Journal of Basic Physical Research ISSN: 2141-8403 Prints, 2141-8411 Online
Vol.15, Issue 1, March, 2026, pp. 198– 204, Available online at https://journals.unizik.edu.ng/jbpr/

A DEEP LEARNING ENABLED FRAMEWORK FOR AUTOMATED CASSAVA CROP DISEASE DETECTION

[1]Blessing N. Iduh, [2]Rotimi O. Folahan and [3]Nonye A. Ezeorachi
[1,2&3]Department of Computer Science, Faculty of Physical Science,
Nnamdi Azikiwe University, Anambra State, Nigeria.
Email: [1]bn.iduh@unizik.edu.ng,[2]folahanr@gmail.com & [3]angelaezeorachi123@gmail.com

Abstract
Cassava is a major staple crop in sub-Saharan Africa and plays a crucial role in food security and rural livelihoods, but its productivity is significantly affected by plant diseases that are often difficult to detect at early stages using traditional visual inspection methods. These conventional approaches are typically slow, subjective, and unreliable, especially in resource-limited farming environments. This study presents the development of a deep learning-based cassava crop disease detection system aimed at improving early diagnosis and supporting effective cassava crop management. The system is built using a Convolutional Neural Network (CNN) enhanced with transfer learning to enable efficient extraction of discriminative features from cassava leaf images. A hybrid development methodology combining Agile, DevOps, and Rapid Application Development (RAD) principles was adopted to support iterative development, continuous testing, and user-centered improvements involving agronomists and farmers. The model was trained using a combined dataset of field-collected images and publicly available plant disease repositories covering five classes: Cassava Mosaic Disease (CMD), which stands for Cassava Mosaic Disease, Cassava Brown Streak Disease (CBSD), which stands for Cassava Brown Streak Disease, Cassava Green Mite (CGM), Cassava Bacterial Blight (CBB), and healthy cassava leaves. To enhance robustness under real-world conditions, preprocessing techniques such as resizing and normalization were applied, alongside data augmentation methods including rotation, flipping, and brightness adjustments. The system was evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis, achieving an overall accuracy of 75%. Strong performance was observed in detecting CMD and CBSD, while lower performance occurred in distinguishing healthy leaves and early-stage infections due to visual similarity between classes. Overall, the results demonstrate that CNN-based approaches are effective for cassava disease detection and can support farmers in early diagnosis, reduce crop losses, and improve agricultural productivity.
Keywords *- Crop disease detection, Deep learning, Convolutional Neural Networks, Precision agriculture, Food security, Cassava Plant.*

Introduction

Crop diseases remain a critical challenge to global agriculture, significantly affecting food production and economic stability. Estimates suggest that over 30% of global crop yield is lost annually due to pathogens and pests, leading not only to food insecurity but also to increased reliance on chemical inputs that may negatively impact the environment (Gai & Wang, 2024). In sub-Saharan Africa, crops such as cassava and maize are essential for daily sustenance and economic activities, yet they are highly vulnerable to diseases like Cassava Mosaic Disease (CMD), Cassava Brown Streak Disease (CBSD), maize blight, and rust (Umeh et al., 2025). Traditional methods of disease diagnosis, which rely on expert observation and laboratory analysis, are often slow, subjective, and inaccessible to rural farmers. Recent studies have emphasized the need for technology-driven solutions to improve early detection and reduce crop losses (Iduh et al., 2024).

Advances in deep learning, particularly Convolutional Neural Networks (CNNs), have opened new possibilities for automated plant disease detection. CNNs have shown remarkable success in image recognition tasks, motivating their application in agriculture. For instance, a wellknown study demonstrated high accuracy using a large dataset of labeled plant images, although performance declined significantly when tested on real-world images due to environmental variability (Umeh et al., 2025). Further research improved real-world applicability by using field-collected data and transfer learning techniques, achieving strong results for cassava disease detection and demonstrating the feasibility of mobile-based diagnostic systems (Paul et al., 2025).

