Soil Classification Using Modified Support Vector Machine

Vijay E V, Navya Ch, Abdul Shabana Begum, Rajaneesh D, Mahesh Babu B

Abstract— Soil classification is a process used to identify the category of soil. Manual soil classification is a time taking procedure and hence on site soil classifications required in many engineering applications. Primary or basic idea is required by on-site engineers regarding the type and structure of soil. In this paper, the conventional techniques used for soil classification are studied and an image processing soil classification technique using an efficient classifier has been developed and tested. Here seven classes of soil were considered for classification, namely Clay, Clayey Peat, Clayey Sand, Humus Clay, Peat, Sandy Clay and Silty Sand. These seven categories are classified using Modified Support Vector Machine. The developed model can be used in real time soil classification applications.

Index Terms— Soil Classification, Image Processing, Modified Support Vector Machine (SVM).

I. INTRODUCTION

Soils are categorized into different groups based on the place of soil availability and based on the dominating size of particle in the soil. Based on location, soil is categorized into red soil, alluvial soil, black soil, desert soil, peaty soil, forest soil and laterite soil, etc. The various types of soils based on dominating particle size are clay, peat and sand. In the same way some soils are classified as a mixture of two soils such as Clayey Peat, Clayey Sand, Humus Clay, Sandy Clay and Silty Sand. Figure 1 shows these seven categories of soils.

The appropriate conditions are to be discovered by an engineer initially before undergoing any engineering process. If this strategy is not maintained then additional expenditure occurs in further processing.

Fig.1: Soils based on dominating particle size

The site needs to be explored by involving in necessary investigations which include in-situ techniques and laboratory experiments. Exploring the soil properties at the ground level or below the surface is an in-situ technique. In surface in-situ investigation, the soil profile is provided by geological mapping while the in-situ density of soil is found by density replacement test.
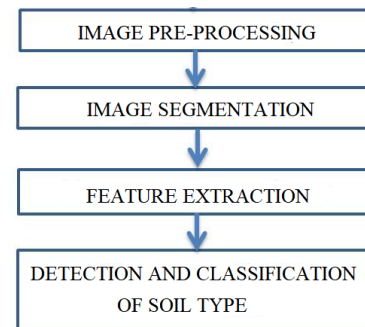

Fig.2: Basic steps for classificationn of an image

The steps involved before classifying a soil image are mentioned in the figure 2. In image pre-processing in soil images are collected and processed, i.e., the contrast of the image is enhanced. The pre-processed image is subjected to image segmentation followed by feature extraction. After extraction of features from the image, classification of the query image is performed to specify soil. This is performed by using Support Vector Machine.

International Journal of Research in Advent Technology, Vol.8, No.9, September 2020
E-ISSN: 2321-9637
Available online at www.ijrat.org

II. LITERATURE SURVEY

Support Vector Machine (SVM) is the most useful technology in machine learning for and computer vision and pattern recognition which refers to statically theory of handle the data efficiently. SVM is mostly applicable to problems like classification of pattern and using non parametric classifiers with the approach of binary classifier for nonlinear regressions. SVM applications are mostly found in classification of medical images, identification, and classification of soil etc.

In the paper [1] different algorithms and filters to obtain and process the colour images of the soil samples are developed.The different features like color, texture, etc. are extracted by this algorithm. Here various soils like red, black, clay, alluvial, etc. are considered.

In [2] the author explained about the method used for feature extraction, i.e., HSV Histogram. This paper explains that histogram conversion plays an important role for extracting features of the image and also explains about the generation and implementation of histogram.

In [3] the author used Support Vector Machines in the estimation of soil properties. Soil compose order in light of known estimations of specific compound and physical properties is presented. It was observed that the achieved results propose that straight techniques are not ready to assess the estimations of physical properties utilizing the officially estimated compound properties.

The paper[4], gives a review on the machine learning algorithms in both supervised and unsupervised learning such as Support Vector Machine(SVM), Decision Trees(DT), K-means clustering (KNN) etc. This also describes about Artificial Neural Networks (ANN), Convolution Neural Networks (CNN) etc in their paper.

