# Multispectral UAV Image-based CNN Model for Cassava Plant Health Classification Using Vegetation Indices

Sufiatul Maryana [1,2,*], Yeni Herdiyeni [1,*], Sri Wahjuni [1], and Edi Santosa [3]

[1] Department of Computer Science, School of Data Science, Mathematics and Informatics, IPB University, Bogor, Indonesia
[2] Vocational School, Pakuan University, Bogor, Indonesia
[3] Department of Agronomy and Horticulture, Faculty of Agriculture, IPB University, Bogor, Indonesia
Email: sufiatulmaryana@apps.ipb.ac.id (S.M.); yeni.herdiyeni@apps.ipb.ac.id (Y.H.); my_juni04@apps.ipb.ac.id (S.W.); edi_santosa@apps.ipb.ac.id (E.S).
*Corresponding author

***Abstract*—This study introduces a Convolutional Neural Network (CNN)-based approach for classifying cassava plant health using multispectral Unmanned Aerial Vehicle (UAV) imagery and vegetation indices. The vegetation indices, including Normalized Difference Vegetation Index (NDVI), Green Normalized Vegetation Index (GNDVI), EVI, Normalized Difference Red Edge (NDRE), and Leaf Chlorophyll Index (LCI), were statistically analyzed using one-way Analysis of Variance (ANOVA) and the Friedman test to evaluate their separability across different growth stages and health conditions. The CNN model was evaluated using datasets collected at 2, 4, and 7 months of cassava growth. The experimental results show that the proposed CNN achieved classification accuracies of 0.84, 0.81, and 0.96 for the 2-, 4-, and 7-month datasets, respectively, indicating that classification performance varied across growth stages. The lower accuracy at the intermediate stage was associated with increased canopy complexity and overlapping spectral characteristics. The proposed CNN was further compared with classical machine learning and pre-trained CNN baselines, including Random Forest, Support Vector Machine (SVM), XGBoost, ResNet50, and MobileNetV2. The proposed CNN demonstrated the most consistent performance across cassava growth stages, achieving the best classification results for the 2- and 7-month datasets, while XGBoost showed comparable performance for the 4-month dataset. The comparison before and after applying the Synthetic Minority Oversampling Technique (SMOTE) also indicates that class balancing improved CNN performance across all growth stages. Furthermore, Local Interpretable Model-Agnostic Explanations (LIME)-based interpretation shows that the model focused on relevant vegetation regions, thereby enhancing interpretability. Overall, the proposed framework demonstrates strong potential for UAV-based cassava health monitoring and provides insights into the relationship between spectral separability, class balancing, and classification performance.**

***Keywords*—cassava health classification, multispectral Unmanned Aerial Vehicle (UAV) imagery, vegetation indices, Convolutional Neural Network (CNN), Local Interpretable Model-Agnostic Explanations (LIME)**

## I. Introduction

Cassava (*Manihot esculenta Crantz*) is widely cultivated in tropical and subtropical regions due to its ability to grow under suboptimal environmental conditions. It plays a crucial role as a primary carbohydrate source for millions of people, particularly in developing countries, making it an essential component of global food security [1, 2]. Therefore, maintaining cassava productivity and plant health is critical for ensuring stable agricultural output.

Despite its resilience, cassava is vulnerable to various biotic and abiotic stresses that significantly affect crop yield and quality. Viral diseases such as Cassava Mosaic Disease (CMD), Cassava Brown Streak Disease (CBSD), and Cassava Common Mosaic Virus (CsCMV) disrupt physiological processes and reduce photosynthetic efficiency [3–6]. In addition, environmental factors such as water stress and nutrient imbalance further contribute to plant health deterioration [7–9]. These challenges highlight the importance of accurate and timely plant health monitoring. Conventional cassava health assessment is primarily based on visual field inspection. While this approach is practical for small-scale observations, it becomes inefficient, subjective, and labor-intensive when applied to large agricultural areas. These limitations restrict its scalability and consistency, emphasizing the need for automated and data-driven monitoring approaches.

Recent advancements in precision agriculture have introduced Unmanned Aerial Vehicles (UAVs) equipped with multispectral sensors as an effective solution for large-scale crop monitoring. Multispectral imagery enables the extraction of Vegetation Indices (VIs), such as

Manuscript received February 26, 2026; revised March 25, 2026; accepted June 30, 2026; published November 12, 2026.

Normalized Difference Vegetation Index (NDVI), Green Normalized Vegetation Index (GNDVI), Normalized Difference Red Edge (NDRE), and Leaf Chlorophyll Index (LCI), which provide quantitative information related to chlorophyll content, plant vigor, and canopy structure [10–13]. However, the effectiveness of these indices may vary across different plant growth stages due to changes in canopy density, spectral mixing, and environmental conditions.

In parallel, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have shown strong potential for plant health classification due to their ability to learn hierarchical spatial features from image data [4, 14–18]. Nevertheless, most existing studies focus on leaf-level images captured under controlled conditions, which differ significantly from UAV-based imagery that involves complex backgrounds, varying illumination, and mixed-pixel effects. Furthermore, limited studies have investigated how vegetation index variability across growth stages influences classification performance, and how model decisions can be interpreted to ensure reliability. This research gap highlights the need for an integrated framework that combines statistical analysis, deep learning, and model interpretability to improve the reliability of cassava health monitoring.

Therefore, this study proposes an integrated CNN-based framework for cassava plant health classification using multispectral UAV imagery and vegetation indices across different growth stages. The novelty of this study lies not in developing a new CNN, Analysis of Variance (ANOVA), or Local Interpretable Model-Agnostic Explanations (LIME) algorithm, but in integrating spectral separability analysis, spatial–spectral deep learning, baseline model comparison, and interpretable model evaluation into a unified framework for UAV-based cassava health monitoring.

This integration is important because plant health classification using UAV multispectral imagery is influenced not only by disease-related spectral responses but also by canopy development, growth-stage variability, mixed pixels, and overlapping vegetation index distributions.

The main contributions of this study are as follows:

- Evaluation of the separability of multispectral vegetation indices across cassava growth stages and health-severity levels using statistical analysis.
- Development and evaluation of a CNN-based classification model using UAV-derived multispectral image patches collected at multiple growth stages.
- Comparison of the proposed CNN with classical machine learning and pre-trained deep learning baseline models.
- Application of LIME-based interpretation to improve the transparency and reliability of cassava plant health classification.

## II. Related Work

Recent studies have demonstrated the effectiveness of deep learning approaches in detecting and classifying crop-leaf diseases based on image data. For instance, Emmanuel et al. developed a model for detecting cassava leaf diseases using a CNN algorithm based on a transfer learning architecture. The dataset consisted of cassava leaf images classified into several disease categories; however, the images were not captured using UAV imagery. The evaluation results demonstrated that the model achieved an accuracy of 99.12%, precision of 98.8%, recall of 98.9%, and F1-Score of 98.8%. These results demonstrate the effectiveness of the CNN-based approach in accurately classifying leaf diseases, even when applied to complex datasets [19].

A previous study evaluated several CNN architectures, including ResNet50, VGG16, DenseNet121, and InceptionV3, for cassava leaf disease classification. Among these models, DenseNet121 demonstrated the best performance, achieving an accuracy of 96.44%, while the remaining models reported accuracies within the range of 92% to 95%. The evaluation, based on accuracy metrics and confusion matrix analysis, indicates that deep learning methods are capable of reliably distinguishing various types of cassava leaf diseases with relatively low computational delay, making them suitable for real-time applications [20]. In a related study, both CNN and ANN models were applied to cassava leaf disease classification tasks. The CNN approach showed better generalization performance, reaching an accuracy of 96% on test data. Furthermore, the study highlighted the role of data augmentation and class balancing strategies in reducing model bias and improving classification robustness under real-world conditions [21–23].

Wahyuni and colleagues developed a simple CNN model for classifying three major cassava leaf diseases. The dataset consisted of images captured directly from the cassava farms in Indonesia. The training results showed an accuracy of 93.2%, a precision of 90.5%, and a recall of 92.1%. This study demonstrated the effectiveness of a CNN-based approach for implementation in mobile device–based agricultural monitoring systems [22].

Guo et al. introduced a rice leaf disease detection model based on a combination of CNN and multispectral image data captured by UAV (drones). The developed model achieved a detection accuracy of 95.67%, outperforming traditional models such as Support Vector Machines (SVM) and random forests. The evaluation was conducted in actual agricultural fields, and this approach demonstrated significant potential for integration into drone-based precision agriculture systems [19].

Although previous studies have shown promising results, there remains a significant gap. Most research on cassava disease detection using deep learning still uses ground-based images and does not take advantage of UAV (drone) images. UAVs can cover larger areas, support real-time monitoring, and are suitable for large plantations. However, their potential for monitoring cassava plant health remains underexplored. Additionally, using deep learning with UAV images, either RGB or multispectral, presents challenges such as varying image quality, occlusions, and complex backgrounds. Therefore, it is important to conduct research that develops and evaluates deep learning models trained specifically on

UAV-acquired images to support more widespread, faster, and more efficient agricultural monitoring. Although previous studies reported promising classification performance, most approaches relied on close-range leaf imagery or conventional single-input models rather than UAV-based multispectral canopy analysis across different growth stages. Based on these limitations, there is still a need for a framework that jointly considers spectral separability, growth-stage variability, spatial–spectral learning, and model interpretability for cassava health classification using multispectral UAV imagery. Previous studies have mainly emphasized classification accuracy, while the relationship between vegetation index separability and model performance across different cassava growth stages has received less attention. In addition, limited work has examined whether the regions used by the model for prediction correspond to meaningful vegetation areas. Therefore, this study contributes by linking statistical vegetation index analysis, CNN-based classification, baseline model comparison, and LIME-based interpretation to provide a more comprehensive evaluation of UAV-based cassava plant health monitoring.

