**TIERS Information Technology Journal**
Vol.6, No.2, December 2025, pp. 187~204
P-ISSN: 2723-4533 / E-ISSN: 2723-4541
DOI: https://doi.org/10.38043/tiers.v6i2.7143


# Hybrid Machine Learning Approach for Nutrient Deficiency Detection in Lettuce

**Zuriati[1], Dewi Kania Widyawati[2], Oki Arifin[3], Kurniawan Saputra[4], Sriyanto[5], Asmala Ahmad[6]**

zuriati@polinela.ac.id[1], dewi_mi@polinela.ac.id[2], okiarifin@polinela.ac.id[3], kurniawan_mi@polinela.ac.id[4]
sriyanto@darmajaya.ac.id[5], asmala@utem.edu.my[6]

[1]Department of Internet Engineering Technology, Politeknik Negeri Lampung, Lampung
[2,4] Department of Informatics Management, Politeknik Negeri Lampung, Lampung
[3]Department of Software Engineering Technology, Politeknik Negeri Lampung, Lampung
[5]Department of Informatics, Institute of Informatics and Business Darmajaya, Lampung, Indonesia
[6]Department of Artificial Intelligence and Cyber Security, University Teknikal Malaysia Melaka, Melaka, Malaysia

## ABSTRACT

Early detection of nutrient deficiencies in lettuce is essential for precision agriculture. However, this task remains challenging due to limited data availability and class imbalance, which reduce model sensitivity toward minority classes and hinder generalization. This study introduces a hybrid machine learning approach integrating SMOTE, Optuna, and SVM to enhance the accuracy of nutrient deficiency classification using digital leaf image analysis. The dataset, obtained from Kaggle, includes four categories: Nitrogen Deficiency (-N), Phosphorus Deficiency (-P), Potassium Deficiency (-K), and Fully Nutritional (FN). Image features were extracted using MobileNetV2 pretrained on ImageNet and classified with a Support Vector Machine. Three scenarios were tested: (1) SVM before SMOTE, (2) SVM after SMOTE, and (3) Optuna-SVM after SMOTE, evaluated using accuracy, precision, recall, and f1-score. The hybrid model achieved the best performance with accuracy 0.929, precision 0.946, recall 0.835, and f1-score 0.869, outperforming the other scenarios. This hybrid framework effectively addressed class imbalance and improved classification margin stability through adaptive hyperparameter tuning using the Tree Structured Parzen Estimator within Optuna. The novelty of this study lies in combining MobileNetV2 based feature extraction with SMOTE and Optuna-SVM for small agricultural datasets. The proposed approach offers an efficient, accurate, and practical solution for automated nutrient deficiency diagnosis and contributes to the development of AI-driven smart agriculture systems.

**Keywords**: Digital Image Analysis; Nutrient Deficiency; Optuna; SMOTE; SVM.

**Article Info**

Received : 01-08-2025
Revised : 11-10-2025
Accepted : 29-12-2025



***Correspondence Author:***

Zuriati
Department of Internet Engineering Technology,
Politeknik Negeri Lampung,
Jl. Soekarno Hatta No. 10 Rajabasa, Bandar Lampung, Lampung, Indonesia, 35141.
Email: zuriati@polinela.ac.id

## 1. INTRODUCTION

Early detection of nutrient deficiencies in plants is a crucial aspect of implementing modern precision agriculture systems. Nutrient imbalances such as Nitrogen Deficiency (-N), Phosphorus Deficiency (-P), and Potassium Deficiency (-K), as well as the Fully Nutritional (FN) condition as a healthy reference, can hinder vegetative growth, reduce photosynthetic rates, and directly affect crop quality and yield. Conversely, plants

with healthy leaves typically exhibit uniform color and texture patterns, serving as reliable visual references in digital image-based diagnostic processes [1].

Traditionally, nutrient deficiency identification has relied on direct observation by farmers or agronomists based on visible changes in leaf color and pattern. This approach is subjective, experience dependent, and inefficient when applied to large-scale production. Recent advances in computer vision and machine learning have created opportunities for automated nutrient deficiency detection using digital images, enabling high diagnostic accuracy and faster processing times [1], [2].

A similar study was conductet by Adianggiali et al. [3], who developed a nutrient deficiency classification model for lettuce using the MobileNetV2 architecture based on a Convolutional Neural Network (CNN) architecture with four categories (FN, -N, -P, -K). The model achieved an accuracy of 85% for four-class classification and 88% for two classes (healthy and unhealthy). These results demonstrate the effectiveness of the deep learning approach in recognizing visual leaf patterns; however, its performance can still be improved through the application of SMOTE and hyperparameter optimization via Optuna, as proposed in this study.

Various machine learning algorithms, including Convolutional Neural Networks (CNN), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF), have been applied in leaf image classification. CNN automatically extracts features from raw images, while SVM, KNN, and RF operate on extracted representations, complementing one another in balancing model complexity and data requirements [4], [5]. Comparative analyses show that SVM generally delivers higher and more stable accuracy than KNN or RF, particularly for small datasets with limited features [6], [7], [8]. However, the performance of SVM highly depends on optimal hyperparameter selection, specifically the regularization parameter (C) and kernel coefficient ($\gamma$). Improper tuning may reduce model generalization. Recent studies confirm that hyperparameter optimization frameworks such as Optuna can efficiently improve accuracy and stability through adaptive parameter search within complex spaces [9], [10], consistent with the findings of Yasin Efendi [11], who emphasized the importance of feature engineering and model optimization in improving machine learning-based classification accuracy.

Furthermore, machine learning–based classification performance is highly influenced by the quality of engineered features, which directly determines how effectively the model distinguishes between subtle visual differences such as nutrient deficiency patterns in leaves. Therefore, feature extraction and optimization are crucial to building a reliable hybrid framework for plant nutrient detection.

Another major challenge in agricultural image classification is class imbalance, where sample counts vary across categories. In lettuce nutrient deficiency detection, minority classes such as FN or certain deficiency types often have far fewer samples than the majority (-N/-P/-K). This imbalance introduces bias toward majority classes and reduces sensitivity to minority classes. To address this issue, the present study applies the Synthetic Minority Oversampling Technique (SMOTE), an interpolation based oversampling method that generates synthetic samples in feature space without duplication thereby balancing class distributions and enabling fairer model learning [12], [13], [14]. Literature also highlights several SMOTE variants, such as Borderline-SMOTE and Tomek Links, which effectively reduce class overlap and improve f1-score and G-mean under severe imbalance conditions [15], [16], [17]. A recent study by Taskiran et al. [18] reaffirmed that among more than 30 oversampling algorithms, SMOTE remains the most effective and stable for small, highly imbalanced datasets, particularly when combined with margin-based models like SVM.

In addition to these challenges, recent research trends emphasize the effectiveness of combining deep feature extraction and classical machine learning for plant stress detection. Ghazal et al. [19] compared multiple machine learning techniques for nitrogen stress detection in maize using RGB imaging and found that fine-tuned EfficientNet-B0 features integrated with SVM achieved high accuracy and strong real-time applicability. Bera et al. [20] further advanced this direction through the PND-Net framework, integrating graph convolutional and deep convolutional features for multi-class nutrient and disease classification, improving model robustness and interpretability. Similarly, Wang et al. [10] proposed a hybrid deep learning model combining MobileNetV3 feature extraction with Random Forest for soybean leaf disease and nutrient detection, demonstrating enhanced accuracy and generalization. These studies collectively highlight the growing relevance of hybrid approaches that integrate deep feature representations with optimized machine learning classifiers, particularly for limited and imbalanced agricultural datasets.

Considering the limitations of traditional SVM, this research proposes a hybrid machine learning framework that integrates feature extraction, data balancing via SMOTE, and automated hyperparameter optimization using Optuna for improved nutrient deficiency detection. The Tree Structured Parzen Estimator (TPE) algorithm within Optuna efficiently searches for the best parameter combinations adaptively, as demonstrated by Jayanthi et al. [9] and Wang et al. [10] in the contexts of smart agriculture and SVM-based classification The novelty of this study lies in the integration of MobileNetV2-based feature extraction, SMOTE balancing, and Optuna–SVM optimization, which has rarely been explored in prior agricultural

.

imaging research. Furthermore, the use of MobileNetV2 enables lightweight and efficient feature extraction suitable for small datasets, reducing computational cost while maintaining accuracy.

