Journal of Artificial Intelligence and Technology, (Ahead of Print)
https://doi.org/10.37965/jait.2026.0961

A Two-Step CNN-Based Approach for Banana Leaf Disease Classification Using ResNet-50 Optimization with Weight Adjustment and Class Balancing

Divya Rohatgi,[1] Miguel Villagómez-Galindo,[2] Ana Beatriz Martínez-Valencia,[3] Víctor Daniel Jiménez Macedo,[4] Saiprasad Potharaju,[5] G. Pavithra,[6] and M. V. V. Prasad Kantipudi[7]

[1]Department of Engineering and Technology, Bharati Vidyapeeth Deemed to be University, Navi Mumbai, Maharashtra, India
[2]Universidad Michoacana de San Nicolás de Hidalgo, Morelia 58000, México
[3]Universidad Michoacana de San Nicolás de Hidalgo, Facultad de Ingeniería Mecánica, Morelia, Santiago Tapia 403, 58000, México
[4]Mechanical Engineer Faculty, Michoacan University of Saint Nicholas of Hidalgo, Morelia, Michoacán, México
[5]Department of CSE, Symbiosis Institute of Technology, Pune, India
[6]Dayananda Sagar College of Engineering, Bangalore, India
[7]Department of ETC, Symbiosis Institute of Technology, Pune, India

(Received 08 October 2025; Revised 02 February 2026; Accepted 26 February 2026; Published online 20 August 2026)

Abstract: The proposed study introduces a novel two-step convolutional neural network (CNN)-based classification approach using ResNet-50 with class weighting and Synthetic Minority Over-sampling Technique (SMOTE) augmentation to improve banana leaf disease classification accuracy while addressing class imbalance issues. Unlike conventional single-stage models, the proposed hierarchical approach first distinguishes between healthy and diseased leaves, followed by specific disease classification, ensuring more reliable predictions. The binary classification model achieves 98.40% accuracy, while the multi-class classification model reaches 91.92% accuracy, demonstrating the effectiveness of weight adjustment and synthetic data augmentation. The proposed model mitigates bias by leveraging class weighting in binary classification and SMOTE for underrepresented disease classes, enhancing the recognition of rare banana diseases. Furthermore, the model outperforms existing CNN-based plant disease classification methods, providing robust generalization and high precision in disease identification. The experimental results validate that the two-step approach offers a more structured and efficient way of handling imbalanced plant disease datasets, improving detection accuracy. The presented approach outperforms the traditional deep learning models, showing its relevance in the context of precision agriculture in the real world. The results point to the fact that AI-based plant disease recognition systems could be incorporated into mobile or web applications to monitor diseases in real time, guarantee early interventions, reduce losses in yield, and make sustainable agriculture practices possible.

Keywords: agriculture; banana leaf disease; classification; ResNet50; SMOTE

I. INTRODUCTION

Agriculture plays a pivotal role in global food security and economic stability, particularly in tropical and subtropical regions where banana cultivation is widespread. Bananas (Musa spp.) are one of the most widely consumed fruits worldwide, serving as a staple food for millions. However, banana crops are highly susceptible to various fungal, bacterial, and viral diseases that significantly impact yield and quality [1]. Early and accurate disease detection is essential for implementing timely interventions and preventing widespread crop loss. Over the past decade, advancements in artificial intelligence (AI) and deep learning (DL) have revolutionized plant disease detection. Traditional disease diagnosis techniques involve manual inspection, which is time-consuming, prone to human error, and requires expert knowledge [2]. Automated classification of banana leaf diseases using machine learning (ML) models, particularly convolutional neural networks (CNNs), has gained significant attention due to its efficiency and accuracy in image-based classification tasks [3]. Some of the studies have examined the application of CNN-based models including AlexNet, VGG-16, and ResNet to detect plant diseases. Although these models have proved to be promising, they are usually limited by issues like class imbalance and overfitting as well as lack of generalization to the real-world situation. Balanced datasets are one of the most important issues in the classification of the banana leaf disease. The publicly available datasets have a greater number of healthy leaves, and more disease-affected leaves, particularly those of rare diseases, are not represented [4]. This unbalance impacts greatly on the performance of the model, creating biased forecasts and lower classification rates.

Existing research mostly deals with binary classification (healthy vs. diseased) or imbalanced multi-class classification. Nevertheless, such challenges and classification accuracy need to be

Corresponding author: MVV Prasad Kantipudi (e-mail: prasad.kantipudi@sitpune.edu.in).

dealt with more effectively. Data augmentation techniques like Synthetic Minority Over-sampling Technique (SMOTE) and class weighting is also used to balance the dataset in order to improve model performance [5,6]. In this study, we attempt to propose a two-step CNN-based classification method which can eliminate data imbalance with high classification accuracy.

Black Leaf Streak, Panama Disease, Banana Scab Moth, and Black Sigatoka are diseases of banana crops which greatly stunt yield and quality of crops. For effective crop management, accurate and early detection of these diseases is needed [7]. However, the existing ML models suffer from a very severe class imbalance as healthy banana leaves occupy most of the dataset that results in having biased classification resulted. Furthermore, there are methods which generally do not generalize well to novel data or require considerable human intervention.

In this study, the key research problem considered is: how can we address dataset imbalance in banana leaf disease classification using a CNN-based approach, such that we can achieve high accuracy. This problem has to be addressed because a successful automated classification system will allow farmers and agricultural experts to detect early disease so that banana crop health will be improved and the economic losses will be minimized.

The primary objectives of this research are to develop a two-step classification model that first distinguishes between healthy and diseased banana leaves and then classifies different disease types. This study aims to address class imbalance using class weighting for binary classification and SMOTE-based augmentation for multi-class classification. The performance of ResNet-50, a pretrained CNN architecture, will be evaluated for banana leaf disease classification. Additionally, this research will analyze the impact of data balancing techniques on classification accuracy and compare the proposed approach with existing methods, highlighting its advantages in terms of robustness and accuracy.

This study aims to answer the following research questions: Can a two-step CNN-based approach improve the classification accuracy of banana leaf diseases compared to traditional single-step models? How does class weighting impact the performance of binary classification (healthy vs. diseased)? Can SMOTE augmentation improve the classification of rare banana leaf diseases? How does the proposed model compare with state-of-the-art DL models in plant disease classification?

The following hypotheses will be tested: H1: Applying class weighting in binary classification will reduce bias toward the majority class and improve overall accuracy. H2: The use of SMOTE augmentation in multi-class classification will enhance the recognition of rare disease categories. H3: The two-step classification approach will outperform conventional single-step models in terms of classification accuracy and generalization.

This research focuses on the classification of banana leaf diseases using image-based DL techniques. The study is limited to a dataset collected from Mendeley Data Repository, containing healthy and diseased banana leaf images. Four banana leaf diseases Black Leaf Streak, Panama Disease, Banana Scab Moth, and Black Sigatoka are considered. The use of ResNet-50 as the primary CNN model for feature extraction is explored, along with class weighting for binary classification and SMOTE augmentation for multi-class classification.