## III. Materials and Methods

This study consisted of four main stages: (1) UAV multispectral data acquisition, (2) vegetation index computation and preprocessing, (3) dataset balancing and labeling, and (4) CNN-based classification with model interpretability. The overall workflow is illustrated in Fig. 1.

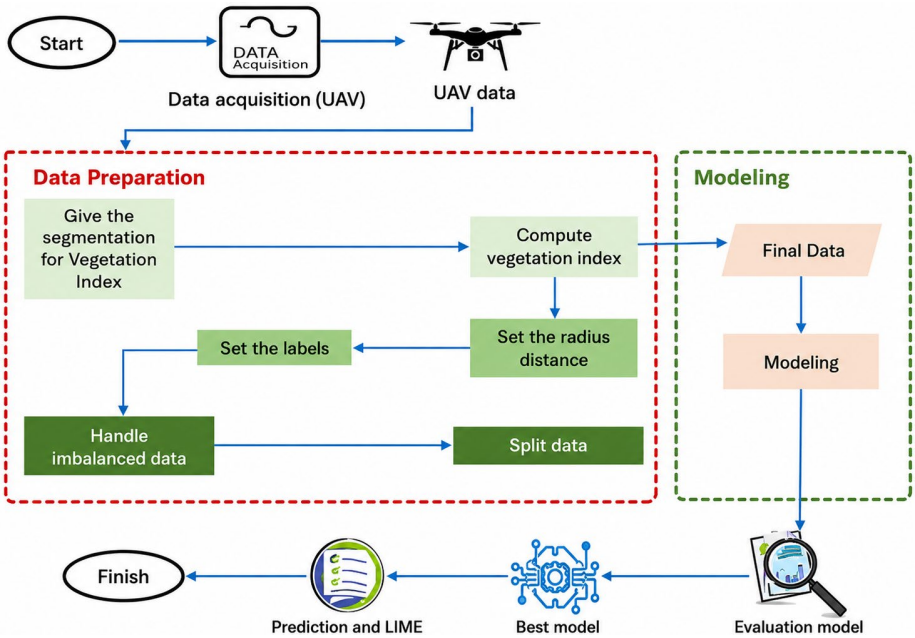



Fig. 1. Overall workflow of the proposed UAV-based framework for cassava plant health classification.

The acquired multispectral imagery consisted of six spectral bands. Each band was processed to separate vegetation pixels from soil and background areas before the vegetation indices were computed. The vegetation indices used in this study included NDVI, GNDVI, EVI, NDRE, and LCI. Image patches were then extracted from the vegetation index maps based on predefined sampling locations. These patches were augmented and labeled according to expert-validated cassava health levels.

To address class imbalance, the dataset was balanced using SMOTE before being divided into training, validation, and testing subsets with a ratio of 70:20:10. The modeling stage involved hyperparameter tuning, CNN training, and performance evaluation using a confusion matrix and classification metrics. Early stopping was applied to reduce overfitting, and the best-performing model was selected based on validation performance. Finally, LIME was used to interpret the CNN predictions and assess whether the model focused on relevant vegetation regions.

### A. UAV Imagery Acquisition and Processing

The research was conducted at a cassava plantation in Telaga Kahuripan, Parung District, Bogor Regency, West Java, Indonesia. The plantation covers approximately 15 ha and is planted with the Adira 1 cassava variety. The selected observation areas consisted of 2-month-old plants covering 14,344 m² and 4-month-old plants covering 14,533 m². These areas were selected because they represent different growth conditions and canopy development stages relevant to cassava health monitoring using UAV multispectral imagery [24–28].

The study site provides a suitable field setting for evaluating UAV-based plant health classification because it contains variations in canopy structure, plant vigor, and field conditions. These variations are important for analyzing how vegetation index responses change across growth stages and how they influence CNN-based classification performance. UAV-based multispectral imaging has been widely used in precision agriculture because it enables high-resolution crop monitoring,

vegetation mapping, and growth-stage analysis over large field areas [24–33]. In cassava-related studies, remote sensing and machine learning approaches have also shown potential for evaluating crop suitability, disease monitoring, and agricultural decision support [24–28]. The research location is shown in Fig. 2.




(a) (b)

Fig. 2. Research location in the Telaga Kahuripan area: (a) regional location map and (b) cassava plantation area used for UAV data acquisition.

Plant health classification in agricultural imagery often encounters limitations in both accuracy and computational efficiency, particularly when dealing with images of low quality or high noise levels. Therefore, there is a strong need for a more robust and reliable image-based detection approach capable of identifying plant health conditions more precisely.

In this study, a segmentation-based framework combined with vegetation index correlation analysis is proposed to evaluate the health status of cassava plants using UAV-acquired imagery. This approach improves the reliability and effectiveness of plant health assessment, particularly for early-stage disease detection. The UAV-based data acquisition process is illustrated in Fig. 3, while the UAV flight configuration and flight path over the cassava plantation are shown in Fig. 4.

After completing the flight mission, the captured UAV data were exported from the onboard system and stored on a local computer for further processing. A preliminary data inspection was conducted to ensure coverage completeness, image quality, and sufficient overlap for reliable processing.

The processing pipeline starts by analyzing the spectral characteristics of the collected imagery, where each band represents reflected signals from different surface objects. This study utilized a multispectral sensor equipped with six spectral bands, including RGB, red edge, and near-infrared channels, each corresponding to specific wavelength ranges.

The image preprocessing stage includes both radiometric and geometric corrections, which are essential for improving data quality, particularly in agricultural mapping applications. Radiometric correction minimizes noise and inconsistencies in pixel values by compensating for atmospheric effects such as light absorption and scattering. Meanwhile, geometric correction addresses spatial distortions caused by factors including rotation, scaling, and positional shifts, ensuring accurate spatial alignment of the imagery.




(a) (b)

Fig. 3. Field conditions during UAV data acquisition showing (a) an early-stage cassava field and (b) a two-month-old cassava plantation.




Fig. 4. UAV flight configuration: (a) flight parameters and (b) flight path over the cassava plantation.

Following this, a Digital Surface Model (DSM) is generated to represent surface features such as vegetation, buildings, and infrastructure. In parallel, a Digital Terrain Model (DTM) is derived to describe the underlying ground morphology, including elevation, contours, and landforms. The orthomosaic generation process then integrates multiple aerial images into a unified two-dimensional representation, while simultaneously producing spectral band layers required for vegetation index computation used in cassava plant identification.

### *B. Data Preparation*

#### *1) Segmentation of image*

The data in this study were processed and analyzed using Google Earth Engine (GEE), a cloud-based platform for large-scale geospatial data processing and analysis. GEE was used because it supports efficient visualization, processing, and analysis of remote sensing imagery. Although it is commonly applied to satellite imagery, its geospatial processing capabilities are also useful for UAV-based vegetation analysis and agricultural monitoring.

In this study, vegetation and soil pixels were separated using k-means clustering implemented in GEE, as shown in Fig. 5. The segmentation process was conducted to ensure that vegetation index values represented cassava canopy conditions rather than mixed signals from soil and background areas. This step is important because mixed pixels can introduce spectral noise and reduce the reliability of vegetation-index-based analysis, particularly in UAV imagery acquired over agricultural fields.

The k-means algorithm grouped pixels based on spectral similarity, allowing vegetation and soil/background classes to be distinguished more clearly. As a result, the extracted vegetation information became more representative of actual cassava canopy conditions. Fig. 5 presents examples of the segmentation results for 2-, 4-, and 7-month-old cassava plants using NDVI vegetation index.

The overall workflow for vegetation segmentation and spectral index extraction is illustrated in Fig. 6. The process begins with the acquisition of the imagery, which is spatially clipped to the defined Region of Interest (ROI). To isolate vegetation from the background, an unsupervised machine learning approach using the k-means clustering algorithm is applied. The algorithm is initialized with two clusters ($k = 2$) to differentiate between the two dominant spectral classes: soil/background and vegetation.

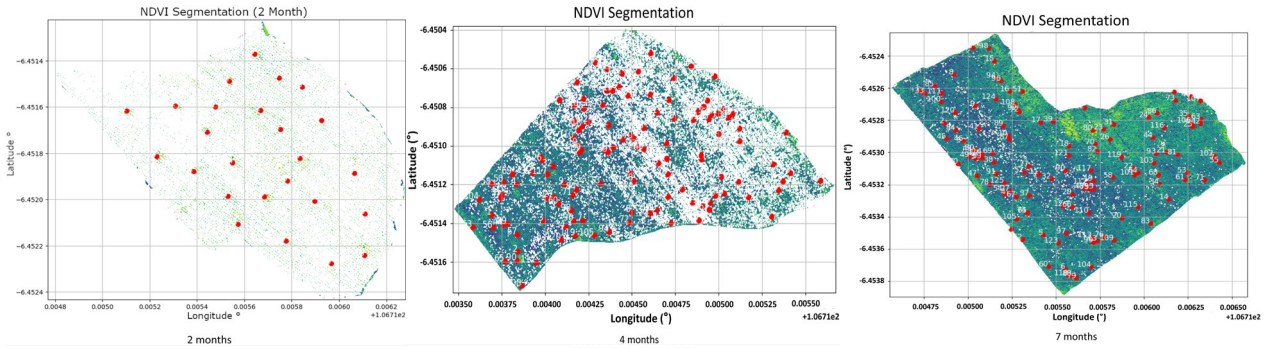



Fig. 5. Segmentation results for 2-, 4-, and 7-month-old cassava plants using the NDVI index.