The author in [6], proposed a feature extraction method in which the soils considered are Alluvial, Silt soil, Clay soil, Red soil, Sea soil, River Soil. This process consists of three steps mainly. Those are
1. The original image is subjected to transformation
2. The features like color, texture and shape on both original image and transformed image using various statistical measurements.
3. The difference is to be found and sorted with the help of Euclidean formula.
4. To find the accuracy

In [8] the author presented a comparison between soil classification based on the Cone Penetration Test (CPT) and the actual soil classification obtained through direct boring and laboratory test results. Soil investigation results are presented in this paper. It is observed from this paper that theoretical approaches does not give satisfactory results.

III. CONVENTIONAL TECHNIQUES USED FOR SOIL CLASSIFICATION

The conventional techniques used for soil classification include Standard Penetration Test, Cone Penetration Test, Pressure Meter Test, and Vane Shear Test, Constraint Clustering and Classification. The test results obtained are used to characterize the properties of soil in test location.

- *Standard Penetration Test (SPT):* This form of testing is introduced from the need to acquire data from subsurface soils. This test was basically designed to determine relative density of cohesion less soils. Standard Penetration Test involves the use of components like percussion drilling rig, hammer and drill rods.
- *Cone Penetration Test (CPT):* This is one of the in-situ geotechnical investigation technique for soil classification. In Cone Penetration Test, we imply the continuous penetration of rods supported by a cone at a constant speed, while some geotechnical parameters (such as tip resistance, sleeve friction, pore water pressure, etc.) are measured after certain distance interval of penetration.
- *Pressure Meter Test (PMT):* A pressure meter is a cylindrical probe that has an expandable flexible membrane designed to apply a uniform pressure to the walls of a borehole. The pressure meter test (PMT) is widely used in weak rocks as an in-situ testing method to determine the stress and deformation relationship of a geo-material.
- *Vane Shear Test (VST):* This is an economical and fast method to determine the shear strength of soft to medium stiff clay. In this technique, four-bladed vane is slowly pushed and constantly rotated into a clay stratum to measure the resisting torque. There are chances of error in Vane Shear Test because of excessive rod friction and poor torque calibrations.
- *Constraint Clustering and Classification (CONCC):* The CONCC is an algorithm is implemented in two steps, segmentation and classification. The Cone Penetration Test data is converted into "J" segments from a single series of data, and then this data is segmented into classes using a fuzzy logic to address the imprecision in the measured data.

IV. PROPOSED METHODOLOGY

This method has two parts. First part is training part in which includes database formation. Second one is testing part where actual classification takes place.
1. *Training Part:* This part is to collect different types of soil sample pictures and then perform pre-processing on that soil images and develop an algorithm to extract the features of soil sample images using features extraction methods like Hsv-histogram, Color moments-Mean and Standard Deviation, Gabor features, Colour auto correlogram. Then develop a database with feature preparation of the sample images. Develop an SVM algorithm for deciding class.
2. *Testing Part:* Select a soil sample image for testing purpose which undergoes processes like pre-processing, feature extraction, feature preparation. Develop an SVM algorithm to find class of soil. Implementation of algorithm is performed by using MATLAB. The block diagram shows the entire steps involved in the system. It includes the following steps:
➢ Soil image
➢ Pre-processing
➢ Feature Extraction
➢ Feature Preparation
➢ Modified Support Vector Machine

International Journal of Research in Advent Technology, Vol.8, No.9, September 2020
E-ISSN: 2321-9637
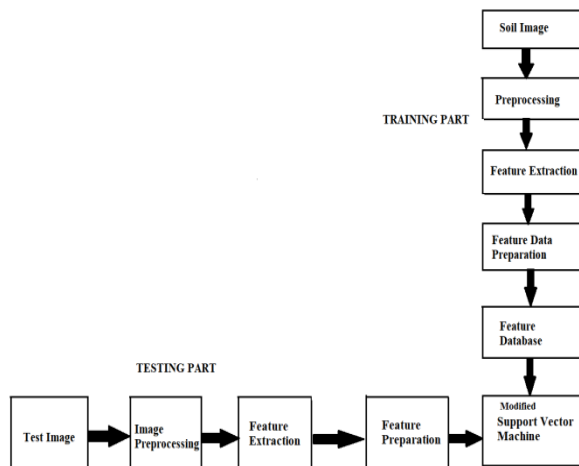*Available online at www.ijrat.org*

Fig. 3: Block Diagram of proposed System

The following is procedure followed in training and testing-
- *Soil Image :* Different types of soil samples are available. For example Clay, Clayey Peat, Clayey Sand, Humus Clay, Peat, Sandy Clay, Silty Sand
- *Pre-processing :* In this section enhancement like processes carried out which will help for feature extraction process.
- *Enhancement:* This is the process of adjusting images so that the images are more suitable for display for further image analysis. Image enhancement process is to remove noise, sharpen or brighten an image and making the identification key features easy.
- *Feature Extraction:* In this section features extraction methods, like hsv-histogram, color moments-mean, standard deviation, gabor features, colour auto correlogram are used to extract features of soil.
- *HSV-Histogram:* The portrayal of the distribution of colors in an image is known as a color histogram. A color histogram denotes the number of pixels that constitute colors. HSV stands for Hue-Saturation-Value. The RGB (Red-Green-Blue) image is converted to HSV image for feature extraction. Hue is the actual image value; saturation is the amount of grey in the image and value means the amount of brightness.
- *Color moment:* The color distribution in an image is characterized by color moments. To compare how similar two images are based on color, one image is compared to a database of digital images to find and retrieve a similar image. These comparisons between images results in a score of similarity. If the score is lower the two images are supposed to be more identical.
  i.Mean: The first color moment is given as the average color in the image.

$$E_i = \sum_{j=1}^{r} \frac{1}{P} q_{ij}$$

  Here r represents the number of pixels in the image and q_i represents the value of the jth pixel of the image at the ith color channel.

ii. Standard Deviation: It is the second color moment which is identified by taking the square root of the variance of the color distribution.

$$\sigma_i = \sqrt{\left(\frac{1}{P} \sum_{j=1}^{r} (q_{ij} - F_i)^2\right)}$$

Ei represents mean value i^{th} color channel of image
- *Gabor Filter :* Gabor filter is a linear filter. It used for edge detection and texture analysis. Gabor filters have frequency and orientation representations similar to those of the human visual system. A set of Gabor filters with different frequencies and orientations can be used in extraction of useful features from an image. Modified Two-dimensional Gabor filters were used for analysis purpose. Modified 2-D Gabor filters have many applications in image processing, mainly in feature extraction for texture analysis and image segmentation. F is the frequency. θ is varied for texture orientation in a particular direction. σ is varied to change the size of the image in the region being analyzed.
- *Colour auto correlation:* Color correlogram is used for indexing and comparison of images. The highlights of this feature are:
  i.Includes spatial correlation of colors
  ii.Describes the global distribution of local spatial correlation of colors
  iii. Easy to compute and
  iv. Feature size is small
- *Feature Data Preparation and Feature Database:* In previous section feature extraction carried out and then different features separation carried out called as Feature Data Preparation. And later with the help of previous two blocks feature database formation is carried out.
- *Modified Support Vector Machine:* For classification of images we have used Modified support vector machine. In machine learning, support vector machines/support vector networks are supervised learning models associated with learning algorithms that analyze data used for image classification.
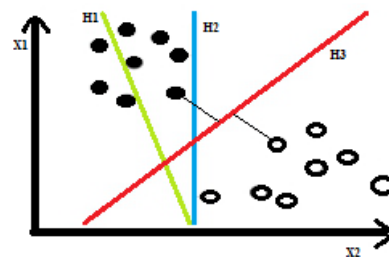-

Fig.4: Example of separation of classes

From the figure 4.3 the plane H1 does not separate the classes but H2 separates the classes but the separation made by H2 is only with a small margin.Here H3 separates the classes with the maximum margin. There are many hyperplanes to classify the data but only one is the best hyperplane. This hyperplane is the one that represents the largest separation or margin between the two classes. Hence the hyperplane is selected in such a way that the distance from it to the nearest data point is maximized on each side.

International Journal of Research in Advent Technology, Vol.8, No.9, September 2020
E-ISSN: 2321-9637
Available online at www.ijrat.org

V. RESULTS AND DISCUSSION

In this research, the image data of various kinds of soils was collected and compiled. The category of the soil is identified and the name is displayed in a dialogue box.And also a GUI is generated which shows the features, i.e., HSV list, Auto Correlogram, Color Moments, Mean Amplitude, Energy, Wavelet Moments along with the accuracy and category of the soil.The proposed features have been given to the Modified support vector machines to classify the soils.

1. Clay

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is clay. A GUI is generated in which the simulation result is given as accuracy is 96.7742%.
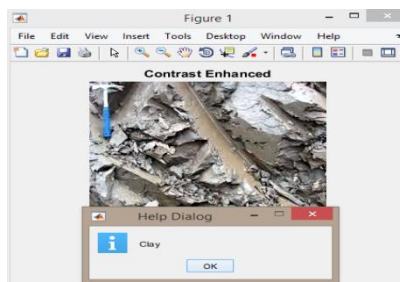
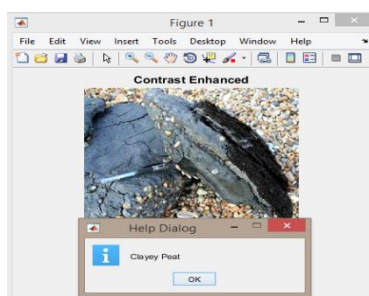**Fig.5: Output of Classified Clay Image**
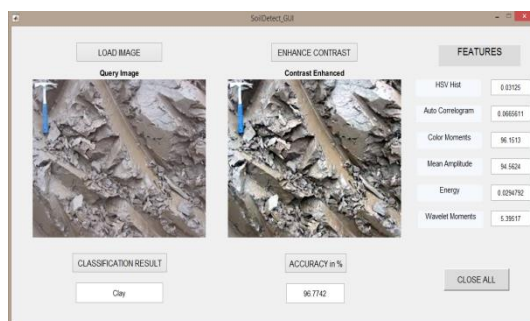
Fig. 6: Features and Accuracy of Query Clay Image

2.Clayey Peat

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is clayey peat. A GUI is generated in which the simulation result is given as accuracy is 96.7742%.

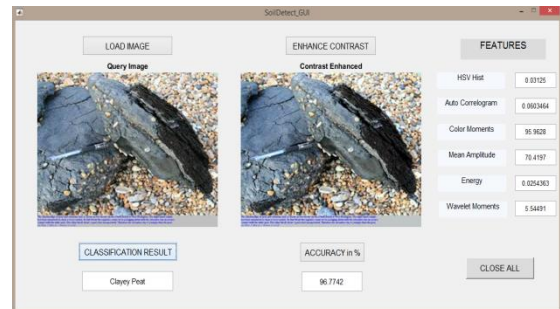**Fig.7: Output of Classified Clayey Peat Image**

Fig.8:Features and Accuracy of Query Clayey Peat Soil

3.Clayey Sand

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is clayey sand. A GUI is generated in which the simulation result is given as accuracy is 98.2%.
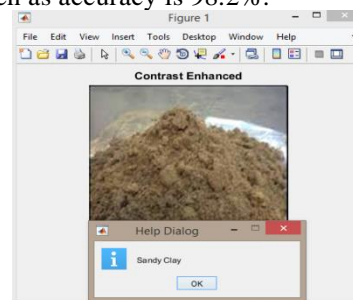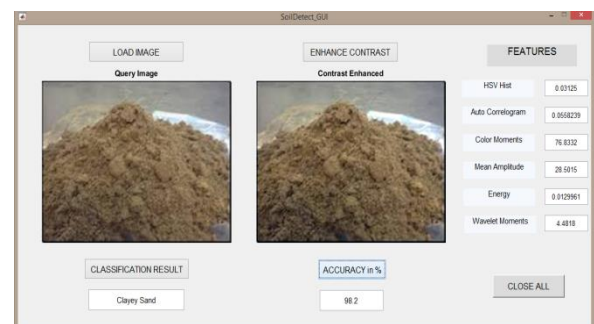

Fig.9 : Output of Classified Clayey Sand Image

Fig.10: Features and Accuracy of Query Clayey Sand

4.Humus Clay

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is humus clay. A GUI is generated in which the simulation result is given as accuracy is 96.7742%.
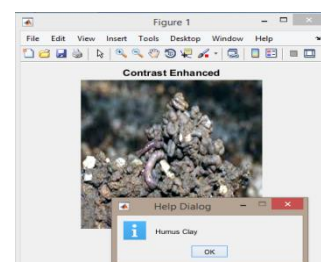

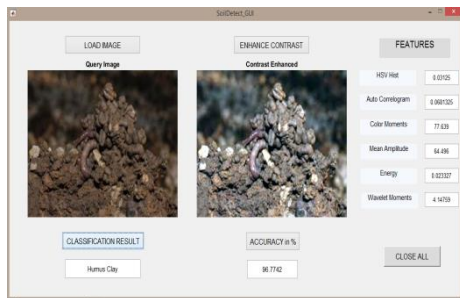**Fig.11: Output of Classified Humus Clay Image**

International Journal of Research in Advent Technology, Vol.8, No.9, September 2020
E-ISSN: 2321-9637
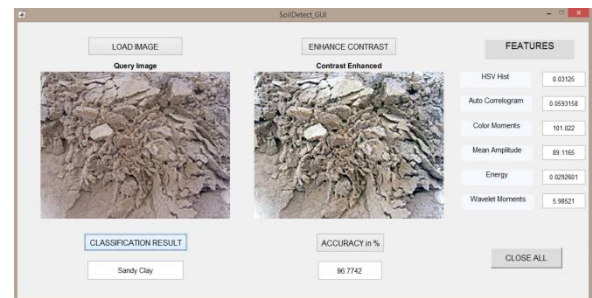*Available online at www.ijrat.org*

Fig.12: Features and Accuracy of Humus Clay Image

5.Peat

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is peat. A GUI is generated in which the simulation result is given as accuracy is 98.3871%.
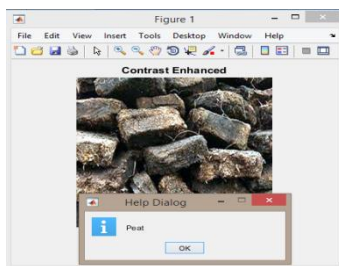

Fig.13 : Output of Classified Peat Image

Fig.14 : Features and Accuracy of Query Peat Image

6.Sandy Clay

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is sandy clay. A GUI is generated in which the simulation result is given as accuracy is 96.7742%.

Fig.15 : Output of Classified Sandy Clay Image

Fig.16: Features and Accuracy of Query Sandy Clay

7.Silty Sand

The Query soil image is pre-processed i.e., contrast is enhanced and shows a dialogue box as query soil image is silty sand. A GUI is generated in which the simulation result is given as accuracy is 98.3871%.
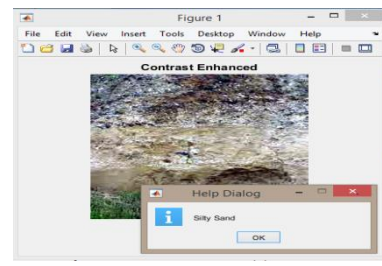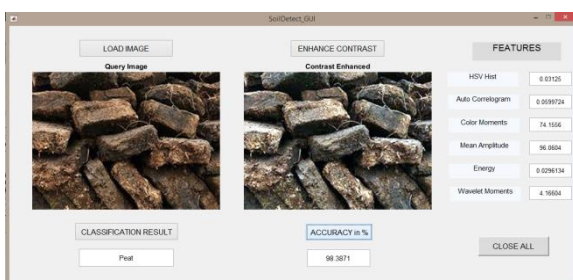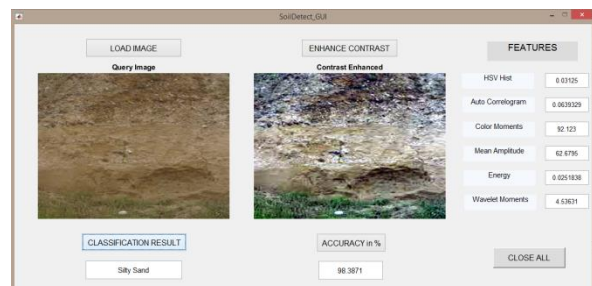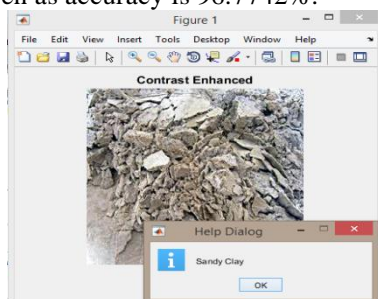