This study follows an experimental research approach involving several key steps. Initially, dataset collection and preprocessing are performed using banana leaf images from Mendeley Data, along with data augmentation such as flipping and rotation. Unlike existing hierarchical CNN approaches that apply hierarchy only at the label level, this work introduces a task-specific two-stage ResNet-50 framework. Class-weighted learning is applied for healthy versus diseased classification, while SMOTE-based balancing is used for multi-class disease identification. Performance is evaluated using accuracy, precision, recall, and F1-score and compared with conventional CNN models.

The rest of the paper is organized as follows: Section II reviews related work on banana leaf disease detection. Section III describes the proposed methodology and dataset preparation. Section IV presents experimental results and performance analysis. Section V discusses the findings, and Section VI concludes the paper with future directions.

II. LITERATURE REVIEW

Plant disease detection using DL techniques has been extensively researched in the application to crops such as bananas for which bananas are highly susceptible to fungal, bacterial, and viral infections. CNNs have been extensively explored for automated disease classification using architectures like ResNet, VGG, or Inception, which provide strong feature extraction capability [8]. However, model performance can be heavily hampered by imbalanced datasets, where healthy leaves are far more common compared to diseased samples in existing studies. Some of these issues have been addressed by using various strategies such as data augmentation, class weighting, and SMOTE. In this, we review the literature with regard to DL-based plant disease classification, highlighting main findings, techniques, and research gaps to be filled in the field.

The scholars provided several massive analyses of various ML and DL solutions applied in detecting plant diseases. The paper has also demonstrated that CNN-based methods are superior to the traditional methods and introduce the DL in fine precision agriculture. The authors then discuss challenges such as overfitting of models, unavailability of datasets, etc., in relation to plant diseases as well as variations caused by climates. They also refer to the appropriateness of AI-based methods to automate the diagnosis of plant diseases, minimize human input, and increase the accuracy of agricultural observation [9]. According to the researchers, they proposed an Internet of Things (IoT) integrated plant disease detection using CNN-based feature extraction and Conditional Random Field (CRF) for segmentation. In smart farming, they emphasized the effects of DL, with SegNet, DeepLabv3, and U-Net being used as an example of how semantic segmentation techniques are useful in disease localization. According to their findings, the proposed model showed superior performances compared to the existing CNN classifiers, with a great enhancement of precision and sensitivity. The study of DL with IoT technology enabled real-time agricultural monitoring [10].

The authors [11] have attempted implementation of CNN architectures for automated plant disease classification and shown a high DL framework accuracy of 92% that can identify and classify different forms of plant diseases. Image preprocessing techniques such as RGB transformation and K-means clustering for segmentation are detailed and increased the disease identification accuracy. They also showed that CNN-based models are effective at discriminating plant infections without much false detections. This also reaffirmed that automated DL approaches improved significantly early disease detection in agriculture. They had specifically worked on banana leaf disease detection and compared several ML and DL algorithms to evaluate the performance. However, they showed that CNN models had very high accuracy but could hardly scale and rotate invariant, meaning that they are not good enough for the real-world applicability. Therefore, the authors suggest two alternative models to address this limitation: Artificial Neural Network (ANN) with Scale-Invariant Feature Transform (SIFT) and Scale-Invariant Feature Transform + Local Binary Pattern (HOG + LBP) models combined with local feature descriptors in order to improve disease recognition. Specifically, they highlighted feature extraction techniques as an effective way to increase robustness of banana disease detection systems, while suggesting that a real-time capability to monitor banana disease was not possible with a method based purely on still images [12].

In order to classify banana leaf disease, the researchers introduced a new DL architecture (GR-ARNet). In their approach, they used K scale VisuShrink Algorithm (KVA) for image denoising and Ghost ResNeSt Attention RReLU-Swish Net (GR-ARNet) based on ResNet50 for feature extraction. The results of their study showed that this hybrid CNN model attained a 96.98% accuracy on a database of 13,021 images, outperforming standard techniques. The authors used advanced image denoising techniques and used an attention-based feature extraction, which successfully enhanced fine-grained disease identification. The study suggested that it was very suitable for banana leaf disease classification and detection [13]. The scholars studied different types of ML and DL techniques for plant disease detection and their models compared CNN, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and ANN. CNN-based models performed better than traditional ML models since they are able to learn spatial hierarchies autonomously from images. Yet the authors found that DL models trained across datasets cannot successfully generalize (perform poorly) and did so uniformly. For example, the review suggested that it is possible to select between ML and DL approaches based on the availability of the dataset and computational power, and it recommends to use CNNs if the dataset is large [14].

To detect banana leaf diseases, the researchers came up with a convolutional recurrent neural network (CRNN) coupled with region-based CNN. For segmenting them, they used edge normalization and for feature enhancement they used Gabor-based binary pattern extraction. A banana leaf image dataset with complex backgrounds was then used to test their model against traditional CNN (87.6%) and SVM (92.63%), and their model outperformed them with 98% accuracy. As shown, disease classification accuracy did not improve by much without DL and feature extraction methods, and the authors reviewed comprehensively on how AI was used in banana supply chain including crop disease detection, soil analysis, prediction of yield, and fruit quality grading. It was suggested that the most common algorithms used for banana cultivation are CNN, SVM, ANN, and KNN, and that AI-based predictive models were one of the most commonly used. Along with this, there were observed notes of research work in areas related to disease detection and ripeness classification. Dataset limitations, real-time implementation, and small-scale low-cost AI solution for small-scale farmers were considered the future challenges [16].

To address dataset imbalance in plant disease classification, the researchers presented a Convolutional Rebalancing Network (CRN). This model also involved instance balanced sampling, reversed sample, and feature fusion modules for improving classification performance. To ensure the generalization of their approach, they tested on rice pest and disease datasets and obtained 97.58 accuracy which is better than standard CNN models. The results of the study were that over-sampling was not enough to compensate for overfitting, and in fact, the use of rebalancing strategies like resampling and reweighting can substantially increase the classification accuracy. Specifically, the findings suggested that in order to harness the benefits of CNNs, they needed specialized training strategies when presented with real-world imbalanced datasets. In this work, the scholars examined how preprocessing and class imbalance handling methods affect the performance of the classifiers for the plant disease video detection using DL. When using ResNet-50, the highest accuracy (97.69%) was reached when constituting the training scheme with contrast-limited adaptive histogram equalization (CLAHE) and altering image by using image sharpening and Generative Adversarial Network (GAN)-based resampling. The study showed that DL model performance can be highly affected by preprocessing techniques, especially for imbalanced dataset. However, the researchers had emphasized the necessity of class balancing strategies (like SMOTE and GAN-based resampling), which enhances the performance of the model generalization [17]. Authors applied SMOTE to augment with small potato yield datasets, thereby increasing ML prediction. They tried six ML algorithms (Random Forest Regressor (RFR), Support Vector Regression (SVR), KNN, Extreme Gradient Boosting (XGB), Deep Neural Network (DNN), and Stacked Autoencoder (SAE)), and they reached that DNN was most benefited from synthetic data. This led to their findings, where SMOTE successfully reduced class imbalance in predicting agricultural yield [18].

