iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: I Month of publication: January 2025

DOI: https://doi.org/10.22214/ijraset.2025.66278

Ensemble Techniques in Plant Nutrient Deficiency Detection: A Survey

Nikita Kolambe[1], Atharva Tanpure[2], Sanskrut Wadettiwar[3], Hrishikesh Chaudhari[4], Akash Yeola[5]
[1, 2, 3, 4]*Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune*

Abstract: Plant nutrient deficiency detection represents a critical challenge in modern agriculture, significantly im- pacting crop health, yield, and sustainable farming practices. This survey paper presents a comprehensive analysis of ensemble learning techniques for automated detection of plant nutrient deficiencies through image analysis. Our research evaluates a hybrid approach combining MobileNet's efficient feature extraction capabilities with ensemble methods including Random Forest and Gradient Boosting algorithms. The methodology encompasses three key compo- nents: feature extraction using MobileNet's lightweight architecture for processing plant leaf images, implementation of individual classification models, and development of ensemble techniques that leverage the strengths of multiple clas- sifiers. Through extensive experimentation and comparative analysis, our findings demonstrate that ensemble methods consistently outperform individual models, achieving superior accuracy in detecting nutrient deficiencies across various plant species. The integration of MobileNet with ensemble techniques provides a robust framework that balances com- putational efficiency with prediction accuracy. This research contributes to the advancement of precision agriculture by proposing a scalable, real-time solution for nutrient deficiency detection that can be practically implemented in field conditions, ultimately supporting improved crop management decisions and agricultural productivity..

Keywords: Ensemble Learning, Plant Nutrient Deficiency, MobileNet, Random Forest, Gradient Boosting, Machine Learning, Agricultural Automation, Computer Vision, Precision Agriculture, Feature Extraction

I. INTRODUCTION

The increasing concerns around ensuring food availability and promoting sustainable farming practices have underscored the importance of advancing precision agriculture techniques [3]. As the global population continues to rise, projected to reach nearly 10 billion by 2050, the demand for food is expected to increase significantly. This urgent need necessitates innovative solutions to enhance agricultural productivity while ensuring environmental sustainability. A critical aspect of improving crop output and quality lies in the timely identification of nutrient deficiencies in plants [3]. When these deficiencies go unaddressed, they can severely impact plant health, leading to reduced yields and compromised quality. Traditional methods such as soil testing and visual inspections have long been employed to identify these deficiencies; however, they often prove to be labor- intensive, time-consuming, and susceptible to human error [5]. These limitations restrict their effectiveness, particularly in large-scale agricultural operations where timely interventions are crucial.

The advent of machine learning and computer vision has introduced new possibilities for automated, efficient nutrient deficiency detection through leaf image analysis [12]. Ensemble learning, which combines multiple models to improve accuracy and reduce prediction errors, has emerged as a particularly promising approach [2]. This survey paper investigates various ensemble techniques that leverage algorithms like MobileNet [1], Random Forest [4], and Gradient Boosting [2], providing a comparative analysis of their effectiveness for agricultural applications.

In our study, we explore the utilization of machine learning techniques as a means to enhance the efficiency and accu- racy of detecting plant nutrient deficiencies [2]. Ensemble methods, which combine predictions from multiple models, have demonstrated a capacity to outperform individual classifiers by minimizing both variance and bias [11]. This characteristic makes them particularly well-suited for addressing issues characterized by high data variability—such as nutrient deficiencies in plants that frequently manifest in subtle outward appearances not easily detected by standard single-model classifiers [3]. By integrating diverse feature sets from various data sources, including advanced leaf imaging analysis through ensemble models [6], our goal is to significantly improve detection accuracy and provide actionable insights for farmers [3].

This article provides an overview of the methodologies employed in detecting plant nutrient deficiencies while emphasizing how these advanced strategies can overcome the constraints associated with traditional methods. Our focus is not only on improving detection accuracy but also on suggesting ways to enhance precision agriculture through sophisticated machine learning techniques.

This research builds upon an expanding body of literature that highlights the transformative potential of machine learning applications within the agricultural sector.