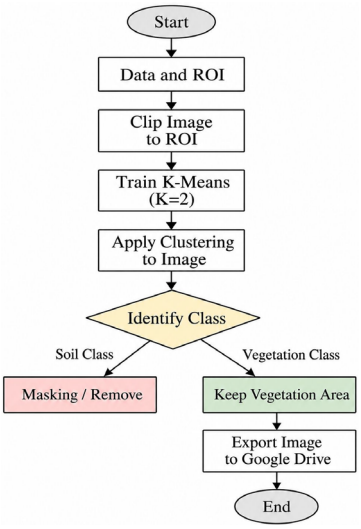



Fig. 6. Workflow for vegetation segmentation and background removal using k-means clustering.

*2) Compute vegetation index*

Vegetation indices, including NDVI, GNDVI, EVI, NDRE, and LCI, were computed to represent complementary spectral information related to cassava plant condition. Each index captures different aspects of vegetation physiology and canopy characteristics. NDVI reflects canopy vigor and chlorophyll-related greenness by comparing near-infrared and red reflectance. GNDVI uses the green band and is more sensitive to chlorophyll concentration and plant greenness. EVI improves sensitivity in high-biomass conditions while reducing background effects. NDRE uses the red-edge band to detect plant stress, particularly when stress symptoms are not yet clearly visible in conventional indices. LCI provides information related to chlorophyll concentration and upper-canopy photosynthetic activity.

By combining these indices, the model obtained a richer spatial–spectral representation than would be available from a single vegetation index. This multi-index representation was used to improve the learning capability of the CNN model for cassava plant health classification.

The multispectral imagery used in this study was acquired using a DJI Phantom 4 Multispectral UAV (Fig. 7(a)–(c)), which is equipped with six imaging sensors. The UAV system captures Blue (450 ± 16 nm), Green (560 ± 16 nm), Red (650 ± 16 nm), Red Edge (730 ± 16 nm), and near-infrared/ Near-Infrared (NIR) (840 ± 26 nm) bands along predefined flight paths, as shown in Fig. 4. Radiometric and geometric corrections were applied to reduce pixel-value inconsistencies and spatial distortions. Orthomosaic generation, DSM, and DTM processing were then conducted to produce georeferenced multispectral layers required for vegetation index computation.

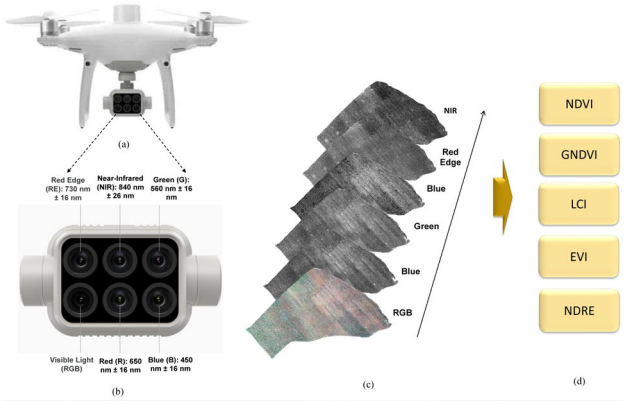



Fig. 7. UAV multispectral imaging system and vegetation index generation process: (a) DJI Phantom 4 Multispectral UAV, (b) multispectral camera with six imaging lenses, (c) real-time images from spectral bands, and (d) vegetation index products.

The five health-severity levels used in this study were defined by agronomists based on field observations of canopy color, leaf turgor, growth uniformity, and visible stress symptoms. These levels serve as the expert-validated ground-truth labels for the supervised classification task. Level 1 corresponds to severely stressed or diseased plants, Level 2 reflects mild stress with reduced vigor, Level 3 represents fairly healthy plants with minor nutrient or water-related symptoms, Level 4 indicates healthy plants with vigorous growth, and Level 5 represents plants in optimal physiological condition. Table I summarizes these health-severity definitions and their corresponding ratings.

TABLE I. HEALTH SEVERITY LEVELS AND DEFINITIONS FOR CASSAVA PLANT CLASSIFICATION

| Severity Impact | Definition | Level |
|---|---|---|
| Sick | Plant shows clear signs of disease, such as yellowing or wilting. Requires attention. | 1 |
| Less healthy | Plant appears weak, with mild symptoms like slow growth or faded leaves. Needs care. | 2 |
| Fairly Healthy | Plant is growing well, with minor stress or nutrient deficiency signs. | 3 |
| Healthy | Plant is thriving, with strong growth and no disease symptoms. | 4 |
| Very Healthy | Plant is in optimal condition, growing vigorously and free of any stress. | 5 |

These categories were defined based on expert assessment and used as ground truth labels for the classification task. In this study, Vegetation Indices (VIs), defined as arithmetic combinations of two or more spectral bands, are utilized to capture the spectral characteristics of vegetation and have been widely applied in phenological monitoring, vegetation classification, and the derivation of biophysical and structural parameters [34].

$$NDVI = (NIR - Red) / (NIR + Red) \quad (1)$$

The GNDVI is similar to the NDVI; however, in GNDVI, the green band is used instead of the red band used in NDVI [35]. The GNDVI calculation formula [36] is given by Eq. (2).

$$GNDVI = (NIR - G) / (NIR + G) \quad (2)$$

The LCI is used to estimate the chlorophyll content in high-level plants and is sensitive to variations in

reflectance caused by chlorophyll uptake. The LCI calculation formula [37] is shown in Eq. (3).

$$LCI = (\text{NIR} - \text{RED EDGE}) / (\text{NIR} + \text{RED}) \quad (3)$$

A NDRE image is obtained by mathematically combining information from the NIR and Red Edge (RE) channels in the electromagnetic spectrum [38].

$$NDRE = (NIR - RE)/(NIR + RE) \quad (4)$$

*3) ANOVA test*

To statistically evaluate whether vegetation index responses differed across plant age and health-severity categories, a one-way ANOVA was applied to each vegetation index, including NDVI, GNDVI, EVI, NDRE, and LCI. ANOVA was used to determine whether the mean spectral responses among multiple groups were significantly different. In this study, the groups corresponded to cassava growth stages of 2, 4, and 7 months and health-severity levels from 1 to 5.

A high F-value indicates that between-group variability is greater than within-group variability, suggesting that the vegetation index captures meaningful physiological differences among categories. As shown in Tables II and III, all vegetation indices exhibited statistically significant differences ($p < 0.05$) for both plant age and health severity. These results confirm that the selected vegetation indices are sensitive to cassava physiological variation and are therefore suitable as input features for CNN-based classification.

The ANOVA results also help explain the differences in CNN performance across growth stages. The 2-month dataset showed clearer spectral separability among health categories, while the 4-month dataset exhibited greater within-class variation, likely due to rapid canopy development, uneven biomass distribution, partial self-shading, and heterogeneous nutrient uptake. These evolving spectral characteristics help explain variations in classification accuracy across growth stages.

TABLE II. ANOVA ON PLANT AGE

| Vegetation Index | F-Value | *p*-Value |
|---|---|---|
| NDVI | 59.47 | <0.001 |
| GNDVI | 56.41 | <0.001 |
| EVI | 444.5 | <0.001 |
| NDRE | 274.16 | <0.001 |
| LCI | 91.66 | <0.001 |

TABLE III. ANOVA ON SEVERITY LEVEL

| Vegetation Index | F-Value | *p*-Value |
|---|---|---|
| NDVI | 84.06 | $1.63 \times 10^{-41}$ |
| GNDVI | 66.45 | $2.08 \times 10^{-34}$ |
| EVI | 17.81 | $8.03 \times 10^{-11}$ |
| NDRE | 9.19 | $7.07 \times 10^{-06}$ |
| LCI | 15.96 | $8.82 \times 10^{-10}$ |

A one-way ANOVA was conducted to examine how vegetation indices—namely NDVI, GNDVI, EVI, NDRE, and LCI—vary across different plant ages and levels of health severity. The statistical results, summarized in Tables II and III, show that all indices differ significantly between groups ($p < 0.001$), indicating a clear sensitivity to both growth stage and health condition. At the early growth phase, particularly at two months, vegetation indices demonstrate strong separability, as reflected by higher F-values and relatively low variability within each class, suggesting that physiological differences among health categories are more distinctly expressed at this stage. In contrast, measurements at four months exhibit greater within-class variation, which is likely associated with rapid canopy development, uneven biomass distribution, partial self-shading, and heterogeneous nutrient uptake. These evolving spectral characteristics help explain subsequent CNN performance, as changes in spectral separability across growth stages are closely mirrored by variations in classification accuracy.

