AN EFFICIENT CLASSIFICATION OF SOIL IMAGES USING GABOR - CNN

Vijaya Nirmala Gera[1*], Rajesh Kumar P[2], Praveen B. Choppala[3]

Abstract –
Soil classification is an emerging field of research in present scenario. It is used to comprehend and evaluate a particular soil's performance and determine whether the soil is suitable for agricultural purpose or particular engineering applications. The typical methods that are employed by farmers are insufficient to meet the rising demands, thus they are forced to impede soil cultivation. There are diverse laboratory and field techniques for classifying soil like statistical, rule-based and traditional learning methods, but many have drawbacks like time consuming and involves domain expert opinion. However, plans are still lacking in providing an accurate classification result. We proposed the novel soil classification technique by pre-processing different soil images and extracting features using Gabor wavelet transform. Further these features are classified using Convolutional Neural Network (CNN) classifier. Recognition rate of 98% has been achieved by using the proposed method.

Keywords - Soil Classification, Pre–processing, Gabor features, Convolutional Neural Networks, Recognition rate

[1*]Dept. of E.C.E, Andhra University TDR-HUB Visakhapatnam, India, vijayanirmala.au@gmail.com
[2]Dept. of E.C.E, Andhra University Visakhapatnam, India, rajeshauce@gmail.com
[3]Dept. of E.C.E, Andhra University Visakhapatnam, India, praveen.b.choppala@gmail.com

***Corresponding Author:** Vijaya Nirmala Gera
*Dept. of E.C.E, Andhra University TDR-HUB Visakhapatnam, India, vijayanirmala.au@gmail.com

DOI: - 10.48047/ecb/2023.12.si10.00187

I. Introduction

Soil is made up of organic matter, minerals, living organisms, water, and air, which are necessary for the growth of plants. All the food we consume originates from the soil. So, it is essential to identify which crop will yield the highest results in specific soil types. This leads to the need for soil classification [1]. Examining soil as a material can be the start of soil classification. In the view of a layman – soil is the dust or detritus on the surface of the earth. For a farmer, soil is the growing medium of the environment for growing plants. The concept of soil as a natural body arose from its potential to encourage and support crops.

Soil properties play an important role in the evaluation of various agricultural tasks and therefore influence agricultural planning. Understanding soil properties could provide useful information for more logical and careful management planning and use in cultivated areas. Biota, geological history, and climate are important considerations that influence soil chemical and physical properties excessively (at continental and regional scales), while the dominant factors would be topography and human activity regulates soil properties on a smaller level. The basic components of the soil are discrete fragments (e.g., plant fragments, clay minerals, quartz grains) that can be seen broken up with an optical microscope. Soil structure is directly related to the sharpness, shape, contrast frequency, size, voids and spatial arrangement of key particles. In addition, many of these functions depend on the alignment of the ingredients as well as the way they are cut and the magnification used. The various types of soils based on dominating particle size are clay, peat and sand. Different combinations of these soils are taken into consideration in this paper.

Fig. 1 Different types of soils

Traditionally, there are different methods to determine soil texture and color. Many methods for use in the laboratory and in the field include the pipette method, the elutriation method, the decantation method, and the Munsell color chart method. The USDA triangle method is also available for soil classification. The disadvantage of these methods is the laborious and time-consuming processes. Therefore, soil classification has attracted the attention of researchers on the use of methods based on computer vision and image processing in soil classification [2].

II. Literature Review

In order to capture and process the color images of the soil samples, various algorithms and filters are created in [3] article. This algorithm extracts the many properties, such as color, texture, etc. Here, different soil types including alluvial, clay, black and red are taken into consideration. Constraint clustering and classification, Vane Shear Test (VST), Cone Penetration Test (CPT), Pressure Meter Test (PMT) and Standard Penetration Test (SPT) are some of the traditional methods used to classify soil. Among these [4] SPT is simple. In this test, classification is done based on the moisture content, visual inspection and even penetrates through the dense layers. The main drawback is costly, time consuming and results from SPT cannot be reproduced. In [5] CPT, the main drawback is requirement of domain expert for analyzing the segmented signals. In strength and compressibility tests like field vane shear test – exclusive to find the undrained shear strength of firm or soft clays. The other side, it is only suitable for clays with undrained strengths up to about 100 kPa. In [6] PMT, a large number of fundamental soil properties are obtained from a single test. To derive these properties no empirical correcting factors are needed. But the instrument will not penetrate gravels and clay stones. Operating in these sands may need borehole to a level one or two meters above the desired test locations. In [7] VST, the shear strength of soft clay at greater depth can be obtained and is relatively quick and easy. Here the main concern is test cannot be conducted on the fissured specimen of clay. In [8] Constraint Clustering and Classification (CONCC) is machine learning based automatic classifier with segmentation and classification are the two phases. In order to resolve the imprecision in the measured data, the Cone Penetration Test data is first converted into "J" segments from a single series of data, and this data is then separated into classes using fuzzy logic. Therefore, from the above methods we understand different types of soil and which crops are suitable for which. We concluded that peat soil is more suitable for agriculture. Then we have known the conventional techniques to classify the soil which are time consuming, more cost and needs human effort. So, from this we can say that there is a need for automated system which can classify soil at low cost in less time.

III. Proposed Method

Convolutional Neural Networks are different from other neural networks by their superior performance with speech or audio signal, image inputs. The core of CNN is convolutional layer where the maximum number of computations occur. It requires a few components, which are input data, a filter, and a feature map. In this classification of soil types, we considered datasets containing different types of soils namely clay, Clayey Peat, Clayey Sand, Humus Clay, Peat, Sandy Clay and Silty Sand. The flow of the proposed model is shown in the Fig. 2 in which the given input soil image is first pre-processed to enhance the image for further analysis. Followed by feature extraction methods like HSV – histogram, color moments-mean, standard deviation, Gabor features, color auto correlogram to extract features of soil. Extracting Gabor features from the given image plays a very important role. From the extracted features, the process of separating the various features from the retrieved characteristics is feature data preparation. Then, with the aid of the above steps, a feature database is created. Finally, Gabor based features are classified using CNN classifier(G-CNN).

