribution, hence indifferent to spectral linearity of data. But it fails to handle large data leading to more iterations, thus greater computation time. Henceforth the ensemble learning approach was adopted which average the predictions, therefore, can be useful in both small and large data handling. Further with large datasets, the mixed problem can be accurately detected and dealt with even without using feature maps.

REFERENCES

[1] J. B. Campell, “Introduction to Remote Sensing”, 2nd ed. London, U.K.: Taylor Francis, 1996.

[2] M.S. Klein Gebbinck and T.E. Schouten, “Decomposition of Mixed Pixels,” SPIE 2579, 1995, pp. 104–114.

[3] Hua Zhang, Wenzhong Shi, Yunjia Wang, Ming Hao, and Zelang Miao, “Spatial-Attraction-Based Markov Random Field Approach for Classification of High Spatial Resolution Multispectral Imagery,”IEEE Geoscience and Remote sensing letters , vol. 11, no. 2, February 2014

[4] Tso, B., Mather, P. M., Classification methods for remotely sensed data, New York: McGraw-Hill, Taylor and Francis, 2001.

[5] G.M. Foody. and D.P. Cox, “Sub-Pixel Land Cover Composition Estimation Using a Linear Mixture Model and Fuzzy Membership Model and Fuzzy Membership Functions,” Int.J. Remote Sens. 15 (3), 1994, pp. 619–631.

[6] Foody, G. M., 1996, Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data. International Journal of Remote Sensing, vol. 17, no. 7,

[7] Lillesand, T.M. and Kiefer. R. W., Remote Sensing and Image Interpretation, Fourth Edition, John Wiley and Sons, 2002, ISBN 9971-51-427-3.

[8] C.C. Borel and S.A.W. Gerstl, “Nonlinear Spectral Mixing Models for Vegetative and Soil Surfaces,” Remote Sens. Environ. 47 (3), 1994, pp. 403–416.

[9] Congalton, R. G., 1991, A review of assessing the accuracy of classifications of remotely sensed data, Remote Sensing of Environment, 37, 35–46.

[10] Adams, J. B., Sabol, D., Kapos, V., Filho, R. A., Roberts, D. A., Smith, M. O. and Gillespie, A. R., 1995, Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. Remote Sensing of Environment, 52, 137 154.

[11] Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics 6(3), 610–621 (1973).

[12] A Color-Texture-Structure Descriptor for High-Resolution Satellite Image Classification Huai Yu , Wen Yang, Gui-Song Xia and Gang Liu

[13] BO Hua, MA Fu-long, JIAO Li-cheng, Research on Computation of GLCM of Image Texture [J], ACTA ELECTRONICA SINICA, Vol.34 No.1, pp.155-158, 2006

[14] “A comparative study of texture measures with classification based on feature distributions”. Pattern Recognition. 29(1), 51–59 (1996)

[15] Erbek, F.S., Ozkan, C. and Taberner, M., 2004, Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. International Journal of Remote Sensing, 25, pp. 1733–1748

[16] Foody, G. M., McCulloch, M. B., and Yates, W. B., The Effect of Training Set and Composition on Neural Network Classification. International Journal of Remote Sensing, 1995, vol. 16 (9), pp. 1707 – 1723.

[17] Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., and Franklin, J., 1999, A neural network method for efficient vegetation mapping, Remote Sensing of Environment, 70, 326–38.

[18] Foody, G. M., McCulloch, M. B., and Yates, W. B., 1995, Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics.Photogrammetric Engineering and Remote Sensing,61,391401

[19] Pal, M., and Mather, P. M., Decision Tree based classification of remotely sensed data. Paper presented at the 22 nd Asian conference on Remote Sensing, 5-9 November 2001, Singapore.

[20] Yang, C-C., Prasher, S. O., Whalen, J., and Goel, P. K, Use of Hyperspectral Imagery for Identification of Different Fertilisation Methods with Decision-tree Technology. Biosystem Engineering, 2002, vol. 83(3), pp. 291 – 298.