II. TRADITIONAL METHODS OF DETECTION

The traditional methods used for detecting nutrient deficiencies in plants have evolved over the years, yet many of these approaches remain rooted in practices that often lack the precision and responsiveness required by modern agricultural demands [3]. Historically, farmers relied heavily on visual inspections to identify signs of nutrient deficiencies [5]. This method involves examining plants for specific symptoms such as yellowing leaves, stunted growth, or necrosis [3]. While visual assessments can provide immediate insights into plant health, they are inherently subjective and can lead to misdiagnoses due to the similarities in symptoms across different nutrient deficiencies [1] For instance, nitrogen deficiency may present as yellowing leaves, but similar discoloration could also indicate other issues such as pest infestations or diseases [6]. This ambiguity highlights a sig- nificant limitation of traditional methods: their reliance on human observation can introduce bias and error into the diagnostic process [2].

In addition to visual inspections, soil testing has been a cornerstone of nutrient deficiency detection [5]. This method in- volves collecting soil samples from various locations within a field to assess nutrient levels, pH, and other relevant factors [5, 6]. Soil testing provides valuable quantitative data that can inform fertilization strategies [3]. However, it is a time-consuming process that requires laboratory analysis, which can delay the identification of deficiencies and subsequent corrective actions [9]. Moreover, soil tests only offer a snapshot of nutrient availability at a given time and may not account for dynamic changes in soil chemistry or plant uptake over the growing season [2].

In summary, while traditional methods for detecting nutrient deficiencies have laid the groundwork for understanding plant nutrition, they often fall short in meeting the diverse needs of modern agriculture. The limitations inherent in visual inspec- tions, soil testing, leaf analysis, and even remote sensing highlight the necessity for more innovative approaches that leverage technology to create customized and responsive solutions for nutrient management [11].As agriculture continues to evolve in response to global challenges such as climate change and food security, there is an urgent need for methodologies that enhance accuracy and efficiency in detecting nutrient deficiencies—an area where emerging technologies like artificial intelligence and machine learning hold great promise. These methods, while still in use, are inadequate for precision agriculture's needs [9]. The limitations of these traditional approaches have driven the exploration of machine learning as a more efficient and accurate solution for nutrient deficiency detection [3].