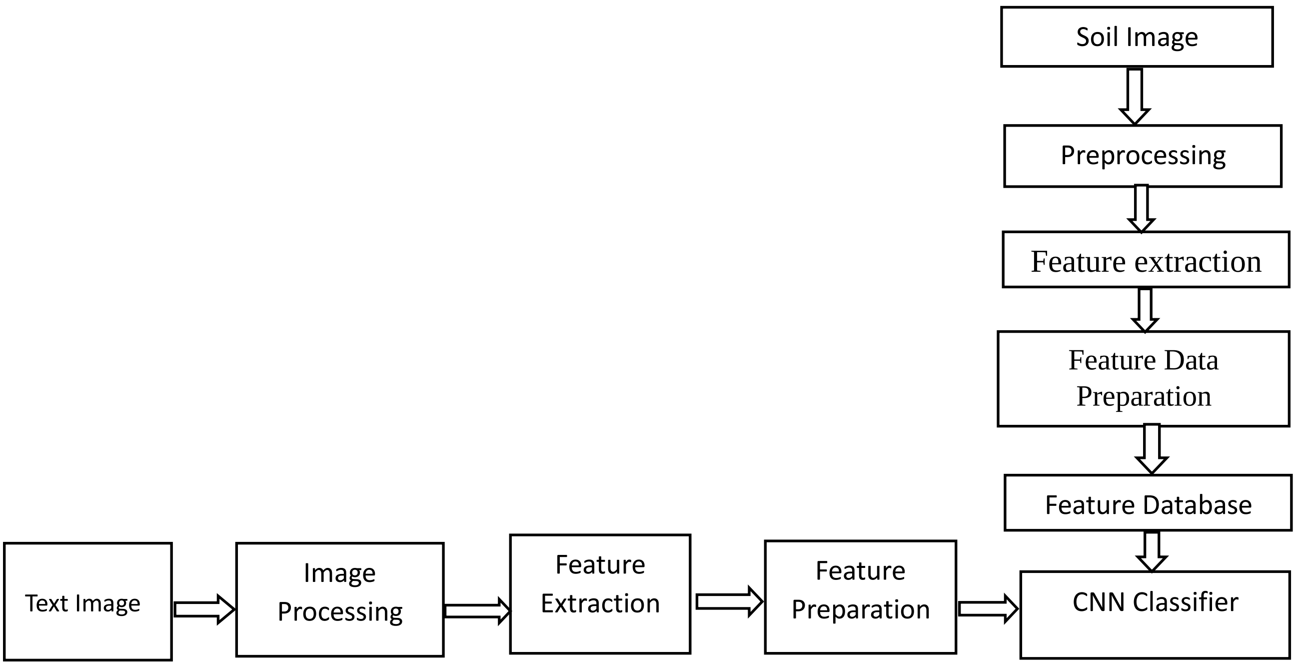

Fig. 2 Block diagram of proposed model

In this method, the dataset is split into training set and testing set. A digital image F(x,y) is represented as matrix which consists of MxN columns and rows. Intensity (color) and Coordinates are the two elements that define picture element(pixels). Before model training and inference, the steps taken to format images are image enhancement, feature extraction and HSV histogram. When the foreground and backdrop are either bright or dark, this procedure is helpful. The histograms of different soil images are shown in Fig.3.

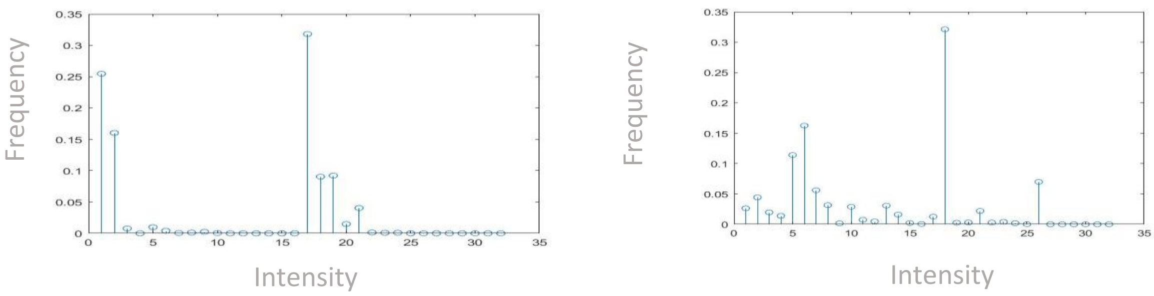

Fig. 3 a)Slity sand histogram b) peat histogram

Color moment of one image is compared to the database of digital images with pre-computed features in order to find and retrieve a similar Image. Each comparison between images results in a similarity score, and lower the score is the more identical the two images are supposed to be. The first and second color moments are computed using equations 1 & 2 respectively.

$$E_i = \sum_{N}^{j=1} \frac{1}{N} P_{ij} \qquad (1)$$

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{N}^{j=1}\left(P_{ij} - E_i\right)^2\right)} \qquad (2)$$

Where number of pixels are represented as N in an image and j-th pixel value in the i-th color channel of an image is represented as P_{ij}. Mean value for the i-th color channel is represented as E_i. The Fig.4 shows the color distribution of soil images.

