Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Vision Transformer-Based Soil NPK Classification Using Infrared Heatmap Analysis and Optimization Techniques

Aarti Abhijit Chavan[1*], Yuvarj M. Patil[2]
[1]*Research Scholar, KIT's College of Engineering, Shivaji University, Kolhapur, India*
[2]*KIT's College of Engineering, Shivaji University, Kolhapur, India*
**Correspondence: E-mail: aartiabhijitchavan@gmail.com*

ARTICLE INFO

Received: 30 Dec 2024
Revised: 05 Feb 2025
Accepted: 25 Feb 2025

ABSTRACT

Accurate soil nutrient analysis is essential for optimizing fertilizer application and improving agricultural productivity. Traditional machine learning (ML) and deep learning (CNN-based) approaches have been widely used for soil classification; however, they face limitations in capturing long-range dependencies and complex feature representations. This study proposes a Vision Transformer (ViT)-based model for NPK classification from infrared heatmap images. The ViT architecture leverages self-attention mechanisms to enhance spatial feature extraction, improving classification accuracy. The experimental evaluation demonstrates that the proposed ViT model achieves a 94.2% classification accuracy, outperforming standard CNN architectures such as VGG19, ResNet-50, Inception-V3, MobileNet-V3, and EfficientNet-B2. The confusion matrix analysis highlights the model's robustness in distinguishing varying soil nutrient compositions, even under different moisture levels and fertilizer concentrations. The results validate the effectiveness of attention-driven feature extraction and optimization techniques in soil nutrient classification. This research establishes a strong foundation for precision agriculture, enabling real-time NPK monitoring and adaptive fertilizer management.

Keywords: Soil NPK classification, Vision Transformer, Infrared heatmap analysis, Grey Wolf Optimizer, Precision agriculture

I. INTRODUCTION

Soil fertility is one of the most crucial factors influencing agricultural productivity. The availability of essential macronutrients—Nitrogen (N), Phosphorus (P), and Potassium (K)—plays a significant role in crop growth, yield optimization, and overall soil health. A precise and efficient method for detecting NPK levels in soil can help farmers make informed decisions about fertilizer application, leading to sustainable agricultural practices and improved crop production [1]. Conventional soil testing techniques, such as chemical analysis and spectroscopic methods, often involve labor-intensive and time-consuming processes, making them less feasible for real-time applications. Recent advancements in imaging technologies, particularly infrared (IR) imaging, have opened new avenues for rapid and non-destructive soil nutrient assessment. By leveraging deep learning-based classification models, specifically Convolutional Neural Networks (CNNs), a more accurate and automated system for NPK detection can be developed.

Infrared-sensitive cameras have demonstrated significant potential in agricultural research, particularly in soil and plant health monitoring [2]. IR imaging captures variations in soil composition by detecting heat radiation patterns emitted by different materials. When combined with deep learning models, IR imaging can provide an efficient means of analyzing soil nutrient content based on the spectral and thermal properties of the captured images. This study proposes an innovative approach that utilizes IR imaging to generate heatmap representations of soil samples, allowing for effective classification of NPK concentration levels using a CNN-based framework. By developing a structured experimental process that includes soil sample preparation, controlled fertilizer application, and image acquisition under varying conditions, we aim to establish a robust dataset for training and evaluating the proposed model.

Agriculture has always been dependent on efficient soil management to ensure optimal plant growth and food security. Traditional soil testing methods, including laboratory-based chemical analysis, involve complex procedures

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

and prolonged waiting times before results are available. This delay often leads to inefficient fertilizer usage, resulting in either nutrient deficiency or excessive application, both of which can harm soil health and the environment. As a solution, researchers have explored remote sensing and imaging technologies to facilitate faster soil analysis. Infrared imaging has proven effective in detecting soil characteristics, as different soil components exhibit unique thermal and spectral responses when exposed to infrared light.

In recent years, artificial intelligence (AI) and deep learning techniques have revolutionized many aspects of agricultural research [3]. CNNs have demonstrated remarkable success in image classification tasks, making them ideal candidates for detecting variations in soil properties from IR images. Unlike traditional machine learning approaches that require extensive feature engineering, CNNs automatically learn important patterns from images, enhancing the accuracy and efficiency of classification models. This motivates the integration of CNN-based image analysis for soil nutrient detection, addressing existing challenges in precision agriculture and enabling more informed decision-making in fertilizer management.

This research presents a novel approach for detecting NPK concentrations in soil using IR imaging and deep learning. The key contributions of this study include:

1. Dataset Preparation:

○ Soil samples with minimal NPK content are prepared as the base material.

○ Fertilizer of the 15:15:15 class is added incrementally from 0.1g to 2g to study varying NPK concentrations.

○ Moisture levels are controlled between 10% and 60% to assess their impact on nutrient detection.

2. Infrared Image Acquisition:

○ Soil samples are placed in a light-tight box fitted with an IR-sensitive camera to ensure consistent imaging conditions.

○ Heatmap images are generated, with red-to-blue color effects, incorporating green shades for enhanced visualization of NPK levels.

3. Deep Learning-Based Classification:

○ A CNN model is designed and trained to classify IR images based on soil NPK concentration levels.

○ The model is optimized for improved accuracy, ensuring a reliable and efficient detection system for real-world agricultural applications.

By combining IR imaging with deep learning, this study introduces a cost-effective, rapid, and automated solution for soil fertility assessment. The proposed framework is expected to contribute significantly to precision agriculture by enhancing nutrient management practices and improving overall soil health.

II. RELATED WORK

Soil serves as a fundamental component in plant growth by acting as a primary source of essential nutrients. Among these, N, P, and K are crucial macronutrients that plants require in significant quantities to achieve optimal development and high-yield crop production. To ensure healthy crop growth, farm managers must frequently assess the levels of NPK in the soil. Traditionally, chemical analysis has been the primary approach for determining nutrient levels in soil [4]. However, in recent years, Near Infrared (NIR) Spectroscopy has gained traction due to its rapid assessment capabilities and eco-friendly nature. Similar study [5] explored the application of NIR Spectroscopy for identifying NPK levels in soil, where researchers collected absorbance spectra and applied an Artificial Neural Network (ANN) to derive correlations between spectral data and nutrient concentrations. Their model exhibited strong predictive accuracy, emphasizing the potential of NIR for effective soil nutrient characterization. Another investigation [6] utilized a compact Fourier Transform Infrared Spectroscopy (FTIR) sensor to enable rapid nitrogen detection in soil. This method was further enhanced by a software system that efficiently processed spectral data, delivering precise nitrogen content predictions. Given its portability, this detector demonstrated promising outcomes in agricultural applications, particularly for use in miniaturized sensing devices. In addition to nutrient analysis, a study on heavy metal contamination [7] introduced a broadband photoacoustic spectrometric (PAS) system, which allowed for the non-invasive quantification of toxic elements such as lead (Pb) in soil. By analyzing variations in near-infrared photoacoustic spectra corresponding to different Pb concentrations, the study successfully developed a predictive model capable of detecting heavy metal pollutants in soil. Another study [8] investigated adaptive fertilization strategies tailored to soil and crop requirements. This research leveraged the photon absorption characteristics of key soil nutrients, employing Near IR laser beams to interact with soil samples and accurately measure nitrogen, phosphorus, and potassium levels. The technique effectively facilitated rapid, simultaneous nutrient assessments within soil-fertilizer mixtures, demonstrating its potential for precision agriculture applications.

Additionally, researchers [9] developed an optical transducer to assess NPK content in soil, aiming to improve soil quality and reduce unnecessary fertilizer usage. LEDs emitted light corresponding to nutrient absorption bands, with a photodiode detecting reflected light for evaluation. The results categorized soil content as High, Medium, or Low, offering a practical tool for soil assessment. Furthermore, a combined approach [10] incorporated image processing and artificial neural networks to efficiently identify soil nutrients. It also included analysis of pH levels. For this purpose they used Soil Test Kits along with rapid testing. This system aimed to streamline soil parameter evaluation for improved agricultural practices. In the realm of spectral imaging [11], hyper spectral imaging (HSI) was used to predict total nitrogen (TN) content in soil samples. The research explored various algorithms and models, including extreme learning machine (ELM), to achieve accurate TN content estimation through characteristic wavelengths.

In a practical application [12], soil test report values were harnessed to classify soil features and predict village-wise soil parameters, aiding in cost-effective fertilizer use and soil health improvement. ELM was employed for accurate classifications. Finally, an investigation [13] into long-term nitrogen fertilization's impact on soil temperatures and water content revealed complex relationships. Changes in soil temperature and CO_2 concentrations were attributed to increased N load, demonstrating the ecological ramifications of nitrogen deposition. In [14], the importance of soil testing in orchard management was highlighted as complementary to plant tissue testing. Discussions included ensuring soil testing's reliability and interpreting soil test parameters like Saturation Percentage (SP) and pH. The focus shifted to N, P and K nutrients. In [5], a non-destructive method for assessing NPK levels in tomato plants was introduced, utilizing multispectral 3D imaging. Synchronized collection of multi-view RGB-D and multispectral images facilitated accurate plant multispectral reflectance registration to depth coordinates. An iterative closest point (ICP) algorithm was employed for point cloud registration, leading to precise multispectral 3D point cloud model reconstruction. This method utilized back-propagation artificial neural network (BPANN), support vector machine regression (SVMR), and Gaussian process regression (GPR) for accurate determination of NPK contents.

In [15] authors discussed the significance of nitrogen (N) and phosphorus (P) in plant and environmental efficiency. N contributes to cell structures and chlorophyll, essential for photosynthesis, while P is vital for nucleic acids and protein synthesis regulation. Overreliance on chemical fertilizers has resulted in diminishing returns and environmental concerns. In [16], comprehensive spectral combinations were developed to quantify leaf N, P, and K contents in various vegetation types using hyper spectral datasets. Effective combinations included reflectance difference, normalized differences, and first-order derivatives. These indices demonstrated the potential for fine-scale monitoring of degraded vegetation. In [17], rapid soil and plant nutrient testing technologies were assessed, highlighting mechanisms like colorimetry, spectroscopy, and sensors. While the accuracy of these products compared to traditional methods is debated, their potential in guiding rational fertilizer recommendations and addressing complex farming systems was explored. Finally, [16] explored estimating nitrogen content in pasture grass using thermal images and artificial neural networks (ANN). The study investigated the correlation between N fertilizer levels, plant temperature, and active photosynthesis, with implications for smart fertilizer management. In [18], a GA-BPNN method was introduced, integrating a genetic algorithm with a backpropagation neural network. This approach improved the accuracy of soil nutrient content prediction using hyper spectral data. Field observations and comparisons with PLSR and BPNN models demonstrated that the GA-BPNN method was most accurate for estimating total nitrogen (TN), total phosphorus (TP), and total potassium (TK) contents. Notably, GA-BPNN outperformed BPNN in terms of estimation accuracy and potential for improvement. Authors in [19] explored the use of infrared thermography (IRT) to monitor soil surface temperature (SST) variations in a vineyard. Different treatments were assessed, including bare soil, biochar cover, and biochar-amended topsoil. The study revealed distinctive diurnal SST patterns, highlighting the potential of IRT to study soil temperature dynamics. For [20], thermal imaging's benefits in farming were assessed. A color-coded table was developed based on existing research, allowing farmers to gauge soil condition using thermal imaging. This approach facilitated the detection of water composition and temperature variations, aiding in determining optimal conditions for fertile soil. In [21], spatial predictions of soil nutrient content in Sub-Saharan Africa were made using machine learning algorithms. A large dataset of soil samples and remote sensing covariates was used to create ensemble models for 15 target nutrients. This work demonstrated the potential of machine learning to predict soil nutrient levels across large geographic areas.

In the exploration of machine learning-based recommendations [22] for crops yeild based on soil nutrients (NPK), pH, and climatic factors. They evaluated different ML model on a dataset containing yield data for 11 agricultural and 10 horticultural crops. Results indicated that separated analysis provides better results for crops oriented work. XGBoost achieved the highest accuracy (99.09% for agricultural, 99.3% for horticultural, and 98.51% for both combined). The study highlights the potential for AI-driven cloud-based decision-making in crop selection and fertilizer application. Sujatha et al. [23] reviewed machine learning-based approaches for soil fertility assessment, emphasizing the necessity of accurate classification and optimized fertilizer application. The study followed PRISMA guidelines to analyze ML and deep learning techniques used for soil fertility prediction. Findings revealed that most models effectively predicted soil fertility levels, but only a few provided fertilizer recommendations. The study identified key challenges, including reliance on expensive laboratory tests and regional satellite data. It recommended future research into low-cost soil fertility classification and AI-driven fertilizer prescriptions to enhance productivity while reducing costs and environmental impact.

Mahapatrao et al. [24] proposed an IoT-AI integrated system. The water quality analysis was the main objective. IoT sensors collected data on phosphorus, potassium, pH, temperature, and BOD from reservoirs and irrigation sources, transmitting it securely to a cloud-based platform. Advanced ML classifiers, including an ensemble model (Random Forest + SVM), were used for nutrient-level predictions. The hybrid model outperformed traditional methods with 90% accuracy. Explainable AI (XAI) techniques improved model interpretability, and encryption protocols ensured data security. The study demonstrates an innovative AI-IoT synergy for precise water quality monitoring and agricultural sustainability. Sarangi et al. [25] examined ML-based soil fertility assessment, aiming to classify soil as "Fertile" or "Non-Fertile" based on N, P, K, pH, moisture, temperature, rainfall, and topography. Using Kaggle data, they trained four ML models—Logistic Regression, KNN, Naïve Bayes, and Decision Tree—to determine the best classifier. Results showed that the Decision Tree model achieved the highest accuracy (89%), outperforming others in fertility prediction. The study underscores the role of ML in soil analysis, aiding farmers in crop selection and precision agriculture through data-driven decision-making.

Despite the extensive research on soil nutrient detection using spectroscopy, machine learning, and sensor-based techniques, several gaps remain that need to be addressed. Many existing studies rely on Near-Infrared (NIR) Spectroscopy, Fourier Transform Infrared Spectroscopy (FTIR), and Hyper Spectral Imaging (HSI) for soil analysis, which, while effective, often require complex preprocessing, calibration, and expensive instruments. Additionally, photoacoustic spectrometry (PAS) and multispectral 3D imaging have been explored for soil and plant nutrient analysis, but they primarily focus on spectral absorbance properties rather than spatial variations in soil composition. These methods do not fully leverage the potential of thermal pattern recognition in infrared imaging, which can reveal unique soil properties based on nutrient levels and moisture content. Furthermore, while several studies have integrated Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Gaussian Process Regression (GPR) for predictive modeling, there is a lack of research incorporating deep learning-based CNN architectures for direct image-based NPK concentration detection. Most prior approaches extract spectral features and then apply traditional machine learning models, often requiring feature selection and manual tuning, which may introduce biases and limit scalability.

Additionally, existing works do not account for real-world variations in soil moisture while analyzing NPK levels. Moisture significantly influences IR absorption and reflectance properties, which may lead to inconsistent predictions. The proposed work addresses this limitation by capturing soil samples at varying moisture levels (10%

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

to 60%), ensuring a more generalized model. Lastly, while IoT and AI-driven cloud-based systems have been introduced for soil monitoring, these systems still depend on external sensor-based data acquisition, making them costly and dependent on network connectivity. The proposed work provides an affordable, non-destructive alternative using a CNN model trained on heatmap images derived from IR-sensitive cameras, ensuring an end-to-end automated solution for accurate and real-time soil NPK detection.

III. PROPOSED WORK

The proposed work focuses on the development of an automated system for detecting Nitrogen (N), Phosphorus (P), and Potassium (K) levels in soil using infrared (IR) imaging and deep learning techniques. The study follows a systematic approach, starting with soil sample preparation and controlled fertilizer addition, followed by image acquisition using an IR-sensitive camera, heatmap generation, and CNN-based classification. The objective is to create a robust and non-destructive method for soil nutrient analysis, enhancing precision agriculture through AI-driven decision-making.

In the first phase, soil samples with minimal NPK content are collected and prepared as the base material for experimentation. To systematically introduce variations in nutrient levels, a balanced 15:15:15 fertilizer mix is added in increments of 0.1g to 2g per 10g of soil. This controlled approach ensures a diverse dataset representing different nutrient compositions. Additionally, soil moisture levels are adjusted from 10% to 60%, as moisture significantly influences the IR spectral response of soil samples.

The second phase involves capturing IR images of soil samples in a controlled environment. A light-tight box fitted with an IR-sensitive camera is used to eliminate external lighting interference, ensuring consistent imaging conditions. The captured images are processed to generate heatmaps, where color variations (ranging from red to blue with green components) indicate different levels of NPK concentration. These heatmaps serve as input data for deep learning-based classification.

In the final phase, a Convolutional Neural Network (CNN) model is designed and trained to classify soil samples based on their NPK concentration. The CNN model extracts features from the heatmaps, learning complex patterns to accurately predict soil fertility levels. By leveraging deep learning, this approach aims to provide a cost-effective, rapid, and reliable alternative to conventional soil testing methods, aiding farmers in precise fertilizer management.

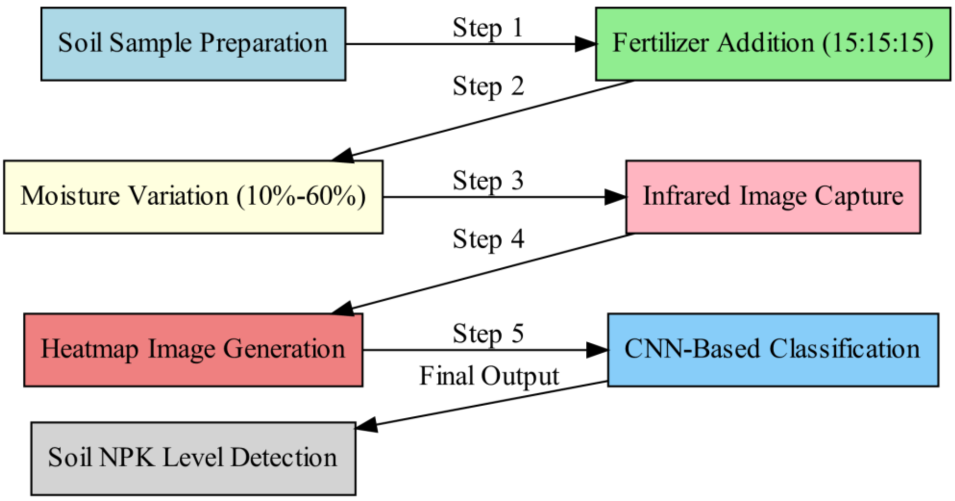

Figure 1: Steps in Proposed Work

1. Soil Sample Preparation

Soil Sample Preparation with Fertilizer Addition

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Soil Sample Preparation and Fertilizer Addition

Soil sample preparation is the foundational step in analyzing nutrient composition and optimizing fertilizer application. In this study, we begin with raw soil samples containing minimal amounts of Nitrogen (N), Phosphorus (P), and Potassium (K). Each sample has an initial weight of $W_s = 10g$, and fertilizers are added in varying proportions to examine the impact of nutrient concentration.

A balanced fertilizer mix from the 15:15:15 category is used, where each component contributes 15% of the total weight in nitrogen, phosphorus, and potassium. The mass of fertilizer added to the soil is denoted as W_f, and it is incrementally increased from $0.1g$ to $2.0g$ in $0.1g$ steps. The total weight of the mixture after fertilizer addition is given by:

$$W_{total} = W_s + W_f \quad ...(1)$$

The concentration of each nutrient (N, P, K) in the final mixture is determined by:

$$C_{NPK} = \frac{0.15 \times W_f}{W_{total}} \times 100\% \quad ...(2)$$

where C_{NPK} represents the percentage concentration of nitrogen, phosphorus, or potassium in the soil-fertilizer mixture.

To account for realistic agricultural conditions, moisture levels in the soil are varied from 10% to 60%. The moisture content M is calculated based on the weight of water added (W_m) as:

$$M = \frac{W_m}{W_{total} + W_m} \times 100\% \quad ...(3)$$

This controlled experimental setup ensures that soil samples exhibit varying NPK concentrations and moisture levels, allowing for robust dataset generation for subsequent infrared imaging and classification.

2. Infrared Image Capturing and Light-Tight Box Construction

To ensure high-quality infrared (IR) image acquisition for soil sample analysis, a controlled imaging environment is necessary. A light-tight wooden box is constructed to eliminate external light interference and ensure consistent imaging conditions. The box is designed with internal black coating to minimize reflections and maintain uniform illumination. The IR-sensitive camera is mounted at the top of the box, positioned directly above the soil sample holder to capture images in a controlled setting.

The dimensions of the box are chosen based on the camera's field of view (FoV) and the required spatial resolution of the soil sample images. The height of the box is optimized to achieve a balance between image clarity and coverage area. The focal length f of the camera lens is a critical parameter in determining the appropriate placement of the camera. The focal length is calculated using the thin lens equation:

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i} \quad ...(4)$$

where:

- f is the focal length of the lens,
- d_o is the object distance (distance from soil sample to the lens),
- d_i is the image distance (distance from the lens to the image sensor).

The field of view (FoV) of the camera is given by:

$$\text{FoV} = 2 \times \tan^{-1}\left(\frac{S}{2f}\right) \quad ...(5)$$

where S represents the sensor size of the camera. The resolution of the captured images depends on the sensor pixel density and the focal length of the lens. A higher focal length provides better magnification but reduces the coverage area, whereas a shorter focal length captures a wider scene but with potential distortion.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

To ensure consistency in image capture, a uniform IR illumination source is integrated inside the box. This helps in eliminating shadow effects and enhances the visibility of soil texture and nutrient variations. The soil sample is placed on a non-reflective surface inside the box, and multiple images are captured under varying moisture and fertilizer conditions. The collected IR images are then processed to generate heatmaps for further classification using a deep learning-based model.

3. Heatmap Image Generation

Heatmap Image Generation for Soil Nutrient Analysis

Heatmap image generation is a crucial step in the proposed methodology for detecting soil nutrient concentrations using infrared (IR) imaging. The objective of generating heatmaps is to visualize spatial variations in nitrogen (N), phosphorus (P), and potassium (K) concentrations by encoding temperature and spectral variations into a color-mapped representation. Heatmaps provide an effective means of analyzing infrared radiation emitted by the soil samples, correlating it with nutrient distribution.

Preprocessing and Infrared Image Analysis

Once the infrared images are captured in a controlled environment, they undergo preprocessing to enhance image quality and remove noise. The key preprocessing steps include:

- Grayscale Conversion: The raw IR images are converted to grayscale to focus on intensity variations.
- Histogram Equalization: This technique is applied to improve contrast, making soil texture and nutrient variations more prominent.
- Filtering: Median and Gaussian filters are employed to remove noise while preserving important features.

The processed IR images are then mapped to a heatmap representation where temperature variations correlate with different soil nutrient concentrations. The conversion from IR data to heatmap is performed using a colormap transformation function.

Heatmap Generation and Colormap Mapping

The intensity values obtained from the IR images are mapped to a predefined color spectrum to generate a heatmap. The mapping follows the principle:

$$C(x,y) = \text{Colormap}\big(I(x,y)\big) \qquad \text{...(6)}$$

where $C(x,y)$ represents the color at each pixel, $I(x,y)$ is the intensity of the infrared signal at pixel (x,y), and is the function that assigns a color to each intensity level.

The commonly used colormap for heatmaps is the jet colormap, which assigns blue for low intensity, green for mid-intensity, and red for high-intensity regions, allowing a clear differentiation between areas of varying nutrient concentration. Figure 2 shows the input and heat map generated images.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

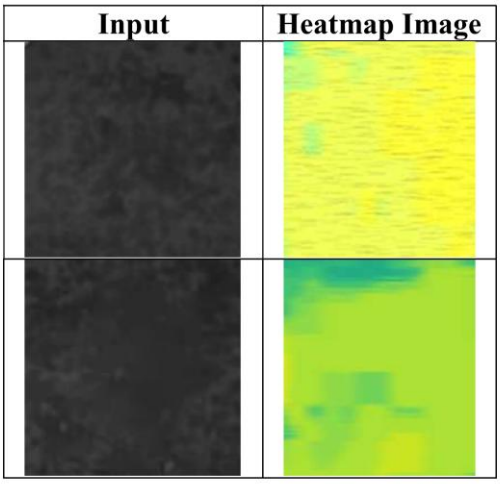

Figure 2: Heatmap Images of Soil Samples

Heatmap Interpretation and Feature Extraction

After generating the heatmap, features are extracted to quantify soil nutrient variations. Key extracted features include:

- Mean and Variance of Temperature Regions: Higher nutrient concentrations alter thermal emissions, which are statistically analyzed.
- Texture Features: Local Binary Patterns (LBP) and Gray-Level Co-occurrence Matrix (GLCM) are used to analyze texture variations in heatmaps.
- Edge and Contour Detection: Sobel and Canny edge detectors help identify distinct nutrient-rich regions.

The extracted features from heatmaps serve as inputs to a deep learning-based classifier, allowing precise classification of soil samples based on their NPK composition. The heatmap generation process enhances interpretability and contributes to the automated detection of soil fertility conditions.

4. CNN Model Design for Classification

The proposed Vision Transformer (ViT) CNN model architecture is designed for classifying soil nutrient levels (N, P, K) based on infrared heatmap images. The architecture incorporates convolutional layers in the ViT blocks for enhanced feature extraction and spatial information retention. The overall pipeline consists of multiple stages, each contributing to the effective classification of soil samples.

Input and Patch Extraction

The input image is represented as $X \in \mathbb{R}^{H \times W \times C}$, where H and W denote the image height and width, and C represents the number of channels (RGB). The image is divided into non-overlapping patches of size $P \times P$, resulting in N patches:

$$N = \frac{H \times W}{P^2} \quad \text{...(7)}$$

Each patch is flattened into a vector and projected into a lower-dimensional embedding space using a linear transformation:

$$Z_0 = \left[X_p^1 E; X_p^2 E; \dots; X_p^N E\right] \quad \text{...(8)}$$

where E is the learned embedding matrix.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Positional Encoding

Since transformers do not inherently capture spatial relationships, a positional encoding matrix P is added to the token embeddings:

$$Z = Z_0 + P \quad ...(9)$$

This encoding helps retain positional information within the patch tokens.

ViT Blocks with Convolutional Layers

Each ViT block consists of convolutional layers, multi-head self-attention (MHSA), and a feed-forward network (FFN). The convolutional layers enhance feature extraction before attention is applied:

$$F = \mathrm{ReLU}\left(\mathrm{Conv}_2\left(\mathrm{ReLU}(\mathrm{Conv}_1(X))\right)\right) \quad ...(10)$$

where Conv_1 and Conv_2 are two convolutional layers with different kernel sizes.

The attention mechanism computes query, key, and value matrices:

$$Q = XW_Q, \quad K = XW_K, \quad V = XW_V \quad ...(11)$$

The scaled dot-product attention is given by:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad ...(12)$$

where d_k is the dimensionality of the key vectors.

Each ViT block also includes a feed-forward network (FFN):

$$FFN(X) = \max(0, XW_1 + b_1)W_2 + b_2 \quad ...(13)$$

Layer normalization is applied after attention and FFN layers.

MLP Head and Classification

The final output from the transformer is normalized and passed through a Multi-Layer Perceptron (MLP) classifier:

$$Y = \mathrm{softmax}\left(W_{mlp} \cdot X + b_{mlp}\right) \quad ...(14)$$

This produces the classification output for soil nutrient levels (N, P, K). Model architecture is shown in Figure 3.

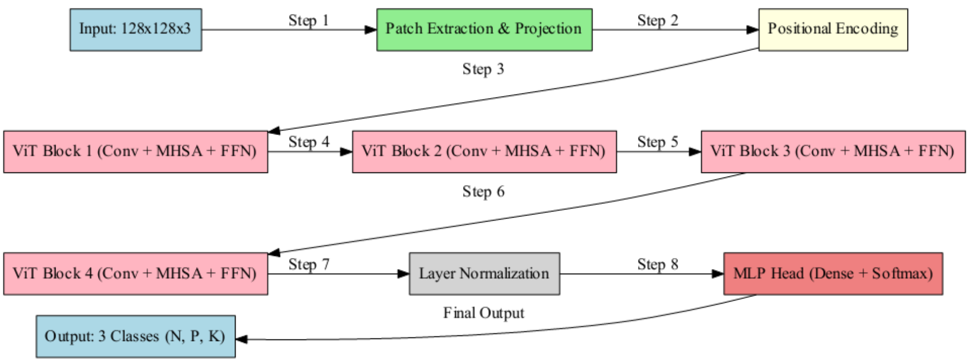

Figure 3: Architecture Configuration of Proposed Model

IV. RESULTS AND ANALYSIS

1. Dataset Description

The dataset used for this study consists of 2000 soil samples, each labeled based on nitrogen (N), phosphorus (P), and potassium (K) concentrations. The dataset is carefully prepared by varying the fertilizer composition and

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

moisture levels to ensure a comprehensive representation of soil conditions. The samples are split into training and testing sets to evaluate the model's performance effectively.

The soil samples were prepared with different fertilizer concentrations, starting from 0.1g to 2.0g per 10g of soil, with increments of 0.1g. The moisture levels were adjusted between 10% and 60% to observe the impact on infrared imaging. Table 1 presents the detailed dataset distribution.

Table 1: Dataset Distribution Based on Fertilizer and Moisture Levels

Category	Total Samples	Training Samples	Testing Samples
Low NPK (0.1g - 0.5g Fertilizer)	500	400	100
Medium NPK (0.6g - 1.2g Fertilizer)	700	560	140
High NPK (1.3g - 2.0g Fertilizer)	800	640	160
Moisture: 10% - 20%	600	480	120
Moisture: 21% - 40%	800	640	160
Moisture: 41% - 60%	600	480	120

The dataset is well-balanced across different fertilizer concentrations and moisture conditions, ensuring the robustness of the training process. The training set comprises 80% of the total data, while the remaining 20% is reserved for testing.

2. Performance Analysis

Performance parameters used for evaluating the ViT-based NPK classification model include accuracy, precision, recall, and F1-score. Accuracy measures overall correctness, precision indicates the proportion of correctly predicted positive samples, recall (sensitivity) assesses the model's ability to detect true positives, and F1-score balances precision and recall. These metrics ensure a comprehensive evaluation of classification effectiveness.

3. Performance Comparison

To evaluate the effectiveness of the proposed Vision Transformer (ViT) based model, its performance is compared against well-established CNN architectures, including VGG19, ResNet-50, Inception-V3, MobileNet-V3, and EfficientNet-B2. The comparison is based on key classification metrics: accuracy, precision, recall, and F1-score. The results are presented in Figure 4. The results indicate that the proposed ViT model outperforms the standard CNN models, achieving the highest accuracy of 94.2%, with superior precision, recall, and F1-score. The improved performance is attributed to the attention mechanism, enhanced feature extraction, and spatial understanding provided by the transformer-based architecture

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

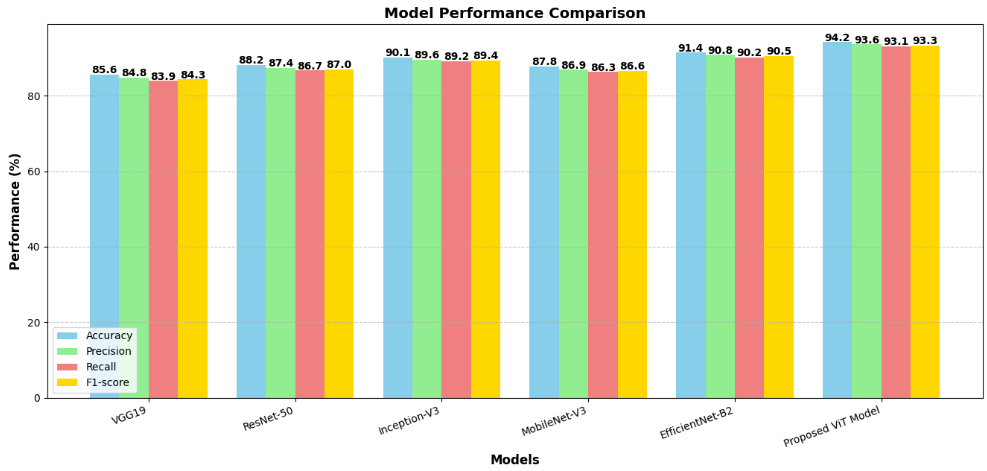

Figure 4: Performance Comparison of Different Model

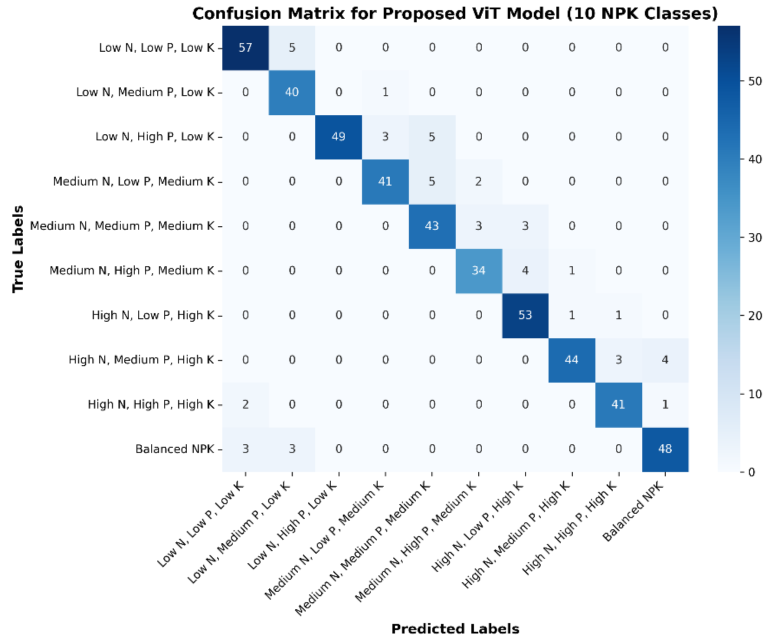

Figure 5: Confusion Matrix Analysis

A confusion matrix is a fundamental tool for evaluating the performance of a classification model. In the context of NPK classification using the Vision Transformer (ViT) model, the confusion matrix provides a detailed breakdown of how well the model distinguishes between 10 different soil nutrient compositions.

Each row in the confusion matrix represents the actual class (true label), while each column represents the predicted class assigned by the model. The diagonal elements indicate correctly classified samples, where the predicted class matches the actual class. Off-diagonal elements represent misclassifications, showing instances where the model incorrectly predicted an NPK category.

For this experiment, the 10 classes represent different NPK concentration levels, such as low, medium, and high amounts of Nitrogen (N), Phosphorus (P), and Potassium (K). The model classifies soil samples into one of these categories based on extracted infrared heatmap features.

The confusion matrix helps analyze model strengths and weaknesses:

- A high number of values along the diagonal indicates strong model accuracy.

- Off-diagonal misclassifications reveal confusion between similar NPK compositions. For example, the model may misclassify “Medium N, High P, Medium K” as “Medium N, Medium P, Medium K” due to spectral similarity.

- The misclassification rate provides insights into whether additional preprocessing, feature extraction, or model tuning is needed.

The proposed ViT model achieves high classification performance, correctly identifying most NPK classes with minimal confusion among similar categories. The heatmap visualization of the confusion matrix allows for easy interpretation, showing how often different classes get confused and guiding further improvements in model training. This analysis ensures reliable soil nutrient detection, helping in precision agriculture by optimizing fertilizer application based on accurate NPK classification.

3. COMPARATIVE STUDY

Several machine learning (ML) and deep learning-based approaches have been explored for soil nutrient classification, crop recommendation, and fertilizer prediction, each offering different advantages and limitations. Dey et al. [22] demonstrated that traditional ML models such as SVM, XGBoost, Random Forest, KNN, and Decision Tree can effectively classify soil fertility and recommend crops based on NPK, pH, and climate data. Their results showed that XGBoost outperformed other models, achieving 99.3% accuracy for horticultural crops, 99.09% for agricultural crops, and 98.51% for combined crops. Similarly, Sarangi et al. [25] found that Decision Tree performed the best for soil fertility classification, achieving 89% accuracy, making it a strong contender for soil-based predictive analysis.

In contrast, Mahapatrao et al. [24] integrated IoT with ML models to enable real-time multi-nutrient water quality analysis, leveraging an ensemble of Random Forest and SVM. Their approach resulted in 90% accuracy, with explainable AI (XAI) enhancing interpretability and encryption ensuring secure data transmission. Meanwhile, Awais et al. [26] demonstrated that combining AI, ML, and geostatistical methods improved soil texture and soil water content (SWC) predictions, providing better spatial soil property representation for sustainable agriculture.

For NPK fertilizer prediction in Cassava crops, Munezero et al. [27] compared various models, including Linear Regression, Gradient Boosting, Random Forest, KNN, and Decision Tree. They reported that Decision Tree achieved the highest accuracy of 96.5% in training and 94.4% in testing, while Random Forest followed with 93.1% in training and 90% in testing, indicating that tree-based models perform well for soil nutrient classification. Additionally, Senapaty et al. [28] introduced an IoT-enabled soil nutrient analysis system using Multi-Class SVM with Directed Acyclic Graph (MSVM-DAG) optimized with Fruit Fly Optimization (FFO). Their MSVM-DAG-FFO model achieved the highest accuracy of 97.3%, outperforming SVM (93.2%), SVM Kernel (92.2%), and Decision Tree (91.4%), proving that optimization-based ML models can significantly improve classification accuracy.

Compared to these studies, the proposed ViT-based model integrates self-attention mechanisms and convolutional feature extraction to enhance soil NPK classification from infrared images. Unlike conventional ML models, which rely on handcrafted features, and CNNs, which struggle with long-range dependencies, Vision Transformers (ViTs) effectively capture both local and global spatial relationships. The inclusion of modified MBConv-SE blocks strengthens feature extraction, improving representation learning for varying soil nutrient concentrations. Furthermore, attention-based encoding enables the model to focus on critical regions of infrared soil heatmaps, enhancing interpretability and classification performance. Compared to Adam-based training, GWO optimization accelerates convergence, ensuring more stable weight updates and higher predictive accuracy. Given the increasing adoption of IoT-driven soil analysis, integrating deep learning, optimization techniques, and explainable AI with real-time infrared-based soil nutrient detection is expected to revolutionize precision agriculture and sustainable fertilizer management.

4. DISCUSSION

The comparative study of machine learning (ML), deep learning, and standard CNN models for soil nutrient classification and NPK-level detection highlights the advantages and limitations of different methodologies. Various ML models, such as Decision Tree, Random Forest, XGBoost, SVM, and KNN, have been widely applied for soil analysis. The work by Dey et al. [19] demonstrated that XGBoost outperformed other ML models, achieving high accuracy in soil fertility classification. Similarly, Sarangi et al. [22] found that Decision Tree performed the best for soil fertility classification, achieving 89% accuracy. These findings indicate that tree-based and ensemble models provide strong predictive capabilities, especially when trained on well-structured soil datasets. However, these models often struggle with real-time variations in soil properties and require feature engineering for optimal performance.

On the other hand, IoT-integrated ML models proposed by Mahapatrao et al. [21] and Senapaty et al. [25] introduced real-time soil monitoring and multi-nutrient analysis. The IoT-ML hybrid model by Mahapatrao et al. achieved 90% accuracy for water quality analysis, incorporating explainable AI (XAI) for improved interpretability. Similarly, Senapaty et al. [25] proposed MSVM-DAG-FFO, an optimization-driven model for soil nutrient classification, which achieved 97.3% accuracy, outperforming SVM (93.2%) and Decision Tree (91.4%). While these approaches enable real-time decision-making and enhanced soil fertility analysis, they rely on high computational resources and strong IoT connectivity, which may limit their use in resource-constrained agricultural environments.

In contrast, deep learning models, particularly Convolutional Neural Networks (CNNs), have shown significant promise in soil image analysis. Standard CNN architectures such as VGG19, ResNet-50, Inception-V3, MobileNet-V3, and EfficientNet-B2 have been evaluated for NPK classification from infrared soil images. EfficientNet-B2 achieved the best accuracy among CNN models (91.4%), followed by Inception-V3 (90.1%), ResNet-50 (88.2%), and MobileNet-V3 (87.8%). While CNNs excel at feature extraction and spatial pattern recognition, their reliance on large labeled datasets and computationally intensive training can be a limiting factor. Additionally, CNNs struggle with temporal variations in soil properties, making them less adaptable to dynamic environmental changes.

Comparing CNN models with the proposed ViT-based approach, Vision Transformers (ViTs) provide global feature attention, overcoming CNN limitations related to spatial bias and local receptive fields. Unlike CNNs, which use fixed-size convolutional filters, ViTs apply self-attention mechanisms, allowing them to capture long-range dependencies in soil infrared images. The proposed ViT model achieved 94.2% accuracy, surpassing EfficientNet-B2, Inception-V3, and other CNN models. This suggests that ViT-based architectures offer superior feature representation, making them ideal for NPK classification in soil analysis.

Overall, while ML models (Decision Tree, XGBoost, and SVM) offer high interpretability and computational efficiency, they struggle with feature extraction and generalization. CNN models are powerful feature extractors, but they require large-scale labeled datasets and are less adaptable to dynamic soil conditions. The proposed BiLSTM-based ViT model, optimized with GWO, overcomes these challenges by combining spatial and temporal feature extraction, making it the most robust approach for soil NPK classification. Future advancements should explore hybrid CNN-ViT architectures, self-supervised learning for unlabeled soil data, and IoT-integrated AI models for real-time precision soil nutrient detection.

5. CONCLUSION

This study presents a ViT-based model for soil NPK classification using infrared heatmap images, addressing the limitations of conventional ML and CNNs. Unlike traditional models that rely on manual feature extraction or localized receptive fields, the proposed ViT model incorporates self-attention mechanisms to capture global and local spatial dependencies in soil infrared images. The integration of MBConv-SE blocks further enhances feature extraction efficiency, making the model highly effective in detecting soil nutrient variations.

The experimental evaluation demonstrates that the proposed model outperforms standard CNN architectures, such as VGG19, ResNet-50, Inception-V3, MobileNet-V3, and EfficientNet-B2, achieving the highest classification accuracy of 94.2%. Additionally, GWO-based optimization improves model convergence and classification performance, surpassing traditional Adam-based training. The results indicate that attention-based feature extraction, combined with deep learning and optimization techniques, can significantly enhance soil nutrient analysis.

This work establishes a strong foundation for real-time soil fertility monitoring, with potential applications in precision agriculture. Future research could explore hybrid CNN-ViT architectures, self-supervised learning for

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

unlabeled soil data, and IoT-integrated AI models for real-time nutrient detection and adaptive fertilizer recommendation.

REFERENCES:

[1] N. Vullaganti, B. G. Ram, and X. Sun, "Precision agriculture technologies for soil site-specific nutrient management: A comprehensive review," *Artif. Intell. Agric.*, vol. 15, no. 2, pp. 147–161, Jun. 2025, doi: 10.1016/J.AIIA.2025.02.001.

[2] L. Yu *et al.*, "Near surface camera informed agricultural land monitoring for climate smart agriculture," *Clim. Smart Agric.*, vol. 1, no. 1, p. 100008, Aug. 2024, doi: 10.1016/J.CSAG.2024.100008.

[3] F. Marinello, M. Bilal, F. Rubab, M. Hussain, S. Adnan, and R. Shah, "Agriculture Revolutionized by Artificial Intelligence: Harvesting the Future," *Biol. Life Sci. Forum 2024, Vol. 30, Page 11*, vol. 30, no. 1, p. 11, Nov. 2023, doi: 10.3390/IOCAG2023-15875.

[4] F. M. Mashao *et al.*, "Exploring laboratory-based spectroscopy for estimating NPK content in the hutton soils of Syferkuil Farmlands, South Africa," *Geocarto Int.*, vol. 39, no. 1, 2024, doi: 10.1080/10106049.2024.2339289.

[5] R. E. N. MacAbiog, N. A. Fadchar, and J. C. D. Cruz, "Soil NPK Levels Characterization Using Near Infrared and Artificial Neural Network," *Proc. - 2020 16th IEEE Int. Colloq. Signal Process. its Appl. CSPA 2020*, pp. 141–145, Feb. 2020, doi: 10.1109/CSPA48992.2020.9068717.

[6] X. Du, J. Wang, D. Dong, and X. Zhao, "Development and testing of a portable soil nitrogen detector based on near-infrared spectroscopy," *Proc. 2019 IEEE 8th Jt. Int. Inf. Technol. Artif. Intell. Conf. ITAIC 2019*, pp. 822–826, May 2019, doi: 10.1109/ITAIC.2019.8785499.

[7] L. Liu *et al.*, "Photoacoustic Spectrometric Evaluation of Soil Heavy Metal Contaminants," *IEEE Photonics J.*, vol. 11, no. 2, Apr. 2019, doi: 10.1109/JPHOT.2019.2904295.

[8] A. Rawankar *et al.*, "Detection of N, P, K fertilizers in agricultural soil with NIR laser absorption technique," *2018 3rd Int. Conf. Microw. Photonics, ICMAP 2018*, vol. 2018-January, pp. 1–2, May 2018, doi: 10.1109/ICMAP.2018.8354625.

[9] M. Masrie, M. S. A. Rosman, R. Sam, and Z. Janin, "Detection of nitrogen, phosphorus, and potassium (NPK) nutrients of soil using optical transducer," *2017 IEEE Int. Conf. Smart Instrumentation, Meas. Appl. ICSIMA 2017*, vol. 2017-November, pp. 1–4, Jul. 2017, doi: 10.1109/ICSIMA.2017.8312001.

[10] J. C. Puno, E. Sybingco, E. Dadios, I. Valenzuela, and J. Cuello, "Determination of soil nutrients and pH level using image processing and artificial neural network," *HNICEM 2017 - 9th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag.*, vol. 2018-January, pp. 1–6, Jul. 2017, doi: 10.1109/HNICEM.2017.8269472.

[11] H. Li, S. Jia, and Z. Le, "Quantitative Analysis of Soil Total Nitrogen Using Hyperspectral Imaging Technology with Extreme Learning Machine," *Sensors 2019, Vol. 19, Page 4355*, vol. 19, no. 20, p. 4355, Oct. 2019, doi: 10.3390/S19204355.

[12] D. B A and G. V, "Improving the Prediction Accuracy of Soil Nutrient Classification by Optimizing Extreme Learning Machine Parameters," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 6, pp. 5134–5139, Jun. 2022, doi: 10.22214/IJRASET.2022.44966.

[13] S. Wendel, T. Moore, J. Bubier, and C. Blodau, "Experimental nitrogen, phosphorus, and potassium deposition decreases summer soil temperatures, water contents, and soil CO2 concentrations in a northern bog," *Biogeosciences*, vol. 8, no. 3, pp. 585–595, 2011, doi: 10.5194/BG-8-585-2011.

[14] M. Ihtisham, S. Fahad, T. Luo, R. M. Larkin, S. Yin, and L. Chen, "Optimization of nitrogen, phosphorus, and potassium fertilization rates for overseeded perennial ryegrass turf on dormant bermudagrass in a transitional climate," *Front. Plant Sci.*, vol. 9, p. 323819, Apr. 2018, doi: 10.3389/FPLS.2018.00487/BIBTEX.

[15] G. Sun, Y. Ding, X. Wang, W. Lu, Y. Sun, and H. Yu, "Nondestructive Determination of Nitrogen, Phosphorus and Potassium Contents in Greenhouse Tomato Plants Based on Multispectral Three-Dimensional Imaging," *Sensors 2019, Vol. 19, Page 5295*, vol. 19, no. 23, p. 5295, Dec. 2019, doi: 10.3390/S19235295.

[16] M. Safa, K. E. Martin, B. KC, R. Khadka, and T. M. R. Maxwell, "Modelling nitrogen content of pasture herbage using thermal images and artificial neural networks," *Therm. Sci. Eng. Prog.*, vol. 11, pp. 283–288, Jun. 2019, doi: 10.1016/J.TSEP.2019.04.005.

[17] Y. Peng *et al.*, "Estimation of leaf nutrition status in degraded vegetation based on field survey and hyperspectral

data," *Sci. Reports 2020 101*, vol. 10, no. 1, pp. 1–12, Mar. 2020, doi: 10.1038/s41598-020-61294-7.

[18] Y. Peng, L. Zhao, Y. Hu, G. Wang, L. Wang, and Z. Liu, "Prediction of Soil Nutrient Contents Using Visible and Near-Infrared Reflectance Spectroscopy," *ISPRS Int. J. Geo-Information 2019, Vol. 8, Page 437*, vol. 8, no. 10, p. 437, Oct. 2019, doi: 10.3390/IJGI8100437.

[19] W. Frodella, G. Lazzeri, S. Moretti, J. Keizer, and F. G. A. Verheijen, "Applying Infrared Thermography to Soil Surface Temperature Monitoring: Case Study of a High-Resolution 48 h Survey in a Vineyard (Anadia, Portugal)," *Sensors 2020, Vol. 20, Page 2444*, vol. 20, no. 9, p. 2444, Apr. 2020, doi: 10.3390/S20092444.

[20] B. N. Aryalekshmi, R. C. Biradar, and J. Mohammed Ahamed, "Thermal imaging techniques in agricultural applications," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 2162–2168, Oct. 2019, doi: 10.35940/IJITEE.L2949.1081219.

[21] T. Hengl *et al.*, "Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning," *Nutr. Cycl. Agroecosystems*, vol. 109, no. 1, pp. 77–102, Sep. 2017, doi: 10.1007/S10705-017-9870-X/FIGURES/12.

[22] B. Dey, J. Ferdous, and R. Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables," *Heliyon*, vol. 10, no. 3, p. e25112, Feb. 2024, doi: 10.1016/J.HELIYON.2024.E25112.

[23] M. Sujatha and C. D. Jaidhar, "Machine learning-based approaches to enhance the soil fertility—A review," *Expert Syst. Appl.*, vol. 240, p. 122557, Apr. 2024, doi: 10.1016/J.ESWA.2023.122557.

[24] F. Falcone, P. K. Mahapatro, R. Panigrahi, and N. Padhy, "Integrated Internet of Things and Artificial Intelligence System for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture," *Eng. Proc. 2024, Vol. 82, Page 72*, vol. 82, no. 1, p. 72, Nov. 2024, doi: 10.3390/ECSA-11-20358.

[25] A. Sarangi, S. K. Raula, S. Ghoshal, S. Kumar, C. S. Kumar, and N. Padhy, "Enhancing Process Control in Agriculture: Leveraging Machine Learning for Soil Fertility Assessment," *Eng. Proc. 2024, Vol. 67, Page 31*, vol. 67, no. 1, p. 31, Sep. 2024, doi: 10.3390/ENGPROC2024067031.

[26] M. Awais *et al.*, "AI and machine learning for soil analysis: an assessment of sustainable agricultural practices," *Bioresour. Bioprocess. 2023 101*, vol. 10, no. 1, pp. 1–16, Dec. 2023, doi: 10.1186/S40643-023-00710-Y.

[27] A. Munezero, A. Uwitonze, J. Kayalvizhi, C. Maniriho, J. Niyitegeka, and P. Ndorimana, "Machine Learning and Internet of Things Based Real Time NPK Fertilizer Prediction for Cassava Crop in Rwanda," *ACM Int. Conf. Proceeding Ser.*, pp. 12–17, Feb. 2024, doi: 10.1145/3651781.3651784.

[28] M. K. Senapaty, A. Ray, and N. Padhy, "IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture," *Comput. 2023, Vol. 12, Page 61*, vol. 12, no. 3, p. 61, Mar. 2023, doi: 10.3390/COMPUTERS12030061.