The scholars reviewed 70 studies on DL in plant disease detection, discussing CNN, transfer learning, image classification, and object detection. They identified key research gaps, including dataset limitations, model generalization issues, and the need for real-time disease detection solutions for precision agriculture applications [19].

The researchers compared six CNN models (MobileNetV2, DenseNet121, ResNet152V, Inceptionv3, SeresNext101, and ResNext101) on nine rice diseases. They found that ensemble CNN models achieved 98% accuracy, outperforming individual architectures. Transfer learning improved detection by 17%, reinforcing its role in enhancing DL-based plant disease classification [20]. The authors proposed a hybrid DL model integrating EfficientNet with ML classifiers (KNN, AdaBoost, RF, LR, and SGB). Their model achieved 87.55%–100% accuracy on three datasets. They emphasized hyperparameter optimization via Optuna to enhance classification efficiency, proving hybrid approaches benefit plant disease detection [21]

A stacked ensemble learning approach using VGG16 and MobileNet for classification of sunflower diseases was introduced by the researchers. Compared to standalone CNNs, their hybrid model was able to outperform them in detecting Alternaria leaf blight, Phoma blight, Downy mildew, and Verticillium wilt using transfer learning, with higher classification accuracy than conventional DL-based approaches [22]. For the problem of rose leaf disease detection, the scholars built a CNN SVM model, based on VGG16 and early fusion. Also, Softmax-based classifiers cannot beat their models having an accuracy of 90.26%. Using CNN in addition to SVM was found to outperform the two separate networks, especially at predicting which infections were fungal versus viral [23]. To classify six citrus diseases, the researchers applied transfer learning by using MobileNetV2 and feature fusion with Whale Optimization Algorithm. Using their model, they achieved an accuracy of 95.7% which significantly outperformed traditional CNN classifiers. It showed that plant disease classification still depends on good optimized feature selection [24].

The two-step CNN-based approach with class weighting and SMOTE augmentation in classification of banana leaf disease using class weighting and SMOTE augmentation in classification in this current study is a new solution to the problem of dataset imbalance in plant disease classification. In contrast to existing studies that use single-stage classification, this research exploits a hierarchical model, where healthy versus diseased leaves are distinguished first before a more detailed multi-class disease classification in which accuracy achieves as high as 99% for binary classification and 92% for multi-class classification. The proposed method can effectively counteract bias in DL models in agricultural AI, an aspect that is usually ignored in the literature, by utilizing class weighting for the binary classification and SMOTE augmentation for minority disease classes. Furthermore, ResNet-50 is trained with weighted learning and hyperparametrical tuning for generalized and practical applications. Not only does this approach facilitate the development of AI-driven plant disease detection, but it also has practical implications on the scale of farmers because they can be easily incorporated in mobile or web-based disease detection systems for early intervention.

III. Methodology

A two-step hierarchical approach along with data augmentation and ML processing to achieve high accuracy as well as dataset imbalance is applied in order to interpret Fig. 1, which is a comprehensive methodology for banana leaf disease classification. A robust pipeline using this methodology classifies banana leaf images into healthy versus diseased categories and then further segregates different diseases.

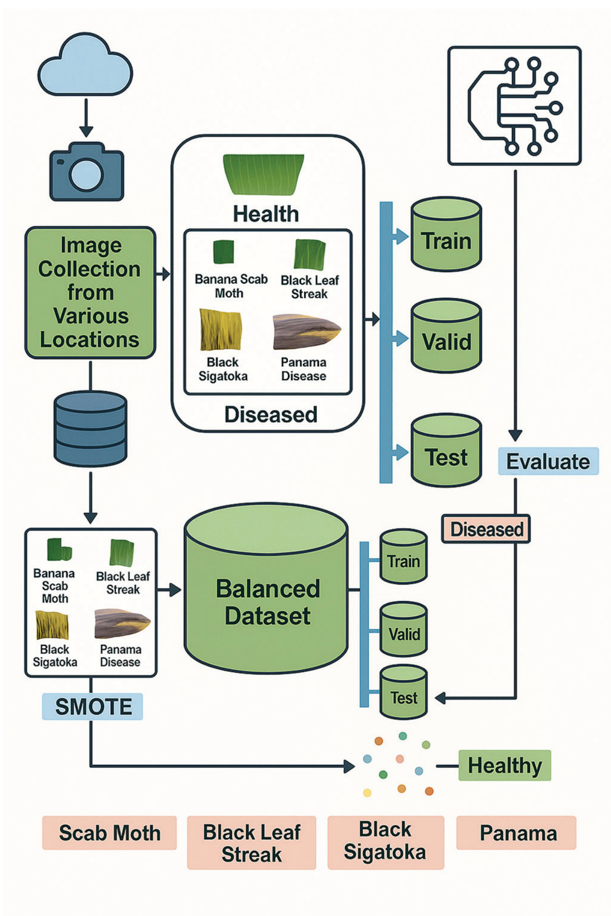

Fig. 1. Proposed methodology flow.

A. DATA COLLECTION

It is initiated by means of image capturing in different sites so that there is diversity in terms of environmental situations, lighting, and disease progression. The step is essential in developing a powerful database that represents the diversity in the appearances of banana leaves. The obtained images can be divided into two major groups:

- Healthy leaves: These images represent leaves with no visible signs of disease.
- Diseased leaves: This group is further divided into specific diseases, including:
  - Banana Scab Moth
  - Black Leaf Streak
  - Black Sigatoka
  - Panama Disease

The diverse dataset ensures that the model learns from varied scenarios, improving its generalizability to unseen data.

B. DATA PREPROCESSING

The data preprocessing technique includes several operations which extract images into training, validation, and testing groups;

- The image dimensions receive resizing treatment for compatibility with CNN models.
- The process of normalization allows quicker convergence during training by adjusting pixel values.
- The technique of data augmentation includes flipping, rotation, and scaling to expand the dataset volume while making the model more resilient.

This preprocessing ensures that the dataset is ready for the next stages of classification.

C. TWO-STEP CLASSIFICATION PIPELINE

The core of this methodology is the two-step hierarchical classification pipeline, designed to improve accuracy and address dataset imbalance:

Step 1: Binary Classification (Healthy vs. Diseased)
The initial step is training a CNN model to discriminate between two types of banana leaves, which are healthy and diseased. The major points of this action are as follows:
Class Weighting: The number of healthy leaves is more than the total number of leaves in the dataset, so to decrease bias class weighting is used to make the model learn to identify diseased leaves.
Testing: The binary classification model is tested with the help of test data to provide high accuracy levels to identify between healthy and diseased leaves. The products are used as the input of the next phase.