*4) Clipping image based on vegetation index*

The clipping process was conducted to extract localized image patches from the vegetation index maps based on predefined cassava sampling locations. In this study, clipping refers to the spatial cropping of vegetation index images into smaller patch regions used as CNN input data, not to histogram or contrast clipping. Each clipped patch represents a local canopy area around a cassava plant and contains vegetation index information derived from multispectral UAV imagery.

This process allowed the CNN model to learn from localized canopy information rather than from the entire orthomosaic image. By using clipped patches, irrelevant background areas were reduced, allowing the classification process to focus on the spatial and spectral characteristics of cassava plants. The clipped patches were generated for each vegetation index and then used as the input dataset for CNN training and evaluation.

Fig. 8 illustrates the spatial clipping process from vegetation index maps into smaller image patches. Differences in color and spatial patterns among the patches indicate variations in vegetation index values, which are associated with canopy condition, plant vigor, and health-severity levels.

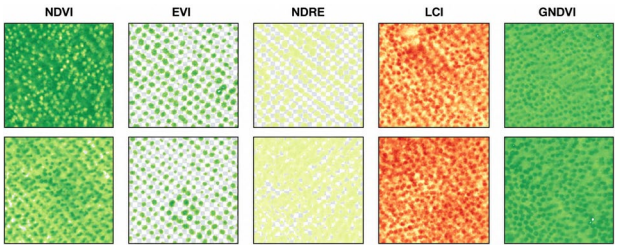



Fig. 8. Spatial clipping process used to extract cassava canopy patches from vegetation index images for CNN-based classification.

*5) Dataset balancing and labeling*

The classification dataset consisted of three cassava growth stages, representing 2-, 4-, and 7-month-old plants, with 625 image patches for each stage before balancing. Because the original dataset showed class imbalance across health-severity categories, expert-guided labeling and data balancing were performed to support reliable supervised learning.

The distribution of samples across health levels was uneven, with some classes being substantially underrepresented. This imbalance could bias the CNN model toward majority classes and reduce its ability to recognize minority health conditions. Therefore, a structured labeling process and a balancing strategy were applied to improve class representation across health-severity categories.

Fig. 9 compares the distribution of cassava plant health samples before and after applying SMOTE. Before oversampling, the dataset showed clear class imbalance across growth stages, with some classes dominating and others being underrepresented. After SMOTE was applied, the class distribution became more balanced because synthetic samples were generated for minority classes. This more balanced distribution allowed the model to learn from all health categories more effectively and improved classification metrics such as precision, recall, and F1-Score.

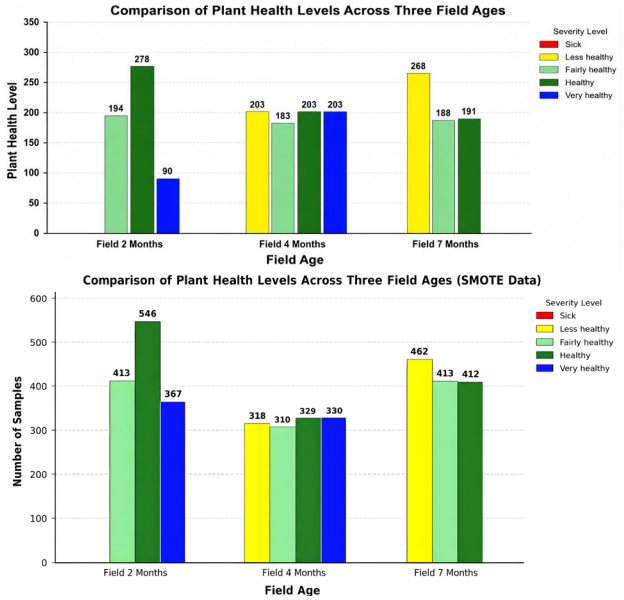



Fig. 9. Class distribution before and after SMOTE-based oversampling.

### *a) Health-level labeling*

Health-level labels were assigned by expert agronomists based on canopy color, leaf turgor, growth uniformity, and visible stress symptoms. Vegetation indices were used only as supporting information, not as automatic labeling criteria. This ensured that the ground-truth labels remained expert-validated and reduced ambiguity in the supervised learning process.

### *b) Oversampling using SMOTE*

Class imbalance is a common issue in agricultural image classification because some plant health categories occur less frequently than others under field conditions. In this study, the original dataset showed an unequal distribution of samples across health-severity levels. If this imbalance was not addressed, the CNN model could become biased toward majority classes and show reduced sensitivity to minority classes, particularly for less frequent health conditions.

To mitigate this issue, SMOTE was applied in the feature representation space after the multispectral image patches were transformed into numerical vectors. Each preprocessed image patch was resized to a uniform resolution and flattened into a high-dimensional feature vector. SMOTE then generated synthetic minority samples by interpolating between neighboring samples within the same class. Therefore, SMOTE was not intended to directly generate new raw UAV images. Instead, it was used as a feature-space balancing strategy to improve class representation before model training.

After oversampling, the generated synthetic feature vectors were reshaped back into the original patch dimensions to maintain compatibility with the CNN input format. This reconstruction step allowed the balanced samples to be used in the same image-based training pipeline as the original patches. However, because feature-space interpolation may not fully preserve all natural spatial structures of real UAV images, an additional comparison was conducted to evaluate the effect of the SMOTE-based balancing strategy.

Specifically, the proposed CNN was trained and evaluated under two settings: without SMOTE using the original imbalanced dataset and with SMOTE-based oversampling. This comparison was conducted to determine whether SMOTE provided a measurable benefit for CNN-based cassava health classification. The final balanced dataset consisted of 1730 samples for the 2-month dataset, 1450 samples for the 4-month dataset, and 1835 samples for the 7-month dataset. The SMOTE-based balancing workflow is illustrated in Fig. 10.

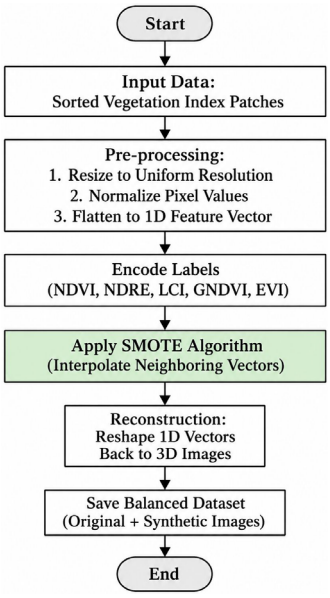



Fig. 10. Workflow of SMOTE-based feature-space oversampling for balancing the cassava plant health dataset.

### *c) Comparison of SMOTE and Non-SMOTE settings*

To further evaluate the effect of the SMOTE-based balancing strategy, two experimental settings were compared: (1) training without SMOTE using the original imbalanced dataset and (2) training with SMOTE-based

oversampling. This comparison was conducted to assess whether feature-space oversampling improved the proposed CNN model's ability to classify cassava health-severity categories across different growth stages. The evaluation focused on accuracy, precision, recall, and F1-Score because these metrics provide a more comprehensive assessment of model performance under imbalanced class conditions.

*d) Final dataset split*

A stratified data-splitting strategy was adopted to ensure that each subset retained a proportional representation of the observed health levels. For the two-month growth stage, the dataset encompassed health categories ranging from levels 3 to 5, while the datasets corresponding to the four- and seven-month stages covered a broader spectrum of conditions, including health levels 2 through 5.

## C. Modeling

To determine the optimal Convolutional Neural Network (CNN) architecture, hyperparameter tuning was performed using a Random Search strategy with a maximum of 20 trials. The tuning process used a consolidated dataset that combined samples from all cassava growth stages to ensure that the selected architecture had sufficient generalization capability. As shown in Table IV, the search space included variations in the number of convolutional filters, kernel sizes, dense-layer units, dropout rates, and learning rates. The learning rate was tuned between 0.001 and 0.0001. The dataset was stratified into training, validation, and testing subsets with a ratio of 70:20:10 to ensure robust model selection and reduce bias toward a specific data distribution. After the optimal configuration was identified, the selected parameters were used to train separate models for each cassava growth stage, namely 2, 4, and 7 months.

Following the tuning phase, the best-performing configuration was selected for final training, as presented in Table V. The model used an input resolution of 256 × 256 pixels to maintain spatial detail while supporting computational efficiency. Model training was conducted with a batch size of 32 for up to 50 epochs using the Adam optimizer, which combines momentum and adaptive learning-rate adjustment to improve convergence. The categorical cross-entropy loss function was used for the multiclass classification task. To reduce overfitting, early stopping with a patience value of 8 was applied, stopping the training process when the validation loss did not improve for eight consecutive epochs.

TABLE IV. CNN MODEL PARAMETERS

| Hyperparameter Component | Parameter Name | Search Space |
|---|---|---|
| Convolutional Layer 1 | Filters | 32, 64 |
| | Kernel Size | 3 × 3 and 5 × 5 |
| Convolutional Layer 2 | Filters | 64, 96, 128 |
| Convolutional Layer 3 | Use Layer? | True, False (Optional) |
| | Filters | 64, 96, 128 |
| Fully Connected (Dense) | Units (Neurons) | 64, 128, 192, 256 |
| Regularization | Dropout Rate | 0.2, 0.3, 0.4, 0.5 |
| Optimization | Learning Rate | 0.001, 0.0001 |

Before model training, data augmentation was incorporated directly into the architecture using Keras preprocessing layers. To increase dataset diversity and reduce overfitting, the augmentation pipeline included pixel rescaling (1/255), random horizontal flipping, random rotation with a factor of 0.1, and random zooming within a range of ±0.2.

