IJAREEIE
e-ISSN: 2278-8875
p-ISSN: 2320-3765

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

Volume 13, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.317

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305055|

Soil Classification Using Convolutional Neural Network

Thilagavathi.K[1], Tarunika.M[2], Sanjana.V[3], Sivani.S.G[4]

Assistant Professor, Department of Electronics, and Communication Engineering, Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India[1]

U.G. Students, Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India[2,3,4]

ABSTRACT: Soil is a natural resource. One of the biggest concerns in agriculture is the quality of the crop, which is determined by the quality of the soil, gradation level of the soil and production of crops. Soil classification helps in identifying the nature of the soil, their characteristics, the pH value, vegetation and also it describes the geographical location where it is present. Since digital image processing has been used for analysis in various fields of agriculture, the purpose of this work is to determine soil color and soil type using image processing and machine learning techniques. The colour of the soil is determined by using classification algorithms and image pre- processing techniques. In image processing, the training data undergoes rescale, flip, zoom, shear, height and width changes and the testing data is rescaled. Convolutional Neural Network (CNN) is used for the classification process. When the images are processed by CNN using 4 convolutional layers and 3 dense layers, the training and test datasets are trained for 30 epochs for error correction and accuracy improvement. Given an input image, CNN identifies the matching probability among the six specified classes and determines the class with the highest probability.

KEYWORDS: Classification, Convolutional Neural Network, Soil, accuracy

I. INTRODUCTION

Assessing soil parameters quickly and accurately is crucial for effective agricultural and natural resource management. Soils are typically difficult to detect using traditional sensing methods. Microwave radiation penetrates only a few centimetres of topsoil, while visible (VIS) and infrared radiations rarely penetrate the soil surface. To study and predict soil, it is essential to understand the properties of soil, as they affect environmental quality and are vital for agricultural purposes. Classifying soil based on their color involves processes such as image processing and machine learning using different algorithms. It also determines the type of soil present in an agricultural land, to cultivate suitable crops using remote sensing. The types of soil mentioned here are - Black Soil, Laterite Soil, Cinder Soil, Peat Soil, Clay Soil, Yellow Soil.

II. RELATED WORK

Deep learning techniques have been employed by earlier algorithms to predict soil. Based on the variables and parameters included in the model, all deep learning techniques are capable of predicting returns. However, the most effective deep learning methods for crop prediction are CNN and Long short-term Memory network (LSTM) based approaches. The second method uses different soil types and a CNN classification model. The classification results of different training samples are analyzed and compared with the support vector machine algorithm. Under the assumption that the Kennard-Stone algorithm divides the calibration set, the classification results of CNN for 6 different land types and 6 individual land types are better than those of SVM.In this paper, Zongyao Sha discussed a wide range of vegetation classification issues using remote sensing images. Although not all possible problems were covered, the basic steps, principles, techniques and methods of vegetation mapping from remote sensing images were discussed. [1]. Xueying Li, Pingping Fan, Zongmin Li, Guangyuan Chen, Huimin Qiu and Guangli Hou develop a convolutional neural classification model. Analysis and comparison of shallow network SVM classification results with -e classification results under different numerical testing conditions. [2]. Researchers Vinay Kumar Gadi, Dastan Alybaev used manual image processing to interpret soil surface moisture using the public ImageJ tool. The difference between the average gray values obtained by the above two approaches was very small [3]. Greema S Raj, Lijin Das S Soil classification is one of the most important problems in engineering. The soil classification process has been greatly

