FMDB *Transactions on Sustainable Health Science Letters*

Explainable-AI-Based Disease Prediction and Classification for Manihot Esculenta Using Meta-Data Enhanced Knowledge Distillation

K. S. S. Sreekrishna[1,*], S. Jasper Sofi[2], Balika J. Chelliah[3]

[1,2]Department of Computer Science and Engineering, Artificial Intelligence and Machine Learning, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[3]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
ks0668@srmist.edu.in[1], js5866@srmist.edu.in[2], ballikaj@srmist.edu.in[3]

Abstract: Cassava (Manihot esculenta) is a critical food staple in most tropical areas, but its productivity is seriously under threat from numerous diseases. Traditional methods of diagnosing cassava diseases are cumbersome and require specialised expertise, underscoring the need for automated disease classification. This work presents an Explainable AI-based disease Prediction and Classification Framework that integrates metadata-augmented knowledge distillation to improve cassava disease detection. The proposed approach combines deep learning-guided image classification with metadata inputs (i.e., climate factors and soil characteristics) to motivate model accuracy. Additionally, Explainable AI (XAI) methods such as SHapley Additive Explanations (SHAP) and Gradient-weighted Class Activation Mapping (Grad-CAM) are employed to provide explainable insights into model predictions, promoting trust and transparency in decision-making. To overcome the limitations of using deep learning models in low-resource settings, knowledge distillation transfers knowledge from a high-performance deep neural network to a compact model. This leads to a computationally effective framework that can perform real-time inference on edge devices. The incorporation of metadata is rigorously tested to measure its effect on disease classification performance. An experimental evaluation using a cassava leaf image dataset and its corresponding metadata shows that the proposed framework significantly improves classification accuracy compared to traditional image-only models.

Keywords: Explainable AI (XAI); Cassava Disease Classification; Knowledge Distillation; Machine Learning; Metadata Integration; Precision Agriculture; Cassava Green Mottle (CGM).

Received on: 13/12/2024, **Revised on:** 21/02/2025, **Accepted on:** 12/04/2025, **Published on:** 07/09/2025

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSHSL

DOI: https://doi.org/10.69888/FTSHSL.2025.000503

Cite as: K. S. S. Sreekrishna, S. J. Sofi, and B. J. Chelliah, "Explainable-AI Based Disease Prediction and Classification for Manihot Esculenta Using Meta-Data Enhanced Knowledge Distillation," *FMDB Transactions on Sustainable Health Science Letters*, vol. 3, no. 3, pp. 179–191, 2025.

1. Introduction

Cassava (Manihot esculenta) is an important food crop that is widely grown in sub-Saharan Africa, Asia, and Latin America. Due to its potential for cultivation in poor soils and in areas of high drought, cassava serves as a crucial source of food security

[*]Corresponding author.

and economic viability in the developing world. Nonetheless, cassava production is severely threatened by a series of leaf diseases, including Cassava Mosaic Disease (CMD), Cassava Brown Streak Disease (CBSD), Cassava Bacterial Blight (CBB), and Cassava Green Mottle (CGM). They cause significant yield losses, compromising the livelihoods and food supplies of millions of people. Early and precise cassava leaf disease identification is essential for disease mitigation and control. Conventional methods using human specialists are usually costly, time-consuming, and impractical in rural areas. Consequently, there has been an increased interest in research aimed at automating disease diagnosis through computer vision and deep learning methods. Recent breakthroughs in deep convolutional neural networks (CNNs), transfer learning, and explainable artificial intelligence (XAI) have opened new possibilities for creating scalable, accurate, and interpretable plant disease classification systems. Nevertheless, most of these systems suffer from key shortcomings, such as model size, limited interpretability, and poor suitability for mobile deployment. In this work, researchers introduce a light, interpretable cassava disease classification system based on a teacher–student knowledge distillation pipeline. Our model combines a highly accurate VGG19-based teacher model with a small CNN student model and a distilled DenseNet201 architecture. To improve interpretability, researchers use two strong XAI methods, Grad-CAM++ and SHAP, that offer localised visual explanations of model predictions. The main contributions of this research are as follows:

- Design of a hybrid deep learning pipeline using knowledge distillation to achieve accuracy and computational efficiency trade-offs.
- Application of two explainability techniques (SHAP and Grad-CAM++) to explain the predictions of student and distilled models.
- Model evaluation on a real-world cassava disease image dataset with extensive experimentation on accuracy, precision, recall, and explainability.
- Proof-of-concept demonstration of deployment on resource-limited devices for real-world agricultural use cases.

With this strategy, researchers address the three key challenges of accuracy, interpretability, and mobility, and provide a scientifically valid, field-deployable, scalable solution for cassava disease management.