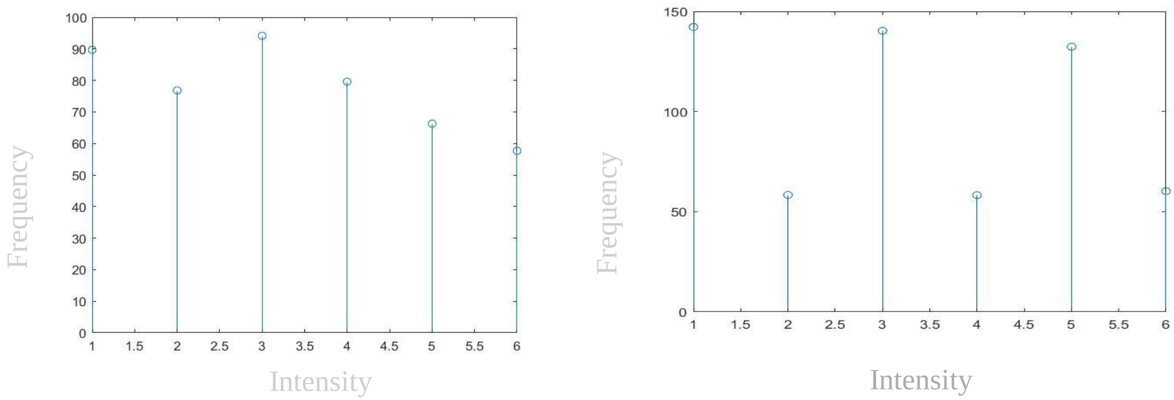

Fig. 4 Color distribution of a)Humus clay b) Silty sand

A Gabor filter is a type of linear filter used in image processing for texture categorization, feature extraction and edge detection. In order to extract required features from an image, a collection of Gabor filters with various frequencies and orientations are applied. Depending on the texture of the soil multi scale Gabor filters were implemented for further analysis. By considering two-dimensional Gabor filter with frequency f, σ for changing the size of an image for analysis depending on its texture and texture orientation is obtained by varying θ in a specific direction is given in Eq.3. For a variety of dilations and

$$f(x, y, \omega, \theta, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma x\ \sigma y} exp\left[\frac{-1}{2}\left(\left(\frac{x}{\sigma x}\right)^2 + \left(\frac{y}{\sigma y}\right)^2\right) + j\omega\left(xcos\theta + ysin\theta\right)\right] \qquad (3)$$

rotations, these filters have a direct relationship to Gabor wavelets. But expansion is typically not used for Gabor wavelets since it requires more time to compute bi-orthogonal wavelets. As a result, a filter bank with different Gabor filters rotations and scales are typically developed. Therefore, Gabor space is created

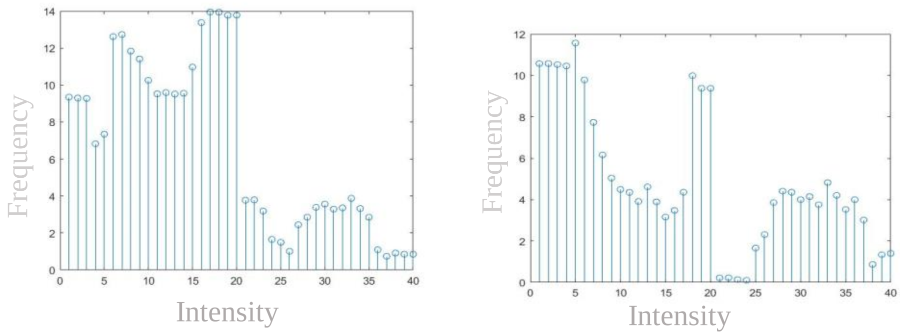

Fig. 5 Gabor wavelet of a) Sandy Clay b) Peat Soil

as a result of the filters being convolved with the signal. The Fig. 5 shows the Gabor wavelet output of sandy clay and peat soil.

Image comparison and indexing are done by using color correlogram. According to the results, this correlogram performs better for picture indexing and retrieval than both the conventional histogram method and the most recently proposed histogram refinement methods. The chance of discovering a pixel of color j at a distance of k from a pixel of color i in the image is defined by the k-th entry of an image color correlogram, which is a table indexed by color pairings. Such an image function demonstrates its ability to survive significant changes in the way the same picture appears as a result of sharp form changes caused by camera zooming, backdrop changes and changes in viewing position etc. The Fig. 6 shows the similarity between pixels at a distance. The clayey sand output value at some places is zero but in clayey peat the value is non zero.