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305055|

improved through the use of machine learning techniques [4]. Gheorghe C.,DeacT.A, Filip N covered the acquisition of satellite images or using aerial devices such as drones and their processing using digital imaging techniques[5]. Kshirsagar, S., S1, L.P. and S2, D. Vibhute A discussed a condition based on a color image processing technique where a digital photograph of the sample was used to determine soil nitrogen, potassium and pH. It can be concluded that by using this technique, we can easily know that NK nutrients are very beneficial for plant growth [6]. Utpal Barma , Ridip Dev Choudhury , Niyar Talukdar , Prashant Deka has carried out the determination of pH level in the soil by digital image processing and Average fractal dimension of soil pH 1.51136 [7].Sudha.R, Aarti.S2, Anitha.S3, Nanthini.K has determined the pH index values 0.0071-0.0451 and 0.0084-0.0239 [8]. Carlo Puno1, J., & Sybingco1, E.,Joel cuello has determined the nutrients and ph level of soil using Image processing and Artificial Neural Network [9]. Kamble, U., Shingne, P. and Kankrayane, R. conducted experiments to record the pH values of seventy soil samples in a database. The process of manual soil testing, if not done correctly, will also affect the initial result. So the software gives the result with 60-70% accuracy [10]. K. Thilagavathi et al. discussed that hyperspectral imaging has hundreds of spectral groups in contrast to the classic digital image of three spectral bands, and therefore it provides a more complete part of the light spectrum for imaging and anatomization. Hyperspectral imaging offers the ability to gather more precise and comprehensive information than any other information ever tasted. Due to the huge amount of data in many spectral bands, the type of hyperspectral remote sensing data is a more sensitive task. To solve this problem, a type based on dimensional reduction is made. Ensemble data is one notable type system among many type styles that works by dividing maturity classes into different small-scale groups. In this work, Ensemble Support Vector Machine (SVM) is used to improve type results. The experiment uses sedateness-standard hyperspectral data, including data from AVIRIS Indian Pines, AVIRIS Salinas, and ROSIS University of Pavia, and it is concluded that the proposed system shows a significant improvement over genres. Indian pines have an image enhancement of 15, University of Pavia of 18 and Salinas of 13. The results are cross validated against realism for sensitivity. [11]. K. Thilagavathi discussed the classification of remote sensing images using surveillance methods in this paper [12]. Thilagavathi, K and Vasuki highlighted various classification Techniques for the Hyperspectral Remote Sensing Images [13]. Rasikapriya et al mentioned that the hyperspectral imagery provides detailed information about the Earth's surface and its object and also pointed out the spectral spatial methods for classifying the images[14]. Hyperspectral Image Classification using MLL and Graphcut Methods' were discussed in [15]. Also a novel multidimensional recurrent deep neural network was proposed for hyperspectral image classification. [16].

III. METHODOLOGY

Image Processing is the methodology used to predict the soil type. It is important to understand what an image is and what exactly constitutes an image before going to image processing. Based on the quantity of pixels, images are represented by height and breadth. For instance, if the image is 500 by 500 (width x height), there are 250000 pixels in the entire image. This pixel is a location on the image with a certain hue, tint, or level of opacity. It typically takes one of the following forms:

- Grayscale: A pixel is an integer having a value between 0 and 255. (where 0 is all black and 255 is all white).
- RGB: A pixel is made up of three numbers ranging from 0 to 255. (the integers represent the red, green, and blue intensities).
- RGBA: An expansion of RGB that includes a second alpha field to reflect the image's opacity. Image processing is the process of digitizing an image and applying various changes to it in order to extract valuable information from it. When using specific preset signal processing techniques, image processing systems typically regard all images as 2D signals. Each pixel of an image must undergo a specific set of processes during image processing. A initial set of actions are carried out pixel-by-pixel by the image processor on the image. It will begin carrying out the second procedure once it has finished the first one in full. Each pixel in the image can compute the output values of these processes.

The steps used in image processing are:

- Visualization - Finding objects not visible in the image
- Recognition - Identifying or recognizing objects in images
- Sharpen and restore - Creating enhanced image from original image
- Pattern Recognition - Measuring various patterns around objects in an image
- Search - Browsing and finding images similar to the original image from a large database of digital images.

A. CONVOLUTIONAL NEURAL NETWORKS:

Deep learning approaches are based on neural networks, which are a subset of machine learning and are often referred to as Artificial Neural Networks (ANNs) or Simulated Neural Networks (SNNs). In order to reflect the communication between organic neurons, their structure and nomenclature are fashioned after the human brain. The network's node

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305055|

layer is made up of an input layer, one or more hidden layers, and an output layer. Each node has an associated threshold and weight, and they are all connected to one another. The node is activated when its output exceeds the threshold, and it provides data to the uppermost layer of the network. Otherwise, no data is transmitted to the network's next tier. The CNN architecture is shown in Fig.1. CNNs are a type of deep learning network designed for processing organised arrays of data, like collections of images. CNN is frequently used in computer vision and has developed to the cutting edge in many visual applications, such as image categorization. Natural language processing has also proved successful in classifying texts. The patterns in the input image, such as lines, gradients, circles, or even eyes and faces, are very well recognized by CNNs. CNN is extremely effective for computer vision because of this quality. CNN does not require any preparation and can operate immediately on a raw image, in contrast to older computer vision methods.