Related Works

Deep learning has significantly advanced the field of plant disease detection, particularly through the use of convolutional neural networks (CNNs). Early research established the potential of these methods by training models on large, well-structured datasets such as PlantVillage, where very high accuracy was achieved under controlled conditions (Mohanty et al., 2016). However, when these models were applied to real-world images, their performance dropped considerably, revealing limitations in generalization (Mohanty et al., 2016). Subsequent studies extended this work by focusing on specific crops like cassava, where researchers applied transfer learning techniques to field images collected in regions such as Tanzania, achieving strong classification performance for diseases like cassava mosaic and brown streak (Ramcharan et al., 2017). Further developments included mobile-based implementations, where disease detection systems were deployed on smartphones; although practical, these systems experienced reduced accuracy in real farming environments due to factors such as lighting variations and differences in symptom severity (Egba et al., 2024).

In maize disease detection, similar progress has been observed with the development of improved CNN architectures and training strategies. For example, enhanced network designs incorporating multi-activation functions and data augmentation have achieved high classification accuracy for diseases such as rust and blight (Zhang et al., 2021). In addition, review-based studies show that CNN models consistently perform well when supported by high-quality labeled datasets and proper preprocessing methods, often reporting accuracy levels above 90% (Gülmez et al., 2024). These works also emphasize the importance of techniques such as image augmentation, spectral analysis, and careful hyperparameter tuning in improving model robustness (Gülmez et al., 2024). More recent innovations have explored hybrid approaches that combine CNNs with transformer-based architectures to better capture complex patterns in field conditions, although such models are still in early stages of evaluation (Xu et al., 2026).

Despite these achievements, several challenges remain across the literature. Many existing models rely on datasets collected under controlled environments or within limited geographic regions, which restricts their applicability in diverse agricultural settings. As a result, performance often declines when models encounter new conditions, such as different crop varieties, backgrounds, or lighting scenarios (Mohanty et al., 2016; Ramcharan et al., 2019). Broader discussions also highlight the significant impact of plant diseases on global crop production, with estimates suggesting substantial yield losses and the urgent need for early detection systems (Gai & Wang, 2024). Comparative studies of lightweight deep learning models for mobile deployment indicate that some architectures, such as EfficientNet, outperform others like MobileNetV2 in accuracy, although many of these evaluations are still based on controlled datasets and lack comprehensive field validation (Abubaker, 2024). Overall, current research underscores the need for more diverse datasets, improved generalization techniques, and practical deployment strategies to ensure reliable performance in real-world farming environments.

Materials And Methods

A hybrid Agile DevOps methodology integrated with Rapid Application Development (RAD) principles was adopted for the system development. The development process was carried out in iterative sprints, allowing continuous refinement of both system functionality and user experience. Key stakeholders, including agronomists and local farmers, were actively involved throughout the development cycle. Their feedback was particularly valuable in refining the user interface and defining system requirements. Based on their input, priority was given to features such as offline functionality, multilingual support, and simplified navigation to enhance usability in rural and low-connectivity environments. To ensure system reliability and rapid iteration, Continuous Integration and Continuous Testing (CI/CT) pipelines were implemented. These pipelines enabled frequent validation of updates to both the Convolutional Neural Network (CNN) model and the mobile application. This iterative, user-centered approach contributed significantly to the progressive improvement of the system, resulting in a robust and practical mobile-based crop disease detection tool.

a) Test Data

The test dataset used for system evaluation consisted of real plant leaf images extracted from the curated cassava and maize disease datasets. These images were strictly excluded from the training phase to ensure unbiased performance evaluation and to accurately measure the generalization capability of the model. The test dataset comprised 800 leaf images distributed across five categories: Cassava Bacterial Blight (173 images), Cassava Brown Streak Disease (165 images), Cassava Green Mite (148 images), Cassava Mosaic Disease (167 images), and Healthy cassava leaves (147 images). The relatively balanced class distribution ensured a fair and comprehensive assessment of the model's classification performance across all disease categories.

b) Training Data Collection

The training dataset was constructed using a combination of field-collected images and publicly available datasets. Field data were obtained in collaboration with agricultural extension officers, who assisted in capturing high-resolution images of cassava leaves directly from farms across several sub-Saharan regions. The images were collected under diverse environmental conditions, including varying lighting (sunlight and shade), backgrounds, and viewing angles. Both singly infected and co-infected leaves (showing multiple disease symptoms) were included to reflect real-world scenarios. In addition, publicly available datasets such as the PlantVillage repository (which contains over 54,000 images covering multiple crops including cassava and maize) were incorporated to enrich the dataset with a wide range of healthy and diseased samples under controlled conditions. Expert agronomists annotated all images according to disease type or healthy status. The cassava dataset included Cassava Mosaic Disease (CMD) and Cassava Brown Streak Disease (CBSD), while the maize dataset included diseases such as bacterial blight and common rust. A co-infection category was also introduced for leaves exhibiting multiple disease symptoms simultaneously.

c) Preprocessing and Augmentation

All images (training and testing) underwent standard preprocessing procedures, including resizing to a uniform resolution, color normalization, and noise reduction to improve model consistency and performance. During training, extensive data augmentation techniques were applied to enhance model robustness and reduce overfitting. These techniques included random rotations, horizontal and vertical flips, brightness and contrast adjustments, and the addition of Gaussian noise. This approach aligns with established best practices in deep learning-based image classification for agricultural applications.

d) Model Performance Evaluation

The cassava disease detection model achieved an overall classification accuracy of 75%. Among the disease categories, Cassava Mosaic Disease (CMD) recorded the highest performance, with a precision of 85%, recall of 89%, and an F1-score of 87%, indicating strong and reliable detection capability with minimal false positives. Cassava Brown Streak Disease (CBSD) also demonstrated strong performance, achieving a precision of 81%, recall of 82%, and an F1-score of 82%. Cassava Green Mite (CGM) showed moderate performance with a precision of 78% and recall of 68%, indicating some level of misclassification. However, Cassava Bacterial Blight (CBB) and Healthy cassava leaves recorded comparatively lower precision values of 75% and 60%, respectively. This suggests occasional confusion between healthy leaves and early-stage disease symptoms, highlighting an area for further model improvement.3.3 Neural Network Model. CNN was implemented using a transfer-learning approach. Starting with a MobileNetV2 backbone (chosen for its efficiency on mobile devices), we replaced the final classification layers and fine-tuned the network on our dataset. This choice balances accuracy and model size. The model architecture consists of convolutional feature extraction layers followed by fully connected (dense) layers for classification. A transfer learning approach was adopted using a base network pre-trained on ImageNet to accelerate convergence and leverage general visual feature representations; initially, the early layers were frozen, and deeper layers were progressively fine-tuned to improve task-specific performance. The model was trained on a GPU server using cross-entropy loss, with early stopping and learning-rate scheduling applied to reduce overfitting and enhance generalization. For deployment efficiency, the trained model was further optimized through quantization and converted for on-device inference using TensorFlow Lite, thereby reducing memory usage and computational requirements while preserving acceptable accuracy. Additional experiments with ensemble methods and attention mechanisms were conducted; however, a single well-tuned CNN model was ultimately selected as it sufficiently met the desired performance objectives.

Results

This section presents a comprehensive evaluation of the detection system using both quantitative performance metrics and visual diagnostic tools. The assessment is based on classification accuracy, precision, recall, F1-score, confusion matrix analysis, and training performance trends for both cassava and maize disease detection models.

The confusion matrix provides a detailed representation of model performance by showing the distribution of correct and incorrect predictions across all disease categories. It is particularly useful for identifying class-specific strengths as well as areas of misclassification and overlap between disease classes. The result is shown in figure 1, and having the following outputs which presents precision, recall, F1-score, and support for each class.

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Cassava Bacterial Blight (CBB)	0.75	0.64	0.69	173
Cassava Brown Streak Disease (CBSD)	0.81	0.82	0.82	165
Cassava Green Mite (CGM)	0.78	0.68	0.72	148
Cassava Mosaic Disease (CMD)	0.85	0.89	0.87	167
Healthy	0.6	0.75	0.67	147
Overall Accuracy			0.75	800
Macro Average	0.76	0.75	0.75	800
Weighted Average	0.76	0.75	0.75	800

The confusion matrix for cassava disease classification is illustrated in Figure 1. The results indicate that Cassava Mosaic Disease (CMD) achieved the best classification performance, with the highest precision, recall, and F1-score, demonstrating the model's strong capability in detecting CMD symptoms accurately. In contrast, the Healthy class and Cassava Bacterial Blight (CBB) recorded comparatively lower precision and recall values, indicating occasional misclassification between healthy leaves and early-stage disease symptoms as well as overlap with mild infection cases.

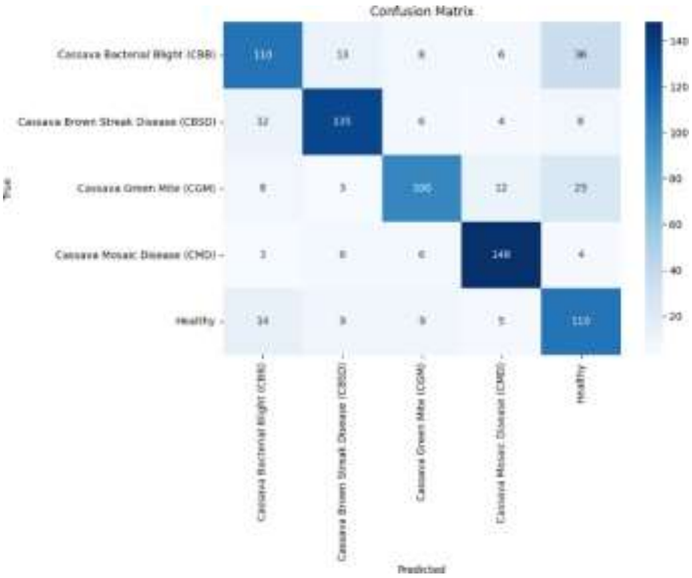

Figure 1- Confusion Matrix for Cassava Disease Classification

Discussion

The results obtained from the evaluation of the proposed plant disease detection system demonstrate that deep learning, particularly convolutional neural networks (CNNs), is effective for automated identification of cassava and maize diseases. Overall, the cassava model achieved an accuracy of 75%, which indicates a reasonable level of performance given the complexity of real-world agricultural image data, variability in environmental conditions, and similarities between certain disease symptoms. From the cassava classification results, it is evident that the model performed best on Cassava Mosaic Disease (CMD), which recorded the highest precision (0.85), recall (0.89), and F1-score (0.87). This strong performance suggests that CMD exhibits more distinctive visual patterns in leaf structure and coloration, making it easier for the model to learn and generalize. Similarly, Cassava Brown Streak Disease (CBSD) also showed strong classification results, indicating that the model can reliably detect this disease under varying conditions. However, the performance for Cassava Green Mite (CGM), Cassava Bacterial Blight (CBB), and especially Healthy leaves was comparatively lower. The reduced recall for CBB (0.64) and CGM (0.68) suggests that the model occasionally struggles to distinguish between visually similar disease symptoms, particularly in early infection stages. The lowest precision observed in the Healthy class (0.60) further indicates a tendency for the model to misclassify healthy leaves as diseased, likely due to subtle visual similarities between healthy foliage and mildly infected samples. The confusion matrix analysis further confirms these observations, highlighting misclassification patterns between Healthy leaves and early-stage disease cases. This is a common challenge in plant disease detection systems, as early symptoms often lack strong visual features and may overlap with natural leaf variations such as discoloration, aging, or environmental stress effects. These findings align with existing research in agricultural image classification, where class imbalance, visual similarity between disease categories, and limited representation of early-stage infections often affect model performance. Despite these limitations, the model demonstrates strong potential for practical deployment, particularly for detecting more visually distinct diseases such as CMD and CBSD.

Conclusion

This study focused on the design, development, and evaluation of a deep learning-based system for cassava plant disease detection using convolutional neural networks (CNNs). The system was developed to support early and accurate identification of major cassava diseases, including Cassava Mosaic Disease (CMD), Cassava Brown Streak Disease (CBSD), Cassava Green Mite (CGM), Cassava Bacterial Blight (CBB), and healthy cassava leaves. A transfer learning approach was adopted to enhance feature extraction efficiency and improve classification performance while reducing training time and computational cost. The dataset used for this study was compiled from both field-collected images and publicly available repositories, and it was carefully preprocessed and augmented to improve model generalization under real-world farming conditions. The evaluation results showed that the model achieved an overall accuracy of 75%, with particularly strong performance in detecting CMD and CBSD. However, lower performance was observed in distinguishing healthy leaves and early-stage infections, indicating challenges related to visual similarity between healthy and mildly diseased samples. Despite these limitations, the findings demonstrate that CNN-based approaches are effective for cassava disease classification and have strong potential for practical agricultural applications. The system can support farmers by enabling faster and more accessible disease diagnosis, thereby contributing to improved crop management and productivity. In conclusion, the proposed cassava disease detection system provides a solid foundation for automated plant health monitoring. Future improvements should focus on expanding the dataset, especially for underperforming classes, and exploring advanced deep learning techniques such as attention mechanisms and ensemble models to further enhance classification accuracy and robustness in real-world conditions.

References

Abubaker, M. (2024). Comparative analysis of lightweight deep learning models for plant disease classification on mobile devices. Computers and Electronics in Agriculture, 215, 108402. https://doi.org/10.1016/j.compag.2023.108402

Egba, A. F., Godspower, A. I., Mayowa, A. S., & Blessing, I. Development of a diabetes mellitus diagnostic system using self-organizing map algorithm: A machine learning approach.

Gai, H., & Wang, Y. (2024). Global crop losses due to plant diseases and pests: Implications for food security. Agricultural Systems, 215, 103871. https://doi.org/10.1016/j.agsy.2024.103871

Gulmez, E., Kaya, S., & Demir, M. (2024). Deep learning-based plant disease detection: A comprehensive review. Computers and Electronics in Agriculture, 218, 108582. https://doi.org/10.1016/j.compag.2024.108582

Iduh, B. N., Umeh, M. N., Anusiuba, O. I., & Egba, F. A. (2024). Development of a predictive modeling framework for athlete injury risk assessment and prevention: A machine learning approach. European Journal of Theoretical and Applied Sciences, 2(4), 894–906.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419. https://doi.org/10.3389/fpls.2016.01419

Paul, R. U., Mbeledogu, N. N., Iduh, B. N., & Okechukwu, O. P. (2025). A Comparative Analysis on diabetes using Ensemble Machine Learning Models on Pima Indian Datasets. Journal of Basic Physical Research, 14(2), 200-213.

Ramcharan, A., McCloskey, P., Baranowski, K., Mwang'ombe, A., & Hughes, D. P. (2017). Deep learning for image-based cassava disease detection. Frontiers in Plant Science, 8, 1852. https://doi.org/10.3389/fpls.2017.01852

Ramcharan, A., McCloskey, P., Baranowski, K., Ahmed, B., & Hughes, D. P. (2019). Mobile-based plant disease detection in field conditions using deep learning. Computers and Electronics in Agriculture, 167, 105035. https://doi.org/10.1016/j.compag.2019.105035

Umeh, M. N., Iduh, B. N., Faseki, F. N., Oboti, N. P., & Nwabudike, O. H. Enhancing public health measures: A deep learning approach for real-time facemask recognition to mitigate the spread of infectious diseases.

Umeh, M. N., Faseki, N. F., Iduh, B. N., Onyedinma, E., & Anyaora, S. C. (2025). Modeling and analysis for cassava aggregation centers within the agricultural zones in Anambra State, Nigeria. Journal of Engineering Research and Reports, 27(6), 395–420. https://doi.org/10.9734/jerr/2025/v27i61234

Zhang, X., Li, W., & Zhao, Y. (2021). Improved CNN architectures for maize disease detection using data augmentation techniques. Computers and Electronics in Agriculture, 190, 106425. https://doi.org/10.1016/j.compag.2021.106425