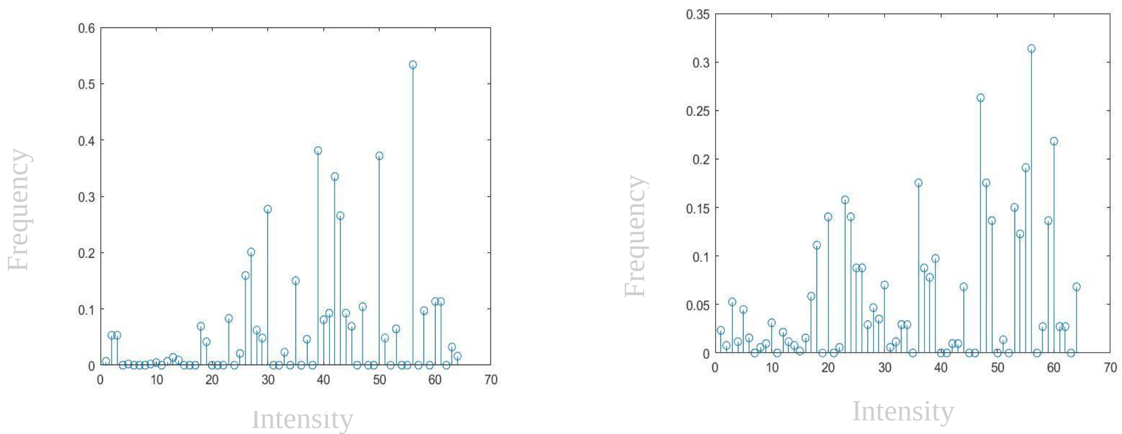

Fig.6 Output waveform of a) Clayey sand and b) Clayey peat

The Gabor features obtained from the soil images is given to the convolutional neural network for further classification. CNN consists of convolutional layers that extract a characteristic from the input data by acting as a filter. This convolutional layer contains many filters that each extract a separate feature while being applied in the same step. The filter's or kernel's size is determined by the anticipated size of a given feature. After the convolution, subsampling is done by the pooling layer where the pixel with the largest intensity value will be transferred into the subsampled image. Now training is carried out through back propagation to adjust the weights correctly. The number of steps and epochs are important parts of the training process as it may lead to under or overfitting.

IV. Results

In the proposed model, the classification process involved a total of 180 samples of different soil images, among which 80% samples are used for training and 20% samples for testing. In the result shown in Fig. 7, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to clay with accuracy of 92.8%.

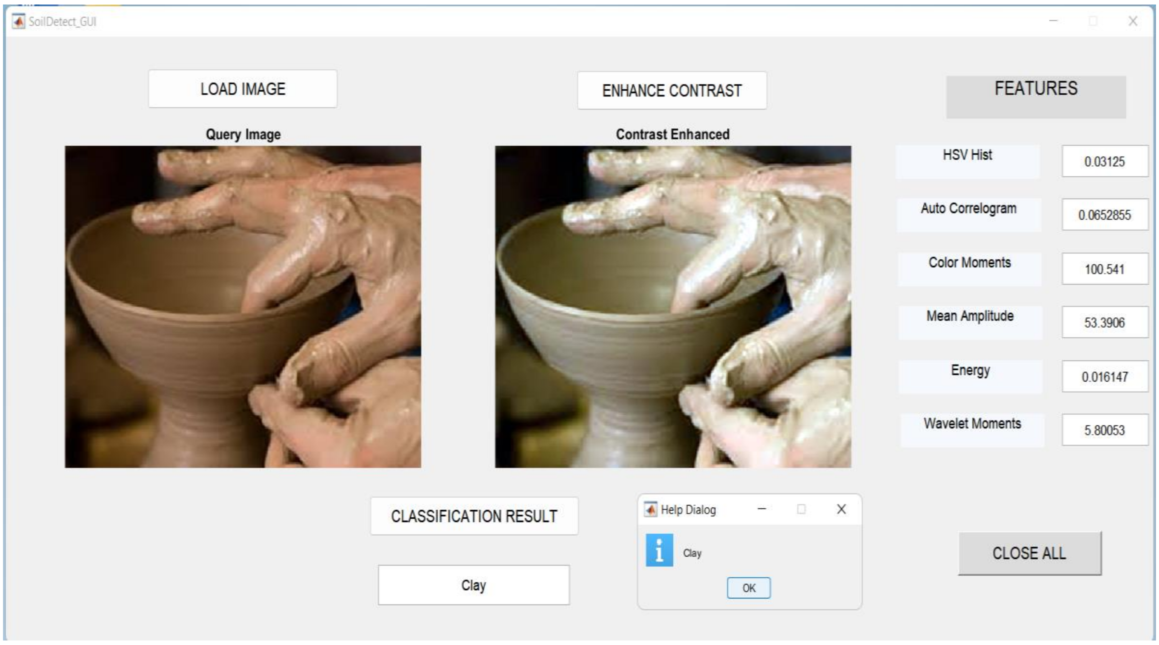

Fig.7 Query Soil - Clay

In the result shown in Fig. 8, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to clayey peat with accuracy of 100%.