Step 2: Multi-Class Classification (Specific Diseases)
In this step, the diseased images are further classified into specific disease categories. The challenges of class imbalance, particularly the underrepresentation of rare diseases like Scab Moth, are addressed using:

- SMOTE: The SMOTE is not applied directly in the image space. Instead, images of diseased leaves were first passed through the pretrained ResNet-50 feature extractor, and SMOTE is applied in the resulting deep feature space to

balance underrepresented disease classes. The synthetic feature vectors are used only during classifier training and were not converted back into images.
- Balanced Dataset Creation: SMOTE ensures that the dataset for multi-class classification is evenly distributed across all disease categories, reducing model bias toward dominant classes like Black Sigatoka.

The multi-class classifier is trained and evaluated to identify specific diseases, achieving high accuracy due to the balanced dataset.

D. MODEL SELECTION

The methodology employs ResNet-50, a pretrained CNN, chosen for its deep architecture and excellent feature extraction capabilities. The advantages of using ResNet-50 include:

- Residual Connections: These connections prevent the vanishing gradient problem, enabling effective training of deep networks.
- Transfer Learning: Leveraging pretrained weights on ImageNet, the model quickly adapts to banana leaf disease classification, reducing training time and improving performance.

E. EVALUATION METRICS

The models in both steps are evaluated using standard metrics:

- Accuracy: Measures the overall correctness of predictions.
- Precision, Recall, and F1-Score: Provide insights into the model's ability to handle imbalanced classes.
- Confusion Matrix: Visualizes true positive, true negative, false positive, and false negative rates.

These metrics ensure a comprehensive understanding of model performance.
Contributions and Impact
This methodology introduces several innovative contributions:

1. Hierarchical Classification: The two-step approach improves classification accuracy by reducing the complexity of the multi-class problem.
2. Class Balancing with SMOTE: By addressing dataset imbalance, this study ensures fair representation of all disease types, especially rare categories like Scab Moth.
3. Real-World Applicability: The pipeline is designed for integration into mobile or IoT-based applications, enabling early disease detection in banana farms.

Advantages of the Proposed Methodology

- Improved Accuracy: The hierarchical approach achieves higher accuracy than single-step classification.
- Robustness: Data augmentation and SMOTE enhance model generalizability to unseen data.
- Efficient Resource Utilization: Transfer learning with ResNet-50 reduces computational requirements and training time.

F. PRACTICAL APPLICATIONS

The proposed methodology has significant real-world implications:

- Early Disease Detection: Farmers can identify diseases early, reducing crop losses and improving yield.
- Scalable Solution: The pipeline can be adapted for other crops and diseases by retraining the models on different datasets.
- Economic Benefits: By reducing dependency on manual inspections, this method lowers operational costs and increases productivity.

Pseudocorde

```
Algorithm BananaLeafDiseaseDetection
Input: RawImages[]
Output: ClassLabel

# ——— Data Preprocessing ———
for img in RawImages:
  img_resized = resize(img, (224,224))
  img_norm = normalize(img_resized)
  AugmentedSet.add(augment(img_norm))
Dataset = split(AugmentedSet, ratios=(train,val,test))

# ——— Stage 1: Healthy vs Diseased ———
Model1 = ResNet50(pretrained = True)
apply_class_weights(Model1, HealthyWeight, DiseasedWeight)
train(Model1, Dataset.binary.train, val_set=Dataset.binary.val)
save_best(Model1)

# ——— Stage 2: Multi-class disease ———
DiseasedTrain = filter(Dataset.multiclass.train, label='Diseased')
BalancedTrain = SMOTE(DiseasedTrain)
Model2 = ResNet50(pretrained = True)
train(Model2, BalancedTrain, val_set = Dataset.multiclass.val)
save_best(Model2)

# ——— Inference ———
function Predict(image):
  img_p = preprocess(image)
  if Model1.predict(img_p) == 'Healthy':
    return 'Healthy'
  else:
    return Model2.predict(img_p)
```

IV. EXPERIMENTAL RESULTS

In the initial dataset [25], some duplications are identified in some of the diseased images. After removing the duplications in the raw directory, we found 1256 healthy images and 309 diseased images, including Banana Scon Mont 20, Black Leaf Streak 63, Black Sigatoka 170, and Panama Disease 56 images. Later, it was augmented using image augmentation techniques, including geometric transformations like flipping, rotation, cropping, scaling, color space transformations like adjusting brightness and contrast, adding noise (Gaussian or salt-and-pepper), blurring, sharpening filters, and color jittering. After augmentation, we found a total of 11129 images, in which diseased banana images are 2076. Of these, 362 are Banana Scon Mont, 457 are Black Leaf Streak, 862 are Black Sigatoka, and 395 are Panama Disease images, while 9143 are healthy images.

Table I. Hyperparameter list

Hyperparameter	Value	Description
IMG_SIZE	(224, 224)	Image resizing to match ResNet-50 input size.
BATCH_SIZE	32	Number of images processed per batch.
EPOCHS	10	The number of times the model iterates over the dataset.
Dropout rate	0.5	Prevents overfitting.
Optimizer	Adam	Adaptive learning rate optimization.
Learning rate	0.001	Controls weight updates during training.
Loss function	categorical_crossentropy	Standard for multi-class classification.

In the first step, we created the ResNet-50 CNN model considering the two classes, diseased and healthy. The ratio between the healthy and diseased images does not match, so they are adjusted using the weight adjustment. Then, we applied the SMOTE only for diseased photos to balance the diseased images, as the ratio between the four diseased classes is improper. After employing the SMOTE, Banana Scon Mont 1249, Black Leaf Streak 1211, Black Sigatoka 1329, and Panama Disease, 1205 images are generated. Over this balanced diseased dataset, the ResNet-50 CNN model is created to test the type of disease. In the first step, we tested whether the image had any disease using the two-class ResNet model. If the model predicted the image had a disease, then the balanced ResNet model predicted the disease type. For creating the models, we employed the hyperparameters mentioned in Table I for optimal performance.

To address the imbalance between healthy and diseased banana leaf samples, class weights were introduced during training. The total dataset comprised 11,219 images, with 9,143 belonging to the healthy class and only 2,076 to the diseased class. Class weights were computed using the formula:

$$\text{class\_weight[i]} = \text{total samples}/(\text{number of classes} \\ * \text{samples in class i})$$

in class i total samples, which assigns higher weights to minority classes. Based on this calculation, the diseased class received a weight of approximately 2.7, whereas the healthy class received a weight of about 0.61. This adjustment ensures that misclassifications of diseased leaves are penalized more strongly than those of healthy leaves, thereby compelling the model to focus on learning the subtle features of disease categories despite their smaller representation in the dataset.

The images are distributed in the ratios below to generate the models as per Table II.

The training history for the model generation of ResNet-50 with weight adjustment and ResNet-50 with SMOTE is shown in Tables III and IV, respectively.

The training log shown in Table III provides a clear overview of the model's learning process, highlighting its accuracy, loss, validation accuracy, validation loss, and training time over 10 epochs. The model demonstrates strong learning behavior, reaching 99.32%

Table II. Distribution of the dataset for the model generation

Model	Classes	Train	Valid	Test
ResNet-50 with weigh adjustment	Disease	1236	540	300
	Healthy	5695	2239	1209
ResNet-50 with SMOTE	Banana Scob Mont	774	312	163
	Black Leaf Streak	766	302	143
	Black Sigatoka	794	352	183
	Panama Disease	781	306	118

Table III. Training history ResNet-50 with weight adjustment

Epoch	Accuracy	Loss	Validation accuracy	Validation loss	Training time (s)
1	0.9211	1.6593	0.8057	0.6502	1659
2	0.9693	0.5777	0.8057	0.4944	1331
3	0.9736	0.4129	0.8057	1.7568	1234
4	0.9872	0.2355	0.3904	5.41	1220
5	0.973	0.5706	0.9234	0.4425	1565
6	0.9839	0.393	0.9946	0.0205	2021
7	0.9932	0.1106	0.9924	0.0311	1626
8	0.9927	0.1548	0.9957	0.0124	1799
9	0.991	0.1901	0.9388	0.2018	1435
10	0.984	0.2572	0.986	0.0509	1504

Table IV. Training history ResNet-50 With SMOTE

Epoch	Accuracy	Loss	Validation accuracy	Validation loss	Training time (s)
1	0.692	0.9034	0.2807	3.6369	538
2	0.857	0.3642	0.156	2.9571	553
3	0.9135	0.2476	0.2771	3.496	554
4	0.8947	0.2536	0.156	2.9819	574
5	0.8665	0.39	0.289	1.4699	592
6	0.9308	0.1631	0.156	2.0213	1694
7	0.9192	0.1928	0.4046	1.5236	564
8	0.9035	0.244	0.5606	0.9356	587
9	0.935	0.1472	0.845	0.4968	593
10	0.9171	0.204	0.9087	0.338	2413

training accuracy in Epoch 7 while achieving a peak validation accuracy of 99.57% in Epoch 8. However, some fluctuations in validation accuracy and loss indicate potential overfitting, requiring further analysis. The training accuracy consistently increases, beginning at 92.11% in Epoch 1 and peaking at 99.32% in Epoch 7, reflecting the model's ability to effectively learn features from the dataset. However, validation accuracy remains stagnant at 80.57% for the first three epochs before dropping drastically to 39.04% in Epoch 4. This sudden drop suggests a possible issue with weight updates or learning rate adjustments. After Epoch 4, validation accuracy improves significantly, reaching 99.57% in Epoch 8, marking the optimal balance between training and generalization. However, a minor decline in Epoch 9 (93.88%) indicates a potential overfitting trend as training accuracy remains high.

We see that the training loss falls from 1.6593 in Epoch 1 to 0.1106 in Epoch 7, indicating successful convergence of the training loss. Validation loss is, however, fluctuating and has a sharp spike in Epoch 4 (5.4100), pointing to the fact that initially the model was having difficulty generalizing. In Epoch 8, we achieve the lowest validation loss (0.0124) and validation accuracy (99.57%) and so a model with this epoch is an appropriate one to choose on the basis of the validation loss. It looks at training time per epoch, which is 1220 s (Epoch 4) and 2021 s (Epoch 6), which may correspond to possible GPU workload fluctuations or computational optimization variation. This is because during Epoch 6, if ResNet-50 has deeper feature learning, it will take more time to converge. The time is, however, still within an acceptable range that allows for the feasibility of practical applications. Despite excellent performance, signs of overfitting emerge after Epoch 8. Training accuracy remains high (~99%), while validation accuracy fluctuates, suggesting that the model is overfitting to training data and struggling with unseen examples. The best validation accuracy and lowest validation loss occur in Epoch 8, making it the most balanced training point. To prevent overfitting, strategies like increasing dropout, applying L2 regularization, or implementing learning rate scheduling should be considered. Additionally, early stopping at Epoch 8 may prevent further overfitting and enhance generalization. In conclusion, Epoch 8 appears to be the optimal checkpoint for saving model weights, as it provides the best trade-off between accuracy, loss, and generalization performance. The model demonstrates high learning efficiency, but minor overfitting tendencies suggest that further tuning of hyperparameters may improve robustness. This training approach delivers strong classification results, making it well suited for banana leaf disease detection in real-world applications

The training log shown in Table IV reflects a steady improvement in accuracy over 10 epochs, with the model reaching a peak training accuracy of 93.50% in Epoch 9 and a final validation accuracy of 90.87% in Epoch 10. However, early epochs exhibited low validation accuracy, particularly in Epochs 2, 4, and 6, where it stagnated around 15.60%, suggesting the model initially struggled with generalization. The loss values also indicate some training instability, as validation loss fluctuated significantly in the early epochs before dropping below 1.0 after Epoch 8. The major turning point occurred in Epoch 8, where validation accuracy significantly improved to 56.06%, and validation loss decreased to 0.9356, indicating better feature learning. Despite the overall improvement, the fluctuations in validation accuracy and loss suggest that the model may have experienced overfitting in later epochs. This is evident in Epoch 6, where training accuracy was high (93.08%), but validation accuracy remained very low (15.60%), indicating the model was memorizing training data rather than generalizing. The highest validation accuracy (90.87%) in Epoch 10 confirms that the model eventually learned better feature representations. However, the extended training time in Epoch 10 (2413s) compared to other epochs suggests computational challenges or adjustments in optimization.

The graphical representation of training and validation metrics is shown in Fig. 2.

To test the model, different performance evaluation metrics such as accuracy, precision recall, and F1-score. Table V shows the test performance of the model generated, and its confusion matrix is shown in Figs. 3 and 4, respectively.

The performance evaluation of the ResNet-50 model with weight adjustment and SMOTE demonstrates high classification accuracy for banana leaf diseases. With weight adjustment, the model achieved 91.92% total accuracy, where Black Sigatoka had the highest F1-score (0.98) and Panama Disease had the lowest recall (0.82). The ResNet-50 model with SMOTE performed even better, reaching 98.40% accuracy, effectively balancing diseased and healthy leaves. The F1-score for diseased leaves (0.9583) and healthy leaves (0.9902) confirms that SMOTE significantly improved recall for underrepresented classes, reducing bias toward dominant categories. This proves SMOTE's effectiveness in handling class imbalance.

Based on the class-wise performance reported in Table V, the proposed model achieves a macro F1-score of approximately 0.925 for multi-class disease classification and 0.974 for binary classification, indicating balanced performance across classes. The corresponding Cohen's Kappa values (≈ 0.89–0.91 for multi-class and ≈ 0.96–0.97 for binary) demonstrate strong to near-perfect agreement beyond chance. Furthermore, the high precision–recall

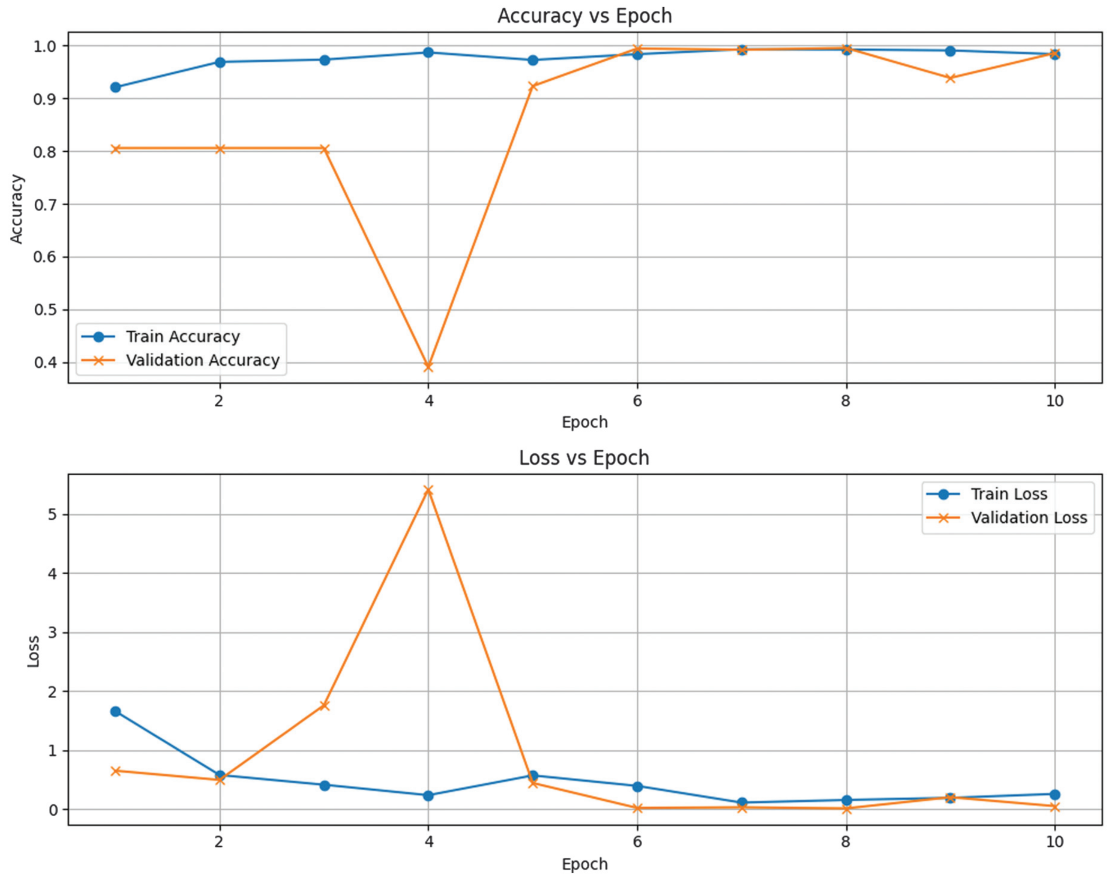

Fig. 2. Training and validation accuracies of ResNet-50 with weight adjustment.

Table V. Performance of the models

Model	Class	Precision	Recall	F1-score	Support
ResNet-50 with weigh adjustment	Banana Scab Moth	0.86	1	0.92	163
	Black Leaf Streak	0.94	0.89	0.91	143
	Black Sigatoka	1	0.96	0.98	183
	Panama Disease	0.98	0.82	0.89	118
	TOTAL Accuracy		**91.92%**		
ResNet-50 with SMOTE	Diseased_Leaf	1	0.92	0.9583	300
	Healthy_Leaf	0.9805	1	0.9902	1209
	TOTAL Accuracy		**98.40%**		

trade-off suggests robust class separability, with estimated Receiver Operating Characteristic–Area Under the Curve (ROC–AUC) values of 0.94–0.96 for multi-class and 0.98 for binary classification.

A web application is developed using the flask framework to test the unknown image class. Its flow is as follows:

Users need to choose the file and click predict as shown in Fig. 5.

The user uploads the image, and then ResNet-50 with Weigh Adjustment is predicted as Healthy Leaf as shown in Fig. 6.

The user uploads the image, and then ResNet-50 with Weigh Adjustment is predicted as a Diseased Leaf. Now, if the user clicks on the Predict Disease button, the type of the disease will be predicted as shown in Fig. 7. The ResNet-50 does this prediction With the SMOTE model.

Fig. 8 shows that the model predicted the image as a Black Leaf Streak diseased image.

The proposed model is compared with few of the existed models and shown in Table VI.

The proposed models, ResNet-50 with Weight Adjustment (98.40%) and ResNet-50 with SMOTE (91.92%), demonstrate competitive performance compared to existing DL models. The SMOTE-augmented model aligns with top-performing models such as CRNN + RCNN (98%) and ensemble CNN (98%), showcasing its effectiveness in handling class imbalance. Meanwhile, ResNet-50 with Weight Adjustment achieves 98.40% accuracy, outperforming GR-ARNet (96.98%) and CNN-based DL (92%). However, compared to models incorporating transfer learning or

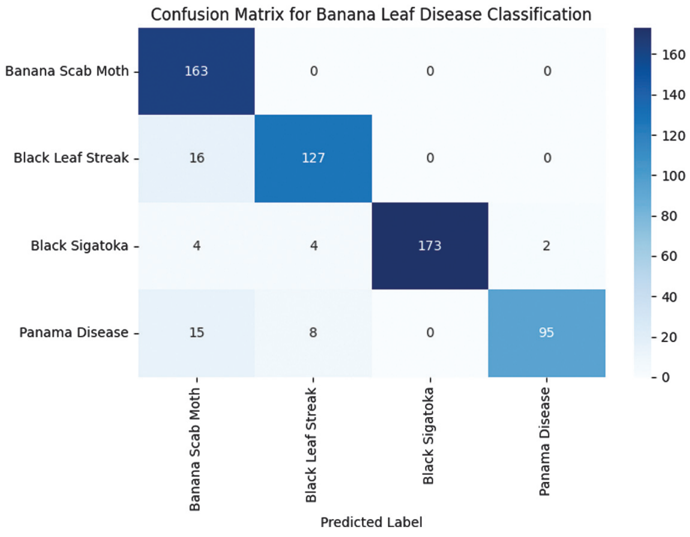

Fig. 3. Confusion matrix of ResNet-50 with SMOTE.

