

International Journal Research Publication Analysis

Page: 01-08

AUTOMATED SOIL NUTRIENT ANALYSIS FOR MACRO AND MICRO ELEMENTS USING IMAGE PROCESSING AND MACHINE LEARNING

***¹Prof. Jayaprasad K. M., ²Sudarshan Kumar Iranna Kudaki, ³Darshan S., ⁴Vishal Chowhan ⁵Prajwal Ugargol**

¹Professor ECE department, Amruta Institute of Engineering and Management Sciences, Bidadi, Bangalore.

^{2,3,4,5}Student ECE department, Amruta Institute of Engineering and Management Sciences, Bidadi, Bangalore.

Article Received: 12 September 2026

Article Revised: 02 October 2026

Published on: 22 October 2026

***Corresponding Author: Prof. Jayaprasad K. M.**

Professor ECE department, Amruta Institute of Engineering and Management Sciences, Bidadi, Bangalore.

DOI: <https://doi.org/10.1555/ijrpa.2280>

ABSTRACT

Soil nutrient analysis is essential for precision agriculture in ensuring proper fertiliser utilisation, increasing productivity, and sustainable farming. Traditional lab tests are tedious, expensive, and not feasible for mass usage. In this paper, four state-of-the-art scientific papers that have implemented ML/DL approaches to address problems related to soil nutrient analysis are reviewed. The first paper used CNN on soil RGB images taken from Karnataka, India, which achieved 95.51% accuracy in classifying three different soil types, namely gravel, sand, and silt. The second paper incorporated IoT-enabled sensor devices to gather soil micronutrient data from mulberry farms in Tamil Nadu. It used an ELM classifier with 88% accuracy after applying k-fold cross validation on the data. In the third paper, hyperspectral cameras (398-1003 nm) were mounted on UAVs to capture spectral images of a 65,000 m² farmland in Shaanxi, China. RF-DE feature selection was employed alongside BP neural networks to predict available potassium with R² up to 0.95. Finally, a Bayesian-optimised 1D-CNN was proposed by the fourth paper on the pan-European LUCAS topsoil dataset (approximately 20,000 samples). It scored R² of 0.75 (nitrogen), 0.66 (potassium), and 0.46 (phosphorus), with 25% faster training times.

KEYWORDS: *Soil Nutrient Classification, Deep Learning, Convolutional Neural Network,*

Hyperspectral Remote Sensing, Internet of Things Agriculture, Extreme Learning Machine, Bayesian Optimisation, Precision Agriculture, Unmanned Aerial.

I. INTRODUCTION

In the world today, agriculture is still the key player in food security and rural development. According to the Indian government, more than 58% of Indians derive their livelihoods from agriculture; however, most small farmers do not have access to proper information about their soil's health status. The nutritional content of the soil, which includes elements such as nitrogen (N), phosphorous (P), potassium (K), organic carbon (OC), and micronutrients such as zinc (Zn), iron (Fe), and boron (B), determines the yield of the harvest.

Traditionally, soil sampling and testing require the farmer to collect soil samples physically and send them to a laboratory where an expert interprets the data collected. Soil health cards in India are an example of efforts made by governments to provide soil-related information to farmers; however, such processes take several days and are quite expensive. New innovations in ML/DL technologies, together with digital imaging and Internet of Things (IoT) sensor devices, could revolutionize soil analysis processes.

The current paper offers an overview and synthesis of four scientific articles published in peer-reviewed journals recently. Each of them presents a different approach towards soil nutrients identification using AI technologies: (1) image recognition by CNN classifiers [1]; (2) IoT sensors data analysis by ELM [2]; (3) hyperspectral remote sensing by UAVs and RF-DE-AHP-BP [3]; and (4) 1D-CNN with Bayesian optimisation on continent-scale spectroscopic database [4].

II. LITERATURE REVIEW & RESEARCH GAP

There is considerable literature on the use of ML in soil science. SVM models have been extensively employed in soil texture classification, with Rahman et al. [5] showing that SVM has better accuracy than KNN models in predicting soil series. RF models have been useful for mapping soil in space, as illustrated by Shukla et al. [6]. ANN models with backpropagation methods have also been used in soil fertility estimation using chemical predictors [7].

