Soil Color as a Measurement for Estimation of Fertility using Deep Learning Techniques

N Lakshmi Kalyani, Kolla Bhanu Prakash
Department of CSE, Koneru Lakshmaiah Education Foundation
Guntur (DST), Andhra Pradesh, India

***Abstract*—Soil Behavior helps the farmer predict performance for growing crops, nutrient movement, and determine soil limitations. The traditional methods for soil classification in the laboratory require time and human resources and are expensive. This analysis examines the possibility of image recognition by artificial intelligence, with a machine learning technique called deep learning, to develop the cases that use artificial intelligence. This study performed deep learning with a model using a neural network. Neural Networks has used to evaluate relationships between the parameters of the three-dimensional coordinates resulting in soil classification and parameters. So Artificial Neural Networks (ANN) can be an effective tool for soil classification. This paper focused on AI techniques used to predict the soil type, advice the crop to yield, and discuss the transformed learning and benefits.**

Keywords*—Artificial neural networks; deep learning; soil classification; soil nutrients; data augmentation; transform learning

I. Introduction

Agriculture stands very quintessential in society. Agriculture is a source of livelihood in most parts of the world. Agricultural produce is of great importance. But in recent years, the farm produce is gradually decreasing. Soil plays a crucial role in agriculture. Soil consists of nutrients that the plants use to grow. There exist various kinds of soils available, and each has different effects. Crop's productivity was mainly based on the type of soil [1]. The possible way to improve productivity is that choose the right crop for the right land type. Soil needs to be investigated primarily before classifying them into distinctive groups. Based on these soil groups and the geographical conditions, one can decide which crop is best suited and is beneficial. The traditional methods are Costly, lengthy process and also time-consuming.

Consequently, there is a necessity for new technologies and methods to enhance the existing system to get faster and better results. Soil texture is due to the percentage of silt, clay, and sand present in it, and change in the fraction may result in different colors due to interaction among soil characteristics. Classification has a significant influence in agriculture for estimating yielding crops, and soil classification will define the relation between soil samples and other natural substances present. Convolutional Neural Network (CNN) is the branch of Deep learning and enables the machine to exhibit self-learning, to show the intelligence to predict by analyzing the input [3]. The essential point here is the privilege of the computer to learn automatically without human interaction.

Soil is the source of the minerals, liquids organisms that produce the foundation for the plant, and classification plays a major role in managing a crop, increasing soil primarily before classifying them into distinctive groups. Based on these soil groups and the geographical conditions, one can decide which crop is best suited and is beneficial [4]. The traditional methods are Costly, lengthy process and also time-consuming.

This research aims to benefit the former in identifying the red soil, and it analyzes the different soil patterns using CNN and recommends particular crops in that soil. And also help the researcher in soil science. The general purpose of this research is to compare the different soils and identify the red soil. The consequence of this may have numerous advantages to agriculture, soil management, and the environment.

II. Literature Survey

In this section, we summarized various works related to machine learning, image processing, and classification models from different papers seen in Table I.

The least median squares regression techniques produce more reliable results than the classical linear regression technique from the set of characteristics [2].

SVM classifier can work efficiently with high level of accuracy. MATLAB software has proved an efficient tool for design and development of classifier and can be used for further development of independent interface for on-site real time soil classification [5].

Time-honored methods of soil assortments are standard penetration test (SPT), cone penetration test (CPT), pressure meter test (PMT), and vane shear test (VST), time-consuming and needed for accurate results. Soil classification can be automated using artificial neural network techniques [6].

Mrs.Saranaya and Ms.A.Mythili researched to analyze the soil types so that it is helpful to farmers to choose a crop that to be cultivated. They have considered the SVM algorithm and chemical effects of soil like pH, salinity, organic matter, potassium, sulphur, zinc, Boron, calcium, Magnesium, Copper, Iron, and Manganese [7].

Classification of soil and quality prediction using a software-enabled solution using machine learning algorithms including decision tree, by considering soil chemical parameters [8].

TABLE I. VARIOUS MODELS USED FOR SOIL CLASSIFICATION

S.No.	Title	Techniques
1	Soil classification and crop suggestion using machine learning	Naive Bayes, J48,JRip algorithms
2	An Intelligent Model for Indian Soil Classification using various Machine Learning Techniques	The Pre-processed images are feature extracted, and the data extracted is used to train the SVM classifier.
3	Recent Trends Of Machine Learning In Soil Classification	Various emerging machine learning algorithms like SVM's, KNN, ANN, DT(decision tress) were discussed.
4	Soil Classification and Crop Suggestion using Machine Learning	Bagging Classifier, SVMachine, and KNN for classification of soil and crop recommendation.
5	Machine Learning in Soil Classification and Crop Detection	Image acquisition, Pre-processing, KNN, Feature extraction, SVM classifier.
6	Prediction Of Soil Quality Using Machine Leaning Techniques	Decision Tree and Random forest algorithms
7	Artificial intelligence system for supporting soil classification	Neural network models are used, CNN, steepest descent method.
8	Soil Classification & Characterization Using Image Processing	Technically proposed system has been based on HSV, Enhancement algorithm, and SVM classification algorithm.
9	Soil Physical Properties	Stokes regulation, bulk density, particle density, mass flow
10	Chemical properties of soils	Different types of chemical organic molecules.
11	Nutrients Detection in the Soil: Review Paper	Soil Nutrient detection,Reflectance sensing, Electro chemical sensing, Electro Conductivity sensing.
12	Soil Classification Methodology: Critical Analysis	Classification Methods, Inventory Methods.
13	Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction	Data mining, classification, regression, soil testing, agriculture , WEKA Tool.
14	Soil Classification & Characterization Using Image Processing	Classification of soil by image processing using SVM technique, Machine Learning.
15	Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series	Classification of soil by image processing using SVM technique based on colour, energy, HSV of soil.
16	Determination of Soil Nutrients and pH level using Image Processing and Artificial Neural Network	Rapid soil Testing, Soil Test kit, Artificial Neural Networks, Image processing are used.
17	Soil Quality Measurement using Image Processing and Internet of Things	IOT and image processing to measure pH, moisture present, Soil Nutrients. Raspberry-Pi and camera are used.
18	Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach	Naive Bayes and K- Nearest Neighbour methods are used.

Prediction soil type was performed by deep learning model using neural network. Research carried out for three classes of soil, specifically for clay, sand, and gravel, an AI model, was constructed that was keen efficient homogeneity of the images utilized and the research has shown that this AI model can be applied to conceive judgments on soil classification [9].

The SVM classifier was found to outperform over all other techniques. Maximum Likelihood classification, sub-pixel classification, and ANN classification were chosen as image classification techniques from various methods [11].

Soil classification and segmentations are necessitated, along with soil nutrients to recommend crops and fertilizers to use. Non-specialists farmers can understand features easily [10].

Researcher T.Abimala proposed a model to predict the variety of soil applying image processing techniques. Classification of soil is by the processing of soil image color and texture patterns [12].

Soil Characteristics have a notable impact on the response of soil in the field of agricultural uses. Soil aeration depends on the soil texture, is a static property of soil quality. Plants' growth depends on soil quality and temperature [13]. Weathering processes may create porous media at the land surface. The correlation between soil and bio-indication may not be strong enough to bring out such suitable changes. Soluble elements are purged into the lower layers of soils where they accumulate. Insoluble chemical elements remain in the upper layers of the soil [14].

Soil nutrients and pH define through image processing and ANN [15]. The process of soil classification systems based solely on grain diameter is capable of misunderstanding because the physical properties of outstanding soil elements depend on many factors other than grain size [16]. There is also the quantity of water (humidity) which is also a non-negligible factor. It is the newest and most used because it is the improved form of all others [17].

The least median squares regression can predict soil is identified to generate abler results than the classical linear regression by analyzing soil sample and cropping pattern [18].

Image processing using SVM techniques classify the soil based on color, energy, and HSV [19]. The researcher recommended a suitable crop for a particular soil by analyzing the PH value, moisture level, and nutrients collected using sensors with IoT technology. Here former need to have a piece of knowledge on hardware [20].

Experimental study says yield prediction is possible with AI Techniques. ANN attained more accuracy than RF, MLR, and NB Techniques [21]. Categorization of soil could be done using Data mining techniques like K-Nearest neighbors and naive Bayes as low, medium, and high. It helps to choose suitable land for more profitable crop production [22]. SVM, principal component analysis can be added further.

The CNN-based regression model is developed based on the characteristics of the TIR images, with required generality and performance. This model, in comparison with DNN, gives

promising accuracy. Remote digital imaging using a drone is advantageous than data acquired using satellite [23].

There are non-exhaustive methods utilizing image processing, machine learning, and deep learning in previous research work related to soil classification [24]. And many researchers focused on traditional methods, image processing, computer vision techniques, and computer-based applications. This paper proposed advanced soil classification techniques using machine learning and deep learning. The primary purpose of soil category is to forerun behavior, determine the uses, assess their productivity, and extrapolate analysis to predict nutrients.

This paper discusses the architecture of the proposed system, methodology, results, concluding remarks, and future scope.

III. METHODOLIGY

In Telangana, 48% of the land covered red soil and types of red soil and color and causes listed in Table II.

Some of these soils look reddish due to the wide diffusion of iron in crystalline and metamorphic rocks, and some are yellow due to diffusion occurring in a hydrated form. Soil is the foothold for plants' roots and holds the necessary nutrients for plants to grow. It filters rainwater and regulates the discharge of excess rainwater, preventing flooding, and it is capable of storing large amounts of organic carbon. All these depend on the characteristics of the soil. The primary skill to succeeding in farming here understands soil type and giving the most fitting plants for that soil [14].

Soil classification could be made as per engineering properties, and it could be free from trouble-free from field survey and the usually, engineers classify soils as per soil characteristics. Classification would commence from the soil image. Convolution Neural Network is an incredible image recognition invention and is a sub-class of deep learning neural networks. CNN generally used to analyze images and image classification.

CNN Architecture Fig. 1 has four layers listed as follow:

- Convolution layers.
- ReLu layers.
- Pooling layers.
- Fully connected layers.

TABLE II. RED SOIL CLASSIFICATION

S.No	Soil Type	Colour	Chemical
1	Red Clay Soil	Red	Iron oxide
2	Red loam soil	Red	Potash
3	Red Laterite Soil	Red	Iron and Aluminium
4	Red –Yellow soil	Red-Yellow	Ferric hydroxide
5	Red Sandy soils	Red	Iron
6	Red Gravel Soil	Red	Iron

A. Convolution Layer

- Convolutional layers: these layers apply a convolution operation to the input image and pass the information on to the next layer.
- Pooling Layer: The next layer is the pooling layer, where outputs combined as cluster of neurons into a single layer. And the next layers are fully connected layers, in which all neurons connect to each neuron in the next layer.

B. Relu

Proposed CNN Model and trained it on a Soil image dataset. The results observed that performance was minimal. Training a model with millions of images takes days to achieve high performance in real-world applications. An alternative is to use a pre-trained model, and it would retain on our dataset as transform learning.

In this Model, an image was directly given to the algorithm, which will classify the given image as either red soil or not, as shown in Fig. 1 and the detailed flow explained in Fig. 2.

C. Pre-Processing

A smoothing (low pass) filter has been used to eliminate high-frequency noise and artifacts from the picture. Smoothing filters employ a moving window operator that adjusts the value of one pixel of an image at a time based on a function of a local area of pixels. As the operator advances over the picture, all of the pixels are affected. As a consequence, the smoothing filter progressively enhances the image over time by eliminating imperfections.

D. Feature Extraction

The feature extraction stage is the essential phase in the process. It encompasses all of the characteristics that are needed to classify the soil type, such as texture, color, and intensity was extracted. As a result, a metric known as color moments was employed to distinguish photographs based on their color characteristics.

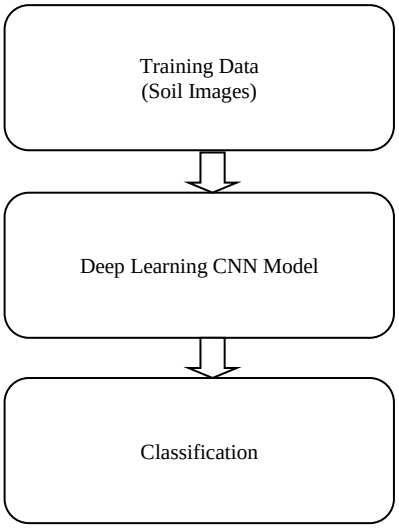

Fig. 1. Abstract Level Soil Classification Architecture.

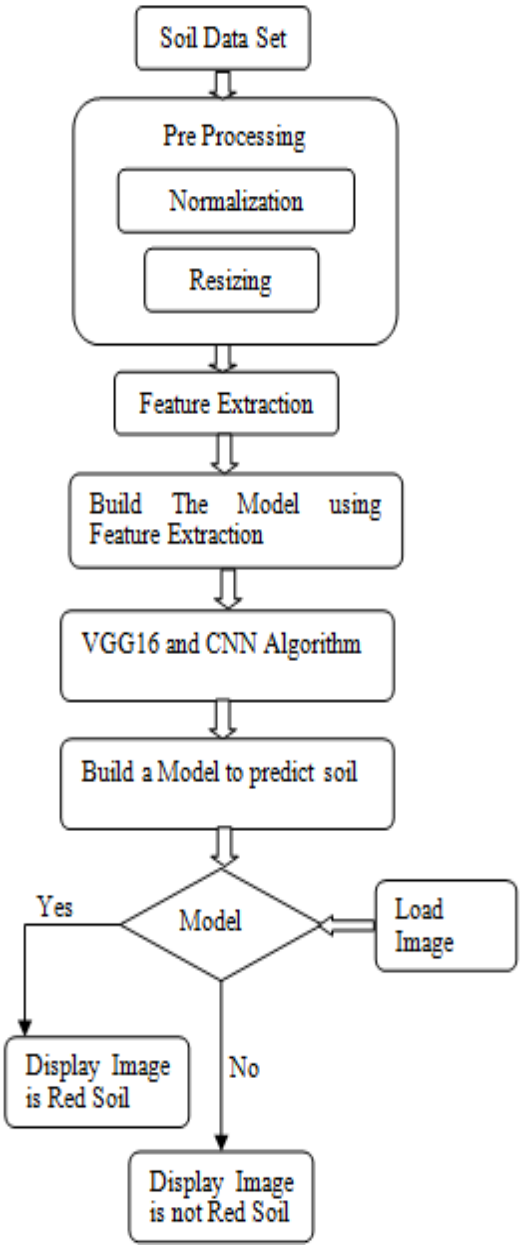

Fig. 2. Detailed Soil Classification Architecture.

E. CNN Classification Model

CNN or ConvNet are deep neural networks used for image recognition, object detection, and classification. Image classification is the process of deciding which class (or combination of categories) best describes an input image. In CNN, we take a picture as an input, assign weight to the image's various aspects/features, and differentiate one from the other. CNN requires much less pre-processing than other classification techniques.

F. Softmax

Softmax is an intriguing activation function since it maps our output to the [0, 1] range and maps each output so that the entire sum equals one. Softmax's output is a probability distribution as a result as shown in Fig. 3.

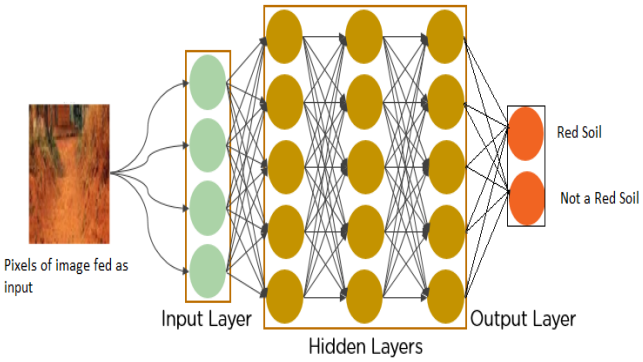

Fig. 3. Soil Classification with CNN.

In-Training the Model, all the images are converted into arrays and stored in Data Variable. Each image is given labels for the detection and stored in Variable Labels.

After preprocessing the steps to be done to build the model Average pooling is applied to down sample the input dimension, Height, and width by taking average value over a window of size 4x4 for each input channel. Strides shift the window along each dimension.

Flatten transforms the pooled feature pictures into a single column, then sent to the fully linked layer. Dense added the fully connected layer to the convolution neural network, and dropout is a strategy for preventing over fitting in a model. Once the model has fitted with the Layers, the trainable parameters have been set.

Fig. 4 shows how many total parameters are there and how many are available for the training.

```
In [19]: bmodel.summary()
Model: "vgg16"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 224, 224, 3)]     0
_________________________________________________________________
block1_conv1 (Conv2D)        (None, 224, 224, 64)      1792
_________________________________________________________________
block1_conv2 (Conv2D)        (None, 224, 224, 64)      36928
_________________________________________________________________
block1_pool (MaxPooling2D)   (None, 112, 112, 64)      0
_________________________________________________________________
block2_conv1 (Conv2D)        (None, 112, 112, 128)     73856
_________________________________________________________________
block2_conv2 (Conv2D)        (None, 112, 112, 128)     147584
_________________________________________________________________
block2_pool (MaxPooling2D)   (None, 56, 56, 128)       0
_________________________________________________________________
block3_conv1 (Conv2D)        (None, 56, 56, 256)       295168
_________________________________________________________________
block3_conv2 (Conv2D)        (None, 56, 56, 256)       590080
_________________________________________________________________
block3_conv3 (Conv2D)        (None, 56, 56, 256)       590080
_________________________________________________________________
block3_pool (MaxPooling2D)   (None, 28, 28, 256)       0
_________________________________________________________________
block4_conv1 (Conv2D)        (None, 28, 28, 512)       1180160
_________________________________________________________________
block4_conv2 (Conv2D)        (None, 28, 28, 512)       2359808
_________________________________________________________________
block4_conv3 (Conv2D)        (None, 28, 28, 512)       2359808
_________________________________________________________________
block4_pool (MaxPooling2D)   (None, 14, 14, 512)       0
_________________________________________________________________
block5_conv1 (Conv2D)        (None, 14, 14, 512)       2359808
_________________________________________________________________
block5_conv2 (Conv2D)        (None, 14, 14, 512)       2359808
_________________________________________________________________
block5_conv3 (Conv2D)        (None, 14, 14, 512)       2359808
_________________________________________________________________
block5_pool (MaxPooling2D)   (None, 7, 7, 512)         0
=================================================================
Total params: 14,714,688
Trainable params: 14,714,688
Non-trainable params: 0
```

Fig. 4. Listing of Parameters in Training.

IV. Result

Every step picks images from the training set at random. Then, each step finds its bottlenecks from the collection. And ultimately, each step puts the images into the final layer to get predictions. Then, the forecasts are equated alongside the actual labels to inform the final layer's weights using the back-propagation procedure.

Each step shows the level of training accuracy and validation accuracy Fig. 5. Training accuracy was the prediction of training images that were classified correctly. Quality of the model measured by validation accuracy and to what extent it guess depending on data it has not seen before. Suppose 86% percentage on the training set and 85% on the validation set. You can expect your model to perform with 85% accuracy on new data.

From Fig. 6 Prediction Accuracy into a percentile, the accuracy of the prediction soil image classifier is 91%.

This model could test using User Interface (UI). Once the model was ready, we developed a web application using the Flask framework. Here users can upload images to predict soil type.

```
Compiling Starts
Epoch 1/10
39/39 [==============================] - 89s 2s/step - loss: 0.6968 - accuracy: 0.5648 - val_loss: 0.5901 - val_accuracy: 0.812
5
Epoch 2/10
39/39 [==============================] - 92s 2s/step - loss: 0.5955 - accuracy: 0.7078 - val_loss: 0.5456 - val_accuracy: 0.737
5
Epoch 3/10
39/39 [==============================] - 92s 2s/step - loss: 0.5172 - accuracy: 0.7922 - val_loss: 0.4793 - val_accuracy: 0.812
5
Epoch 4/10
39/39 [==============================] - 92s 2s/step - loss: 0.4840 - accuracy: 0.8019 - val_loss: 0.4508 - val_accuracy: 0.825
0
```

Fig. 5. Training the Model.

	precision	recall	f1-score	support
Soil_Dataset//Black_Soil	0.93	0.91	0.92	43
Soil_Dataset//Red_Soil	0.89	0.92	0.91	37
accuracy			0.91	80
macro avg	0.91	0.91	0.91	80
weighted avg	0.91	0.91	0.91	80

Fig. 6. The Accuracy of the Model.

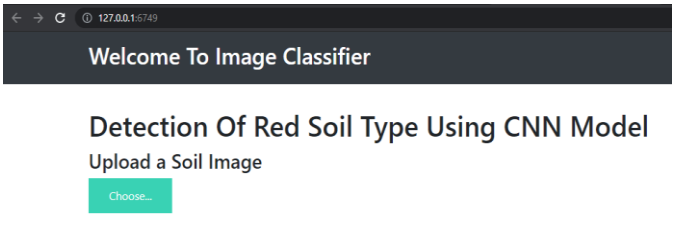

Fig. 7. UI of Web Application.

Fig. 7 demonstrates the web application with one button to upload an image of soil. After uploading an image predicted labeled button will appear on the screen, as shown in Fig. 8. User needs to click on predict button and then display the result as RED Soil as shown in Fig. 9. The exact process is repeated with a different image and displays the result as NOT a RED SOIL, as shown in Fig. 10.

Fig. 8. Uploading Soil Image.

Fig. 9. Prediction of Soil Type.

Fig. 10. Testing Application with other Soil Image.

V. Conclusion

The proposed model categorizes soil using CNN successfully identified the Red soil and tested in the field got 91 percent accurate results. And can estimate the soil's fertility by predicting significant nutrients for specific soil. Introducing cutting-edge technologies in Agriculture can improve the yield by applying adequate application fertilizers, and with an improved dataset, also get 99 percent accurate predictions.

References

[1] N Lakshmi Kalyani, Kolla Bhanu Prakash, "Soil Synthesis and Identification of Nitrogen percentage in Soil using Machine learning algorithms and Augmented Reality – A Typical review", IJETER, 2020(8), PP.5501-5505.

[2] Shravani.V,S.Uday KiranG, "Soil classification and crop suggestion using machine learning," IRJET,2020.

[3] Prakash K.B., Dorai Rangaswamy M.A. Content extraction studies using neural network and attribute generation, 2016, Indian Journal of Science and Technology, 9-22, 1(10).

[4] Prakash K.B. "Information extraction in current Indian web documents",2018,International Journal of Engineering and Technology(UAE),7,2.8,68-71.

[5] Chandan , Ritula Thakur`, "An Intelligent Model for Indian Soil Classification using various Machine Learning Techniques," IJCER, 2018.

[6] Dido, A.A., Krishna, M.S.R., Singh, B.J.K., Tesfaye, K., Degefu, D.T. "Assessment of variability of yield affecting metric characters in barley (Hordeum vulgare) landraces " Research on Crops, 2020, 21(3), pp. 587–594.

[7] Mrs. N.Saranya,Ms. A. Mythili, "Classification of Soil and Crop Suggestion using Machine Learning Techniques," IJERT,2020.

[8] Ashwini Rao ,Janhavi U ,Abhishek Gowda N ,S Manjunatha, Mrs.Rafega Beham "Machine Learning in Soil Classification and Crop Detection," IJSRD, 2018.

[9] Shinya Inazumi, Ph.D. , Sutasinee Intui, M.Eng. , Apiniti Jotisankasa, Ph.D. , Susit Chaiprakaikeow, Ph.D. , Kazuhiko Kojima, "Artificial intelligence system for supporting soil classification," ELSEVIER, 2018.

[10] Hement Kumar Sharma, Shiv Kumar, "Soil Classification & Characterization Using Image Processing," IEEE , 2018.

[11] Priyanka Dewangan, Vaibhav Dedhe, "Soil Classification Using Image Processing and Modified SVM Classifier", IJTSRD, 2018.

[12] Vijaya Nagini, D., Krishna, M.S.R., Karthikeyan, S. "Identification of novel dipeptidyl peptidase-iv inhibitors from ferula asafoetida through gc-ms and molecular docking studies," IJPT, 2020, pp.5072-5076.

[13] Sk Al Zaminur Rahman, Kaushik Chandra Mitra, S.M. Mohidul Islam, "Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series," IEEE, 2018.

[14] A. Balasubramanian,CHEMICAL PROPERTIES OF SOILS,2017.

[15] John Carlo PUNO , Edwin SYBINGCO , Elmer DADIOS , Ira VALENZUELA , Joel CUELLO, "Determination of Soil Nutrients and pH level using Image Processing and Artificial Neural Network," 2017.

[16] Ruwali A., Kumar A.J.S., Prakash K.B., Sivavaraprasad G., Ratnam D.V. Implementation of Hybrid Deep Learning Model (LSTM-CNN) for Ionospheric TEC Forecasting Using GPS Data,2021, IEEE Geoscience and Remote Sensing Letters,18(6),9093827,pp:1004-1008.

[17] Mbuya Mukombo Jr. , Mutonkole Ngomba H., Musimba Kasiya A , Ngoy Biyukaleza B., "Soil Classification Methodology: Critical Analysis," IJSR, 2018.

[18] Kunal Teeda, Nandini Vallabhaneni, Dr.T.Sridevi Jay Gholap, "Comparative Analysis of Data Mining Models for Crop Yield by Using Rainfall and Soil Attributes," IEEE, 2018.

[19] Hement Kumar Sharma, "Soil Classification & Characterization Using Image Processing," IEEE,2018.

[20] Dileep Reddy Bolla, Dr.Shivashankar, Anirudh Sandur, "Soil Quality Measurement using Image Processing and Internet of Things," 2019, IEEE.

[21] Vaishali Pandith , Haneet Kour , Surjeet Singh , Jatinder Manhas , and Vinod Sharma, "Performance Evaluation of Machine Learning Techniques for Mustard Crop Yield Prediction from Soil Analysis," JSR, 2020.

[22] Monali Paul, Santosh K. Vishwakarma, Ashok Verma, "Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach," IEEE, 2015.

[23] Bharadwaj, Prakash K.B., Kanagachidambaresan G.R. "Pattern Recognition and Machine Learning", 2021, EAI/Springer Innovations in Communication and Computing, 105-144.

[24] Prakash K.B. "Content extraction studies using total distance algorithm", 2017, Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016, 7912085, 673-679.