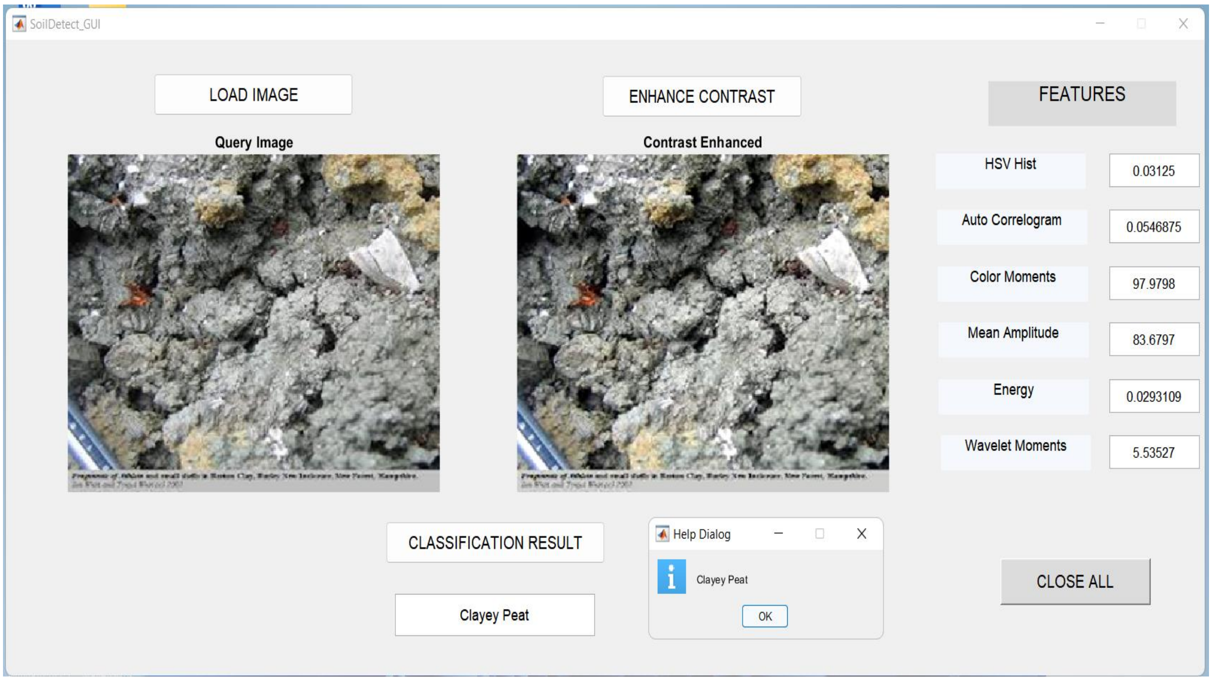

Fig.8 Query Soil - clayey peat

In the result shown in Fig. 9, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to clayey sand with accuracy of 100%.

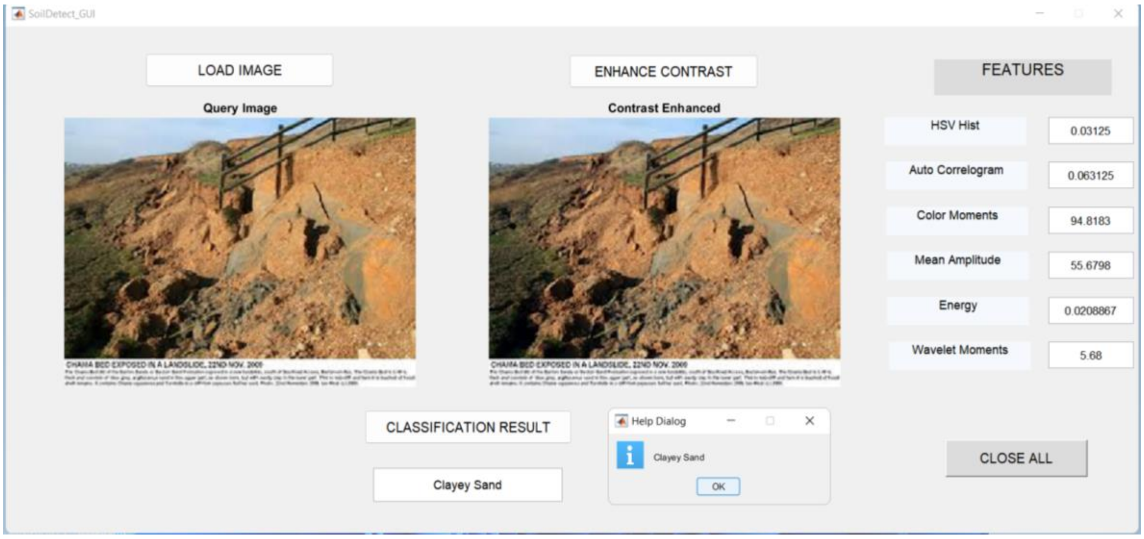

Fig.9 Query Soil - clayey sand

In the result shown in Fig. 10, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to Humus clay with accuracy of 100%.

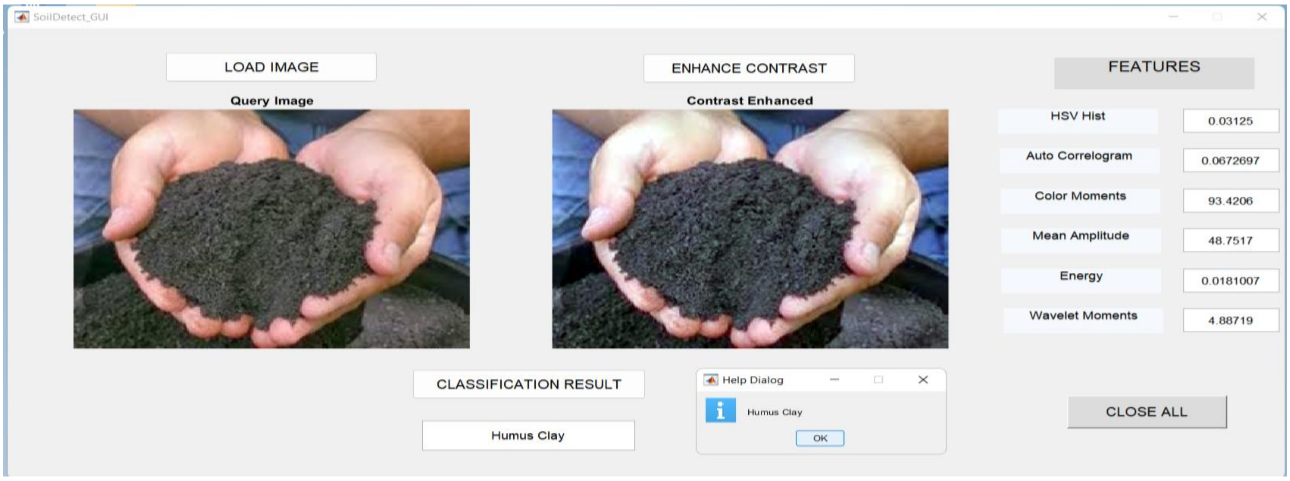

Fig.10 Query Soil - Humus clay

In the result shown in Fig. 11, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to peat with accuracy of 96.1%.