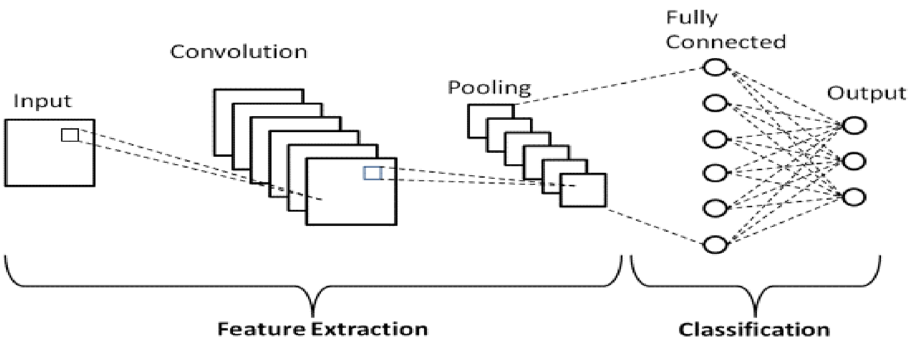


Fig.1- CNN architecture

A feed-forward neural network with up to 20 or 30 layers is known as the CNN which consists of a unique kind of layer that gives CNN its power. In this work, we have used an input layer, convolution layer 1, max pooling layer 1, convolution layer 2, max pooling layer 2 and the output layer. Many convolutional layers are placed on top of one another in CNN, and each layer can identify more complex structures. A convolutional element extracts local features from images and a fully connected component maps the learned functions to output. CNN learns its spatial characteristics by altering its parameters through version observation and no longer relies on manual capabilities.

B. PROCESS FLOW:

In order to get trained weights, which are data patterns or rules from extracted images, we feed a 2D input image to the CNN during training. RELU is the activation function in use here. Apply filters or feature detectors to the RELU activation function to create feature maps or activation maps. They aid in recognising features including edges, bends, edges, and vertical and horizontal lines. The data is subsequently filtered using pooling over the feature maps. In comparison to minimum or average pooling, maximum pooling gives superior performance. The deep neural network receives the input after it has been flattened, and it outputs the object's class. An image's class can be categorical or binary. Since there are five different types of soil, this image's class is categorical.

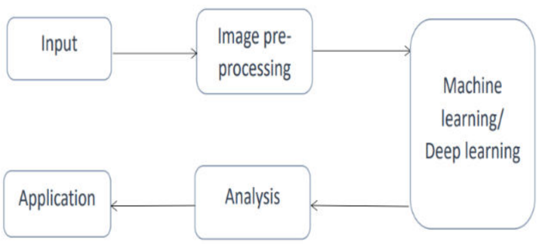


Fig.2 –Processing Steps

CNN adjusts a filter's depth to correspond to the depth of the input image. The three channels are each given a unique conv2D filter. (RGB). A convolutional network's initial layers extract high-level information, whilst the deeper layers learn more computationally costly characteristics. Importing the necessary libraries, creating a dataset (training and test data) for image processing, setting important parameters, rescaling validation data, creating a deep CNN model, and training the data are the first steps in building the CNN. The various processing steps are shown in fig.2.