2. Literature Review

Several studies have utilised transfer learning, employing pre-trained convolutional neural networks (CNNs) for cassava leaf disease classification. Junior et al. [1] showed that fine-tuned deep models perform better than conventional CNNs by leveraging pretrained weights from large-scale datasets. Their findings showed that feature reuse across domains greatly improves classification performance, even for agricultural images with small sample sizes. Likewise, Ayalde et al. [2] used a YOLO-based object detection system to detect post-harvest physiological deterioration in cassava roots. Their use of K-means clustering further improved quantification, highlighting the effectiveness of combining object detection with unsupervised learning. Explainability in plant disease models has also witnessed tremendous strides. Amara et al. [3] highlighted the flaws of black-box CNN models and proposed automated concept-discovery methods to enhance transparency. Their research, together with that of Thakur et al. [4], who outlined PlantXViT, a vision transformer with an interpretable inbuilt, demonstrates that attention mechanisms can serve not just the purpose of high performance but also be used to decipher model reasoning. This contribution emphasises the rising demand for AI interpretability within agriculture. Ngugi et al. [5] provided a broad overview of deep learning for crop disease diagnosis, highlighting both achievements and challenges. Their findings reasserted the need to strike a balance between performance and model explainability, especially in the application of AI to mission-critical agricultural environments. Leveraging this, Gohil et al. [6] proposed a hybrid model using classical image processing coupled with CNNs to facilitate real-time plant disease detection on edge devices, a necessary step towards mobile deployment.

At the application end, Kalezhi and Shumba [7] employed object detection to identify cassava diseases using bounding boxes, providing not only classification results but also actionable knowledge for field applications. Their approach transitioned from pure classification to spatial knowledge, making more specific interventions possible for the crop. The increasing demand for cost-effective, mobile-compatible solutions is also reflected in Upadhyay et al. [8], who discussed deep learning frameworks optimised via metadata and hardware-aware pruning methods. The use of aerial photos and UAVs (Unmanned Aerial Vehicles) has also broadened the scope of cassava disease monitoring. Nnadozie et al. [9] applied CNN-based classification to UAV-collected RGB image datasets and demonstrated robust performance on large cassava fields. Their research established the applicability of low-cost remote sensing in scalable plant monitoring. Conversely, Bajpai et al. [10] examined transfer learning methods in precision agriculture, highlighting the flexibility of deep networks to learn new environmental conditions with minimal fine-tuning. More recently, Li et al. [11] presented an interactive bilinear transformer that learns from user feedback during landscape plant disease classification. Their novel architecture supports on-the-fly retraining and correction, which could be applied in cassava-farming apps for agricultural extension workers. Chhetri et al. [12] also presented a knowledge graph-integrated deep learning architecture that combines domain knowledge with raw data, greatly enhancing classification accuracy and robustness.

Attention-based models remain top contenders for disease classification. Khabusi et al. [13] demonstrated that incorporating attention mechanisms into CNNs improved disease detection accuracy in cassava leaves. At the same time, Thonduri et al. [14] demonstrated MobileNetV2's performance on cassava leaf images, confirming the feasibility of low-weight networks for mobile deployment. Their model obtained high accuracy with low memory usage. Regarding performance optimisation, Pragada et al. [15] used ResNet152 to identify cassava diseases and emphasised preprocessing spectral-domain data to improve learning efficiency. Grieve et al. [16] took this further by using multispectral imaging to identify early-stage plant viruses, opening the way for precision agriculture equipment that operates outside the visible spectrum. Khandagale et al. [17] developed FourCropNet, a generalised model capable of handling disease classification across multiple crops. Their design promotes cross-crop generalisation, which is essential for developing AI tools in regions cultivating multiple species. Complementing this, Tusubira et al. [18] implemented semantic segmentation to assess root necrosis in cassava, a novel approach that goes beyond classification to quantify disease severity. The advantages of ensemble learning in enhancing classification robustness were investigated by Kiruthika et al. [19], who used an ensemble of various deep architectures to improve cassava disease prediction. Finally, Srivats et al. [20] emphasised the need for explainability by incorporating SHAP and Grad-CAM++ visualisations into their model, verifying model predictions and enhancing user trust, particularly important for low-resource farmers. Overall, the papers demonstrate increasing attention toward interpretable, optimised, and mobile-computing-compatible cassava disease classification frameworks. Focus is not just on high accuracy but also on interpreting model behaviour, optimising deployment, and incorporating expert knowledge through explainability and metadata.