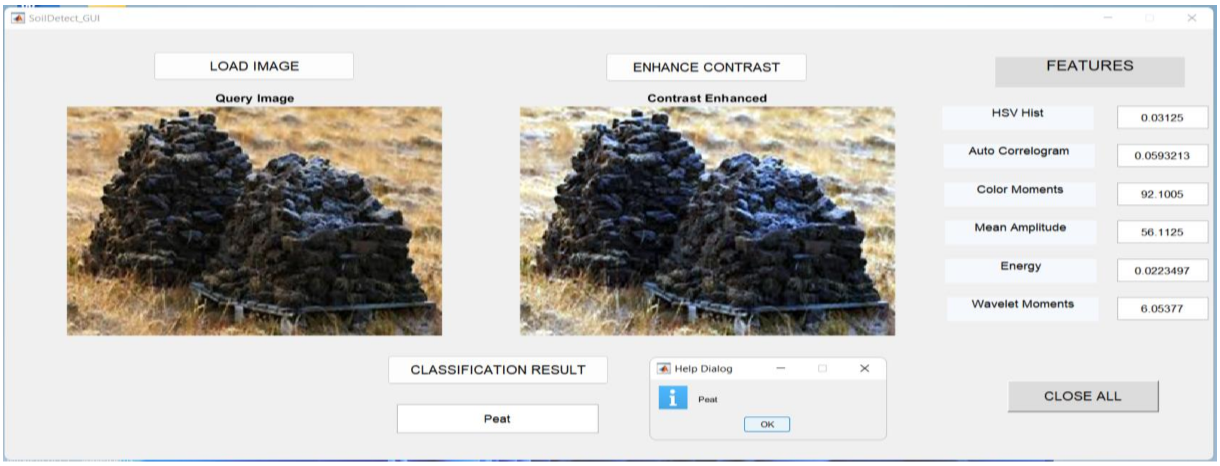

Fig.11 Query Soil - Peat

In the result shown in Fig. 12, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to sandy clay with accuracy of 96.1%.

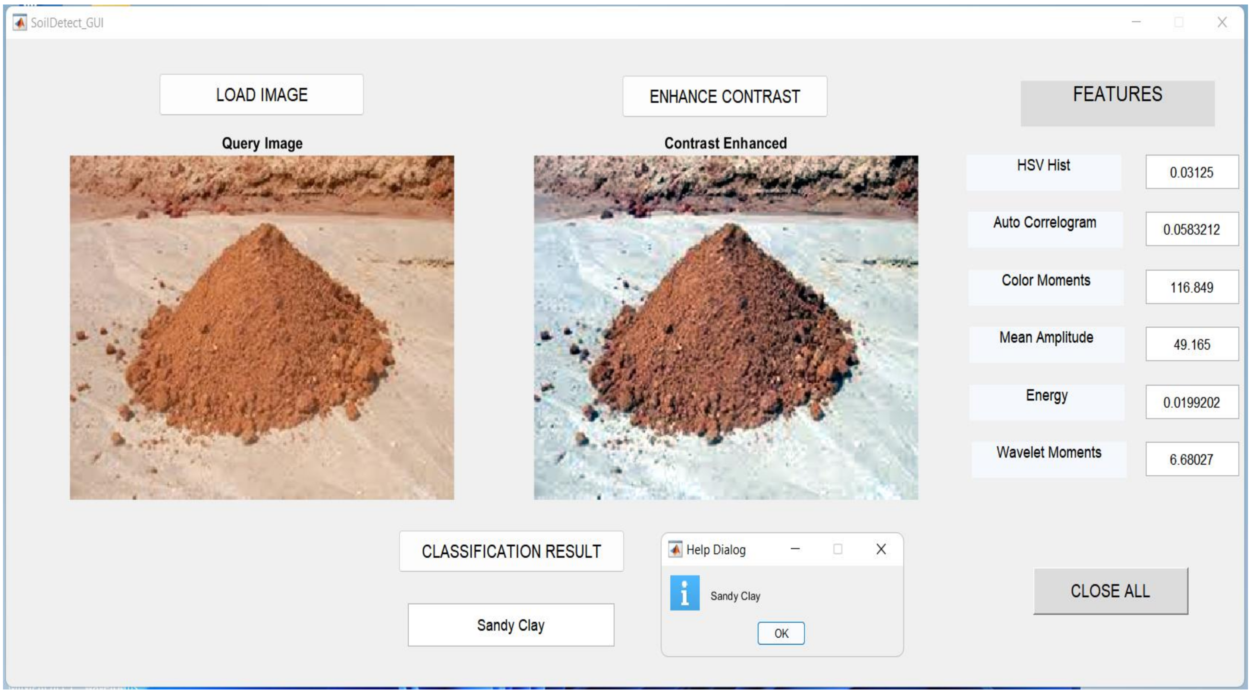

Fig.12 Query Soil - Sandy clay

In the result shown in Fig. 13, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to silty sand with accuracy of 96.1%.