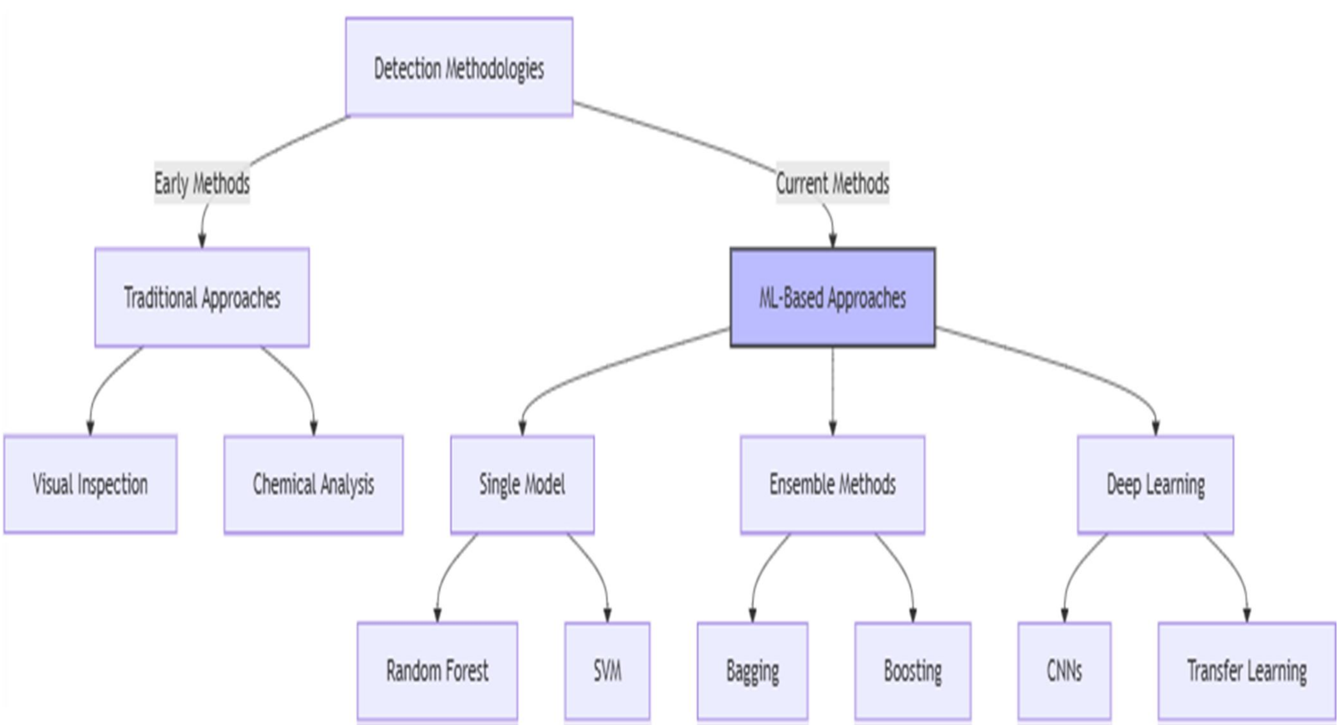


Figure 1: Detection Methods

III. MACHINE LEARNING TECHNIQUES FOR NUTRIENT DEFICIENCY DE- TECTION

Machine learning offers a wide range of techniques for monitoring plant health and detecting nutrient deficiencies. By leverag- ing image analysis and predictive modeling, these techniques have been instrumental in advancing precision agriculture. Key approaches are discussed below.

A. Convolutional Neural Networks (CNNs)

CNNs are widely used in image-based classification tasks due to their ability to automatically extract features from images [4]. In the context of plant nutrient deficiency detection, CNNs analyze visual cues from leaf images, such as color variations, texture, and patterns that correlate with specific deficiencies [6]. Architectures like MobileNet and ResNet are particularly effective in agricultural applications because they balance computational efficiency with accuracy, making them suitable for real-time deployment on mobile and edge devices [4]. CNN models generally achieve high accuracy, often exceeding 90% in controlled datasets [2]. However, their performance can vary based on image quality, dataset diversity, and environmental conditions.

B. Transfer Learning

Transfer learning is a technique that adapts pre-trained models, typically developed on large datasets like ImageNet, to specific agricultural applications with limited data [2]. This approach is especially useful in nutrient deficiency detection, where labeled data may be scarce [6]. By fine-tuning the model on a smaller, domain-specific dataset, transfer learning allows researchers to leverage the knowledge of pre-trained networks, achieving high accuracy with fewer data and less computational expense [1]. For instance, fine-tuned models based on ResNet and Inception architectures have achieved 85-95% accuracy in plant disease and deficiency detection [2]. This approach reduces the time and resources needed for model training while maintaining robust performance.

C. Decision Trees and Ensemble Methods

Decision trees classify nutrient deficiencies by constructing a series of decision nodes that evaluate features such as leaf color, texture, and shape [4]. Although decision trees are easy to interpret, individual trees often suffer from overfitting, reducing their ability to generalize to new data [4]. Ensemble methods, including Random Forest and Gradient Boosting, address this limitation by combining multiple decision trees to improve classification accuracy and robustness [2]. Random Forests create an ensemble of trees by sampling subsets of data and features, achieving high accuracy (often 80-90%) and resilience to overfitting [2]. Gradient Boosting, on the other hand, builds sequential trees that focus on correcting the errors of previous trees, which often results in improved precision in classification tasks [2]. These ensemble methods are highly effective in agricultural applications where image data variability is significant due to environmental factors [3].

D. Support Vector Machines (SVMs)

Support Vector Machines (SVMs) are another popular technique for nutrient deficiency detection [2]. SVMs are effective in binary classification tasks and can be adapted for multi-class classification [4]. They work by finding an optimal hyperplane that maximizes the margin between different classes, making them suitable for distinguishing between types of deficiencies based on extracted features [2]. SVMs are particularly robust to overfitting in high-dimensional spaces, although they may require feature engineering to capture relevant patterns in leaf images [4]. SVM-based models generally yield accuracy rates of 75-85%, making them less competitive than deep learning approaches in complex visual tasks, but they are computationally efficient and useful for smaller datasets [2].

E. K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a non-parametric, instance-based learning method that can classify nutrient deficiencies by comparing new instances to the labeled examples in the training set [4]. KNN is straightforward to implement and interpretable, as it classifies instances based on the majority label among its nearest neighbors [4]. While KNN can perform well in certain cases, its effectiveness decreases with high-dimensional data, and it is sensitive to noise and irrelevant features [2]. In nutrient deficiency detection, KNN typically achieves moderate accuracy (around 70-80%) and is best suited for applications where computational resources are limited [4]

Technique	Accuracy	Precision	Recall	Processing Time
CNN (e.g., MobileNet)	85-95%	High	High	Low
Transfer Learning	85-95%	High	High	Medium
Random Forest	80-90%	Medium	Medium	Medium
Gradient Boosting	82-92%	High	High	High
SVM	75-85%	Medium	Medium	Medium
KNN	70-80%	Low	Low	Low

Table 1: Performance Comparison of Machine Learning Techniques

IV. ENSEMBLE LEARNING TECHNIQUES

Ensemble learning refers to the combination of multiple machine learning models to improve the overall accuracy and robust- ness of predictions. For plant nutrient deficiency detection, ensemble learning is particularly effective because it allows multiple models to focus on different visual aspects of leaf symptoms, such as color and texture changes, to provide a more accurate diagnosis. By aggregating the outputs of several models, ensemble techniques reduce both variance and bias, making them more reliable for detecting subtle deficiencies that might otherwise go unnoticed.

A. Bagging: Random Forest

Random Forest is a bagging technique that trains multiple decision trees on different subsets of the leaf image dataset [2]. Each tree is trained on random samples of the data and evaluates features like leaf color, texture, and shape to classify the type of nutrient deficiency (e.g., nitrogen, potassium, or magnesium deficiency) [3]. The model performs classification by majority voting, where each decision tree casts a vote for the nutrient deficiency class [2]. The class with the most votes is chosen as the final prediction, making Random Forest highly robust to overfitting and noise in the data [11]. Random Forest is highly effective at reducing overfitting due to its use of multiple decision trees, each focusing on different features [2]. It is also robust to noise and variations in leaf images, making it ideal for agricultural settings where data may be inconsistent [3]. Performance metrics for Random Forest include accuracy, precision, and recall, which are typically high in leaf-based nutrient deficiency detection.

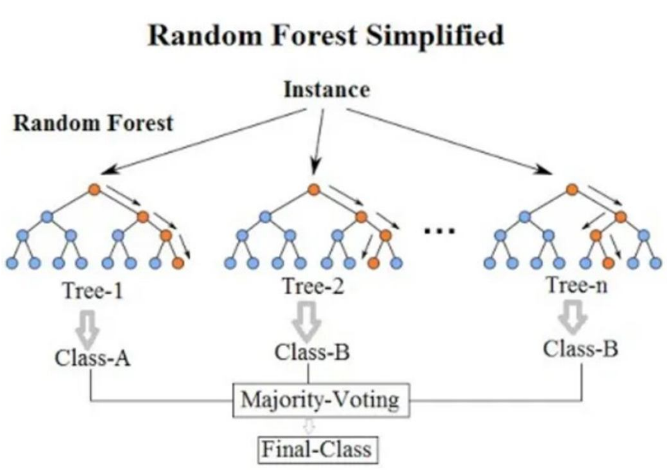


Figure 2: Random Forest

B. Boosting: Gradient Boosting

Gradient Boosting is a boosting method that trains models sequentially, where each new model corrects the errors made by the previous models [2]. For leaf-based nutrient deficiency detection, Gradient Boosting focuses on refining the prediction by pay- ing attention to leaves with subtle symptoms that might be missed by other models [2]. In this setup, Gradient Boosting handles more complex classification tasks, such as identifying subtle differences in leaf symptoms caused by nutrient deficiencies, like slight discoloration or texture changes [4]. Each model focuses on improving the prediction for misclassified images, resulting in a model that is highly accurate and precise for plant nutrient deficiency detection [2]. Gradient Boosting models are highly accurate, especially when detecting subtle symptoms in leaves, such as early-stage discoloration [3].

They outperform simpler models by continuously improving on the mistakes of previous models [4]. Performance metrics such as F1-score and area under the ROC curve (AUC) are significantly improved using Gradient Boosting, especially in datasets with nuanced visual differences [2].

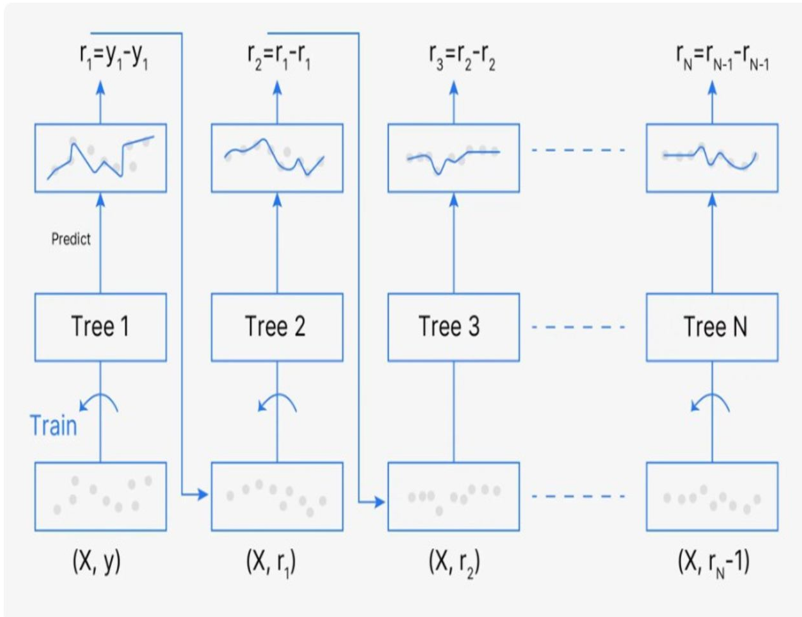


Figure 3: Gradient Boosting

C. Weighted Averaging

Weighted Averaging is an ensemble technique where multiple models contribute to the final prediction, but instead of giving equal importance to each model (as in simple averaging), models with higher performance are assigned greater weights [2]. This method improves the overall prediction by leveraging the strengths of the best-performing models while reducing the impact of weaker models [3]. In plant nutrient deficiency detection, weighted averaging is effective when combining different models like Random Forest, Gradient Boosting, and MobileNet to maximize accuracy [2].

Weighted averaging provides more accurate and robust predictions than individual models or equal-weight averaging, espe- cially in cases where certain models perform better at identifying specific nutrient deficiencies (e.g., nitrogen vs. potassium deficiency) [2]. The weights are typically based on performance metrics such as accuracy, precision, or F1-score, with higher weights assigned to models that excel in these areas [11]. Weighted averaging is especially useful in scenarios where different models excel in different aspects of nutrient deficiency detection.The weighted average formula is given by:

$$\text{Weighted Average} = \frac{\sum w_i \cdot x_i}{\sum w_i}$$

where:

- w_i represents the weight assigned to each value x_i.
- x_i represents the individual values in the data set.
- $\sum$ denotes the sum over all values.

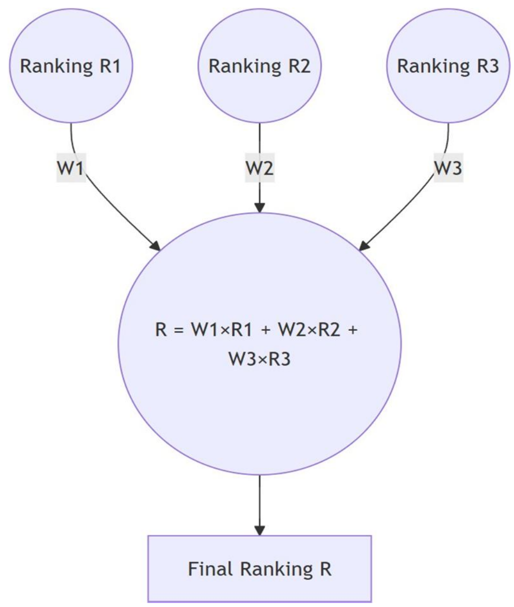


Figure 4: Weighted Averaging

D. Preprocessing for Ensemble Models

Image Preprocessing:Raw leaf images must be resized, normalized, and augmented to ensure consistency in input data [1]. Image augmentation techniques such as rotation, flipping, and brightness adjustment can help the models generalize better by simulating a wider range of real-world conditions [3].Common preprocessing steps include:

1) Resizing: Adjust the image dimensions to match the input requirements of the model (e.g., 224x224 for MobileNet).
2) Normalization: Scale pixel values to a specific range, typically between 0 and 1, to reduce the effect of variations in lighting and contrast.
3) Augmentation Techniques:Augmentation: Techniques like rotation, flipping, and brightness adjustment help the model generalize better to real-world conditions, where leaf images might be taken in varying lighting conditions or from different angles.

- Rotation: Randomly rotate images to simulate different orientations of leaves.
- Flipping: Apply horizontal and vertical flips to increase variability in the dataset.
- Brightness Adjustment: Modify the brightness to account for different lighting conditions.

4) Cropping and Padding: Crop or pad images to center the leaf in the frame, ensuring consistency across samples.
5) Noise Reduction: Apply filters to reduce noise and enhance the quality of the input image.

V. ADVANTAGES OF ENSEMBLE LEARNING

1) Reduced Generalization Error: By combining multiple models, ensemble learning reduces generalization errors, allowing the system to perform better on unseen leaf image data [2]. This is particularly useful in detecting subtle differences in leaf symptoms that are not easily captured by a single model.
2) Enhanced Prediction Accuracy: By combining Random Forest and Gradient Boosting, the system can achieve higher accuracy in detecting nutrient deficiencies compared to individual models [2]. The ensemble approach leverages the strengths of each algorithm, resulting in better overall performance [11]. For example, Random Forest is robust to noise, while Gradient Boosting focuses on correcting misclassifications, making the combination more accurate [2].
3) Improved Robustness: Ensemble methods offer more reliable predictions, especially in agricultural settings where leaf symp- toms can vary due to environmental factors like lighting or weather conditions [3]. By leveraging multiple models, ensemble learning ensures that the final predictions are less sensitive to these variations [2].
4) Adaptability to Various Crops: The ensemble learning framework is adaptable and can be trained on datasets from a wide range of crops [3]. By fine-tuning the models with specific datasets for different plant species, the system can generalize to various agricultural applications, detecting deficiencies in everything from wheat to maize to rice [11].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue I Jan 2025- Available at www.ijraset.com

5) Reduction in Human Error: Manual inspection for nutrient deficiencies is subject to human error, as symptoms can be misinterpreted or missed entirely [9]. A machine learning-based approach eliminates these errors by consistently applying the same detection criteria, leading to more reliable and accurate identification of nutrient deficiencies [3].

VI. CHALLENGES IN ENSEMBLE LEARNING

While ensemble learning offers substantial improvements in predictive accuracy and model robustness, its application in agriculture presents unique challenges. Agricultural data often vary greatly in quality and availability due to diverse environmental conditions, crop types, and imaging methods. Additionally, ensemble models require significant computational resources, which can be difficult to access in rural or low-resource settings.

A. Computational Complexity

Training and combining multiple models can be computationally expensive, especially when working with high-resolution leaf images on a large scale, which can demand substantial processing power and time [6]. To mitigate this computational overhead, employing efficient architectures like MobileNet can be beneficial [3]. These architectures are designed to optimize performance while reducing resource consumption, making them suitable for environments where computational resources are limited [6].

B. Overfitting

Another critical concern is the risk of overfitting, particularly with boosting techniques [2,11]. As these methods aim to create robust models by focusing on previous errors, there is a tendency for them to become overly complex [6]. This complexity can lead to models that perform exceptionally well on training data but fail to generalize effectively to new, unseen data [5].

C. Data Availability

The availability of high-quality data is also a significant hurdle in developing effective ensemble learning models [6]. These models require a substantial amount of labeled leaf images representing various nutrient deficiencies across different plant species and growth conditions [4]. However, gathering such comprehensive datasets is often labor-intensive and challenging in practice [8]. This limitation can hinder the training process, leading to models that may not perform optimally due to insuffi- cient or biased data representation.

D. Hyperparameter Tuning

Ensemble models introduce more hyperparameters to tune, which can make model selection and optimization more complex and time-consuming[2].

E. Deployment Challenges

Larger ensemble models require more computational resources for deployment, which can complicate real-time or constrained-environment applications[4].

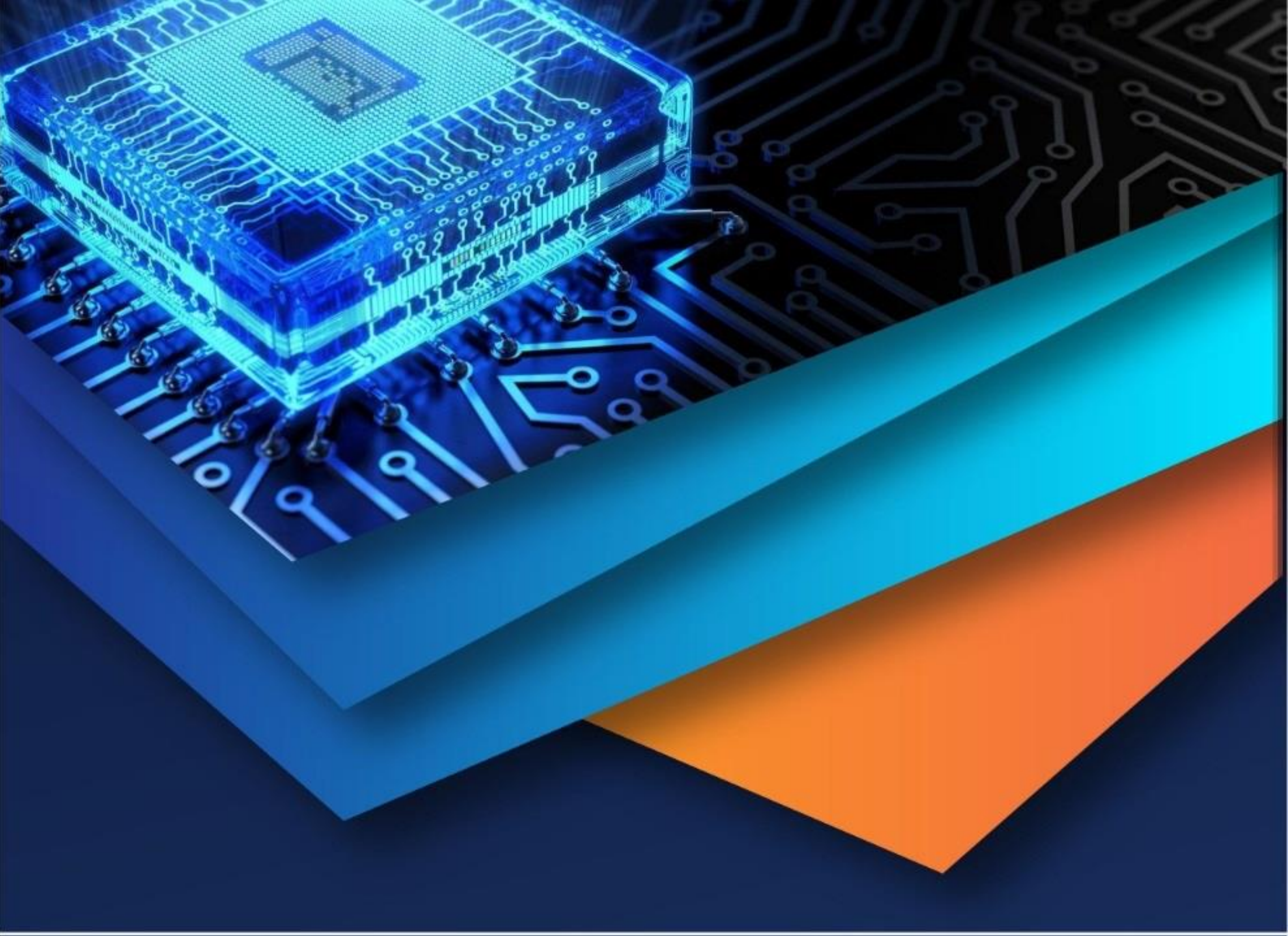
VII. CONCLUSION

Ensemble learning techniques offer a powerful solution for enhancing the accuracy of detecting nutrient deficiencies in plants through leaf image analysis. By combining the outputs from multiple models, these techniques effectively reduce errors and improve prediction robustness, which is crucial in agriculture where data can vary significantly due to different environmental conditions and plant species. The strength of ensemble methods lies in their ability to leverage diverse insights, leading to a more comprehensive understanding of plant health.

As automated systems for real-time monitoring of plant health continue to evolve, ensemble learning will play a vital role in creating reliable and scalable solutions for nutrient deficiency detection in precision agriculture. These techniques enable timely feedback, allowing farmers to make informed decisions that can significantly impact crop yield and quality. Ultimately, the integration of ensemble learning into agricultural practices will contribute to more sustainable farming and improved food security by ensuring that nutrient management strategies are tailored to the specific needs of each crop.

REFERENCES

[1] K. Taji, A. Sohail, T. Shahzad, B. Shoaib Khan, M. Adnan Khan, and K. Ouahada, ”An Ensemble Hybrid Framework: A Comparative Analysis of Metaheuristic Algorithms for Ensemble Hybrid CNN Features for Plants Disease Classifica- tion,” in IEEE Access, vol. 11, pp. 61886-61906, 2024, doi: 10.1109/ACCESS.2024.3389648.
[2] K. Venkatesh and K. J. Naik, ”An ensemble transfer learning for nutrient deficiency identification and yield-loss predic- tion in crop,” Multimedia Tools and Applications, pp. 1-27, 2024.
[3] L. Breiman, ”Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5-32, 2001, doi: 10.1023/A:1010933404324.
[4] N. Patel, A. Kumar, S. Patel, and S. N. Jha, ”Application of Machine Learning Techniques in Plant Nutrient Detection and Classification Using Image Processing,” Agricultural Research, vol. 10, no. 1, pp. 116-128, 2021, doi: 10.1007/s40003- 020-00503-1.
[5] S. Zhu et al., ”EfficientNet for Plant Disease Classification Using Images,” IEEE Access, vol. 10, pp. 3102-3110, 2022, doi: 10.1109/ACCESS.2022.3152705.
[6] S. Mohanty, D. P. Hughes, and M. Salathe´, ”Using deep learning for image-based plant disease detection,” Frontiers in Plant Science, vol. 7, p. 1419, Sep. 2016, doi: 10.3389/fpls.2016.01419.
[7] S. Phadikar and J. Sil, ”Rice Disease Identification Using Pattern Recognition Techniques,” Proceedings of the 11th International Conference on Computer and Information Technology (ICCIT ’08), Khulna, Bangladesh, 2008, pp. 420- 423, doi: 10.1109/ICCITECHN.2008.4803079.
[8] P. Singh, D. K. Vishwakarma, and S. K. Singh, ”Plant Disease Detection Based on Deep Learning and Convolutional Neural Network,” in Proc. International Conference on Innovative Computing and Communications (ICICC), vol. 78, 2020, pp. 261-269, doi: 10.1007/978-981-15-5113-0-25.
[9] S. Muthusamy and S. P. Ramu, ”IncepV3Dense: Deep Ensemble Based Average Learning Strategy for Identification of Micro-Nutrient Deficiency in Banana Crop,” in IEEE Access, vol. 12, pp. 73779-73792, 2024, doi: 10.1109/AC- CESS.2024.3405027.
[10] X. Dong, Z. Yu, W. Cao et al., ”A survey on ensemble learning,” Frontiers in Computational Science, vol. 14, pp. 241-258, 2020, https://doi.org/10.1007/s11704-019-8208-z.
[11] I. D. Mienye and Y. Sun, ”A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects,” in IEEE Access, vol. 10, pp. 99129-99149, 2022, doi: 10.1109/ACCESS.2022.3207287.
[12] E. Elfatimi, R. Eryigit, and L. Elfatimi, ”Beans Leaf Diseases Classification Using MobileNet Models,” in IEEE Access, vol. 10, pp. 9471-9482, 2022.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY