

# HYPERSPECTRAL SENSOR DATA FUSION AT DECISION LEVEL USING SUPPORT VECTOR MACHINE

Y.Sai Praveen<sup>1</sup>, A.Kiranmai<sup>2</sup>, K.Nikitha<sup>3</sup>, V.Sowmya Devi<sup>4</sup>

<sup>1</sup>M.Tech, GITAM University, Hyderabad

<sup>2</sup>M.Tech, GITAM University, Hyderabad

<sup>3</sup>Assistant Professor, GITAM University, Hyderabad

<sup>4</sup>Assistant Professor, GITAM University, Hyderabad

## Abstract

SVM is the major feature extraction technique extensively used in the hyperspectral sensor. First, analysis aimed at understanding and assessing the potentialities of SVM classifiers in data fusion at decision level is presented. This is very significant particularly for improving the algorithms for classification of remote sensing data from different types of sensors operating in hyperspectral or panchromatic mode. The major problem occurs with high dimensionality of hyperspectral sensor data and nonlinear characterization, SVM can overcome these problems and deliver a better result when compared with other classifiers at decision level. The performance and efficiency of SVM depend upon the problem and different kernel functions used in SVM classifier at decision level data fusion. In this paper, problems in the classification of hyperspectral sensor data are analysed and applicability of decision level data fusion as a part of the enhancement of sensor data capability is studied.

**Keywords:** Classification of hyperspectral data, artificial neural networks, Support vector machine, Decision-level Fusion.

\*\*\*

## 1. INTRODUCTION

Hyperspectral sensor images allow users to observe how the surface of the Earth is altering swiftly, at local, regional, national, even global scales. The images produced by remote sensing systems are in digital form. These images must be processed before we use it for processing because the image that we get might get some noise because of the atmospheric effects. So we apply image processing techniques in order to remove atmospheric effects and to get correct reflectance spectra. In Remote sensing, on the basis of a number of spectral bands the imaging systems can be classified as panchromatic that contains single band with grey scale image, multispectral with less than ten spectral bands, Super spectral that possess tens of bands and finally Hyperspectral with hundreds of bands. In this report, we used Hyperspectral data for classification, in details of this Hyperspectral data we considered Hyperion. Hyperion has 242 spectral bands ranging from 355.59-2577.08 nm. Due to poor signal to noise ratio, the only specific number of bands are processed. The sensor operates in 'push broom' mode, which means that the sensor is not scanned across the surface like another sensor. The Spectral bands of Hyperion cover the visible-near infrared region and Short Wave infrared region of Electromagnetic Spectrum. Here we processed 50 bands of VNIR and 148 bands of SWIR. As the satellite progresses along track more lines are built up generation data which is known as a data cube. The main intent of this paper is the classification of remote sensing data using different classification techniques. The technique used for classification of hyperspectral data might be supervised or unsupervised the main aim of is to cluster the

pixels of one group and assign it into one of the several classes. The resulting image that we get after classification is called as a thematic map. Classification can be performed either using Multispectral or Hyperspectral data [11] [12]. Even multispectral data also provide enough information to perform classification but the main advent of hyperspectral data is its spectral resolution. With the use of hundreds of spectral channels, we can discriminate between several classes easily compared to the multispectral data. At the same time Classification of Hyperspectral data is considered as a very challenging task, where problems involved are related to the high dimensionality of the data[11][12][13], diverse noise and high computational cost[2].here in this paper we have specified some results that show the possibility and solutions of the classifiers in different hyperspectral problems.

## 2. LITERATURE

From the past to the now contemporary world, remote sensing and classification techniques have changes a lot as the time keeps on changing these models have achieved high accuracy and turn out to be more robust. The SVM was first introduced by COLT and became famous when classification of different problems is solved using support vector machine. The most common problems like regression and fuzzy are brought to results. SVM can be applied to different areas like patterns recognition and other data classification areas [2] [3]. The Bayesian method helps learning method as well as the mathematical method for classification [4]. Assumes a fundamental probabilistic method and it allows us to note ambiguity about the method

in an upright way by considering probabilities of the given inputs. This method is named after Thomas Bayes (1702-1761), who introduced the Bayes Theorem [5]. The original perceptron model was introduced by Frank Rosenblatt in 1958[5]. Rosenblatt's model consists of three layers, a retina, an association unit and an output unit. Here the input layer data is given to association unit where weights of each input layer are calculated according to the threshold function and these weights are processed and the output is given to the final output layer [6]. The MLP looks like a network designed for the work of classification and regression problems by Weideman et al... It is a three layer one that contains an input layer, a Hidden layer and output layers. The relations in the network designed by Weideman et al. were more difficult than usual MLPs [6].

### 3. STUDY AREA

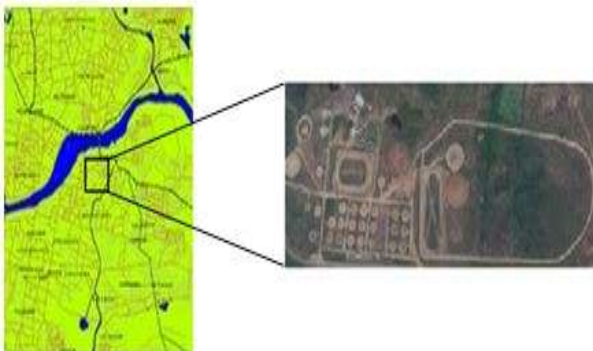


Fig 1: Study Area

The study area is part of Hyderabad city. It falls in the longitude and latitude of 17.3700° N and 78.4800° E. The input datasets are EO-1 Hyperion and Orb view-3 images. The data is radiometrically and atmospherically corrected in order to remove atmospheric effects such as noise and to obtain correct spectral reflectance. Geometric correction is done in order to get the data to standard geographic projection.

### 4. METHODOLOGY

The Hyperion data is atmospherically, radiometrically and geometrically corrected, bad and noisy bands are removed from data. Hyperspectral data contains high dimensions and Principal Component Analysis is used to reduce high dimensionality of data. This dimensionally reduced hyperspectral data is fused with panchromatic data using Gram-Schmidt image fusion technique. Now the data is prepared for feature extraction and classification.

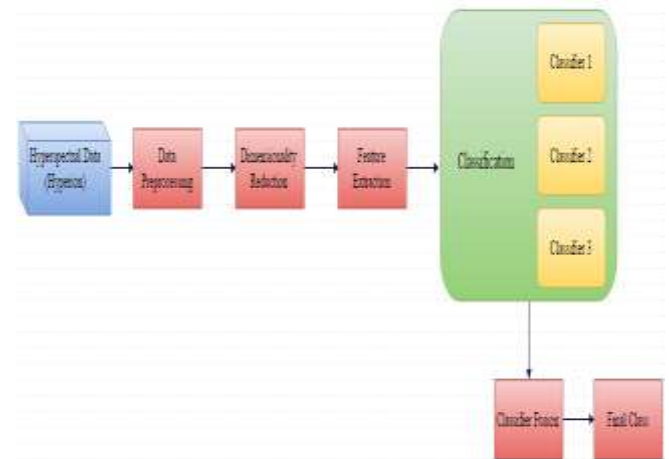


Fig 2: Methodology

From hyperspectral data we have taken some spectral features like NDVI, Simple Ratio, Water Band Index etc. [1]. We have prepared classes like roads, soil, buildings and tanks. The classification techniques such as SVM, Naïve Bayes and MLP are applied separately. Now we tried to increase the classification efficiency of the classifier so we combined all the classifiers to form multiple classifier system (MCC). In this system the first technique is SVM which discriminates classes by fitting an optimal linear hyperplanes. Also Naïve Bayes and MLP are applied as other two classifiers [8] [9].

### 5. CLASSIFICATION MODELS

#### 5.1 Support Vector Machine

The fundamentals of Support Vector Machines (SVM) [2] [3] have been developed by Vapnik (1992) and are gaining reputation due to many attractive features, and shows potential empirical performance. SVMs were primarily developed to perform binary classification; though, applications of binary classification are very restricted. Most of the practical applications involve multiclass classification, especially in remote sensing land cover classification. Basically, SVM is used as a binary classifier in the initial stages later by using many to one and one to one schema which helps SVM to classify data into multiple classes.

SVM belong to the general category of kernel operations. A kernel operation is a mathematical algorithm that can be applied to different kinds of problems. These kernel operations are used as a substitution for regression problems. In hyperspectral classification, SVM can be used with different kernel operations and these results produce complex patterns and these patterns can be reduced by using post-classification techniques. There are four main kernel functions in SVM Linear-kernel, Polynomial-kernel, RBF-kernel and sigmoid kernel.

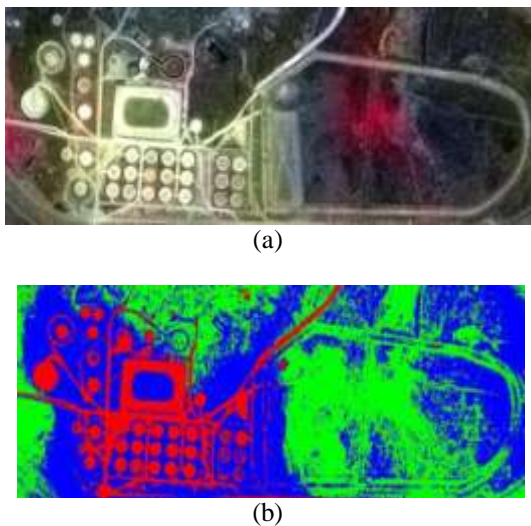


Fig 3: (a)Fused image (b) Classified image using SVM

## 5.2 MLP

In this neural network, we have input layer, hidden layer and output layer. Here the desired data is given to the input layer and the number of inputs should be greater than the hidden layers. Now the input data is mapped to the hidden layer. In the hidden layer, the algorithm works according to the data and the output is given to the final layer. A graphical representation of an MLP is shown below:

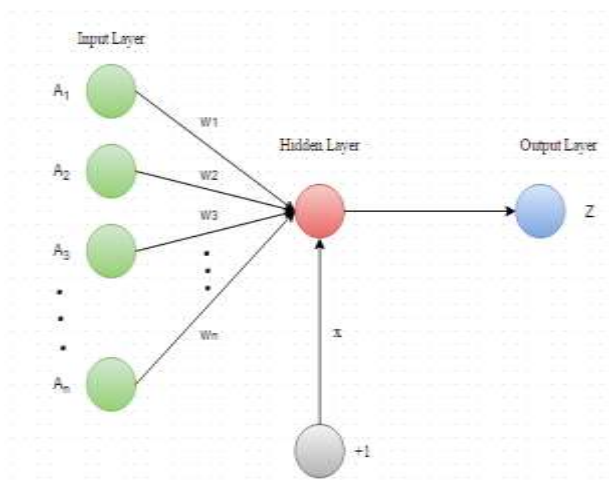


Fig 4: Multi Layer Perceptron

## 5.3 Naïve-Bayes

Naive Bayes is a simple classification technique which works on Bayesian methodology where the probability of occurrence is calculated. In hyperspectral imaging, the data which is to be classified is categorized into classes and attributes. Now using the methodology the attributes of each class is classified and checked for the redundancy of the attributes. Here the redundancy of the attributes can be resolved by classifying into correct classes [4] [5].

$$p(\Theta|X, \mathcal{H}) = \frac{p(X|\Theta, \mathcal{H})p(\Theta|\mathcal{H})}{p(X|\mathcal{H})} \quad (1)$$

## 6. EVALUATION METRICS

### 6.1 Kappa Statistics

Kappa coefficient is a rate which measures the quality of work done; it compares the accuracy of observed and expected outputs. It is considered as a better statistical analyzer for evaluating the ability of a single classifier with the other classifier techniques. Here kappa statistics is denoted by  $\kappa$ .

The equation for  $\kappa$  is

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_e}{1 - p_e} \quad (2)$$

Where  $p_o$  is observed accuracy  
 $p_e$  is expected accuracy

By using observed and expected accuracies, we can calculate values of other categories. Kappa coefficient value varies along with accuracies. The value cannot be negative and in some cases it can be negative by user interpretation. For better-observed accuracy, the Kappa coefficient value can be more than zero and for awful accuracies, the value can be less than zero. Achieving perfect accuracy can be a tough task.

### 6.2 Mean Absolute Error

Mean Absolute Error is the difference between absolute value and its predicted value. These predicted values can be taken after classification of data. By using these predicted values with the help of absolute value, we can classify the data more efficiently with minimum errors.

The notation is given by:

$$MAE = \frac{SAE}{N} = \frac{\sum_{i=1}^N |x_i - \hat{x}_i|}{N} \quad (3)$$

Where  $\{x_i\}$  is the actual observations time series,  $\{\hat{x}_i\}$  is the estimated or forecasted time series, SAE is the sum of the absolute errors (or deviations) and  $N$  is the number of non-missing data points.

### 6.3 Relative Absolute Error

The relative absolute error is the measure which calculates the relative difference of two actual values by their mean value. Mathematically the relative absolute error  $E_i$  can be evaluated by the notation

$$E_i = \frac{\sum_{j=1}^n |P_{(ij)} - T_j|}{\sum_{j=1}^n |T_j - \bar{T}|} \quad (4)$$

### 6.4 Accuracy

Accuracy is the nearest value to the perfect value. The nearest value can be achieved by different methods that are used to perform a task. After the task is completed the result

of that task is compared with the perfect value. For each task, we achieve a value which is closest or nearest to the perfect value and this value is noted as accuracy. A task can be concluded with accuracy if all the parameters and dimensions are achieved according to the perfect value or close to that value.

## 7. RESULTS AND ANALYSIS

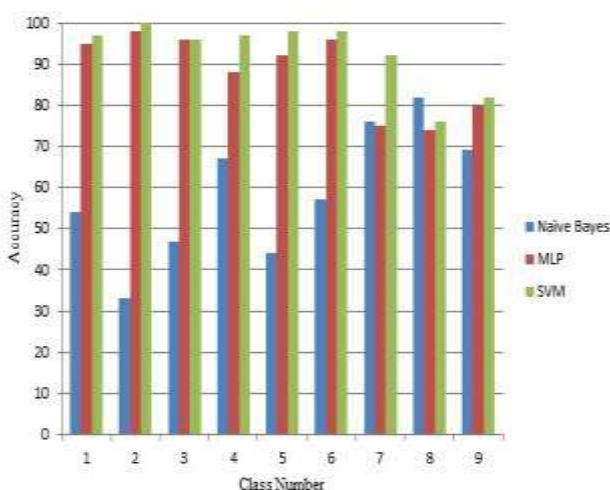
By the results, we can conclude that SVM is more potential and robust classifier than other two classifiers in these particular problems. SVM has classified more instances than the other classifiers because of SVM kernel functions. Other classifiers are better in performance for various problems which are bounded by limitations, where SVM can be applied to different problems and can achieve good consequences.

**Table 1:** Classification accuracy using Single classifier

S.No	Classifier	Accuracy
1	Naive Bayes	80.1
2	Support Vector Machine	92.8
3	Multilayer Perceptron	88.3

**Table 2:** Classification accuracy using Multi Classifier System

S.No	Classifier	Kappa Statistics	Mean absolute error	Relative absolute error	Accuracy
1	Naive Bayes	0.74	0.09	26.62	82.1
2	Support Vector Machine	1.00	0.01	1.00	98.3
3	Multilayer Perceptron	0.95	0.02	7.05	95.2



**Fig 5:** Comparison of classification techniques

## 8. CONCLUSION

The setback in classifying hyperspectral data was addressed and an approach that is based on the fusion of multiple classifiers was proposed. The intended method relies on the decision fusion of individual classifiers trained on a diverse hyperspectral dataset. Finally, we come up with better results from the data where problems like high dimensions and nonlinear classification can be solved. But there are some more problems arise during the process those are fuzzy data and fuzzy classes which can be addressed with the help of neural networks and this study on fuzzy data can be a further opportunistic study.

## REFERENCES

- [1] Fauvel, Mathieu, Jocelyn Chanussot, and Jon AtliBenediktsson. Decision fusion for hyperspectral classification. John Wiley & Sons, New York, NY, USA, 2007.
- [2] C. Cortes and V. Vapnik. Support vector networks. *Machine Learning*, 20:273 – 297, 1995.
- [3] N. Cristianini and J. Shawe-Taylor. *An Introduction to Support Vector Machines*. Cambridge UP, Cambridge, UK, 2000.
- [4] Rish, An empirical study of the naive Bayes classifier, *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*. (Available online: PDF, PostScript)
- [5] Cheeseman, P., & Stutz, J. (1996). Bayesian classification (AutoClass): Theory and results. In *Advances in knowledge discovery and data mining*, 153–180. Menlo Park, CA: AAAI Press.
- [6] Richard, M., & Lippmann, R. (1991). Neural network classifiers estimate Bayesian a posteriori probabilities. *Neural Computation*, 3, 461–483.
- [7] Heermann, P.D.; Khazenie, N. Classification of multispectral remote sensing data using a back propagation neural network. *IEEE Trans. Geosci. Remote Sens.* 1992, 30, 81–88.
- [8] Ceamanos, Xavier, et al. "A classifier ensemble based on fusion of support vector machines for classifying hyperspectral data." *International Journal of Image and Data Fusion 1.4* (2010): 293-307.
- [9] W. Li, S. Prasad, E. W. Tramel, J. E. Fowler, and Q. Du, "Decision Fusion for Hyperspectral Image Classification Based on Minimum-Distance Classifiers in the Wavelet Domain," in *Proceedings of the IEEE China Summit & International Conference on Signal and Information Processing*, Xi'an, China, July 2014, pp. 162-165.
- [10] Roli, F., Giacinto, G., and Vernazza, G., 2001. *Methods for designing multiple classifier systems*. Proceedings of the second international workshop on multiple classifier systems. Cambridge, UK, 78–87.
- [11] M. O. Stitson and J. A. E. Weston. Implementation issues of support vector machines. Technical Report CSD-TR-96-18, Computational Intelligence Group, Royal Holloway, University of London, 1996.

- [12] Bigdeli, Behnaz, Farhad and Peter Reinartz. "Fusion of hyperspectral and LIDAR data using decision template-based fuzzy multiple classifier system." *International Journal of Applied Earth Observation and Geo information* 38 (2015): 309-320.
- [13] Bigdeli, Behnaz and Peter Reinartz. "Feature grouping-based multiple fuzzy classifier system for fusion of hyperspectral and LIDAR data." *Journal of Applied Remote Sensing* 8.1 (2014): 083509-083509.