Recent developments in deep learning have provided additional possibilities: Li et al. introduced the design of 3D-CNN models for hyperspectral soil classification [8], and Steinbrener et al. utilized the CNN model for fruit and vegetable classification from hyperspectral images [9], which can be considered as a similar domain. Smart-farming

models using IoT technology have made it possible to stream soil data in real time [10]. Nevertheless, there are a number of areas that need to be explored in order to fill knowledge gaps. These are, firstly, the lack of image analysis techniques that consider nutrient content and not just soil texture type; secondly, limited geographic scope of IoT-sensor research which does not allow generalizing for other soil ecosystems; thirdly, high costs associated with UAV hyperspectral approach due to high prices of specialized equipment and skilled personnel; fourthly, lack of validation of European-trained ML models (e.g. on LUCAS data) when applied to tropical/subtropical Indian soil; fifthly, absence of comparison of all the four mentioned techniques.

III. PROBLEM STATEMENT & OBJECTIVES

Even with the existence of different AI solutions to measure soil nutrients, there is no consolidated comparative framework that measures these methods based on a common yardstick such as accuracy and hardware requirements. There is a need to determine what technology would be suitable for certain operations such as when using a smartphone camera only, IoT sensors, UAV operations, or massive spectral databases.

The aims of this paper include the following:

1. To conduct a systematic comparison of AI/ML approaches in analyzing the nutrients in the soil as presented in recent literature.
2. To analyze the accuracy, hardware requirements, scalability, and implementation capabilities of the techniques.
3. To analyze the benefits and drawbacks of CNN soil image classification, ELM IoT, UAV hyperspectral BP networks, and Bayesian optimized CNN soil nutrient analysis models.
4. To develop a context-specific framework of deploying these AI models according to farmer capabilities.
5. To develop the basis of soil intelligence platforms for Indian farmers.

IV. METHODOLOGY

A. Overview of the Comparative Framework

For the purpose of this study, systematic literature review and comparative analysis methods have been adopted. For selection of four research papers, the criteria used included: (a) publication in peer reviewed IEEE or indexed journals from 2022 to 2025; (b) use of ML/DL for soil nutrients classification/prediction; (c) presentation of quantitative accuracy measures; and (d) difference in modality of data used

(image, sensors, spectral). Five aspects that were considered while analyzing each paper were as follows:

B. Study 1 — CNN Image Classification (Rathore & Singh, 2022)

In this work, RGB images of the soil surface for three different soils (Gravel, Sand, and Silt) were collected in Karnataka, India, and a classifier was trained using CNN and TensorFlow and OpenCV packages. The CNN model architecture consisted of Convolutional layers, Pooling layers, Flatten layers, and Fully Connected Dense layers with Dropouts (rate 0.1–0.4 in each layer). Data augmentation techniques such as Horizontal Flip, Rotate, and Shift were applied. The training process took place for 15 epochs with categorical cross-entropy loss function. The features were extracted automatically using the learning capabilities of CNN in extracting colour moments, texture variance, and gradients.

The model was tested on five images of each type of soil. The accuracy achieved in the final classification was 95.51% (e.g., Gravel 25.5%, Sand 1.5%, Silt 73.0%). SVM was used as the benchmark algorithm.

C. Study 2 — ELM with IoT Sensors (Sheeba et al., 2022)

Soil samples were taken from 1,000 plots in four mulberry cultivation areas in Tamil Nadu (namely Salem, Coimbatore, Tirupur, and Erode). Four micronutrients were estimated with atomic absorption spectrophotometry using DTPA as an extractor for estimation of Fe, Mn, Cu, and Zn. K was determined by flame photometry, N by Kjeldahl's digestion process, C by titration, while soil pH and EC were measured by taking 1:2.5 soil-to-water ratio suspensions. Training and classification with an Extreme Learning Machine (ELM) classifier involved 10-fold cross-validation (with 75% training and 25% testing samples) using five activation functions including HARDLIM, TRIBAS, SINSQ, TNAH, and Gaussian RBF. Parker's method was used for soil fertility estimation indices.

D. Study 3 — UAV Hyperspectral + RF-DE-AHP + BP Network (Liu et al., 2024)

Samples of soils were taken from 1,000 plots from four mulberry cultivation districts of Tamil Nadu namely Salem, Coimbatore, Tirupur and Erode having 250 samples from each district. Micronutrients were determined by atomic absorption spectrophotometry technique (using DTPA method for Fe, Mn, Cu and Zn); Flame photometry (K); Kjeldahl's digestion (N); Titration (C); and pH and EC from 1:2.5 soil:water suspension. The ELM classification technique was performed using 10-fold cross validation process (with 75% for training and 25% for testing data) using five activation functions like HARDLIM, TRIBAS, SINSQ, TNAH and Gaussian RBF. Soil fertility indexes were obtained using Parker's index.

E. Study 4 — Bayesian-Optimised CNN (Alfin et al., 2025)

In all, 1,000 soil samples were obtained from four mulberry-growing regions in Tamil Nadu

(Salem, Coimbatore, Tirupur, Erode). These samples were obtained at 250 samples each from these districts, and micronutrients were determined by atomic absorption spectrophotometry (DTPA extraction method for determination of Fe, Mn, Cu, Zn), flame photometry (K), Kjeldahl's method of digestion (N), titration (C), and pH/EC by soil/water ratio of 1:2.5. Extreme Learning Machine (ELM) classifier was implemented using ten-fold cross-validation (training sample = 75%, testing sample = 25%) using five types of activation function: hard limThe LUCAS 2009 topsoil dataset of Europe (~20,000 samples with spectra ranging between 400-2500nm in 25 nations) was employed with training, validation, and testing datasets in proportions of 70/15/15 respectively. Down sampling was done at every 10nm giving rise.

The proposed model consists of a 1D-CNN with an adaptive architecture (multiple convolution layers with exponentially growing filters, LeakyReLU activation, batch normalisation, stride-2 downsampling). Optimisation was performed via Bayesian Optimisation (BO) with a Tree-Structured Parzen Estimator (TPE) as the surrogate model for the following hyperparameters: learning rate $\eta \in [10^{-4}, 10^{-3}]$, weight decay λ , filter exponential increase, target downsampling, and pooling type. Every BO iteration took up to 2,000 training steps with early stopping.

V. RESULTS & DISCUSSIONS

A. Classification and Prediction Accuracy

Table I presents the summary of the accuracy measures for each of the four studies. Study 1, where CNN is applied to image classification, had the best single accuracy measure of 95.51%. However, it should be noted that this only holds true for 3 texture categories and not for nutrient content. Study 3, using the UAV and BP neural network, exhibited the best nutrient content prediction in terms of the available potassium nutrient with an R^2 value of 0.95.

Table I: Comparative Accuracy Metrics Across Four Studies.

Study / Paper	Method	Best Accuracy / R^2	Nutrients / Properties Predicted
Rathore & Singh (2022)	CNN + OpenCV	95.51% (Classification)	Soil Type (Gravel, Sand, Silt)
Sheeba et al. (2022)	ELM (TNAH / Gaussian RBF)	~88% CV Accuracy	K, N, P, OC, B, Mn, Fe, Cu, Zn, S, pH
Liu et al. (2024)	RF-DE-AHP + BP Neural Net	$R^2=0.95$ (AK); $R^2=0.90$ (SOM)	SOM, TN, AP, AK — spatial maps
Alfin et al. (2025)	1D-CNN + Bayesian Optimisation	$R^2=0.75$ (N); $R^2=0.66$ (K)	N, P, K, OC, Clay, Silt, CEC, pH

B. Advantages and Limitations

Advantages and Disadvantages for Practical Deployment:

Study 1 (CNN Image): The key advantage is that the equipment is easy to access, requiring merely a smartphone camera. Chemicals and sensors are not required, rendering it feasible for use by farmers directly. The downside is that the scope is restricted to classification of texture (gravel, sand, silt), and nutrients cannot be quantified. Training data is small (15 images per class), and geographically, applicability is confined to Karnataka.

Study 2 (ELM + IoT): Practical deployment of the IoT in real-life makes it the best validated solution. It detects 11 micronutrients at once, yielding valuable district-wise soil fertility maps. The ELM requires little computation and trains much faster than a CNN. The disadvantage lies in the fact that it is applicable to Tamil Nadu mulberry plantation regions and possesses lower cross-validation (~88%) efficacy than deep learning approaches.

Study 3 (UAV+BP Network): This technique gives the best quantitative accuracy ($R^2=0.95$ for AK) along with creating spatial nutrient maps using GIS software. Methodology followed in this case, i.e., the RF-DE-AHP pipeline, is innovative because of its objective selection of characteristic bands in spectra. Limitations associated with this study include expensive UAV hardware amounting to over Rs. 5–10 lakhs, requirement of skillful pilots, and the test and train datasets collected at just one location of area size 65,000m² in China.

Study 4 (Bayesian CNN): This is the most statistically robust technique, which uses data from over 20,000+ soil samples taken from 25 different European countries with confidence intervals of 95%. Bayesian optimization proves to be better than grid search because of its principled nature and reduces training time by 25%. Phosphorus prediction is difficult ($R^2=0.46$), and the LUCAS database is Europe specific.

C. Comparative Discussion

It is clear that none of the methodologies are overwhelmingly dominant on all criteria for evaluation. Resource availability, geographical setting, and nutrient type should be considered when determining methodology. For image analysis of nutrients, CNN is the most feasible for use in the field, while for continuous nutrient monitoring using the IoT, ELM is best suited. UAV with hyperspectral imagery is best suited for large-scale precision agriculture, while Bayesian-optimized CNN is the methodology best suited to research applications that involve large spectra libraries.

The three studies involving IoT and spectral data indicate that nitrogen and potassium can reliably be predicted based on their spectral characteristics compared to phosphorus, owing to

its high spatial variability. This can be seen in Studies 2, 3, and 4. The ability for spatial visualization provided by Study 3 is clearly useful for practical purposes.

VI. CONCLUSION

In this study, a comparative analysis of four different AI-based techniques used in soil nutrient analysis has been discussed. It is clear from the study that although the field of soil nutrient analysis has progressed considerably due to the use of deep learning and machine learning techniques, there is no single technique which can be termed as optimal for all cases. The first technique based on CNN image classification (Study 1) has the highest level of deployability, with low hardware costs, attaining an accuracy rate of 95.51% in soil texture classification. The second technique using IoT and ELM (Study 2) has the maximum level of nutrients covered, along with actual deployment validation.

For the particular scenario of smallholder farmers in India, which usually have limited funds, simple mobile phones, and availment of government programs like the Soil Health Card, a multi-level solution would be suggested, which involves using image-based CNN for rapid soil texture analysis, along with the option of integrating IoT sensors for tracking nutrients, and UAV surveys run by the government or service providers for spatial analysis. Further research must be directed toward building light applications for soil analysis using CNN-based image classification along with edge-computing sensor analysis for Indian soils.

REFERENCES

1. M. Rathore and P. N. Singh, "Application of Deep Learning to Improve the Accuracy of Soil Nutrient Classification," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), 2022, doi: 10.1109/MysuruCon55714.2022.9972666.
2. T. B. Sheeba et al., "Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms," *Journal of Nanomaterials*, vol. 2022, Article ID 5343965, 2022.
3. K. Liu, Y. Wang, Z. Peng et al., "Monitoring soil nutrients using machine learning based on UAV hyperspectral remote sensing," *International Journal of Remote Sensing*, vol. 45, no. 14, pp. 4897–4921, 2024.
4. A. A. Alfin, N. Suciati, and W. N. Khotimah, "Bayesian Optimization of Convolutional Neural Network Hyperparameters for Soil Macro-Nutrient Estimation Using Hyperspectral Imaging," 2025 International Conference on Smart Computing, IoT and Machine Learning (SIML), 2025.

5. S. K. A. L. Z. Rahman, K. C. Mitra, and S. M. Islam, "Soil classification using machine learning methods and crop suggestion based on soil series," 2018 21st International Conference of Computer and Information Technology (ICCIT), 2018.
6. G. Shukla, R. D. Garg, H. S. Srivastava, and P. K. Garg, "An effective implementation and assessment of a random forest classifier as a soil spatial predictive model," International Journal of Remote Sensing, vol. 39, pp. 2637–2669, 2018.
7. Y. Yu, T. Xu, Z. Shen, Y. Zhang, and X. Wang, "Compressive spectral imaging system for soil classification with three-dimensional convolutional neural network," Optics Express, vol. 27, pp. 23029–23048, 2019.