The CNN architecture presented in Table VI consists of seven main stages: input, convolution, pooling, flattening, dense, dropout, and output layers. Each stage includes the corresponding input and output dimensions. The architecture was designed to extract relevant spatial features, reduce feature dimensionality, and perform efficient cassava plant health classification.

TABLE V. CNN BEST PARAMETERS

| Parameter | Value |
|---|---|
| Input Shape | 256 × 256 × 3 |
| Batch Size | 32 |
| Epochs | 50 |
| Optimizer | Adam |
| Learning Rate | 0.001 |
| Loss Function | Categorical Crossentropy |
| Patience (Early Stopping) | 8 |
| Conv Layer 1 Filters | 32 (Kernel 3 × 3) |
| Conv Layer 2 Filters | 64 (Kernel 3 × 3) |
| Dense Units | 128 |
| Dropout Rate | 0.3 |

TABLE VI. CNN MODEL ARCHITECTURE

| Layer (Type) | Configuration | Input Shape | Output Shape |
|---|---|---|---|
| Input | Image Input | (256, 256, 3) | (None, 256, 256, 3) |
| Conv2D | 32 Filters, 3 × 3, ReLU | (None, 256, 256, 3) | (None, 254, 254, 32) |
| MaxPooling2D | Pool Size 2 × 2 | (None, 254, 254, 32) | (None, 127, 127, 32) |
| Conv2D | 64 Filters, 3 × 3, ReLU | (None, 127, 127, 32) | (None, 125, 125, 64) |
| MaxPooling2D | Pool Size 2 × 2 | (None, 125, 125, 64) | (None, 62, 62, 64) |
| Flatten | Flattening 3D to 1D | (None, 62, 62, 64) | (None, 246,016) |
| Dense | 128 Units, ReLU | (None, 246,016) | (None, 128) |
| Dropout | Rate = 0.3 | (None, 128) | (None, 128) |
| Dense (Output) | 4 Units, Softmax | (None, 128) | (None, 4) |

## IV. Results and Discussion

The Results and Discussion section presents five main analyses: statistical evaluation of vegetation indices using ANOVA and the Friedman test, CNN classification performance across cassava growth stages, comparison with classical machine learning and pre-trained CNN baseline models, evaluation of the effect of SMOTE on CNN performance, and model interpretation using LIME. The evaluation was conducted separately for the 2-, 4-, and 7-month cassava datasets to examine the influence of growth stage on spectral separability and classification performance.

The proposed CNN model achieved classification accuracies of 0.84, 0.81, and 0.96 for the 2-, 4-, and 7-month datasets, respectively. The relatively high performance at the early and later growth stages indicates that vegetation characteristics were more distinguishable at these stages, whereas the lower performance at the 4-month stage suggests greater classification difficulty due to increased canopy complexity and overlapping spectral responses.

The statistical analyses support these performance trends. The ANOVA results indicate that vegetation indices differed significantly across plant age and health-severity levels, while the Friedman test further confirms that vegetation index distributions varied significantly across growth stages. These findings suggest that growth-stage-dependent spectral variability directly affects health-severity separability and CNN classification performance. The LIME-based interpretation further supports this relationship by showing that the CNN model focused on relevant vegetation regions and captured meaningful spatial–spectral patterns.

### A. Statistical Comparison of Vegetation Indices Across Growth Stages

To further examine whether vegetation index distributions differed significantly across cassava growth stages, the Friedman test was applied to each vegetation index. The Friedman test was selected as a non-parametric statistical method because the assumption of normality may not always be satisfied in field-based vegetation index data. This analysis complements the ANOVA results and provides additional statistical evidence of growth-stage-dependent differences in vegetation index responses.

Table VII shows that all evaluated vegetation indices exhibited statistically significant differences across cassava growth stages. NDVI, GNDVI, NDRE, and LCI produced $p$-values below 0.05, indicating that their distributions varied significantly among the observed growth stages. The highest Friedman statistic was obtained by GNDVI, followed closely by NDVI, suggesting that these indices were particularly sensitive to growth-stage-related changes in cassava canopy conditions.

TABLE VII. Friedman Test Results for Vegetation Indices Across Growth Stages

| Vegetation Index | Friedman Statistic | $p$-Value | Significance |
|---|---|---|---|
| NDVI | 166.93 | $5.65 \times 10^{-37}$ | Significant |
| GNDVI | 167.90 | $3.48 \times 10^{-37}$ | Significant |
| NDRE | 156.00 | $1.33 \times 10^{-34}$ | Significant |
| LCI | 138.25 | $9.52 \times 10^{-31}$ | Significant |

These results support the ANOVA findings and confirm that vegetation index responses are not constant across plant ages. This statistical evidence helps explain why CNN classification performance varied across the 2-, 4-, and 7-month datasets. Growth-stage-dependent changes in canopy structure, chlorophyll response, and spectral separability may affect the model's ability to distinguish cassava health-severity levels. Therefore, the Friedman test provides additional support for the relationship between vegetation index variability and classification performance across growth stages.

### B. CNN Classification Performance Across Growth Stages

The CNN model was trained and evaluated separately for each cassava growth stage because each stage has distinct canopy structures, spectral responses, and health-level distributions. Figs. 11–16 present the training curves and confusion matrices for the 2-, 4-, and 7-month datasets.

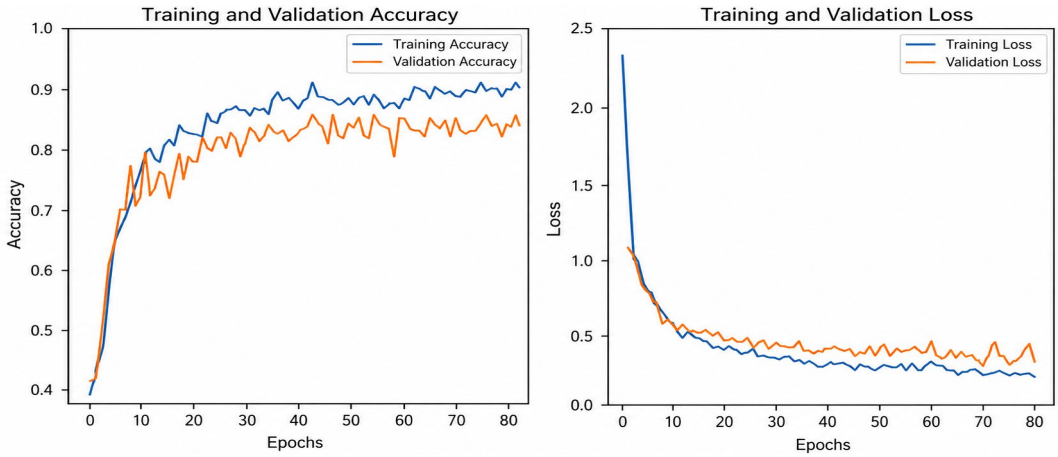



Fig. 11. Training and validation accuracy and loss curves of the proposed CNN model for the 2-month cassava dataset.

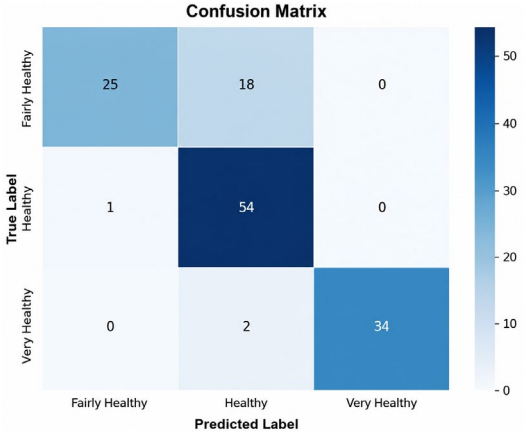



Fig. 12. Confusion matrix of the proposed CNN model for cassava plant health classification at the 2-month growth stage.

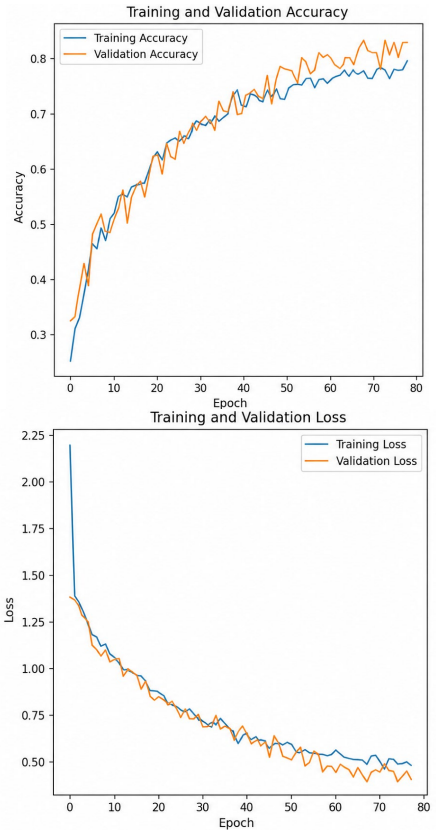



Fig. 13. Training and validation accuracy and loss curves of the proposed CNN model for the 4-month cassava dataset.