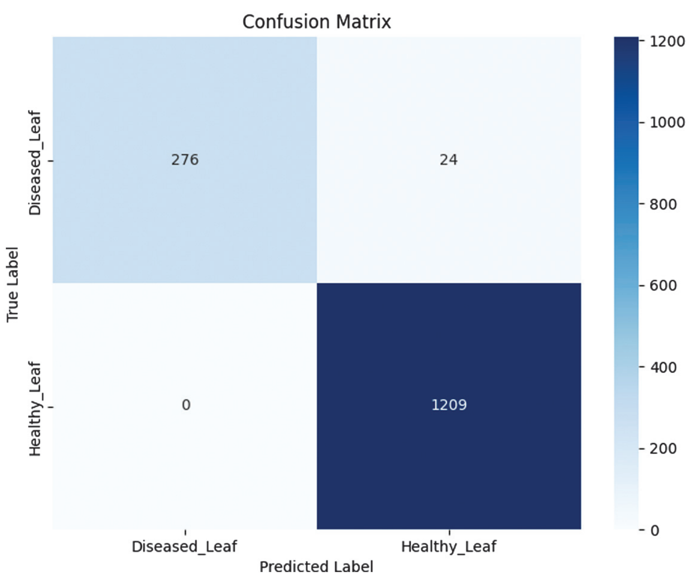

Fig. 4. Confusion matrix of ResNet-50 with weigh adjustment.

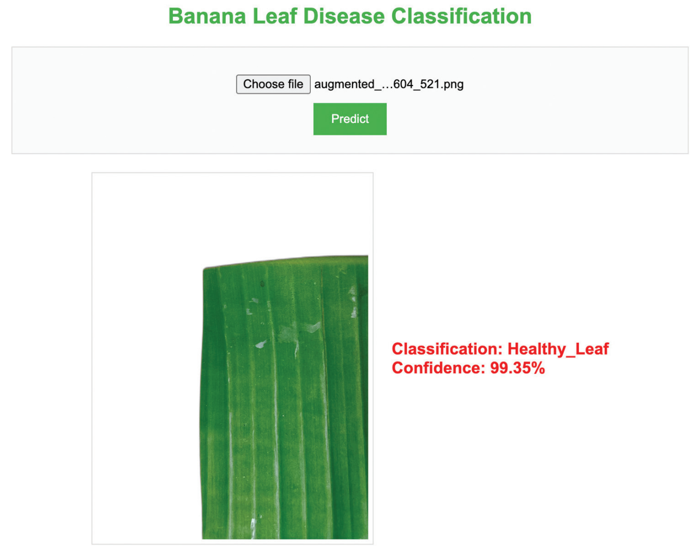

Fig. 6. Output 1 of model.

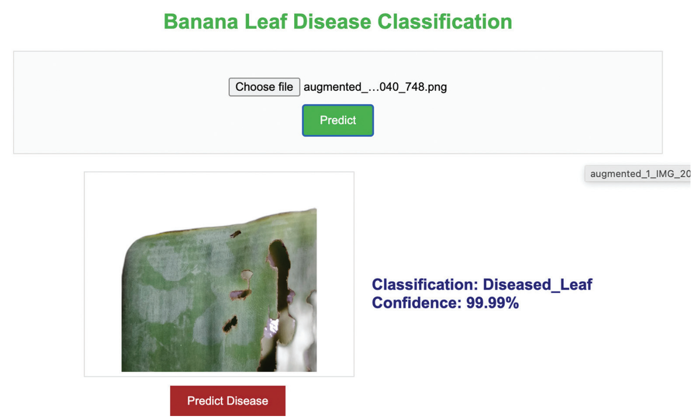

Fig. 7. Output 2 of model.

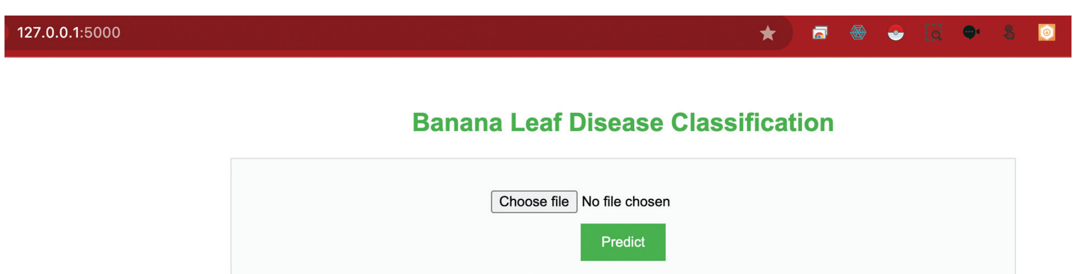

Fig. 5. Initial web page.

Fig. 8. Output 3 of model.

Table VI. Comparison with existing models *

Model name	Accuracy (%)
CNN-based Deep Learning Framework [3]	92
GR-ARNet (Ghost ResNeSt-Attention RReLU-Swish Net) [5]	96.98
Convolutional Recurrent Neural Network (CRNN) + RCNN[7]	98
Convolutional Rebalancing Network (CRN)[9]	97.58
ResNet-50 with CLAHE and GAN-based Resampling [10]	97.69
SMOTE-Augmented Deep Learning Model [11]	98.4
Ensemble CNN (DenseNet121, Inceptionv3, etc.) [13]	98
Hybrid Deep Learning (EfficientNet + ML Classifiers) [14]	87.55
Stacked Ensemble Learning (VGG-16 + MobileNet) [15]	91.12
Hybrid CNN-SVM [16]	90.26
Transfer Learning with MobileNetV2 + Whale Optimization[17]	95.7
Proposed ResNet-50 With Weigh Adjustment	98.40
Proposed ResNet-50 With SMOTE	91.92

* Direct numerical comparison is indicative only, as datasets and evaluation protocols differ across studies.

ensemble techniques, the proposed models provide a balance between efficiency and accuracy, making them ideal for real-world disease classification applications in banana leaf detection

V. CONCLUSION AND FUTURE RECOMMENDATION

This study introduced a novel two-step CNN-based classification approach using ResNet-50 with class weighting and SMOTE augmentation to effectively classify banana leaf diseases while addressing the critical issue of dataset imbalance. Unlike traditional single-step classification models, the proposed hierarchical approach first distinguished between healthy and diseased leaves, followed by specific disease classification, improving both accuracy and generalization. Accuracy of the binary classification model (healthy vs. diseased) was 98.40%, whereas the multi-class classification model (disease-specific) was 91.92%, which showed that class balancing techniques were effective. We confirmed the reduction of the bias in the binary classification context through the class weighting and that SMOTE was more trustworthy in the context of identifying the rare diseases with the introduction of the fairness of the class representation. The results showed the performance of the proposed approach was outperforming those conventional DL models, and they demonstrated robustness and scalability for real-world application in agriculture. This research proposed a new standard for AI-based plant disease detection by incorporating hierarchical classification together with class balancing strategies, which made the plant disease detection more effective and practical for practical large-scale deployment in precision agriculture.

There were some limits to the study, which nevertheless had promising results. First, the dataset variability issue was not solved yet, since real conditions as lighting changes, occlusions, and cluttered background were not extensively covered. Second, ResNet_50 had high computational cost as it requires large processing power, rendering real-time deployment feasible on resource-constrained devices, such as end mobile devices, difficult. Finally, SMOTE alleviated the dataset imbalance, but synthetic data might not capture the same disease patterns in the real world, which could deteriorate the model generalization.