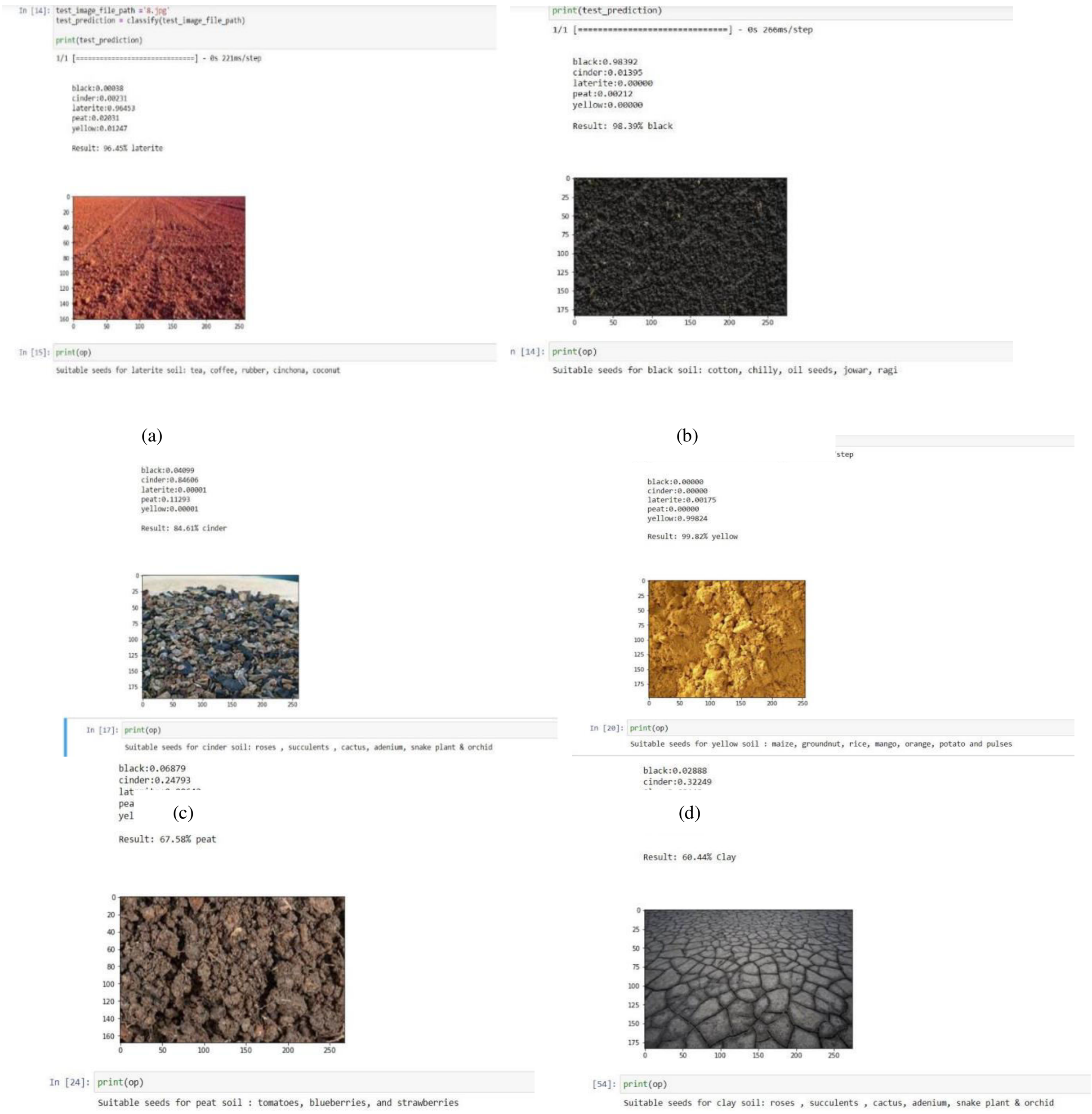
| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305055|

IV. EXPERIMENTAL RESULTS

When the image is processed using CNN, with four convolutional layers and three dense layers, the training (70%) and testing (30%) data sets are trained with 30 epochs, for error correction and for improving accuracy. For the given input image, under the six classes specified, CNN identifies the matching probabilities of each soil and determines the soil with the highest probability. Figure 3 represent the outputs of laterite, black, cinder, yellow, peat and clay soils.



(a) (b)

(c) (d)

(e) (f)

Fig 4.6 Peat soil output

Fig. 3. Soil classification using CNN (a) Laterite soil output (b) black soil output (c) cinder soil output (d) yellow soil output (e) peat soil (f) eat soil output

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

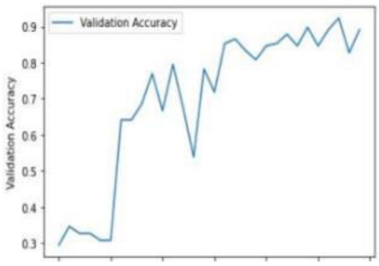
||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305055|

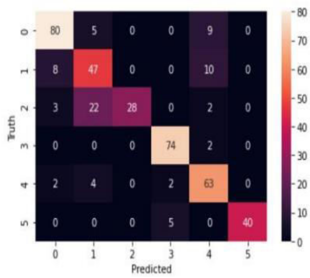
TABLE 1 Classification analysis

TYPES OF SOIL	SUITABLE SEEDS	ACCURACY
Black soil	Cotton, Chilly, oil seeds, jowar, ragi, wheat, castor.	98.39%
Laterite soil	Tea, coffee, rubber, cinchona, coconut.	96.45%
Cinder soil	Roses, succulents, cactus, adenium, snake plant, orchid.	84.61%
Peat soil	Tomatoes, blueberries, strawberries.	67.58%
Yellow soil	Maize, groundnut, rice, mango, orange, potato, pulses.	99.82%
Clay soil	Broccoli, brussel sprouts, cabbage, cauliflower, kale.	60.44%

