



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 6, June 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.118

 9940 572 462

 6381 907 438

 [ijirset@gmail.com](mailto:ijirset@gmail.com)

 [www.ijirset.com](http://www.ijirset.com)



# Machine Learning Based Soil Classification and Crop Recommendation

Miss. Pranali Prashant Panchwagh, Dr. Getika Narang , Prof. Sneha Tirth

Department of Computer Engineering, Trinity College of Engineering, Pune, India

**Abstract:** In a modern agricultural management framework, forecasting soil type profile is acknowledged as critical information for irrigation demand forecasting. Forecasting daily land surface temperatures (LST) is critical for climate, agricultural, environmental, and industrial investigations. For these applications, hybrid data-driven prediction models combining Ensemble Empirical Mode Composition (EEMD) and Machine Learning (ML) algorithms are effective because they reduce modelling difficulties, require less historical data, are simple to create, and are less complex than physical models. To reduce the difficulty of modelling and enhance prediction accuracy, a computationally simple, less data-intensive, fast and efficient unique hybrid data-driven model termed the EEMD Long Short-Term Memory (LSTM) neural network, precisely EEMD-LSTM, is suggested. To begin, the EEMD is used to breakdown the original daily LST data series into a number of Intrinsic Mode Functions (IMFs) and a single residue item. The amount of input data sample points for LSTM models is then calculated using the Partial Autocorrelation Function (PACF). The original soil type series is decomposed into numerous intrinsic mode functions using ensemble empirical mode decomposition (EEMD) in this model (IMFs). The original series are mixed with IMFs after decomposition to create fresh two-dimensional input data for the convolutional neural network (CNN). EEMD-CNN has good performance not just in predicting the next day's soil temperature, but also in predicting the type several days afterwards.

**KEYWORDS:** soil type; forecasting; data-driven; hybrid model; Ensemble Empirical Mode Decomposition (EEMD); Neural Network (NN).

## I. INTRODUCTION

Soil is a technical term that denotes different things to different persons in the geological field. It is used to describe the residue left over from prior surface treatments. This is interpreted by a criminologist as a depiction of the current state of physical and chemical processes in the body. As far as an engineer is concerned, soil is the rock-solid foundation upon which constructions such as houses, factories, buildings, and highways can be built. Depending on their intended use, individuals may describe soils in a variety of ways. It is the study of soil patterns that can be observed from a distance that is called soil studies. It is critical to successfully run a farm that the soil is grouped together. The ability to recognise the features of the soil is critical in decreasing product losses. Countries that export a variety of agricultural products are the ones that are most in need of it. When designing a categorization system in the realm of engineering, the primary emphasis should be on mechanical parameters such as permeability, stiffness, and strength. The classification of a soil can be used to describe the soil. Knowing the type of soil to work with is quite beneficial when it comes to farming, construction, and other endeavours.

The ability of a plant to be planted in line with the specific qualities of a certain soil is crucial to the success of the plant in question. The character of soil is influenced by a variety of parameters, including the power of hydrogen (PH), the percentage of exchangeable sodium, the amount of moisture present, and so on. These plants exhibit distinct traits based on the amount of soil in which they are found, and these characteristics differ from one area to another. A variety of approaches are employed in order to categorise and analyse the many soil types that might be found in any given location. During the preparation stage, the manual segmentation and classification procedure is observed and evaluated. It necessitates a significant investment of time, effort, and money, and it is not for the faint-hearted. The major objective is to increase the efficiency of the process. When the fields of image processing and machine learning were first developed, The soil sample can be efficiently classified into the numerous categories to which it belongs based on the characteristics of the soil sample. It is explained in this study how convolution neural network (CNN) can be used to classify the segments that have been discovered.



## II. PROBLEM STATEMENT

Soils may be described in different ways by different people for their different purposes. Soil study means the knowing of externally identifiable patterns seen on soil. Grouping of soil is particularly basic for reasonable agricultural business. Recognizing the types of soil is the key feature to reduce the product quantity losses. It is crucial for countries that export several agricultural commodities.

## III. LITERATURE SURVEY

Sr. No	Author/ Paper Name	Proposed Work	Implemented Techniques	Advantages
1	Srunitha.k, Dr.S.Padmavathi, "Performance of SVM Classifier For Image Based Soil Classification"	This work explains support vector machine based classification of the soil types. Soil classification includes steps like image acquisition, image preprocessing, feature extraction and classification. The texture features of soil images are extracted using the low pass filter, Gabor filter and using color quantization technique.	1. Support Vector Machine(SVM) 2. Gabor Filter	The classifications of non-sandy soils are better classified withSVM (through WEKA).
2	MałgorzataCharytanowicz, and PiotrKulczycki, "An Image Analysis Algorithmfor Soil Structure Identification"	This work presents an image segmentation approach for detecting the soil pore structures that have been studied by way of soil tomography sections. In so-doing, a research study was conducted using a density-based clustering method, and in turn, the nonparametric kernel estimation methodology. This overcomes the rigidity of arbitrary assumptions concerning the number or shape of clusters among data, and lets the researcher detect inherent data structures.	1. Density-based clustering algorithm 2. Kernel estimation methodology	The presented approach is more objective than classical parametric methods, and can be successfully applied for many tasks in data mining, particularly where arbitrary assumptions concerning the number or shape of clusters among data are not recommended.
3	Richard J. Flavel, Chris N. Guppy, Sheikh M. R. Rabbi, Iain M. Young, "An image processing and analysis tool foridentifying and analyzing complex plant root systems in 3D soil using non destructive analysis: Root1"	The objective of this study was to develop a flexible and free image processing and analysissolution, based on the Public Domain Image platform, for the segmentation and analysis of complex biological plant root systems in soil from x-ray tomography 3D images. Contrasting root architectures from wheat, barley and chickpea root systems were grown in soil and scanned using a high resolution micro-tomography system.	1. Growth and scanning parameters and protocols	High speed, accurate, segmentationand analysis protocols.
4	Karisiddappa, Ramegowda, Shridhara, S. "Soil Characterization Based on Digital Image Analysis"	This paper investigates the development of digital image analysis approach for estimation of physicalproperties of soil in lieu of conventional laboratory approach.The present research deals collecting soil samples for trail pits at designated site as per IS code procedure. The digital	1. Curve fitting technique 2. Image analysis approach	There is a great potential inthe use of fractal dimension for estimating physical soil properties for practical application, with minimum human error.



		image database is prepared for the collected soil sample in the laboratory and physical properties(Y) are determined.		
5	Anita Dixit, Dr. Nagaratna Hedge and Dr. B. Eswar Reddy, " Texture Feature Based Satellite Image Classification Scheme Using SVM"	This work presents a satellite image classification system, which can classify between the vegetation, soil and water bodies. The objective of this work is met by subdividing the works into three important phases, which are satellite image pre-processing, feature extraction and classification. The image pre-processing phase denoises the image by median filter and the contrast is improved by Contrast Limited Adaptive Histogram Equalization (CLAHE) technique.	1. Median filter 2. Contrast Limited Adaptive Histogram Equalization (CLAHE) technique 3. Support Vector Machine (SVM)	The performance of the proposed approach is observed to be satisfactory in terms of accuracy, sensitivity and specificity

#### IV. METHODOLOGY

##### CNN Layers

##### Convolution Layer

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. Convolution of an image with different filters can perform operations such as edge detection, blur and sharpen by applying filters i.e. identity filter, edge detection, sharpen, box blur and Gaussian blur filter.

##### Pooling Layer

Pooling layers would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or down sampling which reduces the dimensionality of each map but retains important information.

##### Fully Connected Layer

In this layer Feature map matrix will be converted as vector (x1, x2, x3, ...). With the fully connected layers, we combined these features together to create a model.

##### Softmax Classifier

Finally, we have an activation function such as softmax or sigmoid to classify the outputs.

##### Modules

User Model

##### Module Description:

##### User Model

In User model user can login when he is already registered.

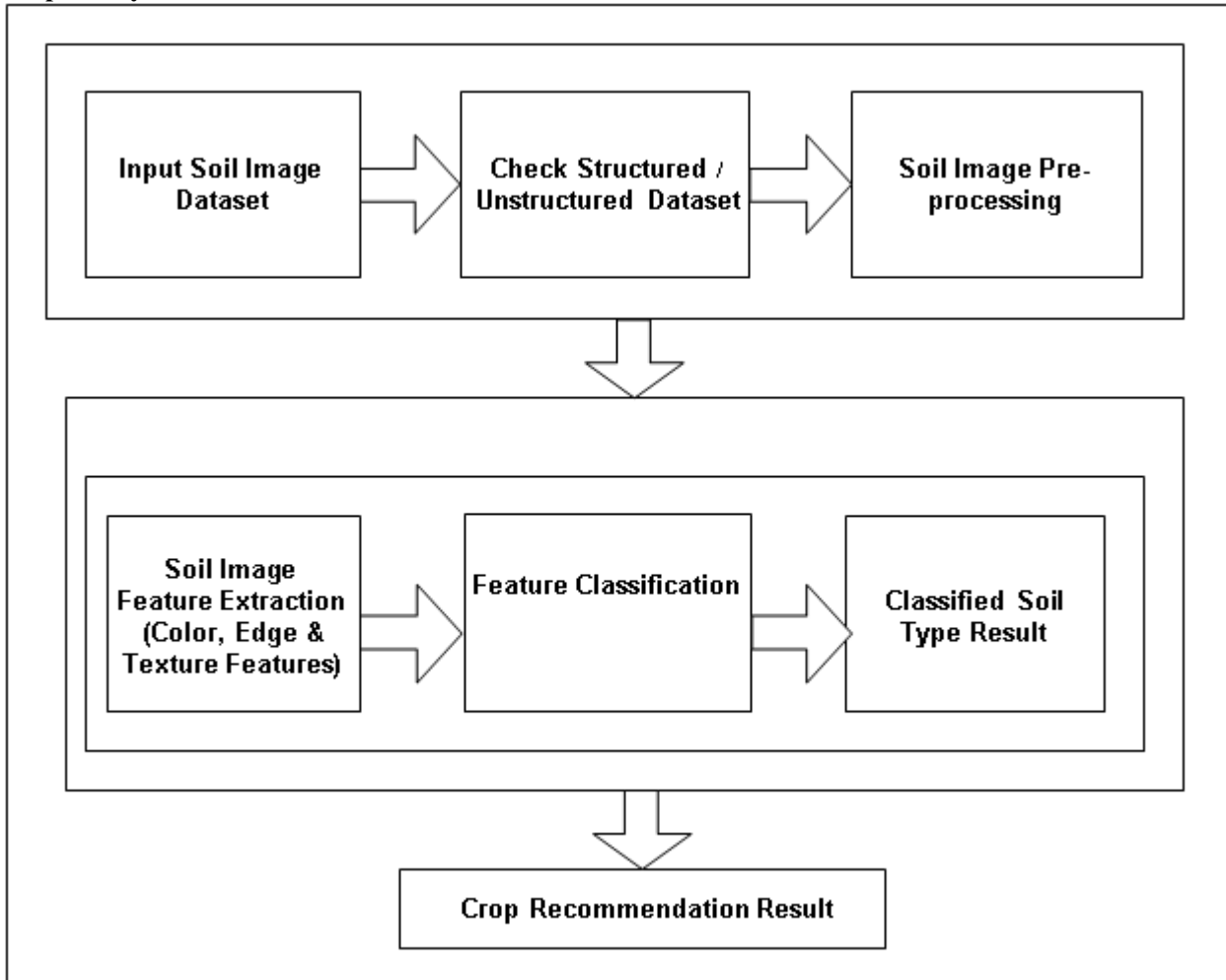
After that the user can upload the soil image and then process to get the classification result that is the type of particular soil and recommendation the crop details using the soil type.

##### Proposed System

In this proposed system we will use the soil image dataset to classify the soil type and Soil temperature. Soil classification includes steps like image acquisition, image preprocessing, feature extraction and classification. In image

preprocessing, the image is converted in grayscale image and then using filtration techniques to remove the noise of that image. After that in feature extraction step, the color features, edge detection features and texture features of soil images are extracted. Using CNN classifies the soil features to find the type of soil. Here we will use the training feature dataset for classification. Using the final soil type we can recommend which crop is suitable for this type of soil which used for farmers and agriculture related companies.

**Proposed System Architecture:**



**Fig 1: Proposed System Architecture**

**Proposed Outcomes**

- Soil Classification result that is exact soil type
- Crop details recommendation result based on particular classified soil type



**V. RESULT AND ANALYSIS**

The Experiments are done by personal computer with configuration: Intel (R) Core (TM) i3-2120 CPU @ 3.30GHz, 4GB memory, Windows 7, MySQL 5.1 backend database and JDK 1.8. The application uses web application tool for design code in Eclipse and execute on Tomcat server.

Positive (P) : Observation is positive.

Negative (N) : Observation is not positive.

True Positive (TP) : Observation is positive, and is predicted to be positive.

False Negative (FN) : Observation is positive, but is predicted negative.

True Negative (TN) : Observation is negative, and is predicted to be negative.

False Positive (FP) : Observation is negative, but is predicted positive.

Accuracy =  $TP + TN / (TP + FP + TN + FN)$

Precision =  $TP / (TP + FP)$

Recall =  $TP / (TP + FN)$

F1-Measure =  $2 * Precision * Recall / (Precision + Recall)$ .

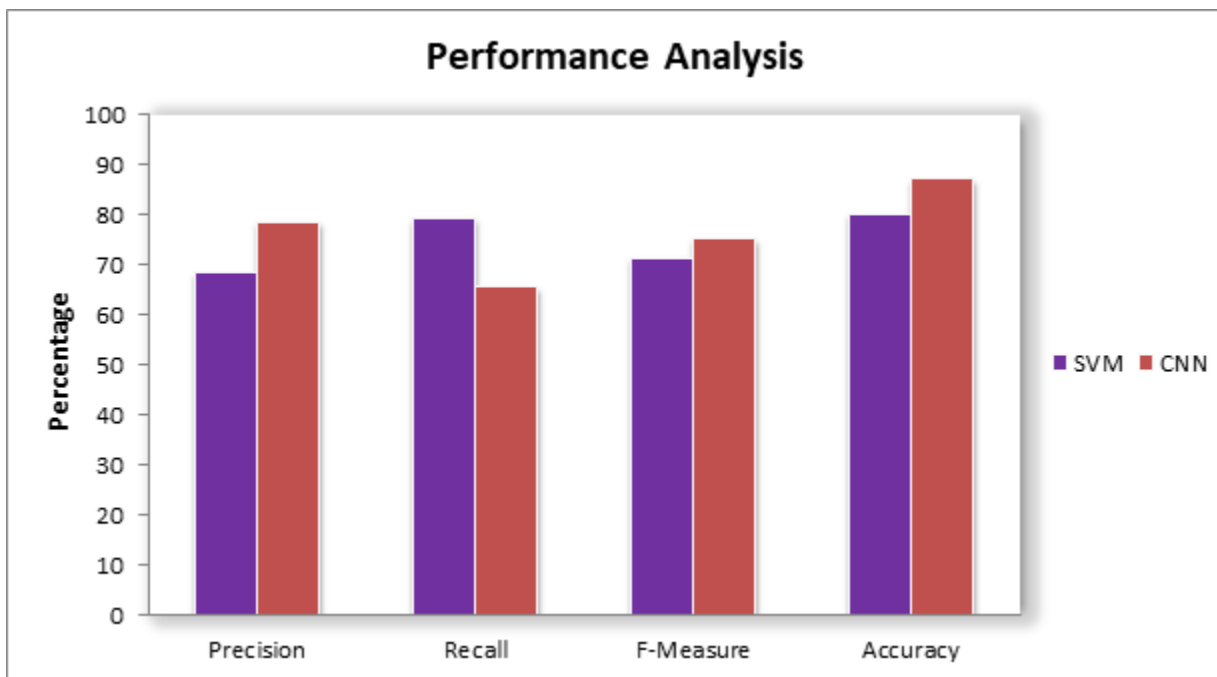


Fig 2: Performance analysis graph

	Existing System	Proposed System
Precision	68.45	78.70
Recall	79.44	65.64
F-Measure	72.11	74.31
Accuracy	80.29	87.26

Table 3: Performance analysis table

**VI. CONCLUSION**

This proposed system we will use the soil image dataset to classify the soil type. Image feature extraction like color features; edge detection features and texture features of soil images are extracted. Using Convolution neural network (CNN) classifies the soil features to find the type of soil. Finally based on soil type we can recommend which crop is



suitable for this type of soil. SO this work will help the farmer regarding crop and increase the productivity soil detection and suggestion according to it will help the farmers.

#### REFERENCES

- 1.J. Deng and L. Li, “Application of parameter optimization stochastic forest in stock prediction,” *Software*, vol. 41, no. 1, pp. 178–182, Jan. 2020.
- 2.M. R. Costa-jussà, A. Allauzen, L. Barrault, K. Cho, and H. Schwenk, “Introduction to the special issue on deep learning approaches for machine translation,” *Comput. Speech Lang.*, vol. 46, pp. 367–373, Nov. 2017.
- 3.H. M. Fayek, M. Lech, and L. Cavedon, “Evaluating deep learning architectures for speech emotion recognition,” *Neural Netw.* vol. 5, no. 3, pp. 23–28, Aug. 2017.
- 4.Tsantekidis, N. Passalis, A. Tefas, J. Kannianen, M. Gabbouj, and A. Iosifidis, “Forecasting stock prices from the limit order book using convolutional neural networks,” in *Proc. IEEE 19th Conf. Bus. Informat. (CBI)*, Beijing, China, Jul. 2017, pp. 10–15.
- 5.U. F. Siddiqi, S. M. Sait, and O. Kaynak, “Genetic algorithm for the mutual information-based feature selection in univariate time series data,” *IEEE Access*, vol. 8, pp. 9597–9609, 2020.
- 6.J. Hu and W. Zheng, “Multistage attention network for multivariate time series prediction,” *Neurocomputing*, vol. 18, no. 383, pp. 122–137, Mar. 2020
- 7.G. Li, P. Liu, C. Le, and B. Zhou, “A novel hybrid meta-heuristic algorithm based on the cross-entropy method and firefly algorithm for global optimization,” *Entropy*, vol. 21, no. 5, p. 494, May 2019.
- 8.S. Jadhav, H. He, and K. Jenkins, “Information gain directed genetic algorithm wrapper feature selection for credit rating,” *Appl. Soft Comput.*, vol. 69, no. 1, pp. 35–37, Aug. 2018.
- 9.D. Wang and J. Fang, “Research on optimization of big data construction engineering quality management based on RNN-LSTM,” *Complexity*, vol. 15, no. 2, pp. 15–20, Jun. 2018
- 10.Y. Kim, J.-H. Roh, and H. Kim, “Early forecasting of rice blast disease using long short-term memory recurrent neural networks,” *Sustainability*, vol. 10, no. 2, p. 34, Dec. 2017



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details