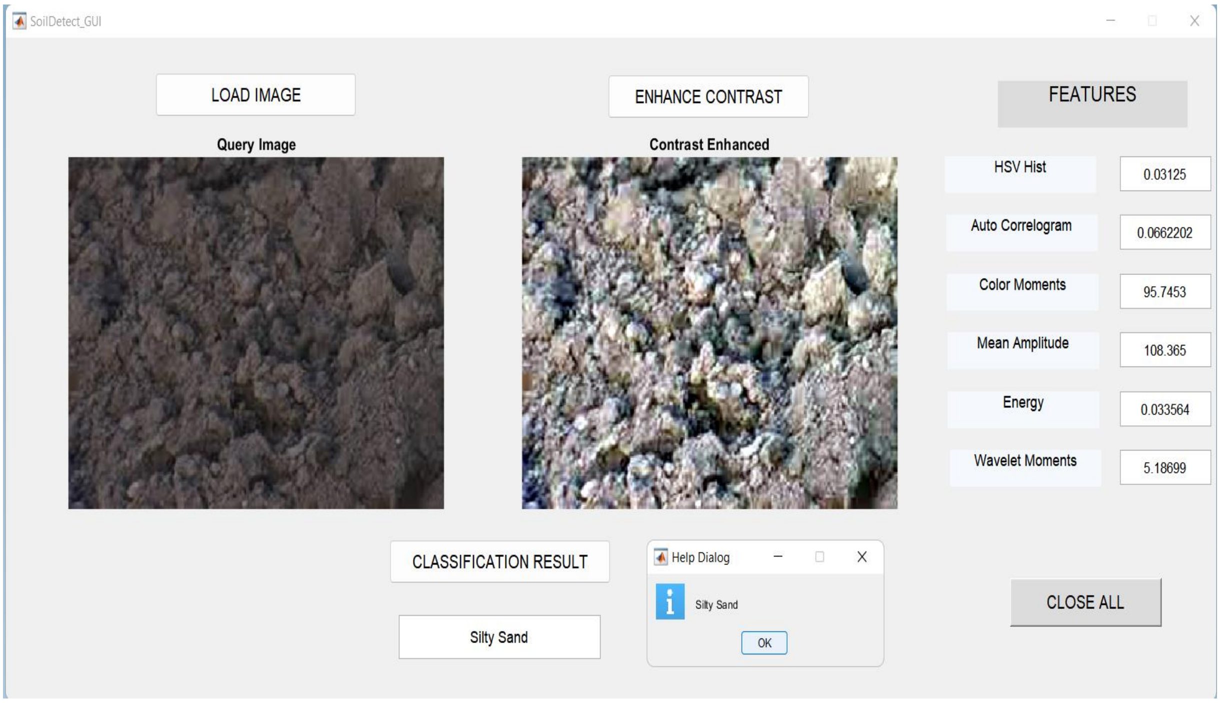

Fig.13 Query Soil - Silty sand

In the result shown in Fig. 14, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to clayey peat with accuracy of 100%.