For the 2-month dataset, the proposed CNN achieved an overall accuracy of 0.84. The model performed particularly well in the "Very Healthy" class, with an F1-Score of 0.97, followed by the "Healthy" class, which showed high recall. In contrast, the "Fairly Healthy" class was more difficult to classify, with lower recall, indicating that intermediate health conditions had greater overlap in spectral and spatial characteristics. The confusion matrix further confirms that most misclassifications occurred between adjacent classes, especially between "Fairly Healthy" and "Healthy". This suggests that the model was more effective in identifying extreme health conditions than intermediate categories.

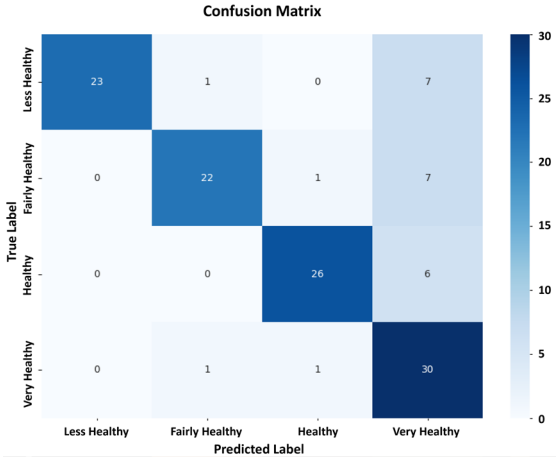



Fig. 14. Confusion matrix of the proposed CNN model for cassava plant health classification at the 4-month growth stage.

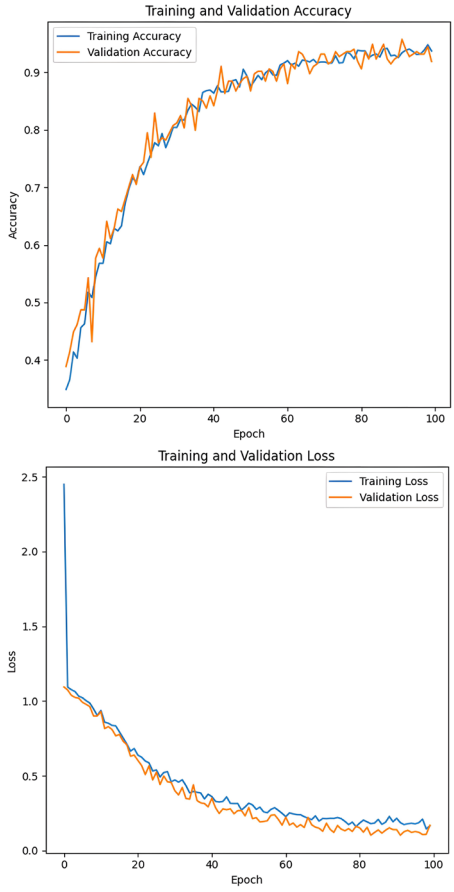



Fig. 15. Training and validation accuracy and loss curves of the proposed CNN model for the 7-month cassava dataset.

For the 4-month dataset, the CNN achieved an accuracy of 0.81. The training and validation curves indicate stable convergence without substantial overfitting. However, the confusion matrix shows more frequent misclassifications between adjacent health classes, particularly involving the "Very Healthy" class. This result suggests that the 4-month growth stage presented greater classification difficulty, likely due to increased canopy complexity, partial overlap among vegetation index distributions, and more heterogeneous spectral responses.

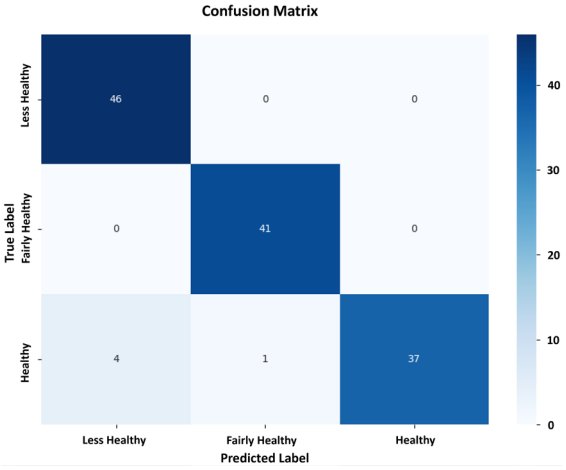



Fig. 16. Confusion matrix of the proposed CNN model for cassava plant health classification at the 7-month growth stage.

For the 7-month dataset, the CNN achieved the highest accuracy of 0.96. The training and validation curves indicate stable convergence with minimal overfitting, while the confusion matrix shows very few misclassifications. These results suggest that health classes at this stage were more clearly separable and that the CNN effectively captured spatial–spectral patterns from mature cassava canopies.

As summarized in Table VIII, the CNN model achieved the best performance at 7 months, followed by 2 months and 4 months. The lower performance at 4 months suggests that this growth stage contains more complex and overlapping spectral characteristics, making health-level discrimination more challenging.

TABLE VIII. SUMMARY OF CNN CLASSIFICATION PERFORMANCE ACROSS GROWTH STAGES

| Growth Stage | Accuracy | Macro F1-Score | Weighted F1-Score |
|---|---|---|---|
| 2 Months | 0.84 | 0.84 | 0.84 |
| 4 Months | 0.81 | 0.82 | 0.82 |
| 7 Months | 0.96 | 0.96 | 0.96 |

### C. CNN Evaluation Results

The detailed classification performance of the proposed CNN model is presented in Table IX using precision, recall, and F1-Score across health-severity levels and cassava growth stages.

TABLE IX. DETAILED CLASSIFICATION PERFORMANCE OF THE PROPOSED CNN ACROSS GROWTH STAGES

| Plant Age | Health Level | Precision (%) | Recall (%) | F1-Score (%) | Acc (%) |
|---|---|---|---|---|---|
| 2 Months | 3 | 96 | 58 | 72 | 84 |
| | 4 | 100 | 94 | 97 | |
| | 5 | 73 | 98 | 84 | |
| 4 Months | 2 | 92 | 73 | 81 | 81 |
| | 3 | 100 | 74 | 85 | |
| | 4 | 60 | 94 | 73 | |
| | 5 | 93 | 81 | 87 | |
| 7 Months | 2 | 98 | 100 | 99 | 96 |
| | 3 | 92 | 100 | 96 | |
| | 5 | 100 | 88 | 94 | |

At the 2-month stage, the model showed strong performance, particularly for higher health levels, where precision and recall values were generally high. This indicates that early-stage vegetation conditions were relatively distinguishable in the vegetation index image patches.

At the 4-month stage, model performance became less stable, with noticeable decreases in recall and F1-Score for some intermediate health levels. This suggests that adjacent health classes were more difficult to distinguish at this stage due to canopy complexity, mixed spectral responses, and overlapping vegetation index characteristics.

At the 7-month stage, the model achieved more consistent performance, with high precision and recall across most classes. This indicates that vegetation conditions became more separable at the later growth stage, allowing the CNN to capture discriminative spatial–spectral patterns more effectively.

Overall, the model performed better in identifying extreme health conditions than intermediate health levels, which tended to share more similar spectral and canopy characteristics. These findings are consistent with the ANOVA and Friedman test results, confirming that spectral separability and growth-stage-dependent vegetation index variability influenced CNN classification performance.

The detailed evaluation metrics further support the overall classification results and highlight the model's ability to distinguish cassava health conditions across different growth stages. These results also provide a basis for further comparison with classical machine learning and pre-trained CNN baseline models.

### D. Comparison with Baseline Models

To further evaluate the effectiveness of the proposed CNN model, additional baseline comparisons were conducted using classical machine learning and pre-trained CNN models. The classical machine learning models included Random Forest, SVM, and XGBoost, while the pre-trained CNN baseline models included ResNet50 and MobileNetV2. This comparison was conducted to assess whether the proposed CNN provides advantages over conventional feature-based classifiers and transfer-learning-based deep learning architectures for UAV-based cassava health classification.

Table X presents the performance comparison between the proposed CNN and several baseline models after applying SMOTE. The proposed CNN achieved the highest accuracy at the 2-month and 7-month growth

stages, with accuracies of 0.84 and 0.96, respectively. At the 4-month stage, XGBoost achieved slightly higher accuracy than the proposed CNN, with accuracies of 0.84 and 0.81, respectively. This result indicates that XGBoost can perform well under certain growth-stage conditions, particularly when the feature distribution is suitable for tree-based decision learning.

The relatively low performance of the pre-trained CNN models, ResNet50 and MobileNetV2, can be further explained by the domain gap between their original pre-training data and the vegetation index image patches used in this study. These pre-trained models are typically trained on large-scale natural RGB image datasets, where discriminative features are learned from object shapes, textures, edges, and color patterns. In contrast, the input data consisted of vegetation index image patches derived from multispectral UAV imagery, in which class differences are primarily characterized by subtle spectral variations, canopy density, and spatial–spectral relationships rather than distinct object boundaries or texture patterns. As a result, the high-level visual features learned from natural images may not transfer effectively to vegetation index image patches. Moreover, the relatively limited number of field samples, together with growth-stage-dependent spectral variability, may reduce the effectiveness of fine-tuning deeper pre-trained architectures. Consequently, the proposed CNN, with its simpler architecture, may be better suited for learning task-specific spatial–spectral features from the available UAV-based cassava dataset.

TABLE X. PERFORMANCE COMPARISON BETWEEN THE PROPOSED CNN AND BASELINE MODELS AFTER SMOTE

| Model | 2 Months | 4 Months | 7 Months |
|---|---|---|---|
| Proposed CNN | 0.84 | 0.81 | 0.96 |
| XGBoost | 0.61 | 0.84 | 0.94 |
| Random Forest | 0.54 | 0.68 | 0.62 |
| SVM | 0.45 | 0.26 | 0.35 |
| MobileNetV2 | 0.48 | 0.49 | 0.53 |
| ResNet50 | 0.43 | 0.38 | 0.38 |

## *E. Effect of SMOTE on CNN Performance*

To evaluate the effect of SMOTE on the proposed CNN model, performance was compared before and after applying SMOTE across the 2-, 4-, and 7-month cassava growth stages. The comparison was based on accuracy, precision, recall, and F1-Score to assess whether class balancing improved the model's ability to classify different health-severity categories.

Table XI shows the effect of SMOTE on the proposed CNN performance across the three cassava growth stages. The results indicate that SMOTE consistently improved CNN performance. At the 2-month stage, accuracy increased from 0.65 without SMOTE to 0.84 with SMOTE, while the F1-Score increased from 0.60 to 0.84. At the 4-month stage, accuracy improved from 0.70 to 0.81, and the F1-Score increased from 0.70 to 0.82. A similar improvement was observed at the 7-month stage, where accuracy increased from 0.84 to 0.96 and the F1-Score increased from 0.84 to 0.96.

TABLE XI. EFFECT OF SMOTE ON THE PROPOSED CNN PERFORMANCE

| Growth Stage | Setting | Acc | Prec. | Recall | F1-Score |
|---|---|---|---|---|---|
| 2 Months | Without SMOTE | 0.65 | 0.69 | 0.64 | 0.60 |
| 2 Months | With SMOTE | 0.84 | 0.84 | 0.83 | 0.84 |
| 4 Months | Without SMOTE | 0.70 | 0.72 | 0.70 | 0.70 |
| 4 Months | With SMOTE | 0.81 | 0.86 | 0.80 | 0.82 |
| 7 Months | Without SMOTE | 0.84 | 0.84 | 0.84 | 0.84 |
| 7 Months | With SMOTE | 0.96 | 0.97 | 0.96 | 0.96 |

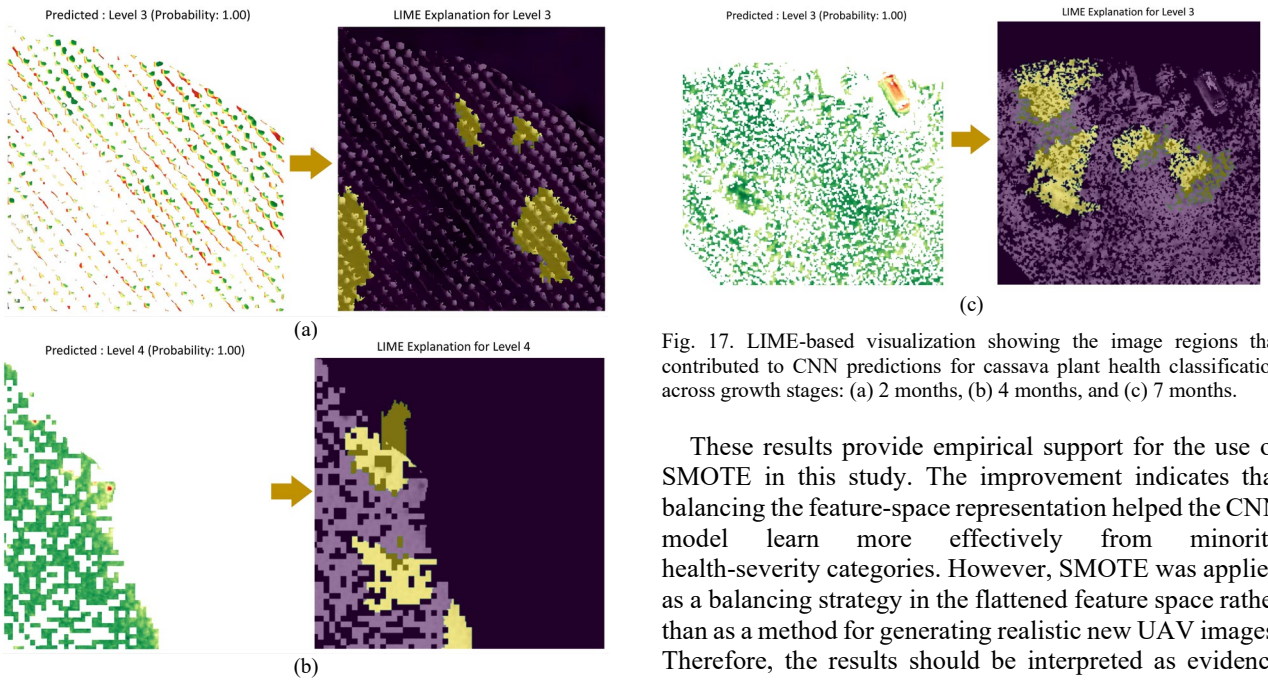



Fig. 17. LIME-based visualization showing the image regions that contributed to CNN predictions for cassava plant health classification across growth stages: (a) 2 months, (b) 4 months, and (c) 7 months.

These results provide empirical support for the use of SMOTE in this study. The improvement indicates that balancing the feature-space representation helped the CNN model learn more effectively from minority health-severity categories. However, SMOTE was applied as a balancing strategy in the flattened feature space rather than as a method for generating realistic new UAV images. Therefore, the results should be interpreted as evidence

that SMOTE improved class representation for CNN training, while acknowledging that feature-space interpolation may not fully preserve the natural spatial structure of UAV image patches.

### F. LIME Evaluation Results

LIME was applied to interpret the CNN model by highlighting the spatial regions that contributed most to its classification decisions across different growth stages (Fig. 17). The resulting visualizations indicate that the model consistently focused on vegetation regions rather than background areas, suggesting that it learned meaningful spatial–spectral features relevant to cassava plant health classification.

Fairly Healthy samples exhibited more scattered and irregular highlighted regions, indicating heterogeneous canopy conditions and transitional physiological states. In contrast, Healthy and Very Healthy samples showed more concentrated and uniform highlighted regions, reflecting stronger and more stable vegetation signals. These patterns help explain the lower classification performance observed in intermediate classes and are consistent with the CNN evaluation results, ANOVA analysis, and Friedman test, which indicate varying levels of spectral separability across growth stages.

## V. Conclusion

This study demonstrates the effectiveness of the proposed CNN model for cassava plant health classification using multispectral UAV imagery and vegetation index image patches. The results show that classification performance varied across growth stages. The highest accuracy was achieved at 7 months (0.96), followed by 2 months (0.84) and 4 months (0.81). This variation indicates that plant age influenced classification performance through differences in canopy structure, spectral separability, and growth-stage-dependent vegetation index responses.

The comparison with baseline models, including Random Forest, SVM, XGBoost, ResNet50, and MobileNetV2, shows that the proposed CNN achieved the most consistent performance across growth stages. Although XGBoost produced competitive performance at the 4-month stage, the proposed CNN performed better at the 2- and 7-month stages and maintained stable classification capability across the observed datasets. The comparison before and after applying SMOTE further indicates that class balancing improved CNN performance across all growth stages. These findings highlight the importance of addressing class imbalance in UAV-based cassava plant health classification.

The integration of ANOVA, the Friedman test, CNN classification, baseline model comparison, SMOTE evaluation, and LIME-based interpretation provides a comprehensive framework for analyzing cassava plant health using multispectral UAV imagery. This study confirms that CNN-based approaches are suitable for UAV-based agricultural monitoring and can effectively capture spatial–spectral patterns related to plant health conditions.

However, this study was limited to data collected from a single geographic location and a single cassava variety. Model performance may vary under different environmental and geographic conditions. Factors such as illumination, seasonal variation, soil background, planting density, canopy structure, moisture availability, and local cultivation practices may influence vegetation index responses and the spectral separability of health classes. Therefore, future work should evaluate larger and more diverse datasets collected from multiple locations and cassava varieties. Future research should also investigate more advanced deep learning architectures to improve model generalization, robustness, and applicability to broader agricultural monitoring scenarios.

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

S.M. conducted the research, including conceptualization, data processing, method implementation, result analysis, and manuscript preparation. Y.H. supervised the study, validated the methodology, and reviewed the manuscript. S.W. contributed to the interpretation of plant health indicators and discussion refinement. E.S. provided agronomic expertise on cassava and supported result interpretation. All authors reviewed and approved the final manuscript.

## Acknowledgment

The authors gratefully acknowledge the contributions of the student team, field personnel, and the staff of the Remote Sensing Laboratory at IPB University for their support in data acquisition and processing.

## References

[1] J. Kalezhi and L. Shumba, "Cassava crop disease prediction and localization using object detection," *Crop Prot.*, vol. 187, 107001, 2025.

[2] M. K. Dharani, D. R. Thamilselvan, D. R. Rajdevi *et al.,* "Analysis on Cassava leaf disease prediction using pre-trained models," in *Proc. 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2022, pp. 1–6.

[3] N. Paul, G. C. Sunil, D. Horvath, and X. Sun, "Deep learning for plant stress detection: A comprehensive review of technologies, challenges, and future directions," *Comput. Electron. Agric.*, vol. 229, 109734, 2025.

