Generalizing In-Field Plant Disease Diagnosis: A Deep Transfer Learning Approach for Multi-Crop and Heterogeneous Imaging Conditions

Khoerul Anwar[1*], Tubagus Mohammad Akhriza[2], Mahmud Yunus[3]
Teknologi Informasi, STMIK PKKIA Pradnya Paramiata, Malang, Indonesia[1, 3]
Sistem Informasi, STMIK PKKIA Pradnya Paramiata, Malang, Indonesia[2]

***Abstract*—Plant diseases pose a serious threat to agricultural productivity, which can cause significant crop losses if not addressed quickly and appropriately. There are significant opportunities for digital image-based treatment with computer vision and artificial intelligence. The main challenges in recognizing image-based plant diseases are: developing a single model capable of diagnosing diseases in various types of plants. Ensuring the model remains reliable even when images are taken under varying lighting conditions, backgrounds, and camera quality. In addition, the challenge in this study is to present a model capable of recognizing leaf diseases of multiple food crops, especially rice and corn. The purpose of this study is to identify leaf diseases of rice and corn crops. This study proposes deep learning and transfer learning for diagnosing plant leaf diseases in various types of plants and unstructured imaging environments. To address these challenges, a selection of VGGNet, ResNet50, InceptionV3, and EfficientNetB0 methods was conducted by testing them using laboratory datasets. Based on the testing, the EfficientNetB0 model performed the best. Then, the selected model parameters were tuned, feature extraction and a new dataset was collected in a real-world domain with varying lighting, changing viewpoints and scales, complex backgrounds, similar symptoms between diseases, and occlusion. The results showed that the proposed model performed very well and robustly, with 98% accuracy and a weighted average F1-score of 98% in identifying food crop diseases: blight, rust, blast, blight, tungro, and healthy leaves. This performance indicates that the developed model is highly reliable in classifying leaf diseases in rice and corn. This model is expected to be applied to precision agriculture technology so that farmers can take timely action regarding treatment without further delay.**

***Keywords*—Plant diseases; deep learning; transfer learning; multi-food crops; precision agriculture**

I. Introduction

Global food security is one of the most pressing challenges of the 21st century. Crop diseases pose a significant threat to agricultural productivity, causing significant crop losses if not addressed promptly and appropriately. Traditional diagnostic methods (expert systems) [1] often rely on agronomic expertise, which is time-consuming, expensive, and not always available in remote locations.

Rapid developments in computer vision and deep learning [2][3], particularly in automated image-based plant disease diagnosis have emerged as an attractive alternative [4]. Using cameras on smartphones or drones [5][6], farmers can take pictures of diseased plants and get instant diagnoses. However, training deep learning models from scratch requires a very large dataset of labeled images, which is difficult to obtain for every combination of plant and disease.

Researchers struggling to obtain such a large dataset for multiple crops and diseases can leverage models trained on complex datasets (ImageNet) and adapt their knowledge to detect crop leaf diseases. This is where transfer learning can be used to address the limitations of large datasets. This approach leverages the "knowledge" a model has learned from large-scale image classification tasks (e.g., recognizing thousands of everyday objects from datasets like ImageNet) and applies it to a new task: identifying crop diseases. This drastically reduces the amount of data required and training time, and most importantly, improves model accuracy.

This study will explore how transfer learning can address two major challenges in crop disease diagnosis. [7][8]: Multi-Crop Diversity: developing a single model capable of diagnosing diseases in multiple crop types [9][10]–[12]. Heterogeneous Imaging Environments: Ensures the model remains reliable even when images are taken under different lighting conditions, backgrounds, and camera quality. The existence of a model with these two capabilities will have a significant impact on the field of digital image-based artificial intelligence, particularly for the task of detecting various leaf diseases across multiple food crops.

This study aims to build a diagnostic model that remains reliable and accurate when food crop images are taken under non-ideal (unstructured) or varying conditions, such as variations in lighting, background, and camera quality. These models are expected to contribute to precision agriculture technology, enabling farmers to take timely action to address problems without further delay.

The study is structured as follows: Section II is a literature review relevant to the topic, containing similar published studies on crop disease detection. Section III describes the proposed methodology, including the dataset and the proposed model. The results and discussion are presented, and their performance is compared with previous research in Section IV. Section V concludes the study with successes and future research.

*Corresponding author.

II. Related Work

A. Deep Learning

Deep learning, especially Convolutional Neural Networks (CNNs)[9], [13], [14], has revolutionized the field of image analysis. CNNs are designed to automatically and adaptively learn a hierarchy of features from images, ranging from simple edges and textures to complex shapes and objects. In the agricultural context, CNNs have been successfully applied to a variety of tasks, including weed identification, crop yield estimation, and disease detection.

Models like Inception3 [15], VGG [16], [17], ResNet [18], and EfficientNetB0 have demonstrated advanced performance in a variety of computer vision tasks. This success is driven by their ability to extract relevant features from raw image data without the need for manual feature engineering.

B. Transfer Learning Concept

Transfer learning is a machine learning technique in which a model developed for one task is reused as a starting point for a model for a subsequent task. In computer vision, this can mean reusing a CNN model that has been trained on a large image dataset like ImageNet. The logic is that the early layers of a trained CNN learn to recognize universal features such as edges, corners, color, and texture. These features are also relevant for other tasks, including identifying disease symptoms in plant leaves. Thus, instead of training the entire network from scratch, we can "freeze" these early layers and retrain only the final layers responsible for task-specific classification. This process is called fine-tuning.

The main advantages of transfer learning are: Less Data Requirement: Since the model already has a basic understanding of image features, it requires fewer specific examples of plant diseases to learn. Faster Training Time: Training focuses on only a small part of the network, making the process faster. Better Accuracy: Knowledge transferred from large datasets can help the model to generalize better, especially on small datasets.

C. Heterogeneous Image

While transfer learning is highly effective, its real-world application faces challenges, including: Cross-Crop Generalization: Similar disease symptoms (e.g., yellow spots) can appear on different plants but are caused by different pathogens. Conversely, the same disease can exhibit different symptoms on different plant varieties. The model must be intelligent enough to distinguish between these contexts. Environmental Variability: Images captured in the field vary widely. Factors such as: Lighting: Direct sunlight, shade, or cloudy conditions can change the appearance of the color and texture of symptoms. Background: Soil, other leaves, weeds, or human hands in the background can confuse the model. Image Quality: Different camera resolutions, focus, and shooting angles can affect model performance.

Current research focuses on developing transfer learning-based models that are robust to variation, often using extensive data augmentation techniques and more sophisticated model architectures. This research is expected to contribute to precision agriculture technology, enabling farmers to take timely action regarding treatment without further delay.

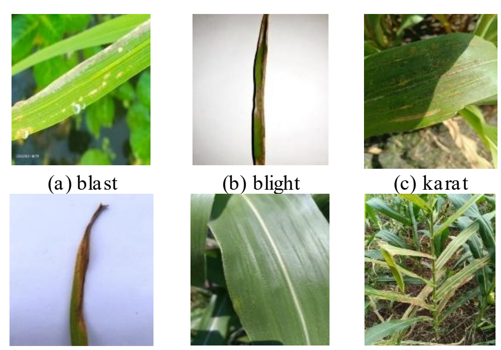
III. Methodology

To build a universal plant disease diagnosis system, this study proposes a transfer learning-based framework consisting of several key stages.

A. Dataset

The dataset, constructed during this crucial stage, consists of images showing leaf diseases in food crops, specifically rice and corn. Both are major commodities produced by farmers and are a source of food. The data sources were obtained from independent field data collection, namely rice leaf images and corn leaf images. The cameras used for field photography were smartphone cameras (Redmi Note 12 Pro 5G, 50 MP, and OPPO A9 2020 edition, 48 MP). The dataset also includes images from previous researchers' repositories available on Kaggle.

This study's novelty is its independently collected dataset, which captures diverse conditions in rice plants, including blast, blight, tungro, and healthy states. Meanwhile, for corn plant types, for types of blight, rust, and healthy diseases. Overall, the dataset consists of six classes. In addition to diversity, the dataset is also designed to have heterogeneity. To ensure heterogeneity, images were taken from three locations (Malang City, Malang Regency, and Pasuruan Regency) to simulate real-world use. To obtain images with different lighting levels, images were taken at three different times, namely: 7 am, 10 pm, and 4 pm. In addition to diverse lighting, the dataset was also collected from various image capture angles and images with complex backgrounds. This dataset design is designed to represent real-world conditions. After collection, the images are pre-processed. Data Cleaning: Removing irrelevant or very poor-quality images. Labeling: Each image was accurately labeled by an agronomist. The labeling of disease types involved experts in the field of food crop leaf diseases from BRMP Malang City. However, in this study, the names of food crop leaf diseases (rice and corn) were addressed as the name of the folder where the dataset was stored, not in each image file for each disease type. Resizing: Adjusting the size of all images to the standard dimensions (224 × 224 pixels) required by the CNN model. An example of a food crop image dataset that will be used for experiments on the food crop disease identification task is shown in Fig. 1.



(a) blast (b) blight (c) karat
(d) tungro (e) sehat (f) hawar

Fig. 1. Example images from the dataset.

Lighting: This dataset design demonstrates good lighting variation. Natural Light (Outdoor): Fig. 1(a), 1(c), 1(e), and 1(f) were taken outdoors in sunlight. There are variations ranging from bright, direct light in Fig. 1(a) to areas with sharp shadows in Fig. 1(c). Artificial Light (Indoor): Fig. 1(b) and 1(d) were taken indoors against a white background. This creates controlled, even lighting conditions, without shadows. Benefits: This variation trains the model to be independent of specific lighting conditions and can recognize diseases in bright sunlight, shade, and indoors.

Image Angle: There is significant diversity in shooting angles. Close-Up and Oblique: Fig. 1(a), 1(c), and 1(e) show the leaf close-up at a slight angle. Top and Flat Lay: Fig. 1(b) and 1(d) show an isolated leaf photographed straight on from above. Wide Angle: Fig. 1(f) is taken from a greater distance, showing multiple leaves in a single clump. Benefits: The model learns to recognize disease symptoms not just from one perspective, but from multiple angles and distances, much like what users will encounter in the field.

Occlusion and Background Complexity: Backgrounds and the presence of occlusions vary significantly. Simple Background: Fig. 1(b) and 1(d) have a clean, white background free of occlusions, allowing the model to focus entirely on the leaf. Complex Background: Fig. 1(a) has a background of other leaves and a water surface. Fig. 1(f) is the most complex, with overlap between the leaf (significant occlusion), the soil, and other surrounding plants. Benefit: This variation trains the model to distinguish between disease symptoms and visual “noise” in the background. This prevents the model from misidentifying shadows or other leaves as diseased.

Pixels, Resolution, and Scale: While difficult to quantify without the original files, the differences in scale are visually noticeable. Macroscale (Close-up): Fig. 1(e) shows the details of the leaf texture and veins very clearly. Medium-scale: Most of the images show a single leaf or part of it. Microscale (Wide): Fig. 1(f) makes each leaf appear smaller and has lower resolution per leaf. Benefits: The model is flexible in recognizing diseases from both detailed close-up photos and long-distance photos that capture the entire plant.

Disease Types and Regions of Interest (ROIs): This dataset includes a wide variety of diseases with very different manifestations (Regions of Interest / ROIs). Small and Scattered: Blast [Fig. 1(a)] and Rust [Fig. 1(c)] appear as small, scattered spots across the leaf surface. Large & Necrotic: Blight [Fig. 1(b)] and Blight [Fig. 1(f)] cover large areas of the leaf, often causing it to dry out and die. Localized at the Tip: Tungro symptoms [Fig. 1(d)] appear concentrated at the leaf tips. No Symptoms: The Healthy class [Fig. 1(e)] serves as an important comparison with no disease ROIs at all. Benefits: This diversity is the essence of training. The model learns to recognize the unique patterns, shapes, colors, and locations of each disease and distinguish them from healthy leaf conditions.

B. Label Inference Automation

In this study, utilizing a subdirectory structure to represent class names is an efficient methodological design in deep learning projects. Its main advantage lies in the automation of label inference directly by the computational framework, eliminating the need for manual data-label mapping. This design inherently offers a structured and intuitive data organization, which is crucial for scalability and management of large-scale datasets. Thus, this approach not only improves efficiency and reduces the potential for human error but also enhances the reproducibility of experiments in image classification.

C. Pre-processing

To improve the model's robustness to image variations, data augmentation techniques are essential. This involves artificially generating modified versions of the training images. Common techniques include: Random rotation = 20, Horizontal flipping = true, Brightness and contrast changes, Zooming, Shearing, width_shift_range = 0.1, height_shift_range = 0.1, and fill mode is nearest.

Data augmentation effectively increases the size of the dataset [19][20] and helps the model learn to focus less on certain orientations or lighting conditions, thereby improving its ability to generalize.

D. Model Selection

Selecting a pre-trained CNN model is an important initial step. In this study, the selection process for several models was carried out as follows:

1) *VGG16:* VGG16 [16] is a Convolutional Neural Network (CNN) architecture that has 16 layers with trainable weights. This architecture is very uniform and simple. Its structure consists of 13 convolutional layers and 3 fully-connected layers. Its main characteristic is the use of very small convolutional filters, namely 3x3, stacked sequentially. These blocks of convolutional layers are interspersed with five max-pooling layers to reduce the spatial dimensionality. Finally, the three fully-connected layers act as classifiers to determine the final output.

2) *ResNet50:* ResNet50 [18] is a 50-layer Convolutional Neural Network (CNN) architecture that introduces a revolutionary concept: residual connections, or skip connections. This architecture not only stacks layers sequentially but also creates "shortcuts" that allow input from previous layers to be added directly to the output of deeper layers. This mechanism overcomes the problem of vanishing gradients in very deep networks, allowing for more efficient training. Its structure consists of a single initial convolutional layer, followed by 16 "residual blocks" (consisting of convolutional layers), and finally with a pooling and fully-connected layer for classification. Using residual connections allows for training very deep networks without the problem of vanishing gradients.

3) InceptionV3: InceptionV3 [15] is a Convolutional Neural Network (CNN) architecture that focuses on computational efficiency without sacrificing accuracy. At the heart of this architecture is an "Inception module" that performs multiple convolution operations (e.g., 1x1, 3x3, 5x5) and pooling in parallel within a single block, allowing the network to capture features at multiple scales. It is computationally efficient by using an "inception module" that performs convolutions at multiple scales. A major update in V3 is convolution factorization, which breaks a large convolution filter (such as 5x5) into a stack of smaller (two 3x3) and asymmetric (1x3 and then 3x1) filters. This step drastically reduces the number of parameters and computational cost, making the network deeper and more efficient.

4) *EfficientNetB0:* EfficientNetB0 [21] is a Convolutional Neural Network (CNN) architecture designed to achieve very high efficiency and accuracy. Instead of randomly changing one dimension of the network (depth, width, or resolution), EfficientNet introduces a compound scaling method. This method intelligently and uniformly balances all three dimensions using predefined scaling coefficients. Its base architecture (B0) is discovered through Neural Architecture Search (NAS) and uses inverted residual blocks (MBConv) similar to MobileNetV2. With balanced scaling, EfficientNetB0 achieves high accuracy with a significantly lower number of parameters and computation (FLOPS) compared to other models.

E. Proposed Model

The architectural modifications proposed in this study aim to improve the accuracy of disease identification. The process steps that will occur during data training are depicted in the framework of Fig. 2.

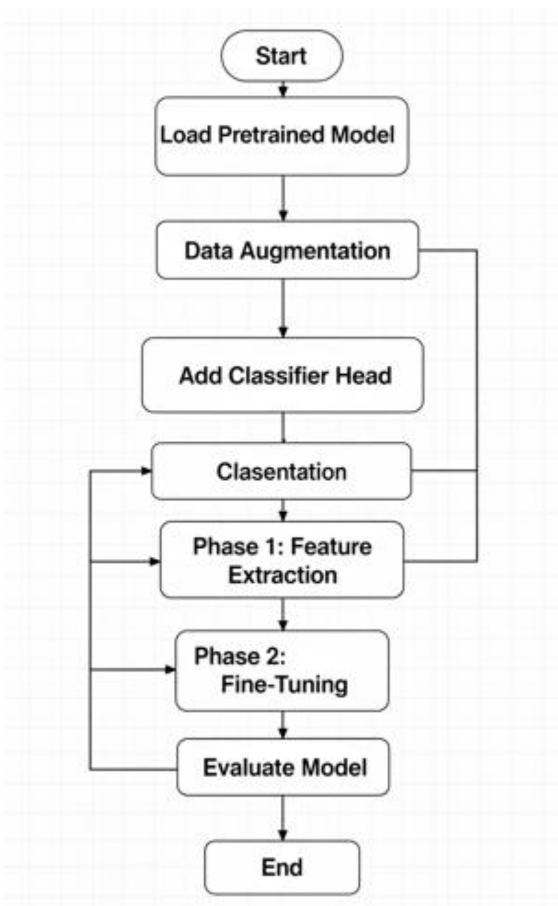


Fig. 2. Proposed method.

1) *Load Pretrained Base Model:* The experimental setup involved importing several convolutional neural networks that had undergone prior training, alongside the weights acquired from the ImageNet dataset. Our investigation specifically incorporated well-known CNN frameworks, namely VGG16, ResNet50, EfficientNetB0, and InceptionV3, each initialized using their respective ImageNet weights. A critical configuration involved disabling the top classification layer (by setting include_top to false), thereby isolating the robust feature extraction capabilities and bypassing ImageNet's original categorization mechanism. This process yielded a collection of highly descriptive feature maps.

2) *Data Augmentation Layer:* The first step will add a data augmentation layer to artificially enrich the training data. This additional layer makes the model more robust and reduces overfitting. Additional augmentation pipelines are Random rotation = 20, Horizontal flipping = true, Brightness and contrast changes, Zooming, Shearing, width_shift_range = 0.1, height_shift_range = 0.1, and fill mode is nearest.

3) *Classifier Head:* Adds an additional dense layer before the output layer with ReLU activation to give the model more capacity to learn complex patterns. Dropout layer (optional, for regularization). Output layer with softmax/sigmoid activation depending on the number of classes.

4) *Two-Stage Training:* Stage 1 (Feature Extraction): Train the classifier head layer by freezing the base model. Freeze all base model layers (trainable=False). Train only the classifier head.

Stage 2 (Fine-Tuning): This layer will "unfreeze" the top layers of the model and retrain the entire model at a very low learning rate to keep the pretrained weights stable but more adaptive to the target dataset. This technique allows the model to adjust the learned features to be more specific to your dataset.

5) *Learning Rate Scheduler:* Uses the ReduceLROnPlateau callback with parameters monitor = "val_loss", factor=0.1 (reduce LR 10x), and patience=3–5 (if validation stagnates for several epochs). To automatically reduce the learning rate when model performance on validation data stagnates.

Evaluate Model: Model performance is evaluated using standard classification metrics, such as: Accuracy: The percentage of correct predictions overall. Precision: Of all positive predictions for a class, how many were correct. Recall (Sensitivity): Of all actual instances of a class, how many were successfully identified. F1-Score: The harmonic mean of precision and recall, providing a balanced measure of performance. Confusion Matrix: A table that visualizes model performance, showing which classes are frequently confused with each other. Evaluation is performed on a test dataset that the model has never seen during training to obtain an estimate of its real-world performance.

IV. Results and Discussion

A. Results

The first experiment was conducted on VGGNet, Resnet50, EfficientNetB0, and InceptionV3 models using homogeneous (laboratory) image data with six disease types and two normal images. A performance comparison of the six models is presented in Table I.

TABLE I. COMPARISON OF INITIAL MODEL PERFORMANCE ON CLASSIFICATION RESULTS WITH LABORATORY DATA

Disease name	VGGNet			ResNet			EfficientNetB0			InceptionV3		
	Precision	Recall	F1- score	Precision	Recall	F1- score	Precision	Recall	F1- score	Precision	Recall	F1- score
Rise leaves	%			%			%			%		
Blast	0.88	0.94	0.91	0.87	1.00	0.93	0.87	1.00	0.93	0.78	0.90	0.84
Blight	0.91	0.63	0.74	0.87	0.65	0.74	1.00	0.75	0.86	0.75	0.75	0.75
Tungro	0.75	0.94	0.83	0.73	0.80	0.76	0.91	1.00	0.95	0.81	0.60	0.72
Healthy	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95	1.00	0.98
Corn leaves												
Hawar	0.90	0.43	0.58	0.67	0.50	0.57	0.64	0.79	0.67	0.80	0.60	0.69
Rust	0.63	1.00	0.77	0.60	0.75	0.67	0.67	0.60	0.63	0.62	0.75	0.68
Healthy	1.00	0.95	0.97	1.00	1.00	1.00	1.00	1.00	1.00	0.90	0.95	0.93
	accuracy	83 %		accuracy	81 %		accuracy	86 %		accuracy	80 %	

Based on the experimental results as written in Table I, it is informed that although classical architectures such as VGG are known to be simple, ResNet50 is effective for deep networks, and InceptionV3 is computationally efficient, EfficientNetB0 shows significant advantages. This model is specifically designed to achieve an optimal balance between accuracy and computational efficiency. With the highest accuracy, EfficientNetB0 is an ideal choice for applications that require high performance on devices with limited resources, such as mobile devices. The training and validation accuracy graph is shown in Fig. 3.

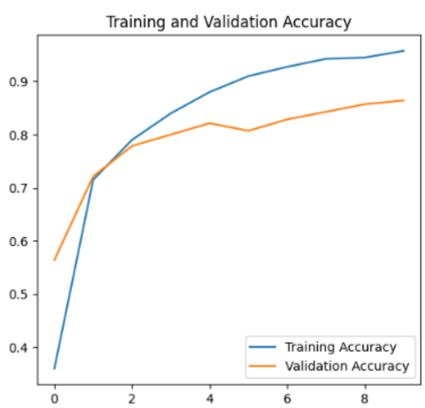


Fig. 3. Graph of accuracy and validation of model training with laboratory data.

```
Classification Report:

              precision    recall  f1-score   support

       Hawar       0.64      0.70      0.67        20
       Karat       0.67      0.60      0.63        20
       Sehat       1.00      1.00      1.00        20
       blast       0.87      1.00      0.93        20
      blight       1.00      0.75      0.86        20
      normal       1.00      1.00      1.00        20
      tungro       0.91      1.00      0.95        20

    accuracy                           0.86       140
   macro avg       0.87      0.86      0.86       140
weighted avg       0.87      0.86      0.86       140
```

Fig. 4. Results of laboratory data image classification.

This classification report analysis (Fig. 4) highlights significant weaknesses in the model's performance, despite an overall accuracy of 86%. The main issue lies in the low F1-scores for the "Blight" (0.67) and "Rust" (0.63) classes.

Specifically, the Recall for "Rust" is only 0.60, meaning that 40% of Rust cases were missed by the model (misclassified as other diseases). The Precision for "Blight" (0.64) is also concerning, indicating a high false positive rate; the model frequently mistakes other diseases for Blight.

Since this dataset is perfectly balanced (Support 20 for each class), this poor performance is not due to imbalance. It strongly indicates a data quality issue: there is likely mislabeling or high visual similarity (ambiguity) between the symptoms "Blight", "Rust", and other classes, which confuses the model.

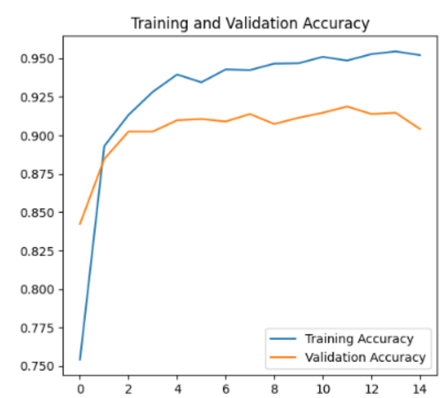


Fig. 5. Accuracy and validation graph of model training with complex images.

The EfficientNetB0 model, which achieved an accuracy of 86%, was selected as the best model. However, there is still room for improvement in accuracy. Therefore, changes were made to increase the complexity and variation of the training data and to avoid data redundancy, especially in healthy image types. The training data was selected for images with high complexity, resembling real-world conditions such as irregular images, varying illumination, images with occlusion, and varying levels of image brightness. After training and validation, the model's performance accuracy increased from 86% to 98% after improvements. The results of data quality improvements on model performance are shown in Fig. 5 as the training graph and Fig. 6 as the classification report.

```
Classification Report:

              precision    recall  f1-score   support

       blast       0.95      0.99      0.97       199
      blight       0.98      0.84      0.91        69
       hawar       0.89      1.00      0.94        25
       karat       1.00      0.97      0.98       225
       sehat       0.91      1.00      0.95        30
      tungro       0.99      1.00      1.00       277

    accuracy                           0.98       825
   macro avg       0.95      0.97      0.96       825
weighted avg       0.98      0.98      0.98       825
```

Fig. 6. The results of the complex data image classification resemble real conditions in the field.

B. Discussion

1) Metric-based performance analysis: Based on the Classification Report of disease classes and health conditions as written in Fig. 6, it is explained as follows:

a) Blast (199 data): Precision reached 0.95: Of all the plants predicted as blast, 95% of them were indeed blast. Recall (0.99): Of all the plants that were actually blast, the model successfully identified 99% of them. F1-Score (0.97): An excellent balance between precision and recall. This indicates very strong model performance for the blast class.

b) Blight (69 data): Precision (0.98): When the model predicted blight, it was correct 98% of the time. This is very high. Recall (0.84): The model only managed to find 84% of all true blight cases. It missed 16% of blight cases (possibly predicted as other classes). This is the lowest point of the model's performance. F1-Score (0.91): While still good, this value is slightly lower than the other classes due to the lower recall value.

c) Hawar (25 data): Precision (0.89): When the model predicted blight, it was correct 89%. Recall (1.00): Perfect! The model successfully identified all plants that were actually blighted. It missed none. F1-Score (0.94): Excellent value.

d) Rust (225 data): Precision (1.00): Perfect! Every time the model predicts rust, it is correct. Recall (97%): The model successfully finds 97% of all rust cases. F1-Score (98%): Excellent performance for this class.

e) Healthy (30 data): Precision (0.91): 91% of healthy predictions were correct. Recall (1.00): Excellent! The model successfully identified all plants that were indeed healthy. This is good because it means the model did not incorrectly label healthy plants as diseased. F1-Score (0.95): Very good.

f) Tungro (277 data): Precision (0.99): Nearly perfect. The model's predictions of tungro are 99% correct. Recall (1.00): Perfect! All cases of tungro were detected. F1-Score (1.00): Perfect performance for the tungro class.

Points for Improvement: The only area of concern is the performance in the blight class. While the precision is high, the recall value (84%) is the lowest. This means the model tends to miss some cases of blight. It may be worthwhile to re-examine the data for the blight class or try other techniques to improve its ability to detect this class. The success in multi-crop diagnosis suggests that the CNN model [21] is able to learn a representation of the underlying disease symptoms (e.g., necrosis, chlorosis) that can be generalized across different crop species. Potential future developments include: Hybrid Models: Combining image information with other data, such as weather or soil type, for more accurate diagnosis. Federated Learning: Training the model on multiple devices (e.g., farmers' mobile phones) without the need to collect data on a central server, to maintain data privacy.

2) Model performance overview: Overall, the model performed very well and robustly. With 98% accuracy and a weighted average F1-score of 98%, the model performed very reliably in classifying the given crop leaf diseases. While these figures indicate good baseline performance, a deeper analysis at the per-class level revealed significant performance variation, suggesting specific challenges in distinguishing between several disease categories (blast, blight, tungro, blight, rust, and health). The dataset appears well-balanced, with an average of 137 samples for each class (support), making the macro average metric a reliable benchmark for average performance without majority class bias. The model's performance was also compared with several single-crop studies, including rice, corn, and similar multi-crop studies (corn and soybean), which yielded quite good results, as shown in Table II.

TABLE II. COMPARISON OF THE PROPOSED METHOD WITH OTHER RESEARCH

Plant Type	Method	Performance %)
Rice [22],	CNN model	84.00
Rice [23],[24]	CNN model	98.86,91.4
Corn [25],[26]	YoloV8 Model, DL	98.00,98,6
Corn [27],[28],[29]	CNN Model	96.30, 84.5,98.3
Multi crop (corn, soybean)[30]	Crop Growth Curve Matching Method	80.00
Multi crop (Rice, corn)	**Proposed method (CNN+transfer learning)**	98.00

The proposed method (CNN + transfer learning) achieves 98.00% accuracy for multi-crop (Rice, Corn) classification. This performance is highly competitive, matching or surpassing many specialized single-crop models, such as those for Rice (which range from 84.00% to 98.86%) and Corn (ranging from 84.5% to 98.6%). Crucially, it significantly outperforms the other listed multi-crop method [30], which only achieved 80.00% using a different technique. This highlights the proposed method's effectiveness and robustness in handling multiple crop types simultaneously with high accuracy.

This 98.00% achievement strongly indicates that the application of transfer learning to the CNN architecture is a determining factor. This technique appears to provide substantial precision improvements compared to standard architectures (such as YoloV8 [25] at 98.00%), optimizing extraction features to accurately distinguish between various types of leaf diseases of rice and corn plants.

V. Conclusion

Plant diseases pose a serious threat to agricultural productivity, potentially causing significant crop losses if not addressed promptly and appropriately. Digital image-based disease management with computer vision and artificial intelligence offers significant opportunities. The main challenges in image-based plant disease recognition are developing a single model capable of diagnosing disease in a wide range of plants and ensuring the model remains reliable even when images are captured under varying lighting conditions, backgrounds, and camera quality.

Transfer learning has emerged as a fundamental methodology in the development of AI-based plant disease diagnosis systems. This study demonstrates the development and validation of an efficient and accurate deep learning model for multi-disease diagnosis in rice and corn crops. Through a comparative evaluation against several leading Convolutional Neural Network (CNN) architectures, the EfficientNetB0 model was shown to exhibit superior performance, achieving an overall accuracy of 98%. This performance significantly outperforms other classic architectures such as VGGNet (83%), ResNet50 (81%), and InceptionV3 (80%), while also confirming its superiority in balancing predictive accuracy with computational efficiency, making it an ideal candidate for implementation on mobile devices in the field.

A more in-depth analysis of the metrics revealed that the model demonstrated excellent capability in identifying healthy plant conditions (F1-score $\approx$ 1.00) as well as diseases with clear visual symptoms, such as blast (F1-score 97%). However, significant challenges were identified in the model's discriminatory ability when faced with diseases with visually overlapping symptomatology. Model performance declined in classes such as blight, which showed a low recall of 84%, the lowest. This indicates that the model tends to miss some blight cases.

These weaknesses are concluded to stem not only from the limitations of the model architecture, but are also fundamentally related to the potential for ambiguity and label inconsistency in the datasets used. Therefore, future research will focus on two strategic avenues: 1) Rigorous dataset curation to improve label quality and consistency as a foundation for more robust training, and 2) Development of hybrid models that integrate visual data with non-image data (e.g., weather and soil type data) to enhance diagnostic context. Further exploration of the Federated Learning paradigm is also proposed as an approach to collaboratively train models without compromising data privacy, ultimately aiming to create more reliable decision support systems for precision agriculture practices.

Acknowledgment

The author would like to thank the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Education and Science, for the research funding support through the 2025 Fundamental Research Grant with contract number 135/LL7/DT.05.00/PL/2025 which is part of DIPA-139.04.1.693320/2025.

References

[1] W. H. Khaled, “an Expert System for Diagnosing Infants’ Diseases Using Clips.,” Int. J. Adv. Res. Comput. Sci., vol. 14, no. 5, pp. 66–72, 2023, doi: 10.26483/ijarcs.v14i5.7025.

[2] A. Behera and S. Goyal, “Plant Disease Detection Using Deep Learning Techniques,” Lect. Notes Networks Syst., vol. 671 LNNS, no. 1, pp. 441–451, 2023, doi: 10.1007/978-3-031-31153-6_35.

[3] P. Chitti, K. Prabhushetty, and S. Allagi, “A Deep Learning and Machine Learning Approach for Image Classification of Tempered Images in Digital Forensic Analysis,” Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 10, pp. 589–593, 2022, doi: 10.14569/IJACSA.2022.0131069.

[4] A. M. S. Kheir, A. Koubaa, V. Kolluru, S. Mungara, and T. Feike, “Smart plant disease diagnosis using multiple deep learning and web application integration,” J. Agric. Food Res., vol. 21, no. April, p. 101948, 2025, doi: 10.1016/j.jafr.2025.101948.

[5] O. Akat, “Drone Use in Agricultural Applications : Advances in Plant Protection through Drone Innovation,” no. January, 2025.

[6] J. Clawson, S. Drake, S. Barstow, and M. L. Pate, “Using Thermal Camera Drones in Beef Cattle Roundup,” no. March, 2025, doi: 10.13140/RG.2.2.20242.98243.

[7] H. T. Vo, K. C. Mui, N. N. Thien, P. P. Tien, and H. L. Le, “Optimizing Grape Leaf Disease Identification Through Transfer Learning and Hyperparameter Tuning,” Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 2, pp. 922–931, 2024, doi: 10.14569/IJACSA.2024.0150293.

[8] M. M. Rahman et al., “Explainable Deep Transfer Learning Framework for Rice Leaf Disease Diagnosis and Classification,” Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 12, pp. 862–884, 2024, doi: 10.14569/IJACSA.2024.0151287.

[9] O. Iparraguirre-Villanueva et al., “Disease Identification in Crop Plants based on Convolutional Neural Networks,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 3, pp. 519–528, 2023, doi: 10.14569/IJACSA.2023.0140360.

[10] M. E. Qureshi, M. H. Ashraf, M. W. Arshad, A. Khan, H. Ali, and Z. U. Abdeen, “Hierarchical Feature Fusion With Inception V3 for Multiclass Plant Disease Classification,” Informatica, vol. 49, no. 27, 2025, doi: 10.31449/inf.v49i27.8208.

[11] K. Sahu, M. K. Singh, M. Niranjan, and P. Joshi, “HACDRAN: A Novel Hybrid Atrous Convolution Model With Densenet and Residual Attenuation Network for Multi-Plant Leaf Disease Classification,” Trans. Emerg. Telecommun. Technol., vol. 36, no. 1, 2025, doi: 10.1002/ett.70046.

[12] Rathiya R, “Multi-Plant Disease Detection and Classification Using Multi-Scale Dynamic Spatial Gating - Squeeze and Excitation (MDSG-SE),” J. Electr. Syst., vol. 20, no. 3, pp. 6755–6764, 2024, doi: 10.52783/jes.7129.

[13] A. A. Gomaa and Y. M. A. El-Latif, “Early Prediction of Plant Diseases using CNN and GANs,” Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 5, pp. 514–519, 2021, doi: 10.14569/IJACSA.2021.0120563.

[14] M. A. Haq, A. Ahmed, and J. Gyani, “Implementation of CNN for Plant Identification using UAV Imagery,” Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 4, pp. 369–378, 2023, doi: 10.14569/IJACSA.2023.0140441.

[15] J. Thomkaew and S. Intakosum, “Improvement Classification Approach in Tomato Leaf Disease using Modified Visual Geometry Group (VGG)-InceptionV3,” Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 12, pp. 362–370, 2022, doi: 10.14569/IJACSA.2022.0131244.

[16] A. A. Alatawi, S. M. Alomani, N. I. Alhawiti, and M. Ayaz, “Plant Disease Detection using AI based VGG-16 Model,” Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 4, pp. 718–727, 2022, doi: 10.14569/IJACSA.2022.0130484.

[17] P. Dinesh and R. Lakshmanan, “Deep Learning-Driven Citrus Disease Detection: A Novel Approach with DeepOverlay L-UNet and VGG-RefineNet: Deep Learning-Driven Citrus Disease Detection,” Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 7, pp. 1023–1041, 2024, doi: 10.14569/IJACSA.2024.01507100.

[18] S. Soni, M. Siroya, R. Shingala, P. Prajapati, M. Ajwalia, and H. Yadav, “A Deep Learning Approach for Early Detection and Classification of Mango Leaf Diseases for Sustainable Agriculture,” Lect. Notes Networks

Syst., vol. 1163, no. 1, pp. 67–74, 2024, doi: 10.1007/978-981-97-8537-7_6.

[19] L. E. Monta, "Assessing the Impact of Data Augmentation on Photovoltaic Module Faults Detection Using Deep Learning Models Assessing the Impact of Data Augmentation on Photovoltaic Module Faults Detection Using Deep Learning Models," no. June, 2025.

[20] M. S. Mayaleh, S. A. Mayaleh, M. S. Mayaleh, S. A. Mayaleh, E. Sentiment, and S. Datasets, "Enhancing Sentiment Classification on Small Datasets through Data Augmentation and Transfer Learning : A Comparative Study To cite this version : HAL Id : hal-05090101 Enhancing Sentiment Classification on Small Datasets through Data Augmentation and Tran," no. June, pp. 0–16, 2025, doi: 10.36227/techrxiv.174970275.53478483/v1.

[21] S. Riyadi, R. Mulya, and A. Nabila Realisti, "Comparison of MobilenetV2 and EfficiennetB3 Method to Classify Diseases on Corn Leaves," E3S Web Conf., vol. 595, pp. 1–11, 2024, doi: 10.1051/e3sconf/202459502006.

[22] M. E. Rana, V. A. Hameed, I. K. Y. Eng, H. K. Tripathy, and S. Mallik, "Harnessing artificial intelligence for sustainable rice leaf disease classification," Front. Plant Sci., vol. 16, no. September, pp. 1–27, 2025, doi: 10.3389/fpls.2025.1594329.

[23] A. Azis, A. Fadlil, and T. Sutikno, "Real-Time Rice Leaf Disease Diagnosis: A Mobile CNN Application with Firebase Integration," J. Tek. Inform., vol. 6, no. 3, pp. 1469–1484, 2025, doi: 10.52436/1.jutif.2025.6.3.4452.

[24] A. K. Abasi, S. N. Makhadmeh, O. A. Alomari, M. Tubishat, and H. J. Mohammed, "Enhancing Rice Leaf Disease Classification: A Customized Convolutional Neural Network Approach," Sustain., vol. 15, no. 20, pp. 1–18, 2023, doi: 10.3390/su152015039.

[25] W. Zhang, S. Wang, F. Wang, and D. Zhu, "Corn leaf disease spot image classification and recognition algorithm based on improved Yolov8," vol. 13562, no. Iccais, p. 130, 2025, doi: 10.1117/12.3062139.

[26] M. Fraiwan, E. Faouri, and N. Khasawneh, "Classification of Corn Diseases from Leaf Images Using Deep Transfer Learning," Plants, vol. 11, no. 20, pp. 1–14, 2022, doi: 10.3390/plants11202668.

[27] M.Chithambarathanu, "Deep Learning for Agriculture: Sweet Corn Disease Detection with Convolutional Neural Networks," J. Inf. Syst. Eng. Manag., vol. 10, no. 19s, pp. 322–340, 2025, doi: 10.52783/jisem.v10i19s.3029.

[28] R. Firmansyah and N. Nafi'iyah, "Identifying Types of Corn Leaf Diseases with Deep Learning," J. Intell. Syst. Comput., vol. 6, no. 1, pp. 18–23, 2024, doi: 10.52985/insyst.v6i1.347.

[29] J. B. Hassan, O. T. Temiloluwa, D. D. Wisdom, and T. D. Ajayi, "Classification of Corn Diseases using Deep Transfer Learning and CNN," 2024) Fjpas, vol. 9, no. 3, pp. 93–108, 2024, doi: 10.55518/fjpas.XBHT4625.

[30] M. S. Pramono and A. P. Wibowo, "Penerapan Convolutional Neural Network Untuk Identifikasi Penyakit Pada Tanaman Padi Dari Citra Daun Menggunakan Model Resnet-101," Djtechno J. Teknol. Inf., vol. 5, no. 3, pp. 415–430, 2024, doi: 10.46576/djtechno.v5i3.5098.