Evaluation metrics are used to measure the quality of a statistical or machine learning model. Evaluating machine learning models or algorithms is essential in all projects. There are many different evaluation metrics available for model testing. These include classification accuracy and log loss. The linear plot and heatmap are shown in fig.4. Classification accuracy is the ratio of the number of correct predictions to the total number of input samples, which is usually referred to when we use the term accuracy. Logarithmic loss works by penalizing misclassification. The confusion matrix gives us a matrix as an output and describes the full performance of the model. Evaluation metrics use a combination of these individual evaluation metrics to test a model or algorithm.



(a)



(b)

V. CONCLUSION

The soil colour of the image is discovered using the RGB values and the appropriate algorithm, and the soil type is consequently identified. The dataset and machine learning algorithms can be used to identify the soil type, and then the pH value can be determined. CNN provided 99.82% accuracy for yellow soil and 98.39% for black soil.It is also possible to classify the crops that are best suited based on the discovered soil type and the appropriate pH. Another future development could be deploying this machine learning process as a webapp where the input given by the user can be tested and then classify them among the classes of soil and display along with its properties and uses.

REFERENCES

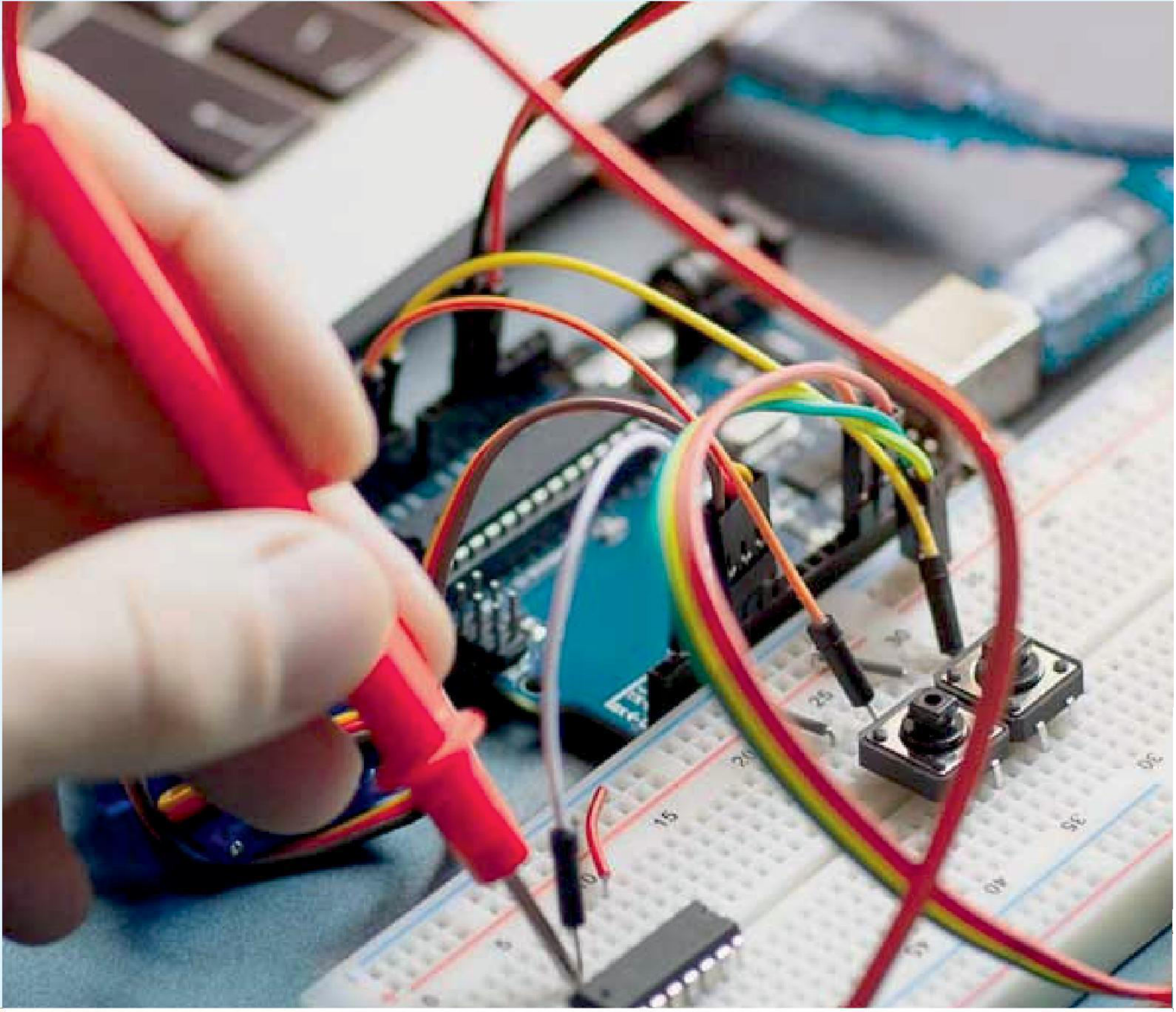
1. Zhao_Liangli, P., Leng, C., & Zhou, K.-S. (2021), Soil moisture retrieval from remote sensing measurements: Current knowledge and directions for the future. Earth -Science Reviews, 218.
2. Li, X., Fan, P., Li, Z., Chen, G., Qiu, H., & Hou, G. (2021), Soil classification based on deep learning algorithm and visible near-infrared spectroscopy. Journal of Spectroscopy, 2021.

