Hybrid Algorithm for Classifying Soil Types Using Hyperspectral Images

P. Bhargavi[1], A. Rajitha[2], S. Jyothi[3]

[1]Assistant professor, [2]Research Scholar, [3]Professor
[1, 2, 3] Department of Computer Science
Sri Padmavati Mahila Visvavidyalayam Tirupati-517 502

Abstract

In field survey, soil type is the key indicator, but these soil classifications are depending on the personal experience. To classify types of soil physically is difficult so hyperspectral image is used to classify soil types quickly and accurately. These hyperspectral image is classified with texture features and spectral information. In this paper types of soils are classified by applying maximum likelihood and K-Nearest Neighbour algorithms to hyperspectral image. Also a Hybrid algorithm which is a combination of Maximum Likelihood and K-Nearest Neighbour was proposed to classify the soil types and applied on HSI. The outcome indicates that hybrid algorithm is the most efficient technology for classification of various types of soils.

Keywords: Hyperspectal Image, Maximum Likelihood, K-Nearest Neighbour,Soil Classification, Hybrid Algorithm.

Introduction

For sustainable utilization of land and soil management soil classification is playing a key role. Soils [13] have different physical and environmental compositions that lead to different soil properties. In present days for distinguish soil categories soil colors are identified to classify soil types [7]. This classification has small sections to acquire large data with modern digital technology. These methods are depending on personal experience that causes errors easily. To overcome these errors remote sensing based Hyperspectral Image (HSI) are considered to classify the types of soils based on soil colors [4]. These HSI based classification also leads to identify different fields like agriculture, food, medical science etc.

To classify HSI, image texture and spectral information are used to obtain the essential information from the image. The image texture is characterized with intense of neighbour pixels to classify the soil types [8]. To our knowledge usages of comprehensive features of image texture and spectral information are reported as seldom for classification of soil types. These HSI generates immense data to contribute relevant information which are classified using data cleaning and data integration models for reducing the noise and dimensionality [5]. The main aim for selecting the HIS [12] is to carry the useful information for classifying the soil types with reduced amount of data, robust with simple model and computational tasks [11].

The unbiased work is to investigate the feasibility of classifying soil types using HSI with maximum likelihood, K-nearest algorithm and proposed hybrid algorithm which is a combination of Maximum likelihood and K-nearest algorithm.

Hyperspectral image

From 1980 on words hyperspectral image is for identifying the minerals, soil tropical, rock types, mapping etc. These HSI is a composition of camera and spectrograph images that are fixed in Hyperion sended to spacecraft for monitoring the earth surface. This hyperion captures the image continuously with several hundred spectral bands of 256 and a

range of 874-1734 wavelengths. Spectroscopic technologies are used for identifying the objects based on their characteristics in Hyperspectral image. The HSI technology is most useful spectral resolution for classifying the soil types. [1], [2], [3].

Methodology

The districts of Chittoor and Tirupati regions were chosen for this study in classification of soil types. Four different hyperspectral image regions were used in the Chittoor and Tirupati districts to classify the types of soils since each region may include only one or two types of soil it is impossible to find all the different types of soil with just one hyperspectral picture. In this work, soft computing techniques like maximum likelihood and K-nearest neighbour algorithm and proposed hybrid algorithm are applied for classification of soil types.

Maximum Likelihood Algorithm (ML)

This algorithm calculates each bands in each classes with probability that belongs to specific class using statistics and probability [9]. The threshold probability is selected for classifying the pixels [14]. The classes that are assigned to each pixel has high probability. This high probability is smaller than a threshold in image pixels remain unclassified. The formula for discriminant function is represented as

$$g_i(x) = inp(\omega_i) - {}^{1}/_{2}\ in|\Sigma_i| - {}^{1}/_{2}\,(x - m_i)^r \sum_{i}^{-1}(x - m_i)$$

Where I represens class, x represents number of bands in dimension, $p(\omega_i)$ represents probabilty of occuring class in image, $|\Sigma_i|$ represents matrix covariance of class, ${\Sigma_i}^{-1}$ represents matrix inverse, m_i represents mean vector.

K- Nearest Neighbour (KNN) Algorithm

It is a non-parametric method widely used for classifying the pattern recognition [10]. The k neighbours take T= $\{(x_1, y_1), (x_2,y_2), \ldots\ldots (x_N, y_N)$ as a training set, the training entities represents as N, the $x_i \epsilon R^d$ represents vector features and $y_i \in y = \{c_1, c_2, \ldots., c_m\}$ represents classification labels i=1,2,…N, x is input, $N_K(x)$ training entities distance represents the classification of hyperspectral image formulated as

$$\mathbf{X_I} = \dddot{\mathbf{min}} \|\mathbf{Y_I} - \mathbf{AX_I}\|\mathbf{F^2}, \mathbf{s.t} \|\mathbf{X_I}\|_{\mathbf{row,0}} \leq \mathbf{T_0}$$

The HSI data the pixel 1 tends to share the characteristics it depicted as $Y_i = AX_i$ where Y_i represents neighbour set, A represents training set and X_i represents matrix coefficient.

The matrix coefficient defines the distance from Y_i with class m is formulated as

$$d(Y_i; A^m) = \left\| Y_i - A^m \overline{X_i^m} \right\|_F^2$$

Here A^m represents the class m in the image, the pixel 1 are label smallest distance.

Hybrid method

In order to classify soil types a hybrid algorithm is proposed which is a combination of maximum likelihood and K-NN. For that distance between a given feature vector and all the samples at each time are classified to suffer with a high complexity in terms of time and space. This is a reasonable approach in terms of complexity estimation with minimum and maximum normalization for classification.

$$M^1 = \frac{M - minimum\ (A)}{maximum\ (A) - minimum\ (A)}$$

Where M represents the attribute A in the true value, M^1 represents attributes normalized value, minimum (A) represents

smallest value, maximum (A) is largest value in attribute A.

Experimental results

Hyperspectral images are collected from Chittoor and Tirupati districts shape files by linking python programing with QGIS. Here in this work four different Landsat images of each district are collected for the classification of different types of soils.

The collected hyperspectral images are pre-processed using data cleaning technique that removes cloud covering in the satellite image, data integration, data transformation and data reduction are also applied for pre-processing. Then the images are segmented into cubes by taking region of interest (ROI) which is measured in the form of polygon, circle and rectangle. Next segmented images are trained and the maximum likelihood algorithm, K-Nearest Neighbour algorithm and proposed hybrid algorithm which is a combination of maximum likelihood and K-Nearest Neighbour algorithm are applied for different soil type classification.

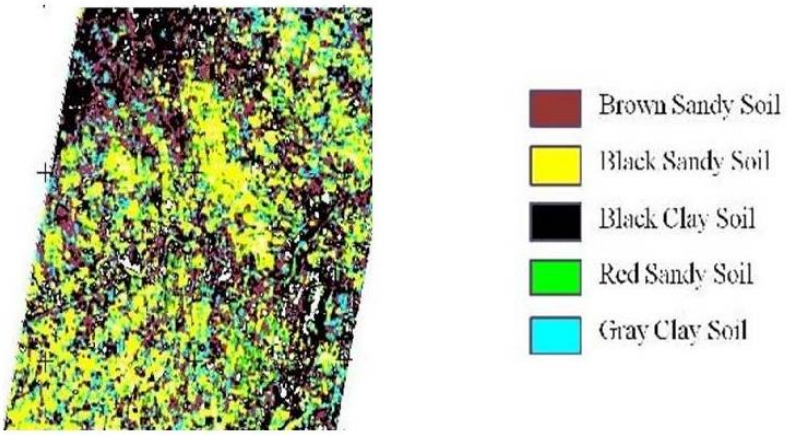
The soil type classified hyperspectral images of four chittoor district regions are show in figure 1 as figure 1 (a), figure 1 (b), figure 1 (c), figure 1 (d). Here each figure describes the classified areas that are represented in different colors and each color classification represents different area of that region like:

In the figure 1 (a), brown colour represents brown sandy soil, yellow color indicates black sandy soil, black color represents black clay soil, green color represents red sandy soil and sky blue represents gray clay soil.

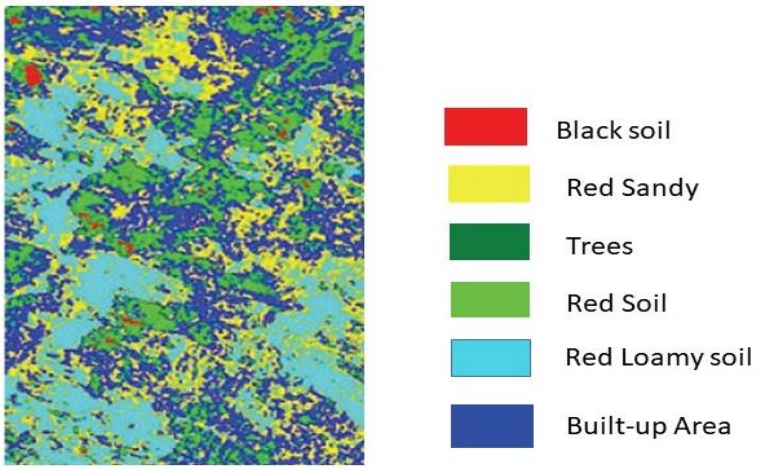
In the Figure 1 (b) red color represents black soils, yellow color represents red sandy soil, dark green represents trees, light green represents red soils, sky blue color represents red loamy soil, blue represents built-up area.

In the figure 1(c) red color represents red clay soil, dark green represents red sandy soil, light green represents black clay soil, yellow color represents forest area with red soil, blue represents built-up area.

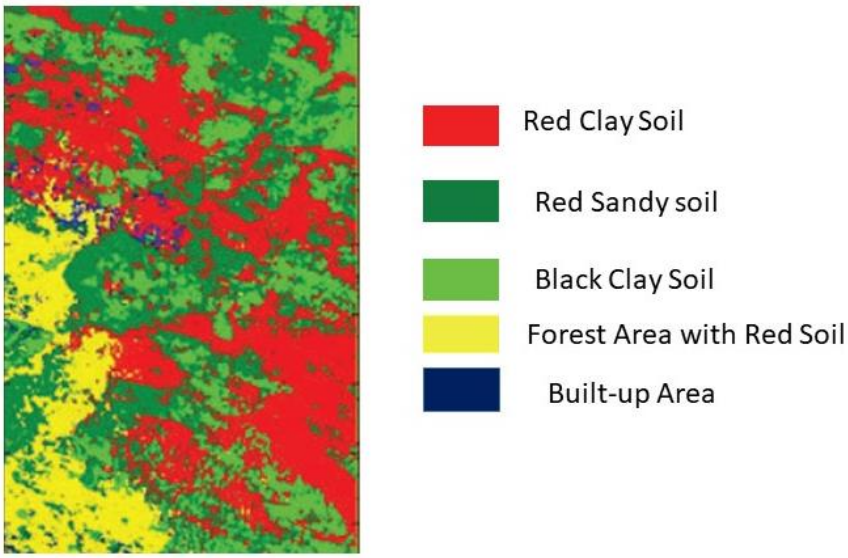
In the Figure 1 (d) shows the classified Chittoor district hyperspectral image and figure describes pink color represents red soil, blue colour indicates water Bodies, green colour represents red sandy soil, sky blue represents grey clay soil and black represents unclassified area.



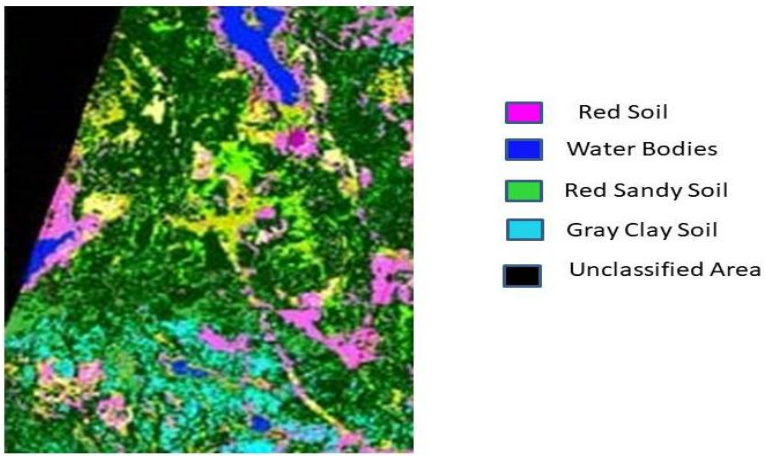
(a)



(b)



(c)



(d)

Figure 1: Chittoor District Regions Hyperspectral Image Classification using ML, K-NN and Hybrid algorithm

The soil type classified hyperspectral images of four Tirupati district regions are show in figure 2 as figure 2 (a), figure 2 (b), figure 2 (c), figure

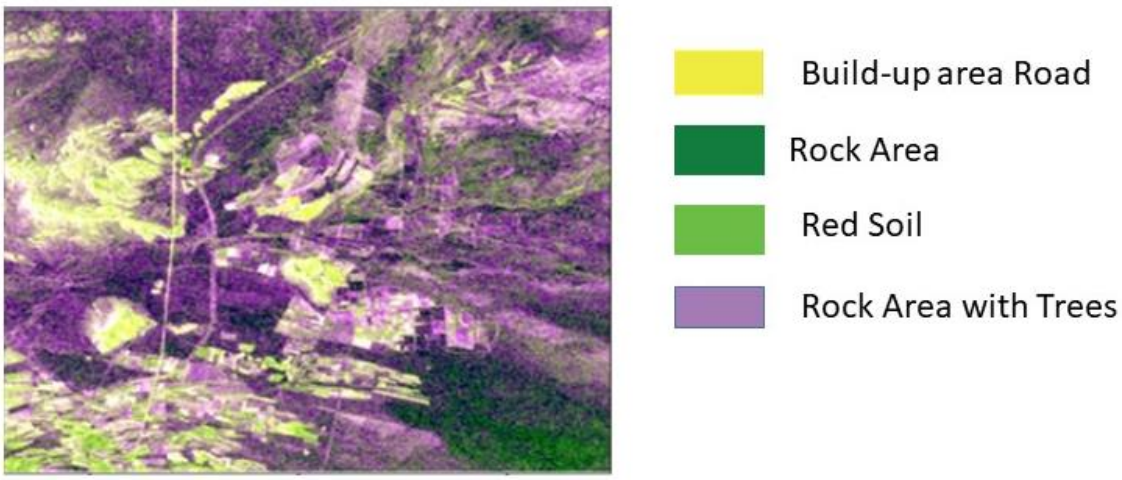
2 (d). Here each figure describes the classified areas that are represented in different colors and each color classification represents different area of that region like:

In the figure 2 (a) yellow color represents built-up area and road, dark green color represents Rock area, light green color represents red soils, violet color represents rock area with trees.

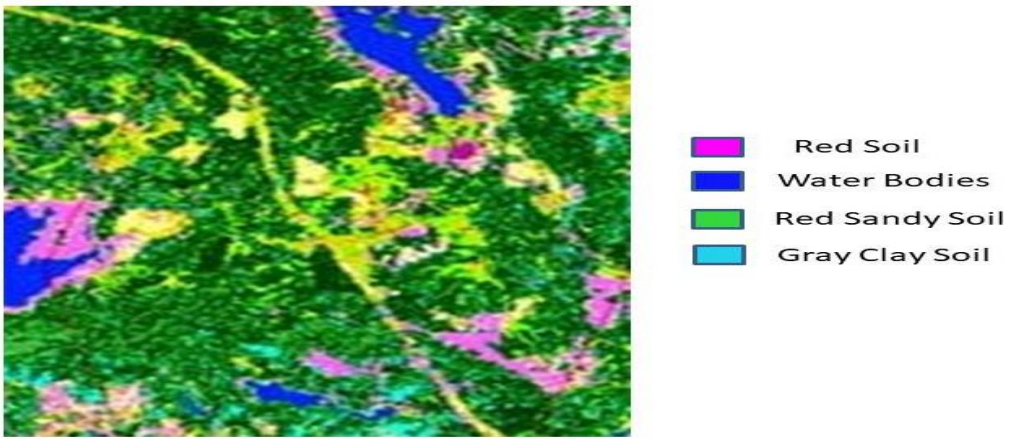
In the figure 2 (b) pink colour represents red soil, blue colour indicates water Bodies, green colour represents red sandy soil and sky blue represents gray clay soil.

In the figure 2 (c) red color represents red clay soil, yellow color represents forest area with red soil, dark green color represents red sandy soil, light green color represents red loamy soil, blue represents built-up area.

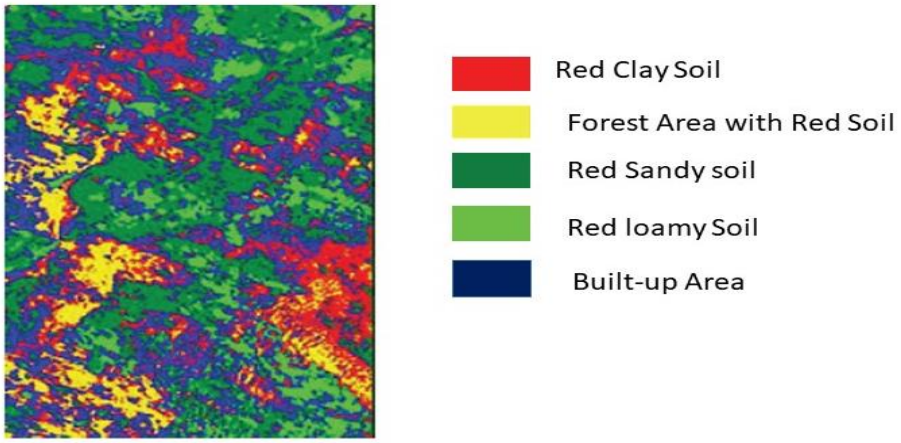
In the figure 2 (d) red color represents red clay soil, dark green color represents red sandy soil, blue represents built-up area, yellow color represents forest area with red soil, light green color represents red loamy soil.



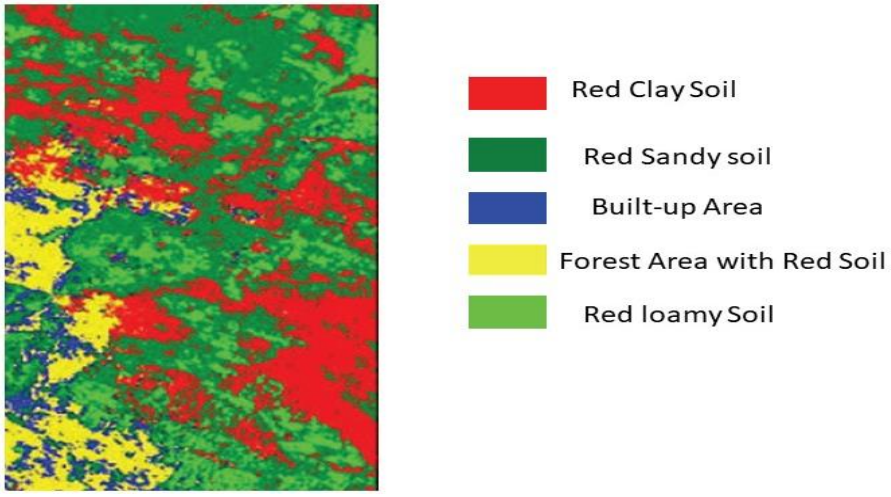
(a)



(b)



(c)



(d)

Figure 2: Tirupati District Regions Hyperspectral Image Classification using ML, K-NN and Hybrid algorithm

The classification accuracies of maximum likelihood algorithm and K-Nearest Neighbour algorithm and proposed hybrid algorithm are shown in table 1. By observing the accuracies hybrid algorithm is the best algorithm with 98.9% for classification of various types of soils in Chittoor and Tirupati regions soil.

Table 1: Comparison of accuracies

S.No.	Model	Accuracies (%)
1.	Maximum likelihood algorithm	92.5
2.	K-Nearest Neighbour	96.8
3.	Hybrid algorithm	98.9

By observing all soil type classification most of the chittoor and tirupati district regions are covered with red soils which are helpful to cultivate the crops.

Conclusion

The soil type classification by collecting site soil samples can be difficult or costly. So by using high resolution and cost effective hyperspectral image is picked to classify the different soil types. Here chittoor and tirupati district regions are picked to know what are the types of soils are present near and around the region. To classify soil type's maximum likelihood and k-nearest neighbour algorithms are applied and then hybrid algorithm is proposed to classify soil types. By observing the accuracies of the algorithm hybrid algorithm is the best to classify soil types. And observing the soil classified hyperspectral image most of the chittoor and tirupati district regions are covered with red soils which are helpful to cultivate the crops like sugarcane, tobacco, wheat, millets.

Reference

[1] Hartemink A E and J Bockheim G 2013 Catena 104 251-256

[2] Viscarra Rossel R A, Minasny B, Roudier P and McBratney A B 2006 Geoderma 133 320-337.

[3] Meyer R and Kirkland A 1998 Ultramicroscopy 75 23-33.

[4] Zhang C, Guo C, Liu F, Kong W, He Y and Lou B 2016 J. Food Eng. 179 11-18.

[5] Gomez C, Gholizadeh A, Borůvka L and Lagacherie P 2016 Geoderma 276 84-92

[6] Neittaanmaki-Perttu N, Gronroos M, Tani T, Polonen I, Ranki A, Saksela O and Snellman E 2013 Laser Surg. Med 45 410-7.

[7] E. Ben-Dor, S. Chabrillat, J. A. M. Demattê et al., (2009), "Using imaging spectroscopy to study soil properties," *Remote Sensing of Environment*, vol. 113, no. 1, pp. 538–555.

[8] G. W. McCarty, J. B. Reeves, V. B. Reeves, R. F. Follett, and J. M. Kimble, (2002), "Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement," *Soil Science Society of America Journal*, vol. 66, no. 2, pp. 640–646.

[9] Shujun Liang, Jing Cheng, Jianwei Zhang, (2020**)** **"**Maximum likelihood classification of soil remote sensing image based on deep learning, Earth Sci. Res. J. vol.24 no.3.

[10] Mrs. N. Saranya, Ms. A. Mythili,(2020), "Classification of Soil and Crop Suggestion using Machine Learning Techniques", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181Vol. 9 Issue 02.

[11] M. Sirish Kumar, B.Kavitha, S.Jyothi [2018]. "Agricultural Land Classification Based on Hyperspectral Images", Journal of Applied Science and Computation, Vol. 5, Issue 7, July, Pages: 454-460.

[12] T. sarath, G. Nagalakshmi, S. Jyothi (2014), A Study on Hyperspectral Remote Sensing Classifications, International Journal of Computer Applications (0975 – 8887).

[13] P.Bhargavi , S. Jyothi. (2011) "Soil Classification Using Data Mining Techniques: A Comparative Study", International Journal of Engineering Trends and Technology, ISSN: 2231-5381 Page55.

[14] T. sarath, G. Nagalakshmi "An Land Cover Fuzzy Logic Classification by Maximum Likelihood" in International Journal for Computer Trends and Technology Vol. 13. Issue 2, July 2014.