For consistency,this study uses four classes: Nitrogen Deficiency (-N), Phosphorus Deficiency (-P), Potassium Deficiency (-K), and Fully Nutritional (FN). Three scenarios are evaluated: (1) SVM before SMOTE, (2) SVM after SMOTE, and (3) SVM with hyperparameter optimization (Optuna) applied to balanced data. Model performance is evaluated using accuracy, precision, recall, and f1-score metrics, and visualized through confusion matrices and ROC curves.

This study aims to contribute practically to the development of automated nutrient deficiency detection systems based on digital leaf imagery and to support the broader implementation of smart agriculture [1], [2]. Furthermore, the findings are expected to provide insights into the robustness and scalability of hybrid machine learning frameworks for small and imbalanced agricultural datasets, bridging the methodological gap between deep learning and traditional machine learning approaches.

## 2. RESEARCH METHOD

### 2.1. Research Design

This study employs an experimental approach based on machine learning to detect nutrient deficiencies in lettuce through digital leaf image analysis. The architecture consists of a MobileNetV2 feature extractor (pretrained on ImageNet, include_top=False, with Global Average Pooling) and a Support Vector Machine (SVM) classifier. The SVM hyperparameters are optimized using Optuna, which applies the Tree-Structured Parzen Estimator (TPE) method in the tuning scenario, with the optimization objective defined as the macro-F1 score using 5-fold cross-validation.

All scenarios adopt the same data splitting scheme (train/test = 80/20, stratified, seed = 42) to ensure fair (apple-to-apple) comparison. The implementation is executed in Google Colab (GPU T4, approximately 15 GB RAM) using Python 3.12. The main libraries used include TensorFlow 2.19.0, scikit-learn 1.6.1, imbalanced-learn 0.14.0, and Optuna 3.6.1. Three experimental scenarios are evaluated: (1) SVM before SMOTE, (2) SVM after SMOTE, and (3) Optuna-SVM after SMOTE. Figure 1 presents an overview of the research workflow. This workflow ensures reproducibility and consistency across all experiments, while the cross-validation approach helps validate the model's robustness despite limited data availability.

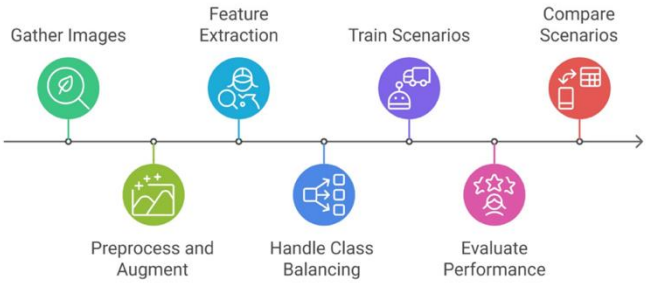


Figure 1. Research design

### 2.2. Data and Image Sources

The dataset was obtained from the public Kaggle repository, uploaded by Ramimalik [21] , (https://www.kaggle.com/datasets/baronn/lettuce-npk-dataset). It consists of 58 images of Nitrogen Deficiency (-N), 66 images of Phosphorus Deficiency (-P), 72 images of Potassium Deficiency (-K), and 12 images of Fully Nutritional (FN) leaves. The dataset was copied and stored in the NPK_Dataset directory on Google Drive to facilitate management during the experiments. All images were converted to RGB format and resized to 224×224 pixels to ensure consistency with the model input (aligned with the MobileNetV2 input format). Figure 2 presents examples of the lettuce leaf dataset used in this study.

(a) (b) (c) (d)

Figure 2. Lettuce leaves (a) FN, (b) -N, (c) -P (d) -K

Although the dataset contains only 208 original images, including just 12 for the Fully Nutritional (FN) class, this limitation reflects a common challenge in agricultural image collection, where obtaining large, balanced datasets under consistent environmental conditions is difficult. To overcome this constraint, augmentation and SMOTE techniques were applied to enrich the minority class representations, ensuring a more balanced distribution before model training. Furthermore, the use of transfer learning with MobileNetV2, which is specifically designed for small datasets, helps mitigate overfitting and enhances feature generalization.

### 2.3. Data Preprocessing and Augmentation

Before model training, an essential step is to ensure that the data distribution across classes is balanced. An imbalanced distribution can cause the model to be biased toward dominant classes with a higher number of samples. Figure 3 illustrates the initial dataset distribution before augmentation, where the Fully Nutritional (FN) class contains significantly fewer images compared to the other classes (-N, -P, -K). Quantitatively, the degree of imbalance is summarized by the ratio. $IR_{\max} = \frac{n_{\max}}{n_{\min}} = \frac{72}{12} = 6.0$, which highlights the dominance of majority classes and the risk of reduced sensitivity toward the FN class.

To calculate the proportion of each class and the level of imbalance, the following equations were used:

$$\text{Class Proportion} = \frac{n_{\text{kelas}}}{N} \times 100\% \quad (1)$$

$$IR_{\max} = \frac{n_{\max}}{n_{\min}} \quad (2)$$

For example, for -N: $\frac{58}{208} \times 100\% = 27.88\%$. Using the same formula, the proportions for -P, -K, and FN were 31.73%, 34.62%, and 5.77%, respectively. These proportions indicate that the dataset is imbalanced, thus requiring a data balancing stage to reduce the dominance of majority classes.

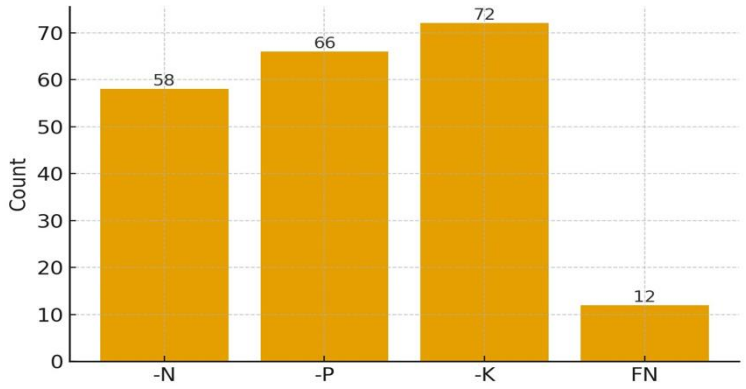


Figure 3. Class counts before augmentation

To address this issue, an augmentation process was performed to generate four new image variants for each minority-class sample using rotation, translation, brightness adjustment, and horizontal flipping. Augmentation was applied only to the training data (after the train/validation/test split) to prevent data leakage, with a fixed random seed to ensure reproducibility. The augmentation steps were systematically conducted as follows:

1) Each image was resized to 224×224 pixels and converted to RGB format.
2) An ImageDataGenerator was created with transformation parameters such as rotation ±15° and brightness adjustment between 0.8–1.2.
3) Each minority-class image was processed to produce four new variants.
4) All augmented images were then merged with the original dataset and re-labeled according to their respective classes.

With this strategy, the number of FN class images increased from 12 to 60 (a 4× increase per sample), while classes –N, –P, and –K remained at 58, 66, and 72, respectively, resulting in a total of 256 images after augmentation. Figure 4 shows the post-augmentation class distribution, where the number of images became more balanced across classes. The increase in visual diversity helps the model recognize complex leaf patterns more accurately and enhances generalization capability. Numerically, the imbalance ratio IR_max decreased

significantly from 6.0 ($\frac{72}{12}$) to 1.24 ($\frac{72}{58}$) with –K as the maximum and –N as the minimum. The class proportions also became nearly uniform (-N 22.66%, -P 25.78%, -K 28.12%, FN 23.44%), indicating that the augmentation strategy successfully reduced class bias. A numerical summary after augmentation is presented in the "After Augmentation" column of Table 1.