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305055|

3. Vinay Kumar Gadi, D. (n.d.). (2020), A Novel Python Program to Automate Soil Colour Analysis and Interpret Surface Moisture Content. International Journal of Geosynthetics and Ground Engineering, 6.
4. Greema, S,Raj., Lijin Das S.(2020), 'Survey on soil classification using different techniques. International Journal of Engineering and Technology, 7, 5233–5236.
5. Gheorghe, C., Deact, A., & Filip, N. (2019), Image processing techniques used in soil moisture analysis. Technical University of Cluj, 58(2).
6. Kshirsagar, S., S1, L. P., & S2, D. Vibhute A. (2018), Soil Nutrients Analysis using Colour Image Processing. International Research Journal of Engineering and Technology.
7. Utpal Barman., Choudhury, R. D., Talukdar, N., & Deka, P. (2018), Predication of soil pH using HIS colour image processing and regression over Guwahati, Assam, India. Journal of Applied and Natural Science, 10(2), 2231–5209.
8. Nanthini, K. (Ed.). (2017), Determination Of Soil Ph And Nutrient Using Image Processing. International Journal of Computer Trends and Technology (IJCTT).
9. Carlo Puno1, J., & Sybingco1, E.,Joel cuello. (2017), Determination of Soil Nutrients and pH level using Image Processing and Artificial Neural Network .
10. Kamble, U., Shingne, P., & Kankrayane, R. (2017) | ISSN (online). Shreyas Somkuwar4 Prof. Sandip Kamble. *Testing of Agriculture Soil by Digital Image Processing*,
11. K. Thilagavathi., R. Nagendran. and I. Thusnavis Bella Mary,2021 "Hyperspectral Image Classification using Ensemble Average Method," IEEE Explore. pp. 1-6, doi: 10.1109/ICAECA52838.2021.9675659
12. K. Thilagavathi, K. Kavitha, S. Dharani., R. Pavithra. and N. Subasri 2021, "Classification of Remotely Sensed Images using Supervised Methods," IEEE Explore pp. 1-5, doi: 10.1109/ICAECA52838.2021.9675511.
13. Thilagavathi, K & Vasuki, A 2019, 'A Comprehensive Review on Classification Techniques for Hyperspectral Remote Sensing Imagery', Caribbean Journal of Science, vol. 53, no. 1, pp. 291-313.ISSN: 0008-6452. Impact factor.0.2.
14. Rasikapriya R,Thilagavathi K and Vasuki A 2017, 'Classification of hyperspectral images using spectral spatial method',IEEE Explore. DOI: 10.1109/ICIIECS.2017.8276129.
15. Priyadharshini.P and K.Thilagavathi, 'Hyperspectral Image Classification using MLL and Graphcut Methods' in IEEE Explore. DOI: 10.1109/GET.2016.7916639 .pp:1-6.
16. Thilagavathi, K & Vasuki, A 2019, 'A Novel Hyperspectral Image Classification Technique Using Deep Multi-Dimensional Recurrent Neural Network', Applied Mathematics & Information Sciences, vol. 13, no. 6, pp. 955-963.ISSN: 2325-0399.



International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering