RICE-TDL: Leveraging on Deep Transfer Learning Architecture for Improved Multi-Class Classification of Rice Leaf Diseases

Adetokunbo MacGregor John-Otumu^{1*}, Emmanuel Osaze Oshoiribhor², Nsisong George Ituen³

^{1,3}Department of Information Technology School of Information and Communication Technology Federal University of Technology, Owerri, Nigeria ²Department of Computer Science Faculty of Physical Sciences Ambrose Alli University, Ekpoma, Nigeria

¹adetokunbo.johnotumu@futo.edu.ng, ²emmaohor@aauekpoma.edu.ng, ³nsisongituen001@gmail.com

Abstract: The paper presents a deep transfer learning-based approach for detecting and classifying rice leaf diseases, driven by the need to improve the accuracy and reliability of traditional visual inspection methods, which are often prone to errors due to overlapping disease symptoms common in tropical regions like Nigeria. The methodology involved collecting 200 images of four major rice diseases like bacterial leaf blight, rice blast, brown spot, and false smut collected from Kaggle and expanding the dataset to 1,960 images using standard image augmentation techniques (flipping, rotation, shearing, and zooming) to enhance the dataset's diversity and generalization potential. The dataset was split into 80% training, 10% validation, and 10% testing. Two models were developed: a custom CNN model and a pre-trained VGG-16 model. The custom CNN achieved 68% accuracy on the original dataset, with False Smut Disease recording the best metrics (68.00% accuracy, 66.70% precision, 65.20% recall, and 65.90% F1-score). The VGG-16 model outperformed it with 74.00% accuracy and a 71.40% F1-score for False Smut. When tested on the augmented dataset, the optimized VGG-16 model demonstrated a significantly improved accuracy of 99.55% for False Smut, highlighting its robustness and effectiveness. The system is implemented as a user-friendly web application, enabling farmers to upload images for instant disease diagnosis, thereby offering a practical and scalable solution for enhancing rice disease management in Nigeria.

Keywords: Deep Learning, Rice Diseases, CNN, VGG-16, Classification.

I. INTRODUCTION

Rice (*Oryza sativa*) is one of the world's oldest and most significant crops, first cultivated about 8,000 years ago along China's Yangtze River. It gradually spread across Asia, becoming a staple in many cultures due to innovations like paddy farming, which allowed rice to thrive in varied environments [1][2]. In Africa, *Oryza glaberrima* was domesticated independently around 1,500 years ago, but *Oryza sativa* eventually became the dominant species across the continent [3]. Rice was introduced to the Americas during the transatlantic slave trade, along with African cultivation techniques [4].

The Green Revolution of the mid-20th century marked a turning point, with the development of high-yield varieties like IR8 significantly boosting production. This period also saw the adoption of synthetic fertilizers, pesticides, and modern irrigation, although it raised concerns about biodiversity loss and environmental damage [5].

Today, rice farming benefits from advanced technologies such as GMOs, precision agriculture, and sustainable practices, with global efforts led by organizations like the FAO to enhance food security [6]. The two main rice species are *Oryza sativa* which is the most widely grown and *Oryza* glaberrima, primarily cultivated in West Africa. *Oryza* sativa includes three key subspecies: (a) **Indica** – Grown in tropical areas like India and Southeast Asia, with long, fluffy grains (e.g., Basmati, Jasmine). It thrives in flooded, humid conditions [8].

(b) **Japonica** – Suited to temperate climates such as Japan and Korea. It has short, sticky grains, ideal for dishes like sushi and risotto (e.g., Arborio, Calrose) [9].

(c) **Javanica** – Also called tropical japonica, it grows mainly in Indonesia. It is less common, with larger, drought-resistant grains, important for local food security [10].

Rice is also classified by grain size long (e.g., Basmati), medium (e.g., Arborio), and short (e.g., sticky rice), and by color: white (highly processed), brown (fiber-rich), red, and black (rich in antioxidants) [11][12].

Despite its global importance, rice faces major threats from diseases like bacterial leaf blight, brown spots, and rice blast, which can severely reduce yield and quality [7].



(c) Rice blast (d) Healthy leaf Figure 1: Sample Images of Selected Rice Leaf Diseases [7]

Rice production faces significant challenges from leaf diseases such as blast, bacterial blight, and brown spots, especially in tropical regions where environmental conditions favor pathogen growth. Countries like China, India, Thailand, and Vietnam have experienced outbreaks resulting in notable agricultural losses [6]. In Nigeria, where rice farming is critical for food security and economic stability, these diseases severely impact yields and farmer livelihoods [13].

Currently, disease identification in Nigeria largely relies on visual inspection and manual assessment, which are often inaccurate, time-consuming, and impractical for large-scale operations. Limited access to agricultural experts and diagnostic tools compounds the problem, leading to misidentification and ineffective treatments [7][8]. Diseases like rice blast (*Magnaporthe oryzae*) and bacterial leaf blight (*Xanthomonas oryzae*) are further exacerbated by Nigeria's climate and poor farming practices, contributing to lower-than-average yields [3]. Visual diagnosis is particularly unreliable due to the subtle similarities among different diseases [14], and the lack of expert support, especially in rural areas, hinders timely and accurate interventions [15].

To address these issues, machine learning (ML) has emerged as a powerful tool for improving plant disease detection. ML algorithms can process vast datasets to uncover patterns and symptoms beyond human perception, enabling more accurate and efficient identification [16]. Deep learning, a subset of ML has shown superior performance in image-based disease classification by training models on large datasets to distinguish between various rice leaf diseases [17].

Among deep learning models, Convolutional Neural Networks (CNNs) stand out for their strength in image recognition tasks. CNNs can automatically extract and analyze features such as discoloration, shape, and texture from leaf images, making them particularly effective for diagnosing rice diseases [18]. Globally, CNN-based models have achieved high accuracy, offering a scalable and efficient alternative to traditional methods and presenting a promising solution for Nigerian farmers [19].



Figure 2: CNN Architecture

Despite the demonstrated potential of Convolutional Neural Network (CNN) models for rice disease detection [22, 28, 30–33], their adoption in Nigeria remains limited. Key barriers include inadequate access to technology, lack of comprehensive datasets, and limited computational infrastructure. Nevertheless, integrating deep learning into Nigeria's agricultural sector holds significant potential to transform rice disease management, enhance crop yields, and strengthen food security [20].

While developed countries have successfully implemented machine learning and deep learning for plant disease detection, Nigeria continues to lag behind due to infrastructural and resource constraints. As a result, Nigerian rice farmers are unable to benefit from the accuracy and efficiency that AI-based approaches provide [20]. To bridge this gap, there is a critical need for an intelligent, contextspecific model capable of accurately identifying multiple rice leaf diseases.

This study addresses the identified research gap by employing a Deep Transfer Learning approach using the Visual Geometry Group 16 (VGG16)-based CNN model for effective multi-class classification of rice leaf diseases. The model will be trained and tested on rice disease datasets tailored to the Nigerian agricultural context, aiming to deliver a practical and precise diagnostic tool. This approach is intended to support early disease detection and effective management, thereby improving rice production and contributing to national food security.

The remainder of this paper is structured as follows: Section 2 presents a brief review of related literature; Section 3 outlines the methodology; Section 4 discusses the experimental results; Section 5 offers a quality discussion of the findings; and Section 6 concludes the study.

II. RELATED WORKS

This section reviews key studies that have applied machine learning (ML) and deep learning (DL) techniques to detect and classify rice diseases, each aiming to enhance diagnostic accuracy.

Kabir, Rana, and Roy [21] developed forecasting models in India linking environmental factors like temperature and humidity to disease outbreaks, emphasizing air temperature as a key predictor. Bari et al. [22] employed Faster R-CNN to detect rice diseases such as blast, brown spot, and hispa, achieving over 98% accuracy. Panchami and Vinod [23] introduced a smartphone-based detection method for five rice diseases with 93% accuracy, including control suggestions. Yang et al. [24] used the DHLC-DETR deep learning model to improve detection accuracy by 17.3% on the IDADP dataset. Patil and Kumar [25] applied an ANN model that predicted disease occurrence with 92.15% accuracy, considering weather patterns. Bandara and Mayurathan [26] used color thresholding along with SVM and k-NN, attaining 89.19% accuracy for leaf blast. Deng et al. [27] proposed a smartphone-compatible ensemble model (DenseNet-121, SE-ResNet-50, and ResNeSt-50) that diagnosed six rice diseases with 91% accuracy. In Northern Nigeria, Habib and Nura [28] achieved 99.34% accuracy using a CNN-based model.

Shwetha and Nagarathna [29] found that KNN outperformed SVM for leaf blast prediction. Sony [30] developed a CNN model in R using UCI image data for rice pest and disease identification. Gülmez [31] highlighted the global use of CNNs and the need for hyperparameter tuning. Kulkarni and Shastri [32] created a CNN model with 95% accuracy, effective even in complex image backgrounds. Udayananda et al. [33] emphasized the role of CNNs in addressing rice yield losses due to diseases, discussing various image processing techniques. Li et al. [34] reviewed multi-scale image processing methods combining ML and DL for better rice disease detection. Tejaswini et al. [35] compared ML and DL models, with a 5-layer CNN achieving 78.2% accuracy and outperforming ML methods.

These studies collectively underscore the potential of AI in improving rice disease detection. However, two major gaps remain: (i) many models struggle with real-time detection in environments, reducing uncontrolled field practical application, and (ii) existing systems often lack user-friendly interfaces, making them difficult for farmers with limited technical expertise to use. Addressing these gaps, the proposed VGG-16-based deep learning model aims to improve real-time adaptability and accessibility, thus enhancing disease detection, classification, and practical adoption in Nigerian agriculture.

III. METHODOLOGY

3.1 Mathematical Formulation of the Proposed Model

This section presents the mathematical formulation underlying the proposed rice leaf disease detection model. The model used a VGG-16 inspired deep learning architecture where the input is a digital image of a rice leaf, and the output is a predicted disease class.

1. Image Resizing:

Let the original image be $I \in \mathbb{R}^{H \times W \times C}$, resized to $I_r \in \mathbb{R}^{128 \times 128 \times 3}$ using bilinear interpolation (1)

$$I_{\rm r}({\rm x},{\rm y}) = \sum_{i=0}^{1} \sum_{j=0}^{1} W_{\rm ij}. I_{\rm r}({\rm x}_{\rm i},{\rm y}_{\rm j})$$

Where:

x_i, y_i are integer pixel coordinates surrounding (x, y), $w_{ii} \in [0, 1]$ are weights based on distance to (x, y)

2. Image Normalization:

 $I_n(x, y, c) = \frac{I(x, y, c) - \mu^c}{\sigma c}, \forall c \in \{R, G, B\}$ (2) Where μc and σc are the mean and standard deviation per color channel

3. Data Augmentation Let T: $R^{H \ x \ W \ x \ C} \rightarrow R^{H \ x \ W \ x \ C}$ be an augmented operator applied to image I

(a) Rotation by θ radians $\mathbf{T}_{rot}(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ (3)

(b) Scaling (zoom) by factor s:

$$T_{zoom}(x, y) = (s.x, s.y), s > 0$$
 (4)

(c) Translation (Shifting) by
$$\Delta x$$
, Δy
 $T_{shift}(x, y) = (x + \Delta x, y + \Delta y)$ (5)

(d) Shear

$$T_{Shear}(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} x + \lambda \ y \\ y \end{bmatrix}, \lambda \in \mathbf{R}$$
(6)

(e) Horizontal/Vertical Flip

 $T_{Flip}(x, y) - (W - x - 1, y)$ (Horizontal) (7)

4. Model Representation: Transfer Learning with VGG-16

Let f Θ : $R^{128 \ x \ 128 \ x \ 3} \rightarrow R^d$ be the feature extractor from pretrained VGG-16

 $\mathbf{Q}\phi: \mathbf{R}^{d} \rightarrow \mathbf{R}^{c}$ be the classifier head C = 4 classes

(a) Model Prediction

$$\hat{y} = G\phi(f\Theta(In)) \in \mathbb{R}^{c}$$
(8)

(b) Softmax output

$$\hat{\mathbf{y}}_{i} = \frac{e^{\mathbf{z}i}}{\sum_{j=1}^{c} e^{\mathbf{z}j}}; \forall i \in \{1, ..., c\}$$
(9)

(c) Loss function (Categorical cross-entropy)

$$\mathcal{L}(\Theta + \phi) = -\sum_{i=1}^{c} y_i \log(\hat{y}_i)$$
(10)

(d) Parameters update (Adam Optimizer)

$$\Theta_{t+1} = \Theta_t - \acute{\alpha}. \frac{m_t}{\sqrt{v}t + \epsilon}$$
(11)

5. Evaluation Metrics

| Given TP, FP, TN, FN \in N | |
|------------------------------|------|
| TP+TN | (12) |

$$\frac{TP}{TP} + FN + FP + TN$$
(12)

$$\operatorname{Precision} = \frac{TP + FP}{TP} \tag{13}$$

$$\operatorname{Recall} = \frac{T}{TP + FN} \tag{14}$$

F1-score =
$$2x \frac{(1+existent x Recall)}{(Pr \ ecision + Recall)}$$
 (15)

$$ROC_{AUC} = \int_0^1 TPR \ d(FPR) \tag{16}$$

3.2 Data Gathering and Preprocessing

This section explains how the data was collected and the various steps taken to preprocess it for this study. Table 1 provides a detailed overview of the dataset used in this study.

Table 1: Dataset Source/Description

| Dataset Size | 5MB |
|----------------|---------------------------------------|
| Dataset Name | Rice_diseases |
| Dataset Source | https://www.kaggle.com/datasets/th |
| | egoanpanda/rice-crop-diseases |
| No of Classes | Four (4) classes with 50 images per |
| | class |
| Names of the | Bacteria blight, blast disease, brown |
| classes | spot and false smut disease |
| Comments | Balanced dataset |

Table 2 outlines the data augmentation parameters applied to enhance the diversity of the training dataset and improve the model's generalization.

| Augmentation Parameter | Value | Description |
|---------------------------|---------|--|
| rotation_range | 40 | Randomly rotate images within a range of 40 degrees. |
| width_shift_range | 0.2 | Randomly shift images horizontally by 20% of the width. |
| height_shift_range | 0.2 | Randomly shift images vertically by 20% of the height. |
| shear_range | 0.2 | Shear intensity for the transformation (0.2 radians). |
| zoom_range | 0.2 | Random zoom by 20% to zoom in or out on images. |
| horizontal_flip | TRUE | Randomly flip images horizontally. |
| fill_mode | nearest | Method to fill in newly created pixels after augmentations |

Table 2: Data Augmentation Parameters

3.3 Data Split Ratio

This section describes the data split ratio adopted in this study for effective training, validation, and evaluation of the rice disease classification model.



Figure 3: Dataset split approach

Figure 3 illustrates the dataset partitioning strategy used in this study, where the dataset is divided into training, validation, and testing subsets. The training set, comprising 80% (1,760 samples), is used for learning patterns within the data. The remaining 20% is split equally between the validation set (10%, 220 samples) for tuning model parameters during training, and the test set (10%, 220 samples) for evaluating performance on unseen data. This structured division ensures effective learning, model optimization, and an unbiased assessment of generalization.

3.4 System Architecture

This section outlines the VGG-16 architecture, emphasizing its core components and the essential training parameters applied in model development. Table 3 details the architectural elements of the VGG-16 model implemented in this study.

| Table 3: | VGG-16 | Model | Building | Block | Parameters |
|----------|--------|---------|----------|-------|-------------|
| Lable C. | 100101 | 1110401 | Dunung | DIOCK | i urumetero |

| Layer Type | Details |
|---------------|--|
| Base Model | VGG16 (pre-trained, weights='imagenet') |
| Input Shape | (128, 128, 3) |
| Conv2D Layers | From VGG16 (pre-trained, frozen layers) |
| Flatten Layer | Converts 3D feature maps to 1D |
| Dense Layer 1 | 128 Neurons, Activation = ReLU |
| Dropout Layer | Rate $= 0.5$ |
| Dense Layer 2 | 4 Neurons (for 4 classes), Activation |
| (Output) | = Soltmax |

Table 4 summarizes the training parameters employed for fine-tuning the VGG-16 model in this study.

Table 4: VGG-16 Training Parameters

| Parameter | Value | Description |
|-----------------------|---|---|
| Model Architecture | VGG16 (pre- trained) | Transfer learning with VGG16 model using 'imagenet' weights. |
| Custom Layers | Flatten, Dense (256), Dropout (0.5), Dense (num_classes, softmax) | Added fully connected and dropout layers to customize the output for classification. |
| Optimizer | Adam (learning rate: 0.0001) | Optimizer used for fine- tuning with a low learning rate for stable convergence. |
| Loss Function | Categorical Crossentropy | Best suited for multi-class classification tasks. |
| Metrics | Accuracy | Accuracy is used as the performance evaluation metric during training and validation. |
| Epochs | 30 | The number of training epochs. Early stopping may terminate the training earlier if needed. |
| Batch Size | 32 | The number of samples processed in one iteration. |
| Image Size | (224, 224) | Input image size for the VGG16 model. |
| Validation Split | 0.1 (10%) | 10% of the training data is used for validation during training. |
| Early Stopping | Monitor: 'val_loss', Patience: 5 | Early stopping callback to stop training when the validation loss does not improve for 5 epochs. |



Figure 4: Proposed System Workflow

Figure 4 illustrates the comprehensive workflow for the rice disease classification model, beginning from data collection to final deployment. The process starts with data collection, where 200 image samples were gathered from Kaggle. This is followed by the data preprocessing stage, which involves two key operations: (a) image scaling to ensure uniform input size, and (b) data augmentation techniques such as rotation, flipping, and shearing to increase data diversity and enhance model generalization.

Subsequently, the dataset is split into three segments based on defined ratios: 80% for training, 10% for validation, and 10% for testing. The training data is further divided internally, reserving part of it for validation during model development. The model selection and development phase employs the VGG-16 architecture, a pre-trained convolutional neural network model tailored for image classification tasks.

After model training, the system performs multi-class classification, predicting one of the four rice disease categories while evaluating performance using metrics such as accuracy, precision, and recall. The model evaluation stage assesses how well the model performs on unseen test data. Finally, the best-performing model is integrated into a user-accessible platform through model deployment, utilizing web technologies such as HTML, CSS, Python, and Flask to enable practical and real-time diagnosis functionality.

3.5 Experimental Setup

This section outlines the specifications of the computer system utilized for the development and training of the deep learning model. Table 5 presents the system configuration employed for developing and training the model.
 Table 5: System Specification for Model Development

| | - | | | |
|-----------------------|-----------------------------------|--|--|--|
| Configuration | Parameters | | | |
| CPU | Intel Core i7-8700 | | | |
| GPU | NVIDIA GeForce GTX 1080Ti 16GB | | | |
| | GDDR5X | | | |
| Memory (RAM) | 16GB DDR3 2133MHz | | | |
| Harddisk | 512GB SSD | | | |
| Operating System (OS) | Microsoft Windows 10 | | | |
| Development IDE | Anaconda3, Jupyter Notebook | | | |
| Programming Language | Python 3.9.0 | | | |
| Package/Library | Tensorflow, Keras, Pandas, Numpy, | | | |
| | Flask framework | | | |

IV. EXPERIMENTAL RESULTS

This section presents the results obtained from the experiments conducted.

4.1 Dataset Visualization



Figure 5: Disease Class Distribution in Original Dataset



Figure 6: Disease Class Distribution in Augmented Dataset



Figure 7: Sample Augmented Dataset by Disease Class

| Table 6: Original and Augmented Dataset Distributed by Disease |
|--|
| Class |

| Datasets | Bacteria Leaf Blight | Rice Blast | Brown Spot | Smut Disease |
|-----------|----------------------------|---------------|---------------|-----------------|
| Original | 50 | 50 | 50 | 50 |
| Augmented | 490 | 490 | 493 | 489 |

4.2 Custom-CNN Model Results



Figure 8: Custom CNN Model Training Accuracy across 30 Epochs Using Original Dataset



Figure 9: Custom CNN Model Training Loss across 30 Epochs Using Original Dataset



Figure 10: Custom CNN Model ROC_AUC Curves by Disease Class Using Original Dataset

Table 7: CNN Classification Results (Original Dataset)

| Class | Accuracy | Precision | Recall | F1-Score |
|--------------------------------|----------|-----------|--------|----------|
| Bacterial Blight Disease | 63.00% | 60.20% | 58.50% | 59.30% |
| Blast Disease | 66.50% | 64.30% | 61.70% | 63.00% |
| Brown Spot Disease | 60.20% | 55.80% | 54.30% | 55.00% |
| False Smut Disease | 68.00% | 66.70% | 65.20% | 65.90% |





Figure 11: VGG-16 Training Accuracy across 30 Epochs



Figure 12: VGG-16 Training Loss across 30 Epochs



Figure 13: VGG-16 ROC_AUC Curves by Disease Class

| Class | Accuracy | Precision | Recall | F1-Score |
|---------------------|----------|-----------|--------|----------|
| Bacterial Blight | 70.20% | 68.50% | 66.70% | 67.60% |
| Blast | 73.10% | 70.40% | 71.20% | 70.80% |
| Brown Spot | 67.50% | 65.80% | 64.90% | 65.30% |
| False Smut | 74.00% | 72.10% | 70.80% | 71.40% |

Table 8: VGG-16 Classification Results (Original Dataset)

4.4 VGG-16 Transfer Learning Model (Augmented Dataset)



Figure 14: VGG-16 Training Accuracy across 30 Epochs using Augmented Dataset



Figure 15: VGG-16 Training Loss across 30 Epochs using Augmented Dataset



Figure 16: VGG-16 Model ROC_AUC Curves by Disease Class using Augmented Dataset

| Table 9: VGG-16 Classification Result (Augmented) | Dataset + |
|--|-----------|
| Hyper-parameter Optimization) | |

| Class | Accuracy | Precision | Recall | F1-Score |
|---------------------|----------|-----------|--------|----------|
| Bacterial Blight | 99.50 | 98.75 | 98.90 | 98.82 |
| Blast | 99.30 | 99.20 | 98.65 | 98.92 |
| Brown Spot | 99.40 | 99.55 | 98.95 | 99.25 |
| False Smut | 99.55 | 98.90 | 99.10 | 99.00 |

4.5 Comparison Evaluation of the Deep Learning Models

Table 10: Model Training Results over 30 Epochs

| Model/Dataset | Optimizer | Learning Rate | Accuracy | Loss |
|----------------------------------|-----------|------------------|----------|--------|
| CNN (Original dataset) | Adam | 0.0001 | 0.9125 | 0.2989 |
| VGG-16 (Original dataset) | Adam | 0.0001 | 0.9435 | 0.1652 |
| VGG-16 (Augmented dataset) | Adam | 0.0001 | 0.9750 | 0.0915 |

| DL Model | Class | Accuracy | Precision | Recall | F1- Score |
|------------------------------------|--------------------------|----------|-----------|--------|--------------|
| CNN Original dataset | False Smut Disease | 68.00 | 66.70 | 65.20 | 65.90 |
| VGG-16 Original dataset | Blast Disease | 73.10 | 70.40 | 71.20 | 70.80 |
| VGG-16 Augment ed dataset | Brown Spot Disease | 99.40 | 99.55 | 98.95 | 99.25 |

 Table 11: Comparison Evaluation of Deep Learning Model

 Performance across Disease Classes



Figure 17: Comparison Evaluation with Previous Studies

V. DISCUSSION

This section offers an in-depth analysis of the proposed study's results, drawing comparisons with related works in terms of dataset characteristics, data preprocessing, model classification performance, model generalization capability, and emerging research directions in rice disease detection and classification.

5.1 Comparison with Previous Studies

Numerous previous works have focused on detecting rice diseases using AI techniques, each with varying datasets, methodologies, and performance results. For instance, Kabir et al. [21] explored environmental predictors like temperature and humidity, showing strong relevance in disease forecasting. Unlike this study, which relied solely on visual features through image datasets, their approach did not leverage image-based CNN models.

Studies like Bari et al. [22], Yang et al. [24], and Deng et al. [27] demonstrated high accuracy using advanced CNN architectures like Faster R-CNN, DHLC-DETR, and ensemble models, respectively. Notably, Bari et al. reported over 98% accuracy, while Deng et al. achieved 91% accuracy by combining three deep learning models. These results surpass the current study's performance using both the custom CNN and VGG-16 models on the original dataset. A contributing factor could be the nature and size of datasets used as most previous works utilized extensive, diverse, and expertly curated datasets, which contrasts with the moderate-sized dataset used in this study.

In contrast, Shwetha and Nagarathna [29] and Tejaswini et al. [35] explored machine learning techniques such as KNN and SVM, achieving varied accuracies. While simpler models showed moderate success, most studies, including this one, support the growing trend that deep learning approaches generally outperform traditional ML models, especially in image-based tasks.

5.2 Dataset Characteristics and Preprocessing Impact

Figures 5 and 6 revealed a well-balanced distribution of disease classes in both the original and augmented datasets, which is essential for unbiased model training. The augmentation process increased the dataset tenfold while maintaining balance, supporting model robustness and reducing the risk of overfitting due to data scarcity.

However, despite these efforts, Figures 8 through 10 indicate that the custom CNN model struggled with generalization, as shown by rising validation loss and poor ROC-AUC scores (ranging from 0.42 to 0.60). This implies that although the dataset was balanced, its size and complexity may not have been sufficient to train a deep CNN effectively without overfitting.

5.3 Model Performance and Generalization

The custom CNN model achieved a peak training accuracy of 67%, while validation accuracy plateaued at 25%, with poor AUC values across all disease classes. These metrics suggest significant overfitting, despite balanced classes and augmentation. Moreover, class-wise performance (Table 7) showed moderate results, with False Smut Disease being the most accurately predicted class.

In comparison, the VGG-16 model (Figures 11–13 and Table 8) showed better generalization, achieving over 90% training accuracy and ~80% validation accuracy, with lower training and validation losses. However, even VGG-16 exhibited limited class separability, with AUC values still close to 0.5, highlighting the challenge of inter-class similarity in rice disease images or potential limitations in image resolution or labeling accuracy.

5.4 Emerging Trends and Implications

This study aligns with current trends that emphasize the importance of deep transfer learning, as evidenced by the superior performance of VGG-16 over the custom-built CNN. Additionally, the integration of data augmentation and balanced class distribution supports robust model training. Yet, the low ROC-AUC values across models suggest that visual similarities among rice disease classes remain a major obstacle in automated classification.

VI. CONCLUSION

This research presents an efficient deep transfer learning model for multi-class classification of rice leaf diseases, leveraging the capabilities of the pre-trained VGG-16 architecture and comparing its performance with a custom-built Convolutional Neural Network (CNN). Through the use of data augmentation techniques, the study successfully addressed the limitations posed by a small dataset, enhancing the model's ability to learn diverse patterns.

Experimental results demonstrated that the VGG-16 model significantly outperformed the custom CNN across all key evaluation metrics, including training and validation accuracy, thereby affirming the effectiveness of transfer learning in scenarios with limited labeled data. However, both models exhibited varying degrees of overfitting, particularly the custom CNN, which struggled with generalization, highlighting the critical importance of larger and more diverse datasets, optimized hyperparameter configurations, and advanced regularization techniques.

Despite these limitations, the study contributes valuable insights into the application of artificial intelligence in precision agriculture. The proposed models lay a foundational framework for developing intelligent, automated disease detection systems that can assist farmers and agricultural experts in monitoring crop health more efficiently, especially in resource-constrained environments. This research underscores the potential of deep learning to transform traditional agricultural practices, offering scalable and timely solutions for managing crop diseases and improving food security.

VII. RECOMMENDATIONS AND FUTURE WORKS

Based on the study's findings and limitations, several future directions are recommended. First, adopting advanced hyperparameter tuning methods like grid search or Bayesian optimization can improve model accuracy and minimize overfitting. Expanding the dataset with real-world images from diverse regions will enhance generalizability.

Integrating hybrid or ensemble models may boost robustness and detection accuracy. Incorporating environmental data such as weather and soil conditions can enable more contextaware predictions. Including Explainable AI (XAI) techniques like Grad-CAM or SHAP will make model outputs more interpretable for users. Developing lightweight models through pruning or quantization is essential for realtime mobile deployment.

Lastly, future research should explore multi-label classification to handle cases where leaves show symptoms of multiple diseases. These steps will help create scalable, accurate, and farmer-friendly AI solutions for smart agriculture.

REFERENCES

- A. Smith et al., "Global rice production and food security," Food Sec. Rev., vol. 8, no. 1, pp. 15–23, 2021.
- [2] J. Doe, "Impact of leaf diseases on rice production," Agric. J., vol. 17, no. 2, pp. 10–18, 2022.
- [3] S. F. Chen et al., "Barriers to effective disease management in rice," Crop Sci., vol. 56, no. 2, pp. 679– 688, 2016.
- [4] H. A. Smith and B. C. Jones, "Machine learning applications in agriculture," AI in Agriculture, vol. 2, pp. 35–42, 2020.
- [5] T. Q. Nguyen et al., "Deep learning for plant disease classification," Comput. Electron. Agric., vol. 162, pp. 484–491, 2019
- [6] R. B. Jones, "Rice farming in Nigeria: Current challenges," Niger. Agric. J., vol. 12, no. 4, pp. 34–40, 2023.
- [7] M. C. Green, "Bacterial leaf blight in rice: Challenges and management," Plant Dis., vol. 99, no. 9, pp. 1058– 1072, 2015.
- [8] P. L. White and Q. T. Tran, "Traditional rice disease identification methods," Int. J. Agric. Sci., vol. 14, no. 3, pp. 75–80, 2019.

- [9] D. L. Ho et al., "Indica rice cultivation and its characteristics," Rice Sci., vol. 25, no. 3, pp. 146–156, 2021.
- [10] M. E. Ling et al., "Japonica rice varieties and their cooking qualities," Food Chem., vol. 134, pp. 452–457, 2016.
- [11] F. C. Santos et al., "Understanding the characteristics of javanica rice," Plant Breed., vol. 137, no. 6, pp. 802– 810, 2018.
- [12] H. N. Odong et al., "Grain size classification in rice," Agron. J., vol. 98, no. 4, pp. 1156–1164, 2020.
- [13] R. Kumar, "AI-based solutions for rice disease detection," Comput. Sci. Rev., vol. 39, pp. 100-107, 2021.
- [14] D. A. Ali and F. A. Hossain, "Challenges in adopting AI for agriculture in Nigeria," J. Agric. Sci., vol. 28, no. 1, pp. 90–98, 2023.
- [15] K. L. Roberts et al., "Economic impact of rice diseases," Eur. J. Agron., vol. 60, pp. 33–41, 2019.
- [16] M. B. Anuar et al., "History of rice cultivation and its global significance," Asian Agric. Hist., vol. 12, no. 3, pp. 215–225, 2021.
- [17] R. S. Srivastava, "Rice in Indian history," J. Indian Hist., vol. 10, pp. 25–36, 2020.
- [18] T. A. Kamara et al., "The domestication of Oryza glaberrima in West Africa," African Crop Sci. J., vol. 26, no. 4, pp. 275–284, 2018.
- [19] D. C. Naylor et al., "Rice cultivation in the Americas: A historical overview," Latin Am. Stud. Assoc. J., vol. 22, pp. 95–109, 2021.
- [20] K. H. Chen, "The Green Revolution and its impact on rice production," Food Policy, vol. 44, pp. 221–228, 2015.
- [21] M. Kabir, M. Rana, and A. Roy, "Insight of rice disease forecasting models," The Pharma Innovation Journal, vol. 10, no. 6, pp. 1060-1069, 2021.
- [22] M. Bari et al., "A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework," PeerJ Computer Science, vol. 7, no. 5, p. e432, 2021. doi: 10.7717/peerj-cs.432.
- [23] R. Panchami and S. S. Vinod-Chandra, "Rice Leaf Disease Detection and Diagnosis Using Convolution Neural Network," Research Square, Preprint, 22 July 2022. [Online]. Available: https://doi.org/10.21203/rs.3.rs-1812823/v1.
- [24] H. Yang, X. Deng, H. Shen, Q. Lei, S. Zhang, and N. Liu, "Disease Detection and Identification of Rice Leaf Based on Improved Detection Transformer," Agriculture, vol. 13, no. 7, p. 1361, 2023. doi: 10.3390/agriculture13071361.
- [25] R. R. Patil and S. Kumar, "Predicting rice diseases across diverse agro-meteorological conditions using an artificial intelligence approach," PeerJ Computer Science, vol. 7, p. e687, 2021. doi: 10.7717/peerjcs.687.
- [26] D. Bandara and B. Mayurathan, "Detection and Classification of Rice Plant Diseases using Image Processing Techniques," presented at the International Conference on Advanced Research in Computing, Belihuloya, Sri Lanka, 2023.
- [27] R. Deng, M. Tao, H. Xing, X. Yang, C. Liu, K. Liao, and L. Qi, "Automatic Diagnosis of Rice Diseases Using Deep Learning," Frontiers in Plant Science, vol. 12, p. 701038, 2021. doi: 10.3389/fpls.2021.701038.

- [28] M. S. Habib and B. M. Nura, "Improving Rice Production by Detecting Diseases Using IoT in North West Nigeria," International Journal of Advanced Academic Research, vol. 7, no. 10, pp. 26-38, 2021.
- [29] M. K. Swetha and Nagarathna, "Leaf Blast Rice Disease Prediction Model Based on Environmental Factors Using Binary Classifiers," International Journal of Computer Science and Information Technologies (IJCSIT), vol. 12, no. 1, pp. 18-23, 2021.
- [30] A. Sony, "Prediction of Rice Diseases Using Convolutional Neural Network (in Rstudio)," International Journal of Innovative Science and Research Technology, vol. 4, no. 12, pp. 595-602, 2019.
- [31]B. Gülmez, "Advancements in rice disease detection through convolutional neural networks: A comprehensive review," Heliyon, vol. 10, no. 1, p. e33328, 2024. doi: 10.1016/j.heliyon.2024.e33328.
- [32] P. Kulkarni and S. Shastri, "Rice Leaf Diseases Detection Using Machine Learning," Journal of Scientific Research and Technology (JSRT), vol. 1, no. 10, pp. 17-22, 2024.
- [33] G. K. V. L. Udayananda, C. Shyalika, and P. P. N. V. Kumara, "Rice plant disease diagnosing using machine learning techniques: a comprehensive review," SN Applied Sciences, vol. 4, p. 311, 2022. doi: 10.1007/s42452-022-05194-7.
- [34] R. Li, S. Chen, H. Matsumoto, M. Gouda, Y. Gafforov, M. Wang, and Y. Liu, "Predicting rice diseases using advanced technologies at different scales: present status and future perspectives," aBIOTECH, vol. 4, pp. 359– 371, 2023. doi: 10.1007/s42994-023-00126-4.
- [35] P. Tejaswini, P. Singh, M. Ramchandani, Y. K. Rathore, and R. R. Janghel, "Rice Leaf Disease Classification Using CNN," IOP Conference Series: Earth and Environmental Science, vol. 1032, no. 1, p. 012017, 2022. doi: 10.1088/1755-1315/1032/1/012017.

Author Profile



Adetokunbo MacGregor John-Otumu is a Lecturer and Researcher in the Department of Information Technology at the Federal University of Technology Owerri (FUTO). He holds a Ph.D. in Computer Science (Artificial Intelligence) from Ebonyi State University and master's degrees in Computer Science and Information Technology. A former

Postdoctoral Research Fellow at Morgan State University, USA (2021–2022), he is actively involved in global research collaborations and leads the Machine Learning Research Group at FUTO. Dr. John-Otumu is a member of several professional bodies, including NCS, CPN, IEEE, ACM, AAAI, and the Internet Society. In 2024, he was a Keynote Speaker at the NCS Climate Change Summit (Imo State) and served as a Technical Committee Member and Session Chair at FUTO's first International ICT Conference. He also sits on editorial boards of various academic journals.



Emmanuel Osaze Oshoiribhor is a Senior Lecturer in the Department of Computer Science at Ambrose Alli University. He earned his PhD in Computer Science from Ambrose Alli University in 2017, following his Master of Science and Bachelor of Science degrees from the University of Benin. As an active member of several professional organizations, including the Nigeria Computer Society (NCS), he has

contributed to the academic community through participation in multiple international conferences. His research interests focus on Investigative Data Mining, Data Science, and Machine Learning.



Nsisong George Ituen is a postgraduate student in the Department of Information Technology, School of Information and Communication Technology, Federal University of Technology Owerri (FUTO). His research interests focus on Machine Learning and Deep Learning.