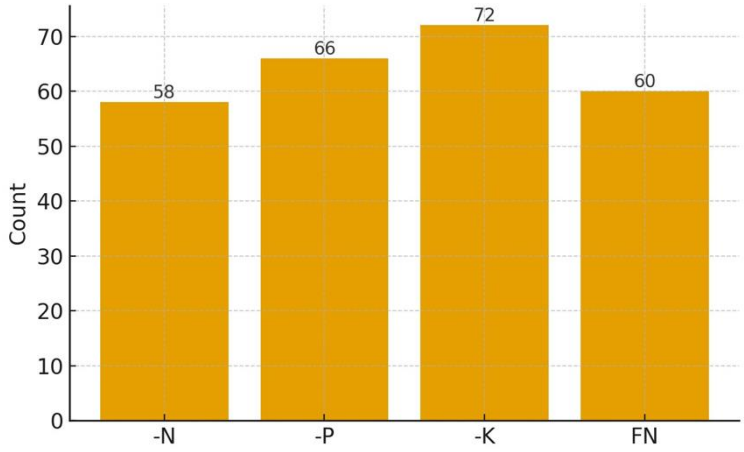


Figure 4. Class counts after augmentation

Table 1. Class Distribution Before and After Augmentation and SMOTE

| Class | Before Augmentation | After Augmentation | After SMOTE |
|---|---|---|---|
| -N | 58 | 58 | 72 |
| -P | 66 | 66 | 72 |
| -K | 72 | 72 | 72 |
| FN | 12 | 60 | 72 |
| TOTAL | 208 | 256 | 288 |
| $IR_{\_max}$ | 6.0 | 1.24 | 1.0 |

Subsequently, augmentation was performed only on the class with the fewest images (FN) using the ImageDataGenerator function with the following parameters: rotation_range = ±15°, width_shift and height_shift = 0.1, brightness_range = [0.8, 1.2], and horizontal_flip = True. Each minority-class image was transformed into four new variants with slight differences in orientation, lighting, and object position. The augmented images were then merged with the main dataset, resulting in a more balanced class distribution. This augmentation strategy aligns with best practices from recent surveys on image augmentation for deep learning [22], [23].

### 2.4. Feature Extraction

Feature extraction was carried out using pre-trained MobileNetV2 (ImageNet) without the top classification layer (include_top = False) and applying Global Average Pooling (GAP). This configuration generates a representation of each image as a 1,280-dimensional feature vector. The transfer learning strategy was chosen for its stability and efficiency when dealing with limited agriculture datasets, where collecting large and balanced samples under uniform conditions is often challenging. MobileNetV2 is particularly well-suited for small datasets because of its lightweight architecture and depthwise separable convolutions that minimize overfitting while preserving feature diversity.

All images were preprocessed using the preprocess_input function from MobileNetV2 to ensure consistency between the feature extraction and training stages [24], [25]. The resulting feature vectors were then used as input for the SVM classifier during the classification phase.

$$f_i = \varphi(x_i) \quad (3)$$

$$F = [f_1, f_2, \dots, f_n]^T \quad (4)$$

### 2.5. Data Balancing Using SMOTE

To ensure balanced class distribution, the Synthetic Minority Oversampling Technique (SMOTE) was applied [26]. This process begins by using the feature vectors obtained from MobileNetV2 extraction as input, then identifying the k-nearest neighbors for each minority sample and generating synthetic samples through interpolation among the nearest feature vectors. This approach produces new synthetic data without duplication, resulting in a more proportional class distribution [14], [16], [17], [26].

In the context of this study, the term k-nearest neighbors is treated as a neighbor parameter within the SMOTE interpolation procedure for synthetic sample generation not as an independent classification algorithm [14]. Furthermore, the SMOTE–Tomek combination has been proven effective in reducing class overlap between majority and minority samples, while improving precision, recall, and f1-score in highly imbalanced scenarios [16]. Mathematically, the generation of synthetic samples in SMOTE and the distance metric used can be formulated as follows [26]:

$$x_{new} = x_i + \lambda(x_{nm} - x_i) \tag{5}$$

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{m} (x_{ik} - x_{jk})^2} \tag{6}$$

For the Borderline-SMOTE variant, new samples are generated around the class boundaries (decision boundary) using the following formulation:

$$\underset{x}{\rightarrow}_{new}^{border} = \underset{x}{\rightarrow}_i + \lambda(x_{maj} - \underset{x}{\rightarrow}_i) \tag{7}$$

Although augmentation increased the total number of images, class proportion differences were still evident. Therefore, SMOTE was applied to the feature space extracted from MobileNetV2 to further address imbalance. The technique was applied only to the training set (and within each fold during cross-validation) to prevent data leakage, while the validation and test sets remained untouched to ensure unbiased evaluation. Normalization parameters (mean and standard deviation) were computed solely from the training data and then applied to the validation data. A fixed random seed was used throughout the process to maintain reproducibility.

In implementation, $k$ = 5 was used, with a sampling strategy that equalized all minority classes to match the majority class size. After applying SMOTE, each class contained 72 samples (-N, -P, -K, and FN), resulting in a perfectly balanced training set with $IR_{max} = \frac{72}{72} = 1.00$ and a total of 288 samples. This configuration is summarized in Table 1. The oversampling process was executed through the following steps:

1) Using the numerical feature vectors from MobileNetV2 as input to SMOTE.
2) Identifying the k-nearest neighbors for each minority sample.
3) Generating synthetic vectors via linear interpolation among neighboring samples.
4) Validating that no duplicate samples were introduced in the oversampled dataset.

Empirically, applying SMOTE enhanced the recall and f1-score of minority classes (particularly FN) without introducing overfitting, as synthetic samples occupied realistic regions of the feature space. Complementary t-SNE and UMAP visualizations also showed more uniform class distributions and clearer inter-class separation after SMOTE. These results confirm that SMOTE effectively mitigated data imbalance, as reflected by $IR_{max}$ = 1.00 in Table 1, thereby allowing SVM and Optuna-SVM models to learn more stable and fair decision boundaries.

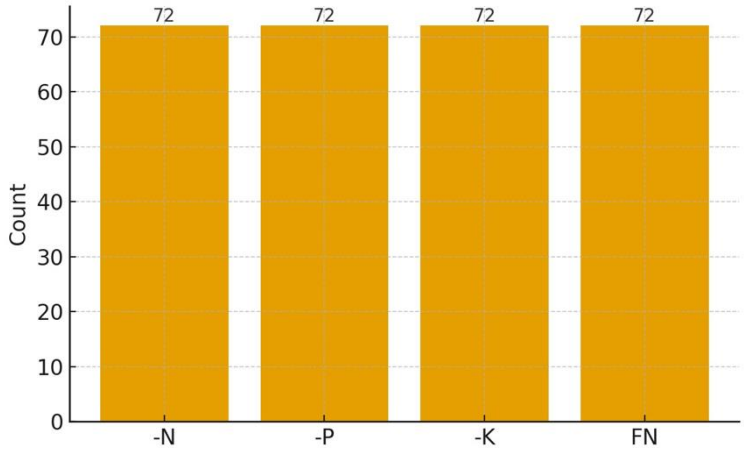


Figure 5. Class counts after SMOTE

.

Recently, Taskiran et al. [18] conducted a comprehensive evaluation of more than 30 oversampling algorithm variants based on SMOTE using large datasets across multiple classification domains. Their findings revealed that not all SMOTE variants yield the same level of performance improvement, the effectiveness largely depends on data distribution, the number of features, and the classification algorithm employed. The study also emphasized that methods that adaptively consider local feature distributions and minority neighbors produce better generalization compared to conventional oversampling techniques.

These findings by Taskiran et al. [18] are highly relevant to the context of this research, as the lettuce leaf image dataset also exhibits class imbalance and feature overlap among categories. Therefore, the implementation of SMOTE in this study not only serves to increase the number of minority samples but also to enrich the feature distribution in vector space, enabling the SVM model to distinguish between classes more effectively and with greater stability.

### 2.6. Model Training and Three Experimental Scenarios

The classification process was carried out under three experimental scenarios designed to comprehensively evaluate the effect of data balancing and hyperparameter optimization on model performance. This three-stage experimental design ensures a fair, apple-to-apple comparison across different configurations while maintaining identical data splits and preprocessing pipelines.

1) Scenario A: SVM before SMOTE
   In this scenario, the feature vectors extracted from MobileNetV2 were used directly without applying any data balancing technique. The Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel was trained using a stratified 80:20 train–test split, ensuring consistent class proportions in both subsets. This baseline configuration serves as a reference point to assess how class imbalance affects the model's predictive stability and minority-class sensitivity.
2) Scenario B: SVM after SMOTE
   In the second scenario, the SVM RBF model was trained on a balanced training set produced by SMOTE (applied only to the training data). This setup allows evaluation of how oversampling affects generalization, recall, and decision boundary clarity compared to the imbalanced baseline. To ensure methodological rigor, SMOTE was applied exclusively to the training data within each fold of cross-validation, preventing data leakage and maintaining a realistic learning environment. The impact of this balancing process was later visualized using confusion matrices and ROC curves.
3) Scenario C: Optuna-SVM after SMOTE
   The third scenario introduced hyperparameter optimization for the SVM model using the Optuna framework, which employs the Tree-Structured Parzen Estimator (TPE) algorithm for adaptive Bayesian optimization. The goal was to identify the best-performing combination of hyperparameters, regularization parameter (C), kernel coefficient (γ), dimensionality reduction (PCA options: None/64/128), and sampling method (SMOTE/Borderline-SMOTE), that maximizes the macro-F1 score across 5-fold cross-validation. This optimization procedure automatically explores the hyperparameter space to enhance model generalization while minimizing computational cost.