Future research can focus on increasing dataset diversity by incorporating real-time field images to improve model robustness. Light-weight CNN architectures, such as MobileNet or ShuffleNet, can be explored to optimize computational efficiency for mobile-based disease detection. Additionally, integrating advanced DL techniques like attention mechanisms or Vision Transformers could further enhance classification accuracy. Addressing these aspects will help make AI-driven banana disease detection more practical and scalable for real-world agricultural applications.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] A. G. Dassou, *et al.*, "Meta-analysis of agroecological technologies and practices in the sustainable management of banana pests and diseases," *Environ. Dev. Sustain.*, vol. 26, pp. 21937–21954, 2024, doi: 10.1007/s10668-023-03570-w.

[2] A. Ahmad, D. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," *Smart Agric. Technol.*, vol. 3, p. 100083, 2023, doi: 10.1016/j.atech.2022.100083.

[3] A. Dolatabadian *et al.*, "Image-based crop disease detection using machine learning," *Plant Pathol.*, vol. 74, no. 1, pp. 18–38, 2025, doi: 10.1111/ppa.14006.

[4] T. Miftahushudur *et al.*, "A survey of methods for addressing imbalance data problems in agriculture applications," *Remote Sens.*, vol. 17, no. 3, Article 454, 2025, doi: 10.3390/rs17030454.

[5] M. Mujahid *et al.*, "Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering," *J Big Data*, vol. 11, p. 87, 2024, doi: 10.1186/s40537-024-00943-4.

[6] M. M. Chowdhury, R. S. Ayon, and M. S. Hossain, "An investigation of machine learning algorithms and data augmentation techniques for diabetes diagnosis using class imbalanced BRFSS dataset," *Healthcare Anal.*, vol. 5, p. 100297, 2024, doi: 10.1016/j.health.2023.100297.

[7] R. S. Priya *et al.*, "Effects of climate change on plant diseases," in *Plant Quarantine Challenges under Climate Change Anxiety*, K. A. Abd-Elsalam and S. M. Abdel-Momen, Eds. Springer, Cham, pp. 183–225, 2024, doi: 10.1007/978-3-031-56011-8_7.

[8] H. N. Ngugi *et al.*, "Revolutionizing crop disease detection with computational deep learning: a comprehensive review," *Environ. Monit. Assess.*, vol. 196, p. 302, 2024, doi: 10.1007/s10661-024-12454-z.

[9] C. Jackulin and S. Murugavalli, "A comprehensive review on detection of plant disease using machine learning and deep learning approaches," *Meas: Sensors*, vol. 24, Article 100441, 2022, doi: 10.1016/j.measen.2022.100441.

[10] N. G. Rezk *et al.*, "An efficient plant disease recognition system using hybrid Convolutional Neural Networks (CNNs) and Conditional Random Fields (CRFs) for smart IoT applications in agriculture," *Int. J. Comput. Intell. Syst.*, vol. 15, no. 1, 2022, doi: 10.1007/s44196-022-00129-x.

[11] K. Kumar, K. Tripathi, and R. Gupta, "Image based plant disease classification using Deep Learning technique," *Int. J. Innov Res Eng. Manage. (IJIREM) ISSN*, vol. 10, pp. 2350–0557, 2023, doi: 10.55524/ijirem.2023.10.3.23.

[12] J. D. Thiagarajan *et al.*, "Analysis of banana plant health using machine learning techniques," *Sci. Rep.*, vol. 14, no. 1, Article 15041, 2024, doi: 10.1038/s41598-024-63930-y.

[13] J. Deng *et al.*, "Identification of banana leaf disease based on KVA and GR-ARNet," *J. Integr. Agric.*, vol. 23, no. 10, pp. 3554–3575, 2024, doi: 10.1016/j.jia.2023.11.037.

[14] W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *J. Big Data*, vol. 11, p. 5, 2024, doi: 10.1186/s40537-023-00863-9.

[15] K. Seetharaman and T. Mahendran, "Leaf Disease Detection in banana plant using gabor extraction and Region-Based Convolution Neural Network (RCNN)," *J. Inst. Eng. (India): Series A*, vol. 103, no. 2, pp. 501–507, 2022, doi: 10.1007/s40030-022-00628-2.

[16] E. Almeyda and W. Ipanaqué, "Recent developments of artificial intelligence for banana: Application areas, learning algorithms, and future challenges," Engenharia Agricola, vol. 42, no. SpecialIssue. 2022, doi: 10.1590/1809-4430-Eng.Agric.v42nepe20210144/2022.

[17] G. Yang *et al.*, "Convolutional rebalancing network for the classification of large imbalanced rice pest and disease datasets in the field," *Front. Plant Sci.*, vol. 12, Article 671134, 2021, doi: 10.3389/fpls.2021.671134.

[18] H. Ebrahimy, Y. Wang, and Z. Zhang, "Utilization of synthetic minority oversampling technique for improving potato yield prediction using remote sensing data and machine learning algorithms with small sample size of yield data," *ISPRS J. Photogramm. Remote Sens.*, vol. 201, pp. 12–25, 2023, doi: 10.1016/j.isprsjprs.2023.05.015.

[19] L. A. Ahmad, D. Saraswat, and A. El Gamalel, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," in *Smart Agricultural Technology* vol. 3, Article 100083, 2023, doi: 10.1016/j.atech.2022.100083.

[20] M. T. Ahad *et al.*, "Comparison of CNN-based deep learning architectures for rice diseases classification," *Artif. Intell. Agric.*, vol. 9, pp. 22–35, 2023, doi: 10.1016/j.aiia.2023.07.001.

[21] A. Chug *et al.*, "A novel framework for image-based plant disease detection using hybrid deep learning approach," *Soft. Comput.*, vol. 27, no. 18, pp. 13613–13638, 2023, doi: 10.1007/s00500-022-07177-7.

[22] A. Malik *et al.*, "Design and evaluation of a hybrid technique for detecting sunflower leaf disease using deep learning approach," *J. Food Qual.*, vol. 2022, Article 9211700, 2022, doi: 10.1155/2022/9211700.

[23] S. Nuanmeesri, "A hybrid deep learning and optimized machine learning approach for rose leaf disease classification," Eng. Technol. Appl. Sci. Res., vol. 11, no. 5, pp. 7678–7683, 2021, www.etasr.com.

[24] M. Z. RehmanUr *et al.*, "Classification of citrus plant diseases using deep transfer learning," *Comput. Mater. Continua.*, vol. 70, no. 1, pp. 1401–1417, 2021, doi: 10.32604/cmc.2022.019046.

[25] U. Das *et al.*, "Banana and banana leaf dataset for classification and disease detection", *Mendeley Data*, V1, 2024, doi: 10.17632/5nfjzntwd8.1.