Fig.17: Output of Classified Silty Sand Image

Fig.18 : Features and Accuracy of Silty Sand Image

VI. CONCLUSION

On site soil classification is required in an automated procedure instead if manual classification to decrease the effort to determine the soil. The image processing concepts are proven as efficient methods for automating this task. Different algorithms exist for soil classification but still soil classification with high accuracy and with less expenditure is challenging process. In this paper, the automation is performed as per the discussion. Here seven different categories of soil were considered and these soils were processed. And the accuracy is also calculated. In this paper, it is observed that Modified Support Vector Machine can work in an efficient manner with high accuracy level. MATLAB software worked as an efficient tool for the development of proposed classifiction method and this can be used for further development of on-site real time soil classification using independent interface.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Ashwini Rao Janhavi, Abhishek Gowda, Manjunath, Mrs. Rafega Beham, Soil Classification and Crop Detection using Machine Learning. IJSRD - International Journal for Scientific Research & Development, 2016.

[2] Shamik Sural, Gang Qian and Sakti Pramanik, Segmentation and histogram generation using the hsv color space for image retrieval, Dept. of Computer Science Engineering, Michigan State University, USA.

[3] V. Rajeswari and K. Arunesh, Soil Data Analysing using Data Mining Classification Techniques. Indian Journal of Science and Technology, 2016.

[4] Ayon Dey, Machine Learning Algorithms: A Review, International Journal of Computer Science and Information Technologies, Vol. 7(3), 2016, 1174-1179.

[5] Utpal Barman, Ridip Dev Choudhury, Soil Texture Classification Using Multi Class Support Vector Machine, Information Processing in Agriculture, August 2019.

[6] R.Shenbagavalli and K.Ramar, Feature extraction of soil images for retrieval based on statistics, International Journal of Computer Applications, Volume 88 – No.14, February 2014.

[7] B. Bhattacharya, D. P. Solomatine, Machine Learning in Soil Classification, International Joint Conference on Neural Networks, Montreal, Canada, July 31 - August 4, 2005.

[8] Osman, Mohammed, Ahmed Khalid, Elfatih. Evaluation of cone penetration test classification methods for some local soils, Building and Road Research Journal, 5, 2003.

AUTHORS PROFILE

E V Vijay is working as Assistant professor in the department of Electronics and Communication Engineering (ECE) at G E College and pursuing Ph.D in Wireless Sensor Networks.

Ch Navya is a final year engineering student from Electronics and Communication Engineering (ECE) department and studying at G E College. She has co-authored one journal publication till date. Her areas of interest include Arduino and MATLAB.

Abdul Shabana Begum is a final year student from Electronics and Communication Engineering (ECE) department and studying at G E College. She has co-authored one journal publication till date. Her areas of interest in research includes MATLAB and Arduino.

D Rajaneesh is a final year student from Electronics and Communication Engineering (ECE) department and studying at G E College. He has co-authored one journal publication till date. His areas of interest in research includes MATLAB and Arduino.

B Mahesh Babu is a final year student from Electronics and Communication Engineering (ECE) department and studying at G E College. He has co-authored one journal publication till date. His areas of interest in research includes MATLAB and Arduino.