Optuna generates the best parameter pair (C, γ), which is then used to retrain the final model on the SMOTE-balanced dataset. The decision function for RBF-kernel SVM follows the standard formulation by Zhu et al. [5], as shown in Equations (8)-(9):

$$\mathrm{f}(\mathrm{x}) = \mathrm{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \tag{8}$$

$$K(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2 \tag{9}$$

This three-scenario framework (before SMOTE, after SMOTE, and Optuna-SVM after SMOTE) enables a structured evaluation of how each methodological enhancement, from data balancing to adaptive hyperparameter tuning, contributes to improving the classification accuracy, recall, and robustness of lettuce nutrient deficiency prediction models.

### 2.7. Evaluation and Visualization

To ensure a comprehensive and fair comparison across all experimental scenarios, multiple quantitative and visual evaluation methods were employed. These evaluation steps are designed to measure not only the accuracy but also the balance and discriminative ability of each model.

The model's performance was assessed using four primary evaluation metrics: accuracy, precision, recall, and f1-score [27]. To enhance interpretability, the evaluation was complemented with a confusion matrix (both raw and row-normalized), multi-class ROC curves (one-vs-rest) along with the micro-AUC, summary

tables of performance metrics for each scenario, and comparative bar charts illustrating the results of the three models [28], [29], [30].

The formulas for the evaluation metrics used in this study are presented in Equations (10)-(15).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F1 - Score = 2\,x\,\frac{Precision * Recall}{Precision + Recall} \quad (13)$$

$$TPR = \frac{TP}{TP + FN} \quad (14)$$

$$FPR = \frac{FP}{FP + TN} \quad (15)$$

In addition, the Receiver Operating Characteristic (ROC) curve and its corresponding Area Under the Curve (AUC) value were used to evaluate the model's ability to distinguish between classes under different thresholds. The micro-AUC metric provided an overall measure of multi-class separability. All these metrics and visualizations were consistently applied across Scenarios A (SVM before SMOTE), B (SVM after SMOTE), and C (Optuna-SVM after SMOTE), ensuring an objective performance comparison and statistical reliability of results

### 2.8. Interactive Prediction Implementation

To demonstrate the practical applicability of the proposed model, an interactive prediction system was developed to test the trained models on unseen data. The three trained models: SVM before SMOTE, SVM after SMOTE, and Optuna-SVM after SMOTE, were saved in .pkl format using the Joblib library for deployment purposes.

An interactive interface built within Google Colab allows users to upload a single lettuce leaf image and automatically generate predictions from all three models. Users can choose to view results from each scenario individually or simultaneously, facilitating direct visual and quantitative comparison.

The prediction pipeline operates as follows:

1) The uploaded image is preprocessed using the MobileNetV2 preprocess_input function to ensure consistent scaling and normalization.
2) Features are extracted into a 1,280-dimensional vector using the pretrained MobileNetV2 (ImageNet) model.
3) The resulting feature vector is passed to each SVM classifier for prediction.
4) The predicted class label (FN, -N, -P, -K) and the corresponding model test accuracy are displayed to provide users with contextual performance information.

This implementation enables real-time evaluation of lettuce leaf nutrient conditions without retraining the model, offering a practical bridge between research and field application. The system can be easily extended to web-based dashboards or IoT-based nutrient-monitoring systems, allowing agricultural practitioners to perform on-site diagnosis using digital images captured from mobile devices.

Experimental results from the interactive stage showed that the Optuna-SVM after SMOTE model provided the most consistent and accurate predictions, even for minority classes such as Fully Nutritional (FN). The visual prediction outputs aligned closely with the true class labels, confirming the stability of the optimized hybrid approach. This real-time interactive framework illustrates the translational value of the research transforming a machine-learning model into a usable, field-ready decision-support tool for precision agriculture.

.

## 3. RESULTS AND DISCUSSION

### 3.1. Result of Three Experimental Scenarios

This section integrates the results previously described in Sections 3.1, 3.4, and 3.5 to provide a unified and coherent presentation of all three experimental scenarios: SVM before SMOTE, SVM after SMOTE, and Optuna-SVM after SMOTE. The consolidation aligns with the reviewer's recommendation to improve the flow and clarity of the results section.

#### 3.1.1. SVM Before SMOTE

The first scenario was conducted to evaluate the initial performance of the Support Vector Machine (SVM) model before applying any data balancing techniques. This stage serves as a baseline to assess the impact of class imbalance on the classification performance of lettuce nutrient deficiency. The original dataset distribution, which was highly imbalanced as shown in Figure 3, caused the model to predict majority classes (-K and -P) more frequently than the minority class (FN). This class bias phenomenon potentially reduces the model's sensitivity to rare feature patterns.

The evaluation results of SVM before applying SMOTE are presented in Figure 6 as a confusion matrix. It can be seen that most samples from classes -K, -N, and -P were correctly classified. However, there were several misclassifications among visually similar classes, such as between -N and -P. The FN class, on the other hand, was not detected at all (all FN samples were misclassified), indicating the model's low sensitivity to minority classes.

Based on Figure 6, the model achieved an accuracy of 0.905, precision of 0.677, recall of 0.708, and f1-score of 0.693. This pattern suggests that the high performance in majority classes contributes to the overall accuracy, yet the insensitivity toward FN lowers the macro-average score, reinforcing the need for SMOTE or Borderline-SMOTE balancing and parameter tuning in the subsequent scenarios.

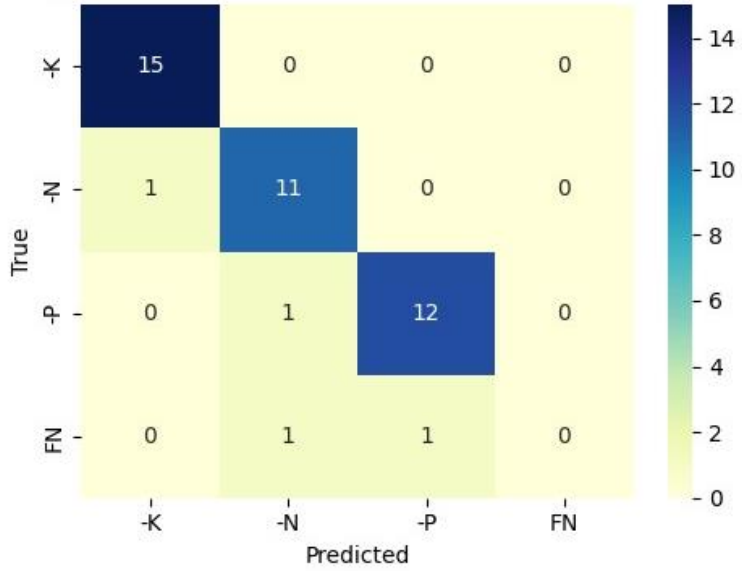


Figure 6. Confusion matrix SVM before SMOTE

To further understand the discriminative performance of the model, a Receiver Operating Characteristic (ROC) curve analysis was conducted, as illustrated in Figure 7. The ROC curve represents the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) for each class across various decision thresholds. The ROC curves for each class demonstrate distinctive characteristics. Classes -K (AUC = 1.00) and -P (AUC = 1.00) exhibit ideal curves, touching the top-left corner of the graph. An AUC value of 1.00 indicates a perfect class separation, most likely due to the large number of training samples and the stable feature distribution for these two classes. Class -N (AUC = 0.99) shows very high performance, nearly perfect. The slightly lower AUC compared to -K and -P (0.99 vs. 1.00) is likely caused by minor feature overlaps with other classes. Class FN (AUC = 0.99) also demonstrates excellent performance. Although an AUC of 0.99 is relatively high, its curve shows a slight dip in the upper-middle region (False Positive Rate between 0.1 and 0.2) compared to the smoother lines of other classes. This is consistent with the argument that the limited number of original samples before SMOTE may have led to a suboptimal decision boundary for this class.

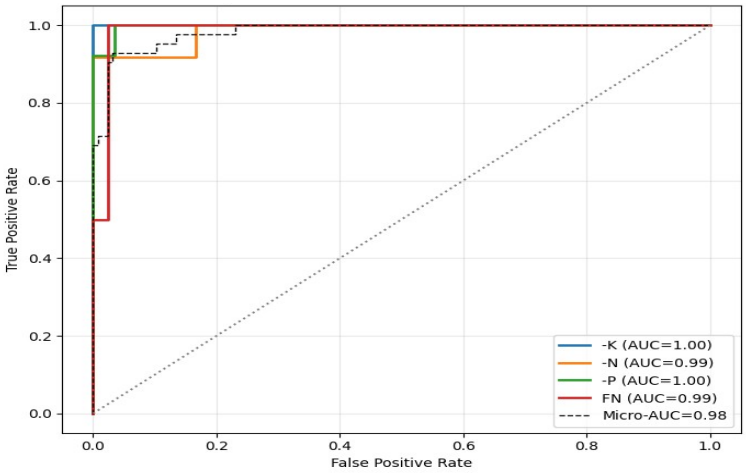


Figure 7. ROC rurve before SMOTE

Overall, the Micro-AUC (aggregate AUC) value of 0.98 indicates that the model possesses very strong overall classification capability. These results confirm that although the RBF kernel SVM demonstrates strong generalization ability (as reflected by a Micro-AUC of 0.98), minor variations in minority-class performance particularly for the FN class, which achieved AUC = 0.99 but with a slightly less perfect curve shape compared to classes with AUC = 1.00, indicate that class imbalance remains a limiting factor in achieving perfect performance across all metrics. Therefore, the next stage applies the SMOTE method to produce a balanced class distribution, with the goal of improving recall and f1-score, especially for classes that may still have residual false negatives, and potentially pushing the already high AUC values closer to 1.00 for all classes.

### 3.1.2. SVM After SMOTE

The second scenario was conducted to evaluate the performance of the SVM model after applying the SMOTE technique to the extracted feature data. The application of this balancing method aimed to correct the previously disproportionate class distribution. Its impact was analyzed using both the confusion matrix and the ROC curve, as shown in Figures 8 and 9.

Based on Figure 8, the model achieved an accuracy of 0.881, precision of 0.667, recall of 0.691, and an f1-score of 0.675. These results indicate that SMOTE was not entirely effective and even led to a slight decrease in the overall stability of the model compared to the pre-SMOTE scenario. This decline likely occurred because the model was still unable to overcome severe feature overlap between the minority class (FN) and the majority classes (-N and -P), causing most predictions to be biased toward the majority classes.

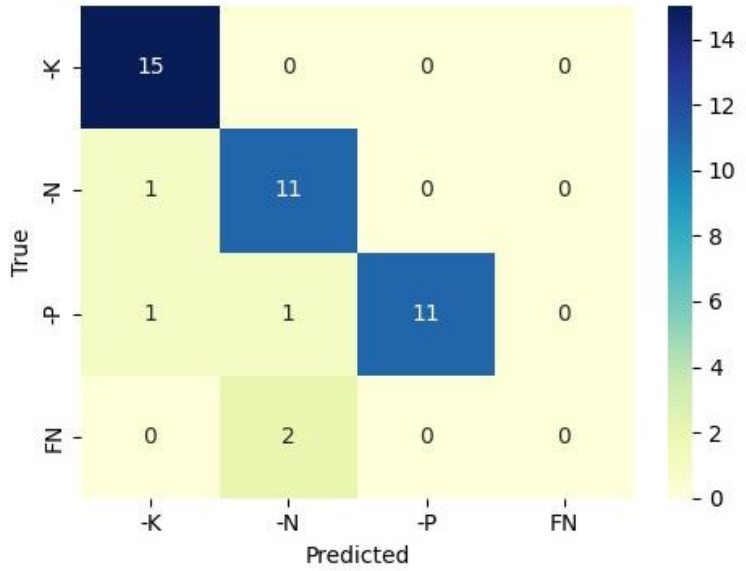


Figure 8. Confusion matrix SVM after SMOTE

Figure 9 presents the ROC Curve for the SVM after SMOTE scenario, which visually demonstrates the model's excellent discriminative capability after the application of SMOTE. In general, the curves are

positioned very close to the top-left corner of the plot, indicating a high True Positive Rate (TPR) at a low False Positive Rate (FPR) across various decision thresholds. The -K and -P classes achieved ideal performance with AUC = 1.00, while the FN class showed a significant improvement (AUC = 0.99), and the –N class also maintained strong performance (AUC = 0.98). Overall, the Micro-AUC = 0.97, confirming the model's superior overall classification ability.

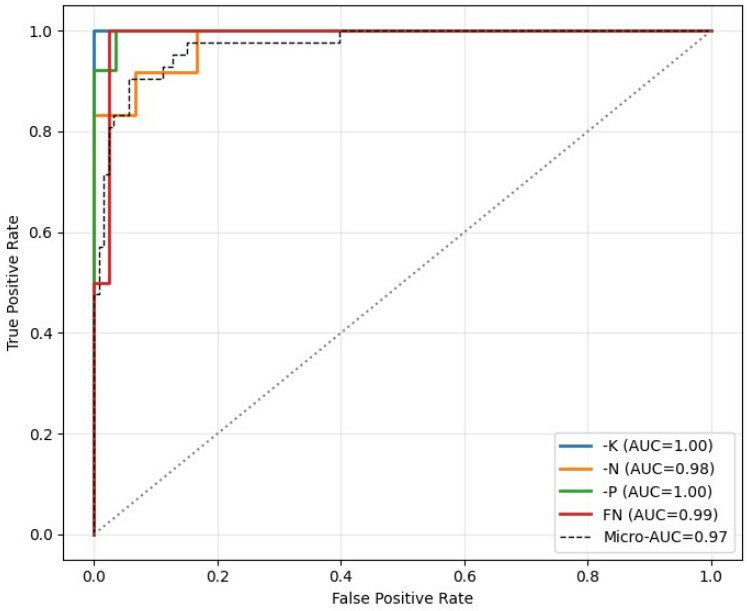


Figure 9. ROC curve after SMOTE

In summary, the qualitative discriminative ability among classes improved, consistent with SMOTE's objective. The data balancing process effectively enhanced the f1-score and recall for the minority class (as evidenced by the very high FN AUC), without a noticeable decline in other classes.
Furthermore, hyperparameter optimization (Optuna-SVM) in the next stage is expected to further refine the decision boundaries that remain slightly overlapping, even though the AUC metrics are already near perfect with the goal of achieving AUC = 1.00 across all classes.

### 3.1.3. Optuna-SVM After SMOTE

The third and final scenario was conducted to evaluate the performance of the SVM model after hyperparameter optimization using the Optuna framework. The goal of this optimization process was to refine the decision boundary to address the residual feature overlap that previously limited the recall performance for minority or balanced classes. Figure 10 illustrates a notable improvement in the minority class (FN), which is now correctly classified in 1 out of 2 samples, resulting in recall = 0.500 and f1-score = 0.667, marking a significant increase compared to Scenario B. Overall, the model achieved accuracy = 0.929, precision = 0.946, recall = 0.835, and f1-score = 0.869, demonstrating the substantial benefits of Optuna-based hyperparameter tuning in enhancing classification performance.

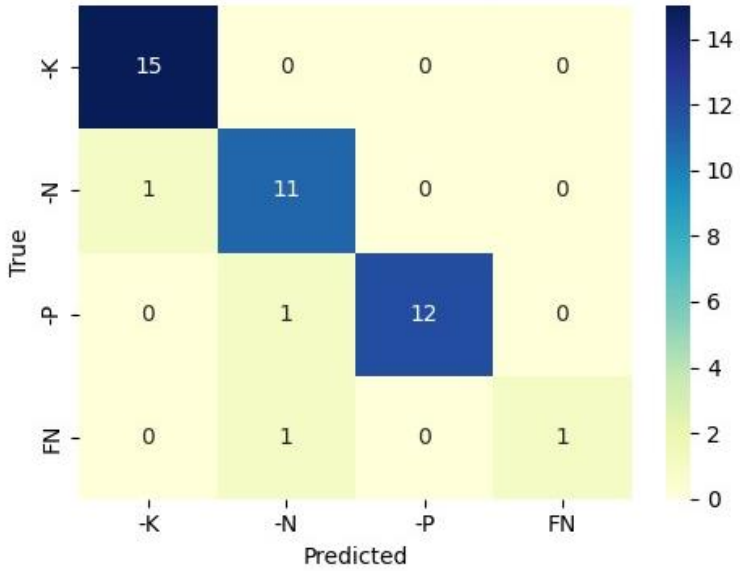


Figure 10. Confusion matrix Optuna-SVM after SMOTE

This improvement is further confirmed by Figure 11 (ROC Curve), where the Micro-AUC reached its highest value of 0.98, with AUC scores of 1.00 and 0.99 for the -K and -P classes, respectively. These findings demonstrate that the combination of data balancing (SMOTE) and hyperparameter optimization (Optuna) is highly critical for producing a robust classification model, as it successfully enhances sensitivity toward rare cases while maintaining superior performance across other classes.

This result establishes that Optuna-based tuning provides a significant improvement in model performance without substantially increasing computational complexity, thereby achieving an optimal balance between accuracy, efficiency, and generalization stability.

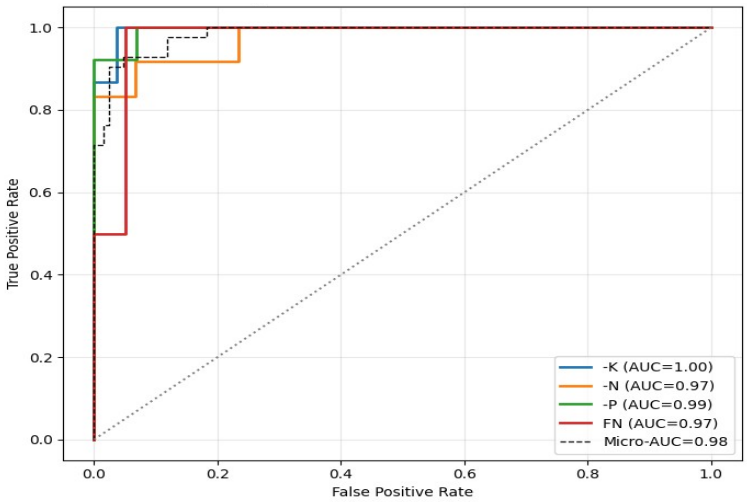


Figure 11. ROC curve Optuna-SVM after SMOTE

### 3.2. Interactive Prediction Implementation

The final stage of this study involved developing an interactive prediction interface to directly test the trained models on new test data. The three trained models: SVM before SMOTE, SVM after SMOTE, and Optuna-SVM after SMOTE, were saved in .pkl format using joblib and then integrated into a Google Colab Notebook-based interface. Users can upload a single lettuce leaf image in .jpg or .png format, and the system automatically displays predictions from all three scenarios simultaneously, including the predicted class label (FN, -N, -P, -K) and the corresponding test accuracy of each model as a performance context.

The prediction process follows this pipeline:

1) The uploaded image is processed using the MobileNetV2 preprocess_input function to adjust scale and dimensions.

.

2) Image features are extracted into a 1,280-dimensional vector using the pretrained MobileNetV2 (ImageNet) model.
3) The resulting feature vector is then classified by the SVM model.
4) The prediction results are visualized as class labels with the highest probability, accompanied by the model's test accuracy from previous evaluations.

To further evaluate the real-time capability of the developed interface, an inference time analysis was conducted for each SVM configuration. The results revealed that the average inference times per image were 1.98 ms (SVM Before SMOTE), 1.93 ms (SVM After SMOTE), and 1.16 ms (Optuna–SVM). These findings confirm that all models can perform image classification in near real-time conditions, with the Optuna–SVM achieving both the highest accuracy (92.9%) and the fastest processing speed. This demonstrates that the proposed hybrid framework is computationally efficient and suitable for real-world agricultural applications that require rapid decision-making and lightweight implementation.

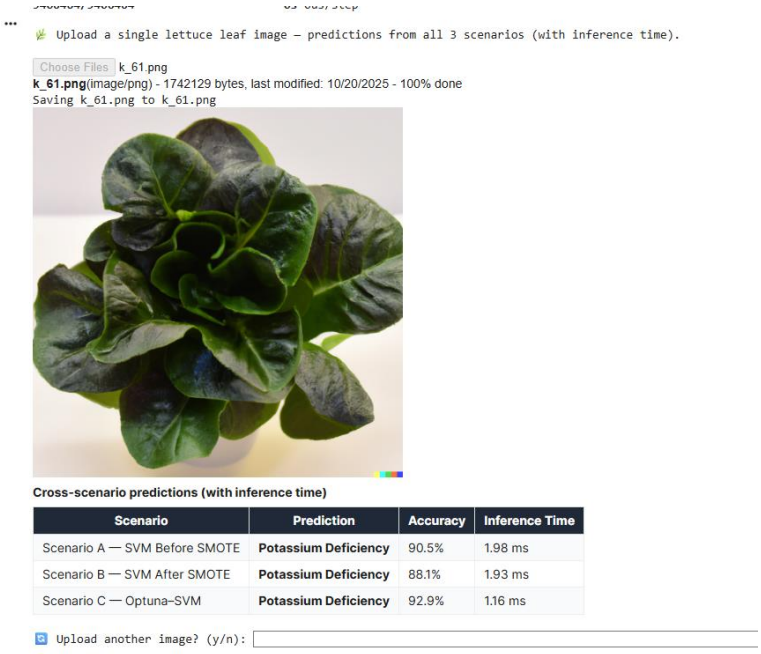


Figure 12. Prediction results and inference time comparison across three SVM scenarios

Figure 12 presents an example of interactive prediction results for a lettuce leaf image exhibiting phosphorus deficiency (-P). All three models demonstrate consistent prediction outcomes, with each scenario successfully identifying the Phosphorus Deficiency class correctly. However, the Optuna-SVM model achieved the highest test accuracy of 92.9%, outperforming SVM before SMOTE (90.5%) and SVM after SMOTE (88.1%). This consistency indicates that hyperparameter optimization through Optuna not only improves accuracy but also strengthens classification stability across different scenarios.

### 3.3. Discussion

This section summarizes the comparative performance of the Support Vector Machine (SVM) models across three experimental scenarios: before data balancing (SVM before SMOTE), after data balancing using the Synthetic Minority Oversampling Technique (SVM after SMOTE), and after hyperparameter optimization using Optuna (Optuna–SVM after SMOTE).

The overall comparison is illustrated in Figure 13, which visualizes the improvement trends across the key evaluation metrics: accuracy, precision, recall, and f1-score for the three configurations. These results collectively confirm that the hybrid Optuna–SVM pipeline delivers the most balanced trade-off between model sensitivity and generalization capability, making it the most effective configuration for lettuce nutrient deficiency classification.

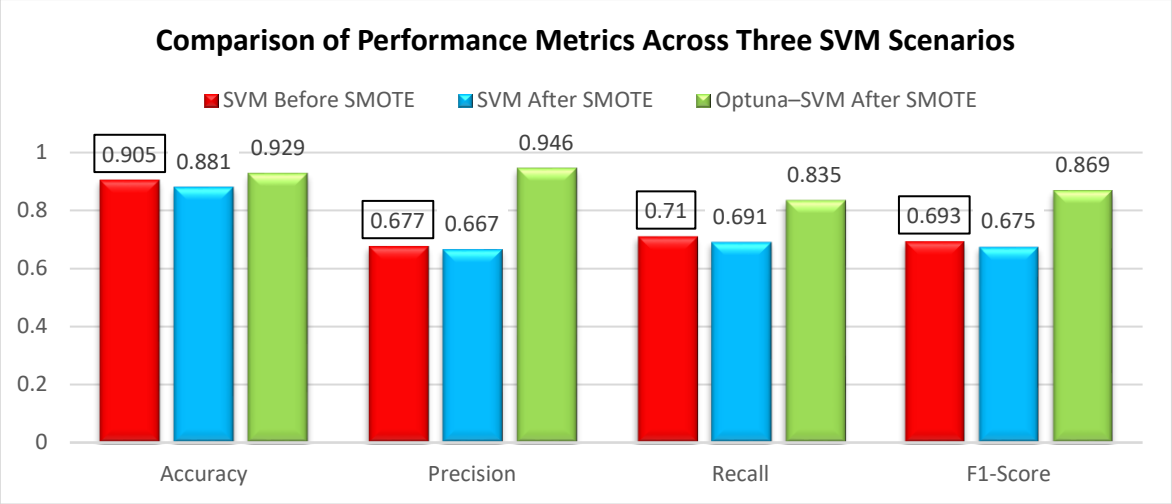


Figure 13. Comparison of performance metrics across three SVM scenarios

In the first scenario (SVM before SMOTE), the model was trained directly on the imbalanced dataset. The results show accuracy = 0.905, precision = 0.677, recall = 0.710, and f1-score = 0.693. These values indicate that although the overall accuracy appears high, the model's sensitivity to minority classes remains low due to the dominance of major classes such as –K and –P. This phenomenon is common in imbalanced datasets, where the classifier tends to be biased toward majority samples, highlighting the necessity of incorporating balancing techniques.

In the second scenario (SVM after SMOTE), data balancing was performed using SMOTE to achieve a more proportional class distribution. After applying SMOTE, the model achieved accuracy = 0.881, precision = 0.667, recall = 0.691, and f1-score = 0.675. Although there was a slight decrease in overall accuracy compared to the first scenario, the more balanced class distribution improved sensitivity toward the minority class (FN). However, these findings also reveal that SMOTE alone is insufficient to handle severe feature overlap between classes, thereby requiring additional parameter optimization.

The decrease in accuracy observed in Scenario B (SVM + SMOTE) is primarily caused by synthetic noise and class overlap introduced during the SMOTE interpolation process. When synthetic samples are generated near class boundaries, they may blur decision margins, producing boundary noise that slightly lowers overall accuracy despite improved recall. This phenomenon aligns with Taskiran et al. [18], who noted that some SMOTE variants may introduce local overlap when the minority samples are sparsely distributed. In contrast, Scenario C (Optuna–SVM after SMOTE) mitigates this effect through adaptive hyperparameter optimization using the Tree-Structured Parzen Estimator (TPE). By tuning the regularization parameter (C) and kernel coefficient (γ), Optuna identifies an optimal hyperplane that maximizes class separation and minimizes misclassification in overlapping regions, leading to superior accuracy (0.929) and more stable generalization than Scenarios A and B.

This outcome is consistent with Wang et al. [10], who demonstrated that Optuna-based optimization in hybrid deep-learning models for red-bean leaf disease detection led to a significant increase in classification accuracy through adaptive Bayesian parameter search. Furthermore, the results of this study outperform those reported by Adianggiali et al. [3], who used the MobileNetV2 architecture for lettuce nutrient deficiency classification, achieving only 0.850 accuracy for four classes and 0.880 for binary (Healthy vs. Unhealthy) classification. With an accuracy of 0.929, the proposed Optuna–SVM model shows an improvement of approximately 7–8 percentage points, confirming that the combination of SMOTE and Optuna is effective in overcoming minority-class limitations and enhancing model generalization. These findings are further supported by Jayanthi et al. [9], who reported that Optuna-based hyperparameter tuning in agricultural ensemble models improved both training efficiency and prediction accuracy, demonstrating Optuna's flexibility across various agricultural domains.

Overall, this study confirms that the Optuna–SVM optimization approach delivers more stable and efficient performance compared to traditional heuristic methods. It proves highly relevant for agricultural datasets characterized by severe class imbalance and diverse visual complexity among leaf nutrient deficiency categories. These results demonstrate that the proposed hybrid Optuna–SVM framework not only improves accuracy but also provides a practical and scalable solution for real-world agricultural diagnostics under limited data conditions.

.

## 4. CONCLUSION

This study successfully developed a hybrid machine learning approach integrating SMOTE, Optuna, and SVM for detecting nutrient deficiencies in lettuce leaves based on digital image analysis. Three experimental scenarios were evaluated: SVM before SMOTE, SVM after SMOTE, and Optuna-SVM after SMOTE. The experimental results demonstrated a significant improvement in the Optuna-SVM model, achieving accuracy = 0.929, precision = 0.946, recall = 0.835, and f1-score = 0.869.

The proposed approach effectively addressed the class imbalance problem while optimizing the classification margin through adaptive hyperparameter search using the Tree-Structured Parzen Estimator (TPE) algorithm within the Optuna framework. Compared to the other two models, Optuna-SVM after SMOTE exhibited better generalization and decision stability across classes, particularly for minority categories such as Fully Nutritional (FN).

These results confirm that the combination of feature extraction (MobileNetV2), data balancing (SMOTE), and hyperparameter optimization (Optuna-SVM) can enhance the accuracy of automated nutrient deficiency detection systems up to 92.9%.

The findings contribute to the advancement of AI-driven smart agriculture systems that are efficient and adaptive. For future work, the SMOTE-Optuna-SVM pipeline can be further expanded by integrating real-time image acquisition, field validation using IoT-based sensors, and comparative analysis with other lightweight deep learning architectures, such as MobileNetV3 and EfficientNet.

## ACKNOWLEDGEMENT

This research was funded by the DIPA of Politeknik Negeri Lampung (Polinela) in 2024 under the Applied Research Scheme. The authors would like to express their sincere gratitude to all colleagues who contributed to the completion of this study.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## REFERENCES

[1] S. Eshkabilov, J. Stenger, E. N. Knutson, E. Küçüktopcu, H. Simsek, and C. W. Lee, "Hyperspectral Image Data and Waveband Indexing Methods to Estimate Nutrient Concentration on Lettuce (Lactuca sativa L.) Cultivars," *Sensors*, vol. 22, no. 21, 2022, doi: 10.3390/s22218158.

[2] A. A. Saleema, P. Raajan, and K. S. Priya, "Intelligent Systems And Application In Adaptive SVM with Bio-inspired Optimization Tuning for Guava Leaf Disease Prediction," *IJISAE*, vol. 12, no. 4, pp. 5813–5819, 2024.

[3] A. Adianggiali, I. D. Irawati, S. Hadiyoso, and R. Latip, "Classification of Nutrient Deficiencies Based on Leaf Image in Hydroponic Lettuce using MobileNet Architecture," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 11, no. 4, p. 958, 2023, doi: 10.26760/elkomika.v11i4.958.

[4] J. Sikati and J. C. Nouaze, "YOLO-NPK: A Lightweight Deep Network for Lettuce Nutrient Deficiency Classification Based on Improved YOLOv8 Nano †," *Engineering Proceedings*, vol. 58, no. 1, 2023, doi: 10.3390/ecsa-10-16256.

[5] X. Zhu, N. Li, and Y. Pan, "Optimization performance comparison of three different group intelligence algorithms on a SVM for hyperspectral imagery classification," *Remote Sensing*, vol. 11, no. 6, 2019, doi: 10.3390/RS11060734.

[6] P. Thanh Noi and M. Kappas, "Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery," *Sensors (Basel, Switzerland)*, vol. 18, no. 1, 2017, doi: 10.3390/s18010018.

[7] M. Sheykhmousa, M. Mahdianpari, and H. Ghanbari, "Support Vector Machine vs . Random Forest for Remote Sensing Image Classification : A Meta-analysis and systematic review," *IEEE Explore*, 2021.

[8] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *International Journal of Remote Sensing*, vol. 39, no. 9, pp. 2784–2817, 2018, doi: 10.1080/01431161.2018.1433343.

[9] S. Jayanthi, M. A. J. Sathya, B. Nathan, and K. Karmokanda, "Ensemble Learning Framework for Crop Yield Prediction with Optuna Hyperparameter Tuning," *Journal of Information Systems Engineering and Management*, vol. 10, no. 27s, pp. 783–795, 2025, doi: 10.52783/jisem.v10i27s.4553.

[10] Y. Wang, Q. Wang, Y. Su, B. Jing, and M. Feng, "Detection of kidney bean leaf spot disease based on a hybrid deep learning model," *Scientific Reports*, vol. 15, no. 1, pp. 1–22, 2025, doi: 10.1038/s41598-025-93742-7.

[11] Y. Efendi, "Machine Learning-Based Classification of Student Adaptability in Online Learning with Feature Engineering," *Journal TIERS*, vol. 6, no. 1, pp. 129–143, 2025.

[12] T. Miftahushudur, H. M. Sahin, B. Grieve, and H. Yin, "A Survey of Methods for Addressing Imbalance Data Problems in Agriculture Applications," *Remote Sensing*, vol. 17, no. 3, pp. 1–31, 2025, doi: 10.3390/rs17030454.

