International Journal of Computer Science Research (IJCSR)

FOREX Publication

Multi-disciplinary | Open Access | ISSN: 2347-47XX

Date of Publication: 10 June 2017 | ID: IJCSR-050204 |Volume 5, Issue 2 | Pages: 16-19

# Analysis of Supervised Image Classification Method for Satellite Images

**P. Sathya**, Ph.d Research Scholar, Department of Computer Science, Bharathidasan University, Tamilnadu, India, p.sathya.karur@gmail.com

**V. Baby Deepa**, Assistant Professor, Department of Computer Science, Bharathidasan University, Tamilnadu, India, deepamct@gmail.com

**ABSTRACT:** Remote sensing is the acquisition of information about an object without creating physical contact with the object and thus in contrast to on-site observation. Remote sensing is used in many fields, including geography, land surveying and most Earth Science disciplines and it also has military, intelligence, commercial, economic, planning, and humanitarian applications. Remote sensing analyses cover way for satellite image classification which facilitates the image interpretation of large amount of data. Satellite Images covers large geographical spread and results in the utilization of more data which includes classifying into different division. More than one classification algorithms are already available for satellite image classification, but with the broad range of applications an algorithm with better Performance and accuracy is needed. In this paper to analyses four types of supervised classification such as parallelepiped, Minimum distance, Maximum likelihood and k-Nearest neighbor and provide a comparative study on their efficiency.

**Keywords:** Classification, Supervised classifiers, K-nearest, Training sites, Satellite image.

## 1. INTRODUCTION

In 1960 the remote sensing technique was introduced to analysis and interpretation for data. Remote sensing gathers large number of satellite images. Remote sensing refers to the activities of observing objects at faraway places. In this technology the sensors are not in direct contact with the objects being observed. The information needs a physical carrier to travel from an object to the sensors through an intervening medium. Remote sensing normally refers to the method of acquiring information about the earth surface. Images of remote sensing are actually in the form of digital images. To extract useful information from the images, image processing techniques may be hire to enhance the image to help visual interpretation, and to restore the image if the image has been divided to geometric blurring, degradation or distortion by other factors. Based on the requirements of the specific problem many image analysis methods and techniques are available.

Image classification techniques group of pixels used to represent land cover features. Land cover may be urban, forest, agricultural and other types of features. Three main image classification techniques are supervised, unsupervised and object based image classification.

Different classification techniques are used for data extraction from remote sensing images. Satellite image classification technique is the most useful technique for image information extraction and interpretation. Satellite image classification is used to divide the pixels of the image into number of different defined sub classes. The pixels are grouped together based on the digital values extracted from the satellite images. The pixel values are extracted or taken from the satellite images; it could be a single value as in case of multi-modal image. The classification is used to extracting the information contained in different bands of the satellite sensor and the information is extracted in terms of digital numbers which is then converted to a category.

In supervised classification the image processing software is directed by the analyst, the information is passed by analyst to specify the land cover classes of interest. In supervised image classification required input from an analyst; it is called as training set. In unsupervised image classification does not provide training set. The analyst indicates how many classes want to generate and which bands to use. Clustering technique mechanism is used in unsupervised classification to group of satellite image pixels into unlabelled classes. Then assign meaningful labels to the classes by analyst and provide well classified satellite image. This paper provides a comparative study on the accuracy of different supervised classification algorithms and techniques.

.

## 2. SUPERVISED CLASSIFICATION

In supervised classification the image processing software is directed by the analyst, the information is passed by analyst to specify the land cover classes of interest. In supervised image classification required input from an analyst; it is called as training set. In this method one of the most important factors is training samples taken for training. Accuracy of the classification only depends on the training sample. Training samples are two types, first one is used for classification and second one is used for supervising classification accuracy. In this method training is required, so this technique become time consuming. So supervised classification is not suitable for some areas. In order to determine a decision rule for classification, it is necessary to know the spectral characteristics or feature with respect to the population of each class.

Date of Publication: 10 June 2017 | ID: IJCSR-050204 |Volume 5, Issue 2 | Pages: 16-19

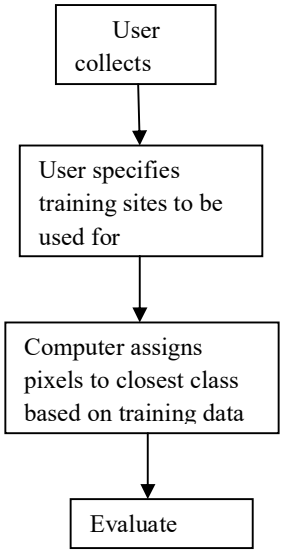


Figure 1: Process of supervised classification

### 2.1 Advantages

- Generates information classes
- Self-assessment using training sites
- Training sites are reusable

### 2.2 Disadvantages

- Information classes may not match spectral classes
- Signature homogeneity of information classes varies
- Signature uniformity of a class may vary
- Difficulty and cost of selecting training sites
- Training sites may not encompass unique spectral classes

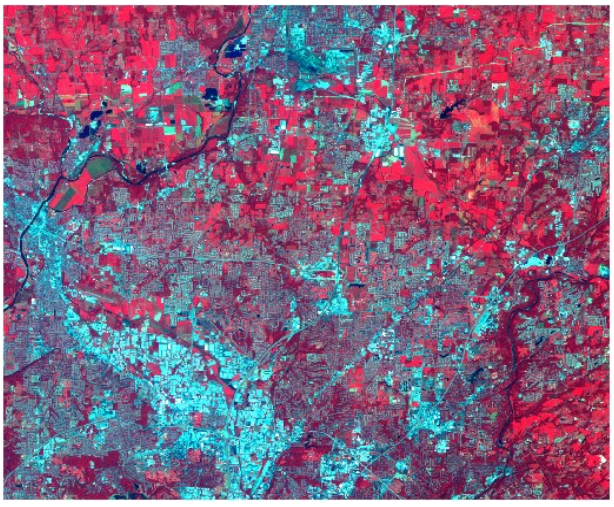

Figure 2: Input four-band Landsat TM image

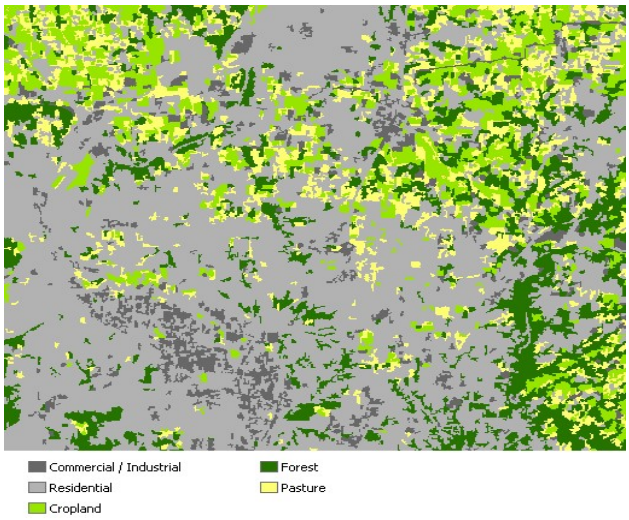


Figure 3: Output supervised classified land use map

## 3. SUPERVISED CLASSIFICATION ALGORITHMS

### 3.1 Parallelepiped classification

Parallelepiped classification is a simple classification method based on a decision rule for multispectral data. Decision boundaries for the parallelepiped algorithm are formed based on a standard deviation threshold which is chosen from the mean of each selected class in the training set. The decision boundaries form an interval between two pixel values with a hyper rectangle region in feature space. A pixel is classified based on whether the value of that pixel lies above the lower threshold value and below the higher threshold value of the interval.

The mean value Mt of all the pixels for a class C for band M is taken for all the N classes of the training set and the variation (standard deviation) of the training data class C of band M of all the N classes be St. The mean and the standard deviation forms the parallelepiped boxes as decision boundaries or intervals for assigning the pixels .If the digital value Dv of the pixel lie inside the parallelepiped decision boundaries means a pixel will be assigned to a particular class.

$$Mt - St \leq Dv < Mt + St$$

Mt – Mean value
Dv – Digital value
St – Standard deviation

The pixel value lies in between the lower and the upper threshold value means the pixel will be assigned to the class. Parallelepiped classification is easy to understand and very simple. In this method the computing time will be a minimum to compare with other classification.

FOREX Publication
Multi-disciplinary | Open Access | ISSN: 2347-47XX

Date of Publication: 10 June 2017 | ID: IJCSR-050204 |Volume 5, Issue 2 | Pages: 16-19

### 3.2 Maximum Likelihood Classification

Maximum likelihood classification calculates the probability to a given pixel to each class and then the pixel will be allocated to a particular class with the highest probability. It is used to calculate the mean and covariance matrix to the training sets and assumes that the pixel values are normally distributed. The class can be characterized by the mean value and covariance matrix. A probability density function is defined and the input pixels are mapped based on the likelihood that the pixel belongs to that particular class. The likelihood Expressed for normal distribution can be calculated by

$$Lk(X) = \frac{1}{2\pi^{\frac{n}{2}}|\sum k|^{\frac{1}{2}}} \text{Exp-}\frac{1}{2}(X - Mk)\sum{}^{-1}{}_{k}(X - Mk)^{t}$$

X indicates the image data of n bands
Lk (X) represents the likelihood of X belonging to class k
Mk is the mean vector of class k
K is the variance covariance matrix of class k

This classifier is a sophisticated classifier with good separation of classes, but the training set should be strong to sufficiently describe mean and covariance structure. Also the algorithm is computationally intensive.

### 3.3 Minimum Distance Classifier

This method is a simple supervised classifier which uses the centre point to represent a class in training set. This technique uses the distance measure, where the Euclidean distance is considered between the pixel values and the centroid value of the sample class. The pixel with the shortest distance with the class is assigned with that class. The minimum distance classifier is fast in execution and computation time is minimum as it depends on the training dataset and all pixels will be classified, but the algorithm may be prone to errors resulting in misclassification of pixels as it will classify a pixel even if the shortest v distance is far away. Spectral distance is calculated for all values of a class mean, the unclassified pixel is assigned to the class with the lowest spectral distance resulting in classification of all pixels.

The minimum distance classifier algorithm is based on the minimum distance from the mean value Mt of each class of the training data to the digital value Dv of each pixel in the imagery. The minimum distance is calculated by using the Euclidean distance measurement.

$$\text{Sqrt}(Dv - Mt)^2$$

The class mean with the minimum distance with the pixel will be assigned as the class of the pixel.

### 3.4 K-nearest neighbor algorithm

This method is a non-parametric mining technique which uses K nearest training samples to determine the pixel class. Here the K samples are chosen based on a similarity measure. Commonly used similarity measure is the distance function. The classification uses Euclidean, Manhattan or other distance measure to calculate the distance between a pixel and the different training samples. A class I assigned to a pixel based on the majority voting from the K training samples which would be to assign the most common class among the training samples. This method is simple to process, but computationally expensive to select K nearest neighbors when the training dataset is large.

Euclidean distance = $\sqrt{(Dv - Mt)^2}$
Manhattan distance = $|Dv - Mt|$

Mt measures the digital value of each training sample and Dv represents the digital value of each pixel in the imagery

**Table 1: Comparison between supervised classification algorithms**

| ALGORITHMS | ADVANTAGES | DISADVANTAGES |
|---|---|---|
| Parallelepiped | 1.Computationally efficient<br>2.Fast execution | 1.Pixels not classified<br>2.Pixels in several classes |
| Maximum likelihood | 1.Provide good separation from classes | 1.Requires well trained training set<br>2.Computationally intensive |
| Minimum Distance | 1.Fast execution<br>2.All pixels are classified | 1.Prone to commission errors |
| K-nearest neighbour | 1.Low cost<br>2.Effort for learning process | 1.Performance depends on the number of dimensions<br>2.Computationally expensive to find the K neighbours when sample dataset is large |

.

## 4. CONCLUSION

This paper analyses different supervised classification approaches and methods. Satellite image classification is a field which has great significance for different socio-economic, environmental applications. Through classification of satellite imagery, the information as cadastral information, land cover type, vegetation type, soil properties could be obtained. Different methods are discussed and provide advantages and disadvantages of supervised classification algorithms. In future work we will overcome the disadvantages of the supervised algorithms.

## REFERENCES

[1] Jipsa Kurian, V. Karunakaran, "A survey on image classification methods", Department of Computer Science, Karunya University, Coimbatore, India, IJARECE, Volume 1, Issue.4, Oct 2012.

Date of Publication: 10 June 2017 | ID: IJCSR-050204 |Volume 5, Issue 2 | Pages: 16-19

[2] Wei-jiu Zhang, Li Mao and Wen-boXu. "Automatic image classification using the classification ant- colony algorithm", Jiangnan University, Wuxi, China, 2009 International Conference on Environmental Science and Information Application Technology.

[3] D. Lu and Q. Wend, "A survey of image classification methods and technology for improving classification performance", Department of Geography, Geology, and Anthropology, Indiana State University, USA, IJRS, Vol.28, No.5, 10 March 2007.

[4] W. G. Rees, Physical Principles of Remote Sensing, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[5] P. Mather and B. Tso, Classification Methods for Remotely Sensed Data,2nd ed. Boca Raton, FL, USA: CRC Press, 2009.

[6] Sunitha Abburu , Suresh BabuGolla Satellite Image Classification Methods and Techniques: A Review, International journal of computer applications, Volume 119 No.8, June 2015.

[7] [D. Liu, Q. Weng;"Survey of image classification methods and techniques for improving classification performance",InternationalJournal of Remote Sensing (Vol. 28, No. 5,
March 2007).

[8] G.D. Sunayana, Z. Tanish; Comparative Analysis of Unsupervised and Supervised Image Classification Techniques, National Conference on Recent Trends in Engineering & Technology.