[4] W. Mbinda and A. Mukami, "Breeding for postharvest physiological deterioration in cassava: problems and strategies," *CABI Agric. Biosci.*, vol. 3, no. 1, pp. 1–18, 2022.

[5] S. Prakadeswaran and A. B. Banu, "Gaussian regressed generative adversarial network based hermitian extreme gradient boosting for plant leaf disease detection," *Biomed. Signal Process. Control*, vol. 106, 107761, 2025.

[6] V. A. Chavez, A. E. Milne, F. van den Bosch, J. Pita, and C. F. McQuaid, "Modelling cassava production and pest management under biotic and abiotic constraints," *Plant Mol. Biol.*, vol. 109, no. 3, pp. 325–349, 2022.

[7] L. V. Satyanarayana, D. C. Rao, and S. K. Nayak, "Multi leaf disease identification and classification using efficient capsule convolutional shuffle attention network," *Comput. Electron. Agric.*, vol. 237, 110662, 2025.

[8] A. Chaiyana, N. Khiripet, S. Ninsawat, W. Siriwan, M. S.

Shanmugam, and S. G. P. Virdis, "Mapping and predicting cassava mosaic disease outbreaks using earth observation and meteorological data-driven approaches," *Remote Sens. Appl. Soc. Environ.*, vol. 35, 101231, 2024.

[9] P. M. D. Nguezet *et al.*, "Determinants of the profitability and producers' preferences for quality cassava planting material in Cameroon," *J. Agric. Food Res.*, vol. 22, 101971, 2025.

[10] C. Synthesis, J. Wattanathorn, and S. Muchimapura, "Multiple Bioactivities of Manihot esculenta Leaves: UV Filter ," *Molecules*, vol. 27, no. 1556, pp. 1–17, 2022.

[11] J. V. Laere *et al.*, "Stable isotope composition of long and short term carbon pools can screen drought tolerance in cassava," *F. Crop. Res.*, vol. 318, 109586, 2024.

[12] F. Calleja, B. Ondiviela, C. Galván, M. Recio, and J. A. Juanes, "Mapping estuarine vegetation using satellite imagery: The case of the invasive species Baccharis halimifolia at a Natura 2000 site," *Cont. Shelf Res.*, vol. 174, pp. 35–47, 2019.

[13] S. Maryana, Y. Herdiyeni, S. Wahjuni, and E. Santosa, "UAV-Based Segmentation and Correlation Analysis of Vegetation Indices for Cassava Crop Health Assessment," *Int. J. Informatics Vis.*, vol. 9, no. 4, pp. 1381–1389, 2025.

[14] A. Ramcharan *et al.*, "A mobile-based deep learning model for cassava disease diagnosis," *Front. Plant Sci.*, vol. 10, pp. 1–8, 2019.

[15] P. Nisha and J. Vijayakumar, "Recognition and Classification of Cassava Leaf Diseases Using Machine Learning Techniques," *Int. J. Nov. Res. Dev.*, vol. 7, no. 4, pp. 783–791, 2022.

[16] R. A. Purnamasari, T. Ahamed, and R. Noguchi, "Land suitability assessment for cassava production in Indonesia using GIS, remote sensing and multi-criteria analysis," *Asia-Pacific J. Reg. Sci.*, vol. 3, no. 1, pp. 1–32, 2019.

[17] J. K. Mbugua and W. Suksa-ngiam, "Predicting Suitable Areas for Growing Cassava Using Remote Sensing and Machine Learning Techniques," *Issues Informing Sci. + Inf. Technol.*, vol. 15, pp. 43–56, 2018.

[18] R. S. Rahmawati, D. Sukma, S. W. Ardie, and S. Sudarsono, "Postharvest physiological deterioration in cassava: potential problems, possible inhibition, and resistant level identification Postharvest physiological deterioration in cassava: potential problems, possible inhibition, and resistant level identifi," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 694, no. 1, 2021.

[19] G. Sambasivam, G. P. kanna, M. S. Chauhan, P. Raja, and Y. Kumar, "A hybrid deep learning model approach for automated detection and classification of cassava leaf diseases," *Sci. Rep.*, vol. 15, no. 1, pp. 1–24, 2025.

[20] C. Goyal, K. S. Gill, D. Upadhyay, and S. Devliyal, "Assessing the Effectiveness of Deep Learning Approaches in Forecasting Cassava Leaf Diseases," in *Proc. 2024 Int. Conf. Artif. Intell. Emerg. Technol. Glob. AI Summit 2024*, 2024, pp. 108–112. doi: 10.1109/GlobalAISummit62156.2024.10947791

[21] A. Gopi, L. R. Sudha, and S. I. T. Joseph, "Disclosing the Potential of Deep Learning in Cassava Leaf Disease Analysis by Using CNN and Neural Networks Approach," in *Proc. 7th Int. Conf. Inven. Comput. Technol. ICICT 2024*, 2024, pp. 188–193. doi: 10.1109/ICICT60155.2024.10544880

[22] A. A. John, "Identification of Diseases in Cassava Leaves using Convolutional Neural Network," in *Proc. 2022 5th Int. Conf. Comput. Intell. Commun. Technol. CCICT 2022*, 2022, pp. 1–6. doi: 10.1109/CCiCT56684.2022.00013

[23] S. Bhagat *et al.*, "Advancing real-time plant disease detection: A lightweight deep learning approach and novel dataset for pigeon pea crop," *Smart Agric. Technol.*, vol. 7, 100408, 2024. doi: 10.1016/j.atech.2024.100408

[24] P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, and I. Moscholios, "A compilation of UAV applications for precision agriculture," *Comput. Networks*, vol. 172, 107148, 2020. doi: 10.1016/j.comnet.2020.107148

[25] T. R. Alabi, A. T. Abebe, G. Chigeza, and K. R. Fowobaje, "Estimation of soybean grain yield from multispectral high-resolution UAV data with machine learning models in West Africa," *Remote Sens. Appl. Soc. Environ.*, vol. 27, 100782, 2022. doi: 10.1016/j.rsase.2022.100782

[26] H.-T. Thai, K.-H. Le, and N. L.-T. Nguyen, "FormerLeaf: An efficient vision transformer for Cassava Leaf Disease detection," *Comput. Electron. Agric.*, vol. 204, 107518, 2023. doi: 10.1016/j.compag.2022.107518

[27] A. H. T. E. De Silva, J. Katupitiya, and A. V Savkin, "UAV-UGV Collaborative Localisation with Ambiguity Aversion by UAV Re-positioning," *IFAC-PapersOnLine*, vol. 55, no. 32, pp. 101–106, 2022. doi: 10.1016/j.ifacol.2022.11.122

[28] Y. Zhang *et al.*, "Unsupervised domain adaptation semantic segmentation method for wheat disease detection based on UAV multispectral images," *Comput. Electron. Agric.*, vol. 236, 110473, 2025. doi: 10.1016/j.compag.2025.110473

[29] F. Noromiarilanto, K. Brinkmann, M. H. Faramalala, and A. Buerkert, "Assessment of food self-sufficiency in smallholder farming systems of south-western Madagascar using survey and remote sensing data," *Agric. Syst.*, vol. 149, pp. 139–149, 2016. doi: 10.1016/j.agsy.2016.09.005

[30] Y. Ji, Z. Liu, R. Liu, Z. Wang, X. Zong, and T. Yang, "High-throughput phenotypic traits estimation of faba bean based on machine learning and drone-based multimodal data," *Comput. Electron. Agric.*, vol. 227, 109584, 2024. doi: 10.1016/j.compag.2024.109584

[31] Y. Yang *et al.*, "Advanced technologies for precision tree fruit disease management: A review," *Comput. Electron. Agric.*, vol. 229, 109704, 2025. doi: 10.1016/j.compag.2024.109704

[32] W. Qian *et al.*, "UAV and a deep convolutional neural network for monitoring invasive alien plants in the wild," *Comput. Electron. Agric.*, vol. 174, 105519, 2020. doi: 10.1016/j.compag.2020.105519

[33] F. Ouyang *et al.*, "Automatic delivery and recovery system of Wireless Sensor Networks (WSN) nodes based on UAV for agricultural applications," *Comput. Electron. Agric.*, vol. 162, pp. 31–43, 2019.

[34] S. Adebayo, H. O. Aworinde, A. O. Akinwunmi, A. Ayandiji, and A. O. Monsir, "Convolutional neural network-based crop disease detection model using transfer learning approach," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 29, no. 1, pp. 365–374, 2023. doi: 10.11591/ijeecs.v29.i1.pp365-374

[35] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Front. Plant Sci.*, vol. 8, pp. 1–7, 2017. doi: 10.3389/fpls.2017.01852

[36] I. Sangbamrung, P. Praneetpholkrang, and S. Kanjanawattana, "A novel automatic method for cassava disease classification using deep learning," *J. Adv. Inf. Technol.*, vol. 11, no. 4, pp. 241–248, 2020. doi: 10.12720/jait.11.4.241-248

[37] S. Shi, Y. Du, and W. Fan, "An Extension of LIME with Improvement of Interpretability and Fidelity," arXiv preprint, arXiv:2004.12277, 2020.

[38] S. Shi, X. Zhang, and W. Fan, "Explaining the Predictions of Any Image Classifier via Decision Trees," arXiv preprint, arXiv:1911.01058, 2019.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).