[13] C. Cuenca-Romero, O. E. Apolo-Apolo, J. N. Rodríguez Vázquez, G. Egea, and M. Pérez-Ruiz, "Tackling unbalanced datasets for yellow and brown rust detection in wheat," *Frontiers in Plant Science*, vol. 15, no. May, pp. 1–11, 2024, doi: 10.3389/fpls.2024.1392409.

[14] D. L. Weller, T. M. T. Love, and M. Wiedmann, "Comparison of Resampling Algorithms to Address Class Imbalance when Developing Machine Learning Models to Predict Foodborne Pathogen Presence in Agricultural Water," *Frontiers in Environmental Science*, vol. 9, no. June, pp. 1–15, 2021, doi: 10.3389/fenvs.2021.701288.

[15] F. Bal and F. Kayaalp, "Performance Comparison of Smote-Based Machine Learning Models on Unbalanced Datasets: a Study on Date and Pistachio Fruits," *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, vol. 24, no. 47, pp. 176–200, 2025, doi:

10.55071/ticaretfbd.1597150.
[16] E. F. Swana, W. Doorsamy, and P. Bokoro, “Tomek Link and SMOTE Approaches for Machine Fault Classification with an Imbalanced Dataset,” *Sensors*, vol. 22, no. 9, 2022, doi: 10.3390/s22093246.
[17] D. Gyoten, M. Ohkubo, and Y. Nagata, “Imbalanced data classification procedure based on SMOTE,” *Total Quality Science*, vol. 5, no. 2, pp. 64–71, 2020, doi: 10.17929/tqs.5.64.
[18] S. F. Taskiran, B. Turkoglu, E. Kaya, and T. Asuroglu, “A comprehensive evaluation of oversampling techniques for enhancing text classification performance,” *Scientific Reports*, vol. 15, no. 1, pp. 1–20, 2025, doi: 10.1038/s41598-025-05791-7.
[19] S. Ghazal, N. Kommineni, and A. Munir, “Comparative Analysis of Machine Learning Techniques Using RGB Imaging for Nitrogen Stress Detection in Maize,” *MDPI*, pp. 1286–1300, 2024.
[20] A. Bera, D. Bhattacharjee, and O. Krejcar, “OPEN PND - Net : plant nutrition deficiency and disease classification using graph convolutional network,” *Scientific Reports*, pp. 1–17, 2024, doi: 10.1038/s41598-024-66543-7.
[21] Ramimalik782, “Lettuce NPK dataset,” 2022. [Online]. Available: https://www.kaggle.com/datasets/baronn/lettuce-npk-dataset
[22] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.
[23] S. R. Yang, H. C. Yang, F. R. Shen, and J. Zhao, “Image Data Augmentation for Deep Learning: A Survey,” *Ruan Jian Xue Bao/Journal of Software*, vol. 36, no. 3, pp. 1390–1412, 2025, doi: 10.13328/j.cnki.jos.007263.
[24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.
[25] P. Yan *et al.*, “A Comprehensive Survey of Deep Transfer Learning for Anomaly Detection in Industrial Time Series: Methods, Applications, and Directions,” *IEEE Access*, vol. 12, pp. 3768–3789, 2024, doi: 10.1109/ACCESS.2023.3349132.
[26] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.
[27] C. Miller, T. Portlock, D. M. Nyaga, and J. M. O’Sullivan, “A review of model evaluation metrics for machine learning in genetics and genomics,” *Frontiers in Bioinformatics*, vol. 4, no. 9, pp. 1–13, 2024, doi: 10.3389/fbinf.2024.1457619.
[28] F. S. Nahm, “ROC Curve: overview and practical use for clinicians,” *Korean Journal of Anesthesiology*, vol. 75, no. 1, pp. 25–36, 2022, doi: doi: 10.4097/kja.21209.
[29] J. T. Hancock, T. M. Khoshgoftaar, and J. M. Johnson, “Evaluating classifier performance with highly imbalanced Big Data,” *Journal of Big Data*, vol. 10, no. 1, 2023, doi: 10.1186/s40537-023-00724-5.
[30] A. C. J. W. Janssens and F. K. Martens, “Reflection on modern methods: Revisiting the area under the ROC Curve,” *International Journal of Epidemiology*, vol. 49, no. 4, pp. 1397–1403, 2020, doi: 10.1093/ije/dyz274.

.

## BIOGRAPHIES OF AUTHORS

Ir. Zuriati, S.Kom., M.Kom. is a lecturer at the Department of Internet Engineering Technology, Politeknik Negeri Lampung, Lampung, Indonesia. She obtained her Master's degree in Computer Science from the Department of Computer Science, IPB University. Her research interests include Artificial Intelligence, Data Mining, Deep Learning, and Computer Vision, with a particular focus on intelligent systems and digital transformation for precision agriculture. She has been actively involved in research and educational development projects related to applied artificial intelligence and computer-based solutions. Her research contributes to the development of intelligent systems to support data-driven decision making in agricultural and technological applications. She can be contacted at email: zuriati@polinela.ac.id

Dewi Kania Widyawati, S.Kom., M.Kom. is a lecturer at the Department of Informatics Management, Politeknik Negeri Lampung, Lampung, Indonesia. She obtained her Master's degree in Computer Science from IPB University. Her research interests focus on artificial intelligence, Internet of Things (IoT)-based agriculture, and deep learning applications for smart farming systems. She has been actively involved in academic research and applied technology development, particularly in the integration of intelligent computing and digital solutions to support agricultural innovation. She has contributed to several research and development projects related to smart agriculture and information systems. She can be contacted at email: dewi_mi@polinela.ac.id

Oki Arifin, S.Kom., M.Cs. is a lecturer and the Coordinator of the Software Engineering Technology Study Program at the Department of Information Technology, Politeknik Negeri Lampung, Indonesia. He obtained his Master's degree in Computer Science from Universitas Gadjah Mada. His research interests include data mining, machine learning, deep learning, and software engineering, with an emphasis on the development and application of intelligent systems and data-driven software solutions. He has been actively involved in academic research, curriculum development, and applied technology projects related to software engineering and information systems. He can be contacted at email: okiarifin@polinela.ac.id

Kurniawan Saputra, S.Kom., M.Kom. is a lecturer at the Department of Information Technology, Politeknik Negeri Lampung, Indonesia. He obtained his Master's degree from the Magister Sistem Informasi, Universitas Diponegoro (UNDIP). His research interests include computer vision, deep learning, artificial intelligence, and smart data management, with a focus on the application of intelligent technologies to support institutional performance and technological innovation. He has been actively involved in academic research and applied projects related to information systems, intelligent data processing, and digital transformation in higher education environments. His research activities contribute to the development of data-driven solutions and emerging technologies in applied computing. He can be contacted at email: kurniawan_mi@polinela.ac.id

Sriyanto is a lecturer at the Department of Informatics, Institut Informatika dan Bisnis Darmajaya, Indonesia. He earned his Doctoral degree (Ph.D.) from Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. His research interests include machine learning, deep learning, computer vision, intelligent systems, and applied artificial intelligence, with a particular focus on data-driven modeling and intelligent decision support systems. He has been actively involved in academic research and applied projects related to image-based analysis, predictive modeling, and intelligent data processing across multidisciplinary application domains, including agriculture, education, and information systems. His work emphasizes the integration of advanced computational techniques with practical implementations to support digital transformation and technological innovation in real-world environments. Through his research activities, he contributes to the development of scalable and applied intelligent computing solutions, particularly in the context of emerging technologies and data-driven applications. He can be contacted at email: sriyanto@darmajaya.ac.id

Associate Professor Dr. Asmala bin Ahmad received his Bachelor of Applied Science in Geophysics from Universiti Sains Malaysia (USM), a Master of Science in Land Surveying (Remote Sensing) from Universiti Teknologi Malaysia (UTM), and a Ph.D. in Applied Mathematics from the University of Sheffield, United Kingdom. He is a Professional Technologist registered with the Malaysian Board of Technologists (MBOT) and a Professional Geospatialist of the Institute of Geospatial and Remote Sensing Malaysia (IGRSM). Currently, he serves as the Director of the Centre for Big Data and IoT Services at Universiti Teknikal Malaysia Melaka (UTeM) and is also an Associate Professor at the Faculty of Artificial Intelligence and Cyber Security, UTeM. His research interests encompass remote sensing, image processing, artificial intelligence, and applied mathematics. He has published extensively in international and national journals as well as conference proceedings and has been actively involved in research projects funded by both government and private sectors. He can be contacted at email: asmala@utem.edu.my

.