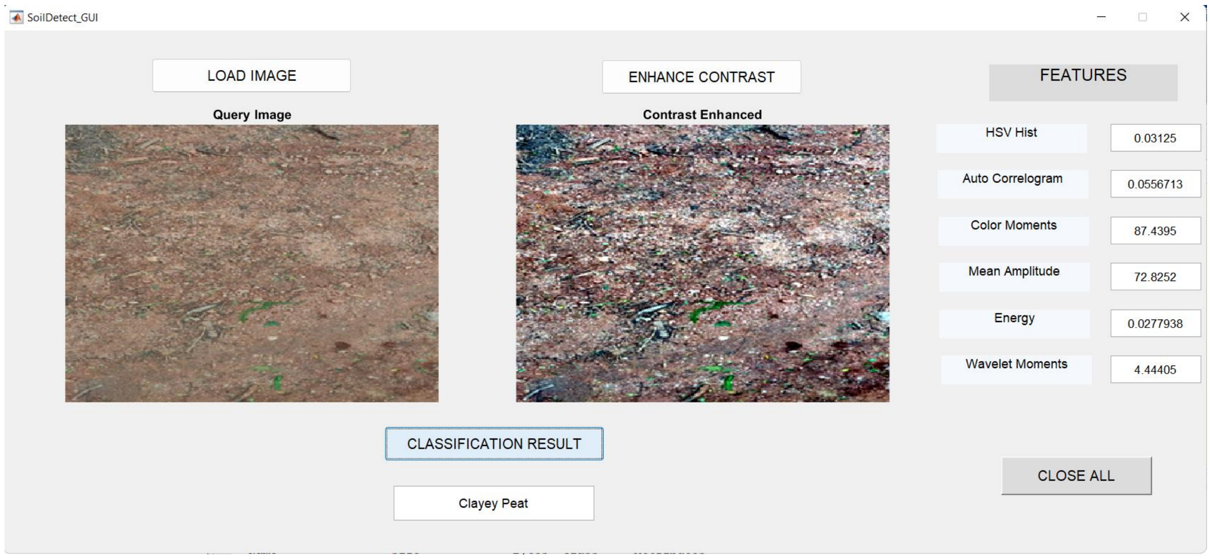

Fig.14 Query Soil - Clayey peat

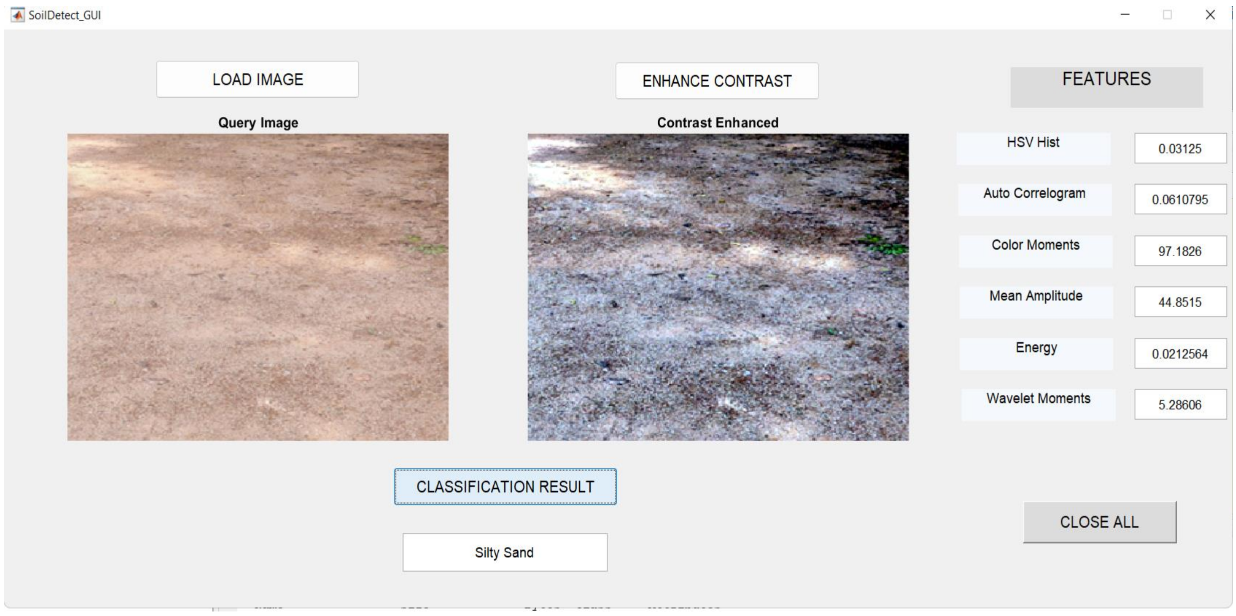

Fig.15 Query Soil - Silty sand

In the result shown in Fig. 15, First, preprocessing (contrast enhancement) is done on the query soil picture, features are extracted using Gabor and these features are classified using CNN showing the given test soil image belongs to silty sand with accuracy of 100%.

Table 1. Training accuracy results for different soil types

SOIL SAMPLES	TOTAL NO OF SAMPLES	ACCURACY (Correct samples/total samples) *100			
		Artificial Neural Network (ANN)	K-means Clustering (KNN)	Linear Support Vector Machine (SVM)	G-CNN (Proposed Method)
Clay	28	90	90.5	91.3	92.8
Clayey Peat	25	93.2	93.8	98	100
Clayey Sand	25	94.2	95.7	98	100
Humus Clay	25	95.2	96.8	98	100
Peat	26	91.5	92.8	95	96.1
Sandy Clay	25	96.1	96.8	98	100
Silty Sand	26	93	92.4	94.8	96.1

The Table 1 shows the accuracy of our proposed method. For every soil we have taken approximately about 20-25 samples to know how much error it contained. We can see that for some soils it was able to give almost all samples correct which is superior when compared with the existing methods like linear SVM, Fine KNN and ANN. Thus, by applying G-CNN for the soil classification, we have achieved an overall accuracy of 98%.

V. Conclusion

The soil classification is obtained efficiently with the help of Gabor Convolutional Neural Networks. In this paper, the task of automation is carried out with seven different types of soil images. These images are collected, pre-processed by using HSV – histogram, color moments and further feature extraction by using Gabor filter and finally through the CNN classifier. Hence, the results obtained by using Gabor CNN classifier exhibits high level of accuracy when compared with the methods by using ANN and linear SVM. Different other features can be extracted for further real time soil classification in real time.

REFERENCES

1. Sharma, Hement Kumar, and Shiv Kumar. "Soil classification & characterization using image processing." In *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 885-890. IEEE, 2018.
2. Barman, Utpal, and Ridip Dev Choudhury. "Soil texture classification using multi class support vector machine." *Information processing in agriculture* 7, no. 2 (2020): 318-332.
3. Srivastava, Pallavi, Aasheesh Shukla, and Atul Bansal. "A comprehensive review on soil classification using deep learning and computer vision techniques." *Multimedia Tools and Applications* 80 (2021): 14887-14914.
4. Biswas, Suparna, Tarek Ibne Faysal, Arshi Siddiqui Promiti, Md Sazzad Hossain, Lubaba Bazlul, Abdullah Md Sarwar, Sayeed Md Shaiban, and Rashedur M. Rahman. "Analysis of soil and various geo-technical properties using data mining techniques." In *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, pp. 288-293. IEEE, 2020.
5. Wang, Yu, Yue Hu, and Tengyuan Zhao. "Cone penetration test (CPT)-based subsurface soil classification and zonation in two-dimensional vertical cross section using Bayesian compressive sampling." *Canadian Geotechnical Journal* 57, no. 7 (2020): 947-958.
6. Zhang, Wengang, Chongzhi Wu, Haiyi Zhong, Yongqin Li, and Lin Wang. "Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization." *Geoscience Frontiers* 12, no. 1 (2021): 469-477.
7. Masud, Mohammad, Jing Gao, Latifur Khan, Jiawei Han, and Bhavani M. Thuraisingham. "Classification and novel class detection in concept-drifting data streams under time constraints." *IEEE Transactions on knowledge and data engineering* 23, no. 6 (2010): 859-874.
8. Bhattacharya, Biswanath, and Dimitri P. Solomatine. "Machine learning in soil classification." *Neural networks* 19, no. 2 (2006): 186-195.
9. Xiao, Te, Hai-Feng Zou, Ke-Sheng Yin, Yu Du, and Li-Min Zhang. "Machine learning-enhanced soil classification by integrating borehole and CPTU data with noise filtering." *Bulletin of Engineering Geology and the Environment* 80 (2021): 9157-9171.
10. Azizi, Afshin, Yousef Abbaspour Gilandeh, Tarahom Mesri-Gundoshmian, Ali Akbar Saleh-Bigdeli, and Hamid Abrishami Moghaddam. "Classification of soil aggregates: A novel approach based on deep learning." *Soil and Tillage Research* 199 (2020): 104586.
11. Rahman, Sk Al Zaminur, Kaushik Chandra Mitra, and SM Mohidul Islam. "Soil classification using machine learning methods and crop suggestion based on soil series." In *2018 21st International Conference of Computer and Information Technology (ICCIT)*, pp. 1-4. IEEE, 2018.