3. Dataset and Preprocessing

This research uses the well-known "Cassava Leaf Disease Classification" dataset, which is publicly available on Kaggle. The dataset is a robust collection of high-quality RGB images of cassava leaves and can serve as a strong benchmark for evaluating disease classification systems in precision agriculture. The dataset is especially useful owing to the diversity in disease classes, image quality, and adequate representation of each disease state, including healthy leaves. It mirrors the complexity of real-world disease manifestations and contributes significantly to the development of AI systems in crop pathology. The data consists of 21,397 labelled images across five classes. The classes represent the most common cassava leaf conditions, consisting of four diseases and one healthy label. They are Cassava Bacterial Blight (CBB), Cassava Brown Streak Disease (CBSD), Cassava Green Mottle (CGM), Cassava Mosaic Disease (CMD), and Healthy. The data set has an evenly distributed sample across classes, so deep models trained on it avoid class imbalance and achieve stable generalisation across various forms of disease. The images are stored in a flat directory tree, and each image is referenced by a separate identifier in the CSV file train.csv, which maps image filenames to their respective class labels (Figure 1).

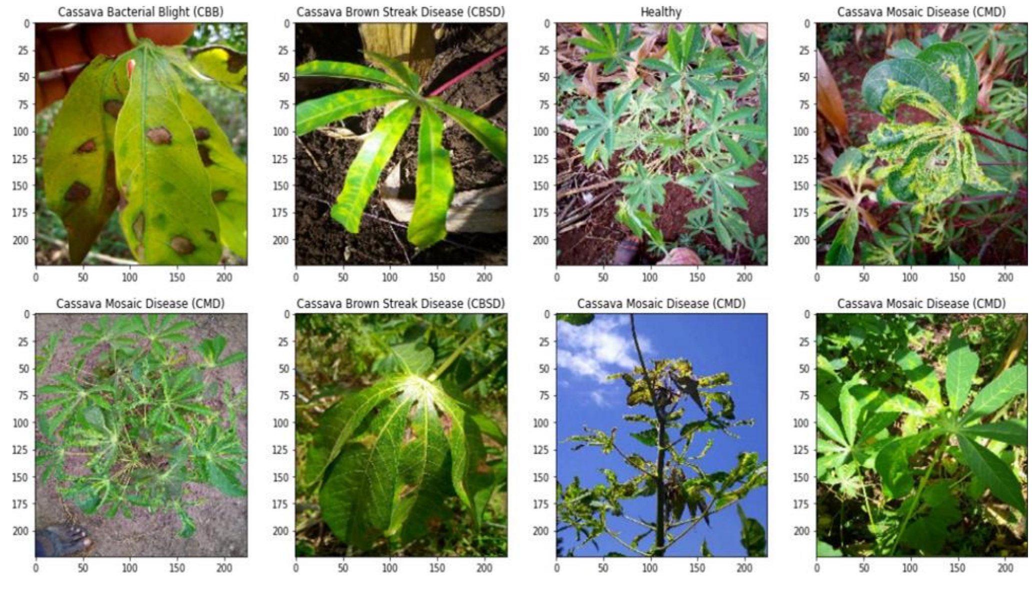

Figure 1: Sample dataset images of 5 classes

To aid interpretability and readability, a JSON mapping file maps numeric class labels (0-4) to their human-readable names. For instance, label "0" is "Cassava Bacterial Blight," and label "4" is healthy leaves. This mapping is essential at both the model training and result visualisation stages, enabling smooth conversion between model outputs and the real-world disease names used by pathologists and agronomists. The script-based implementation allows these mappings to be dynamically incorporated, helping maintain consistent labelling throughout the pipeline. Before feeding the images into the deep models, a strong preprocessing pipeline was used. All images were resized to a standard resolution of 224×224 pixels, in accordance with the input requirements of most state-of-the-art convolutional neural networks (CNNs). The image pixel values were normalised to the [0, 1] interval to speed up convergence and reduce variance across training samples. To further promote generalisation and resilience, the dataset was subjected to rigorous data augmentation using Keras's ImageDataGenerator. This involved random horizontal and vertical flips, small-angle rotations, zooming, width and height shifts, and shear transformations. These transformations mimic real-world variability, such as changes in camera orientation, lighting conditions, and minor leaf deformations. The dataset was divided into training, validation, and testing sets. That is, 90% of the images were used for training and validation (with a 90:10 split), and the remaining 10% constituted the test set for end evaluation. The flow_from_dataframe() function was used to match images to labels dynamically during training. This function also ensured that augmentation was performed in real time during model fitting. Moreover, both TensorFlow and PyTorch frameworks were employed at various stages of experimentation. TensorFlow's ImageDataGenerator enabled SHAP-based interpretability and TFHub's cassava classifier, whereas PyTorch's ImageFolder enabled Grad-CAM++ and knowledge distillation. This two-framework approach provided flexibility in investigating complementary features in both ecosystems (Figure 2).

Figure 2: Training class distribution

4. Proposed Methodology

The suggested methodology combines a teacher-student learning architecture with knowledge distillation and two explainability methods, SHAP and Grad-CAM++, to develop a correct, explainable, and deployable cassava disease classification system (Figure 3).

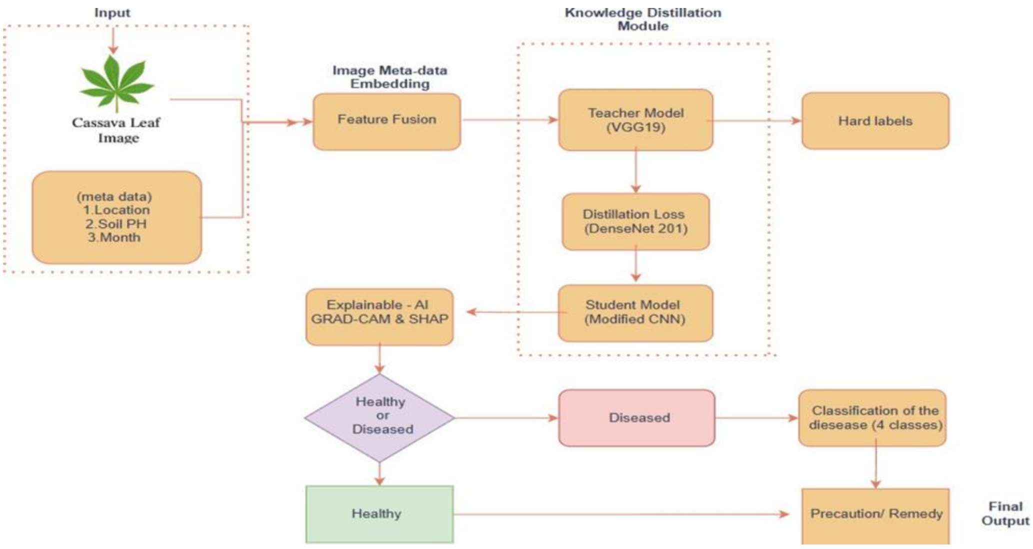

Figure 3: General architecture diagram

The system's goal is not just to provide high classification accuracy but also to make the decision-making process clear and the model footprint light enough for real-time mobile deployment. The method unites the best representational ability of a VGG19-based teacher model, a simple custom student CNN, and the explainability powers of SHAP and Grad-CAM++ in a single framework. At the core of the system is the high-capacity teacher model, based on VGG19, an existing convolutional network renowned for its depth and efficient feature extraction. The teacher model is tasked with learning intricate visual features from the cassava disease dataset (Figure 4).

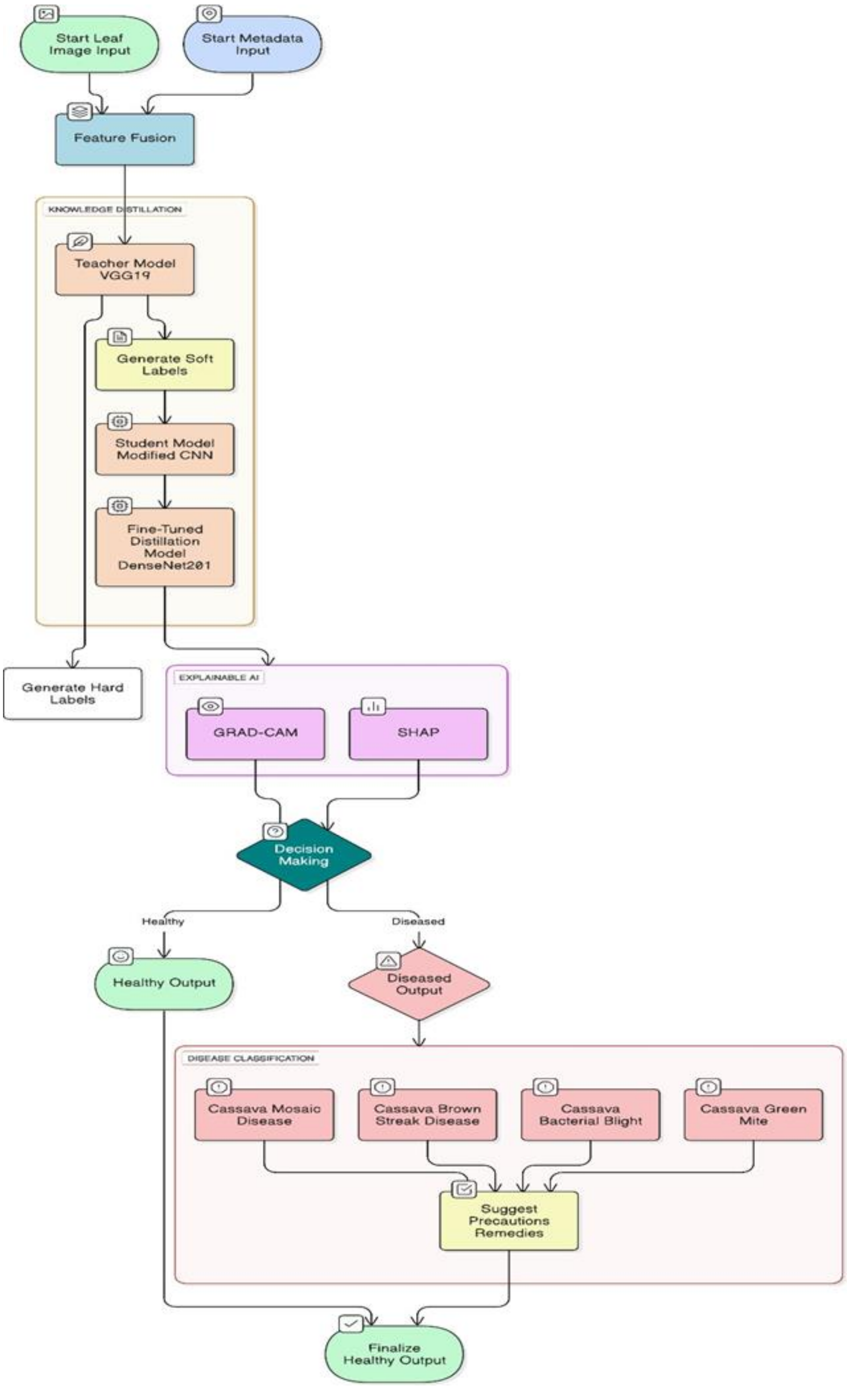

Figure 4: Flow diagram

The trained model's output logits are used as soft labels to supervise knowledge distillation-based training of a lightweight student model. The student model distilled has fewer parameters, yet should replicate the teacher's predictive pattern very closely. Also, the system includes interpretability modules that can visualise how the model makes decisions. This enables domain experts and agronomists to understand and trust the system's predictions (Figure 5).

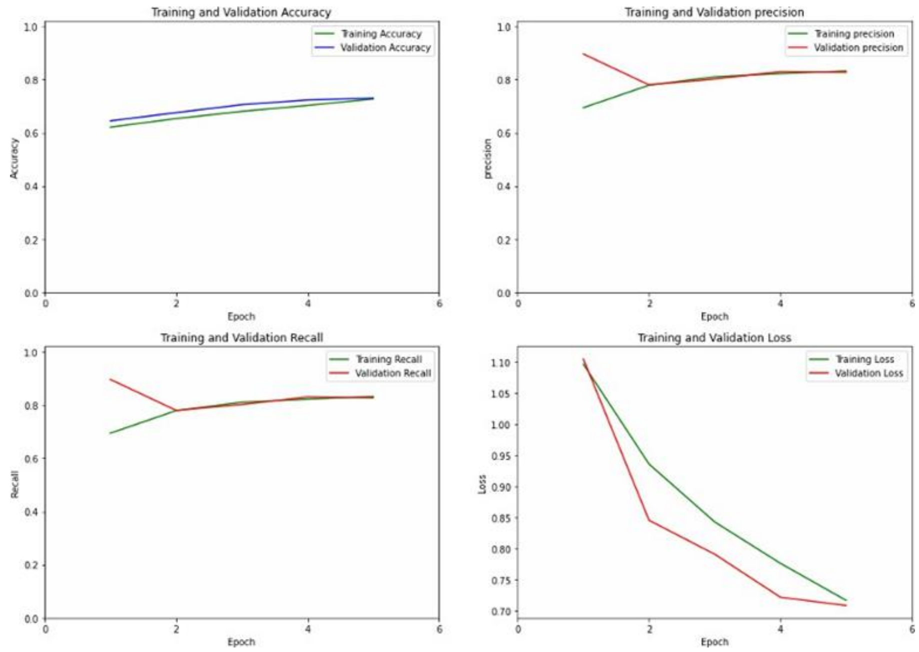

Figure 5: Teacher model performance

The teacher model based on VGG19 was modified from its ImageNet-pretrained state. The convolutional layers were kept as frozen feature extractors, and a specialised classifier head was added. This head included fully connected layers of 512, 256, and 128 neurons each, followed by ReLU activation and dropout layers for regularisation. A final dense layer with five output neurons and softmax activation was included to perform classification. Batch normalisation after intermediate dense layers helped stabilise learning and accelerate convergence. The network was trained with the categorical cross-entropy loss and the Adam optimiser. The student model was designed specifically for efficiency. It had four layers of convolutional blocks, increasing in depth (32 to 256 filters), with ReLU activation and max pooling to reduce spatial resolution, followed by a flattening layer, a two-stage dense head with 1024 and 512 neurons, and finally a softmax classification layer. The dropout layers were incorporated after the dense blocks to help reduce overfitting. Despite the relative simplicity of this architecture compared to the teacher model, the student model performed similarly, thanks to knowledge distillation during its training (Figure 6).

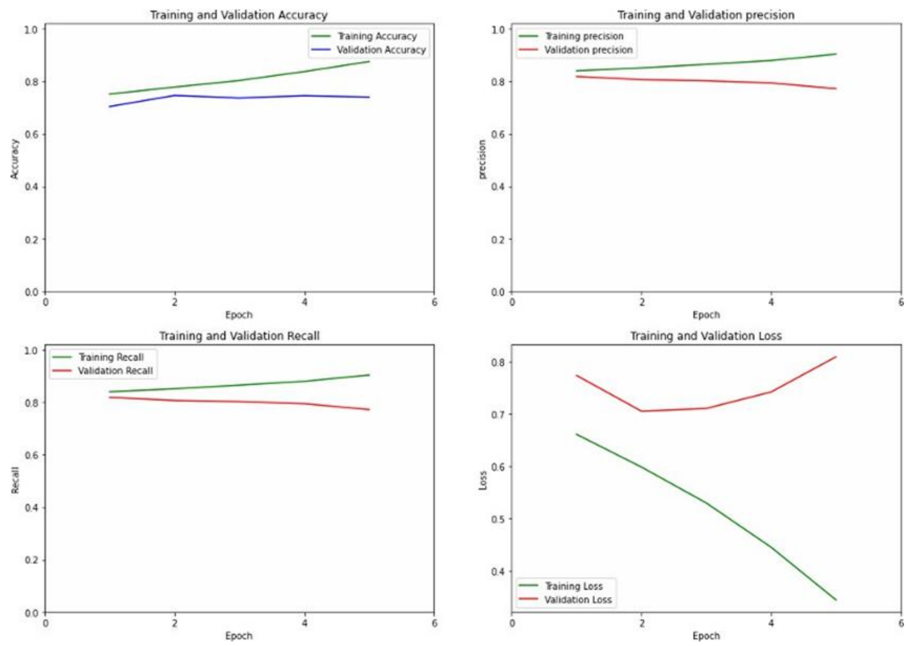

Figure 6: Student model performance

Knowledge distillation was implemented using a dual-loss framework that combined the standard categorical cross-entropy loss with a Kullback-Leibler divergence loss computed between the teacher and student softmax outputs. These softmax outputs were softened with a temperature scaling parameter (T = 10). The contribution of soft vs. hard labels during training was controlled by an alpha parameter ($\alpha = 0.1$). The distillation technique enabled the student model to learn the teacher's inter-class relations, thereby improving its generalisation performance on borderline cases (Figure 7).

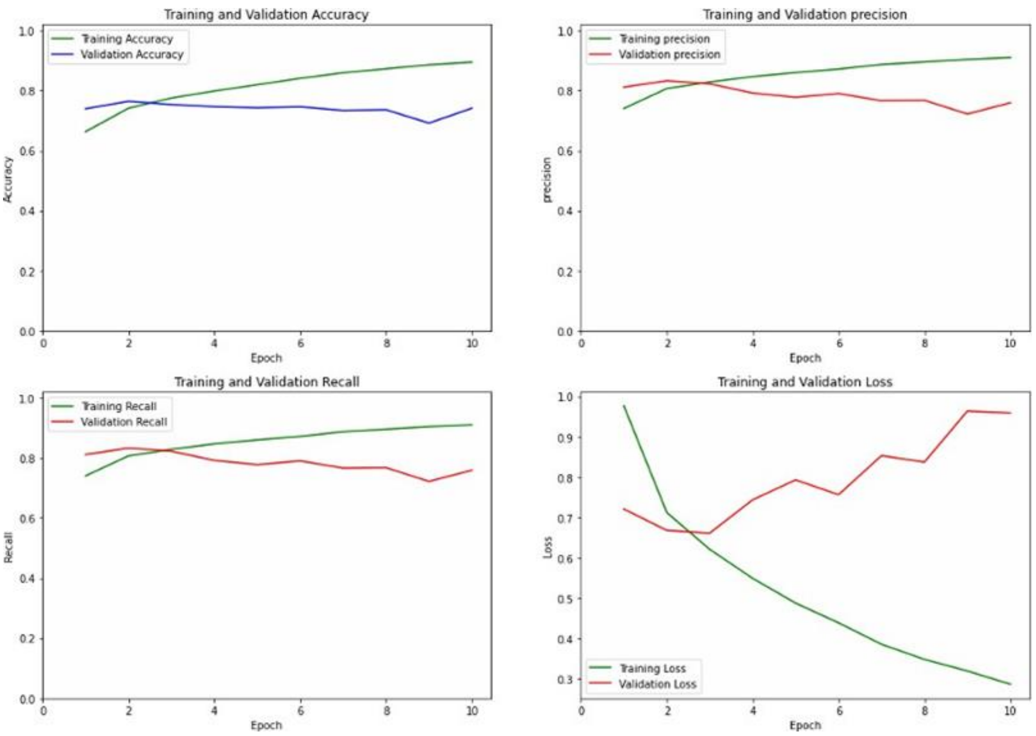

Figure 7: Distilled model performance

All the models were trained with the Adam optimiser and a learning rate of 1-e-4, and a batch size of 32. Training lasted up to 25 epochs, with early stopping based on validation accuracy and loss. The ReduceLROnPlateau callback was used to decrease the learning rate as validation performance plateaued dynamically. These training parameters were kept constant across all experiments to ensure a fair comparison among the baseline, distilled, and student-only models. To aid interpretability, two recent explainability techniques were employed. SHAP was used for the TensorFlow-trained models to indicate the pixel areas that positively or negatively contribute to a class prediction. It was applied via the PartitionExplainer and image masker for pixel-wise high-resolution attribution. Grad-CAM++ was used instead with the PyTorch-trained models. It generated gradient-based class activation maps by capturing higher-order derivatives with respect to the target class output. These maps were superimposed on the initial input images so domain specialists could confirm whether the model was focusing on the affected regions of the leaf. The suggested pipeline strikes a good balance between accuracy, interpretability, and efficiency. It integrates a strong teacher model for learning sophisticated features, a light student model for fast inference, and robust explainability tools to guarantee model transparency. Such an architecture is best suited for real-time deployment in mobile apps for cassava disease diagnosis in field environments.

5. Implementation and Results

This section describes the technical realisation of the put forth cassava leaf disease classification model. It presents the tools and libraries involved, the end-to-end training approach adopted for the model, and the results for each of the three architectures: teacher model, student model, and distilled model. It further addresses interpretability by combining Grad-CAM++ and SHAP visualisations, providing robust visual support for the system's decision-making processes. The suggested system was developed in Python using both TensorFlow and PyTorch libraries to leverage their complementary strengths. TensorFlow was used for its mature model training environment and compatibility with SHAP, whereas PyTorch enabled greater control for explainability via Grad-CAM++. The student model was developed in PyTorch due to its modular architecture and support for lightweight deployment. The distilled and teacher models were trained with TensorFlow, with early stopping and learning rate scheduling configured via standard callbacks. The experiments were performed using a Google Colab GPU runtime with 16 GB RAM and NVIDIA Tesla T4 GPU acceleration. The system environment included TensorFlow 2.14, PyTorch 2.0, OpenCV, Matplotlib, and ancillary libraries such as scikit-learn and pandas. The platform provided a perfect environment for experimenting with medium-scale datasets, such as the Cassava Leaf Disease dataset. The dataset used in this research is from the publicly released Cassava Leaf Disease Classification Challenge on Kaggle. It has a total of 21,397 RGB cassava leaf images, labelled into five classes: Cassava Bacterial Blight (CBB), Cassava Brown Streak Disease (CBSD), Cassava Green Mottle (CGM), Cassava Mosaic Disease (CMD), and Healthy. Every image is tagged with a numeric ID, which was translated to its respective class name using a JSON-based mapping schema. To properly train and test the models, the dataset was split into 90% for training and 10% for validation. The images were resized to 224×224, normalised to [0, 1], and augmented with random transformations, including flipping, rotation, zooming, and shearing. These random transformations enhanced dataset diversity and lessened the likelihood of overfitting, particularly in the student model with fewer learnable parameters.

Three variants of the model were trained and tested with the same hyperparameters to compare them fairly. EarlyStopping stopped training when validation accuracy plateaued, and ReduceLROnPlateau progressively lowered the learning rate. The teacher model, built with VGG19 and added dense layers, achieved a training accuracy of 94%. It achieved 73% accuracy, 77% precision, and 70% recall on the validation set. Though it performed better in feature extraction and generalisation, its large parameter count (over 140 million) made it unsuitable for deployment on memory-limited devices such as smartphones or drones. The student model, a lightweight CNN composed of four convolutional layers and dense layers, achieved training accuracy of 94% and test accuracy of 74%. The slightly lower precision and recall from the teacher were offset by the model's lightweight nature, which enabled deployment on extremely mobile devices. Its inference speed was approximately 3× that of the teacher model. The distilled model, trained on a DenseNet201 backbone via knowledge distillation, used a hybrid loss that blended categorical cross-entropy and KL divergence. The teacher model's soft target distribution helped guide the student toward more refined inter-class boundaries. With a temperature of 10 and an alpha of 0.1, the distilled model had a test accuracy of 74%, precision of 76%, and recall of 72%. The results confirmed that the model retained the teacher's abstraction capabilities while remaining compact and inference-efficient (Figure 8).

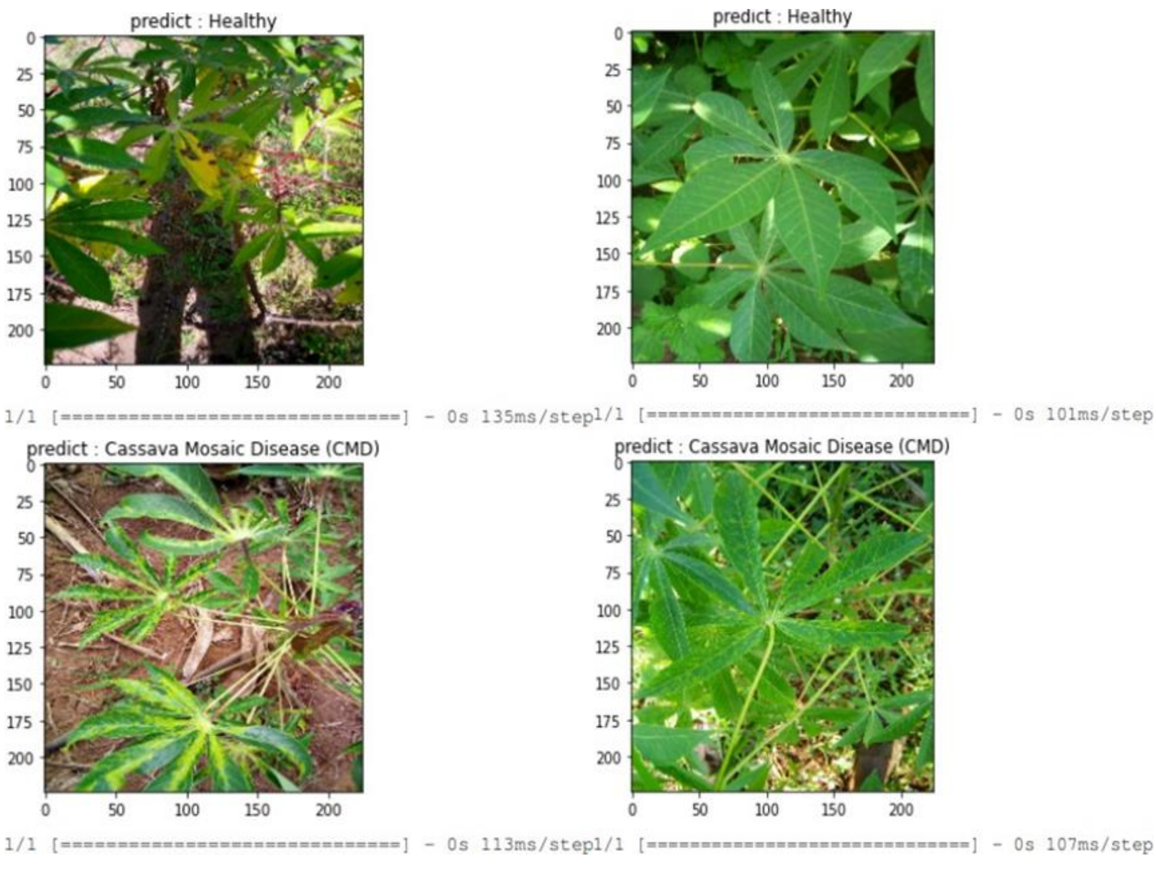

Figure 8: Sample prediction output

For interpretability, both Grad-CAM++ and SHAP were integrated to explain the predictions intuitively. SHAP was applied to the distilled TensorFlow model. The SHAP plots emphasised which regions of pixels played the most important role in the class prediction. For instance, in CMD cases, the high SHAP values were focused on leaf veins and mosaic-pattern edges.

These outputs gave a valid biological explanation to the classifier's predictions and were informative to domain experts (Figure 9).

Figure 9: SHAP results

Grad-CAM++ was used with the PyTorch-trained student model. Grad-CAM++ heatmaps were superimposed on the original images to identify class-discriminative regions. For disease-positive images, the model consistently highlighted diseased areas such as lesions, blotches, or discolourations, indicating that it was learning appropriate features (Figure 10).

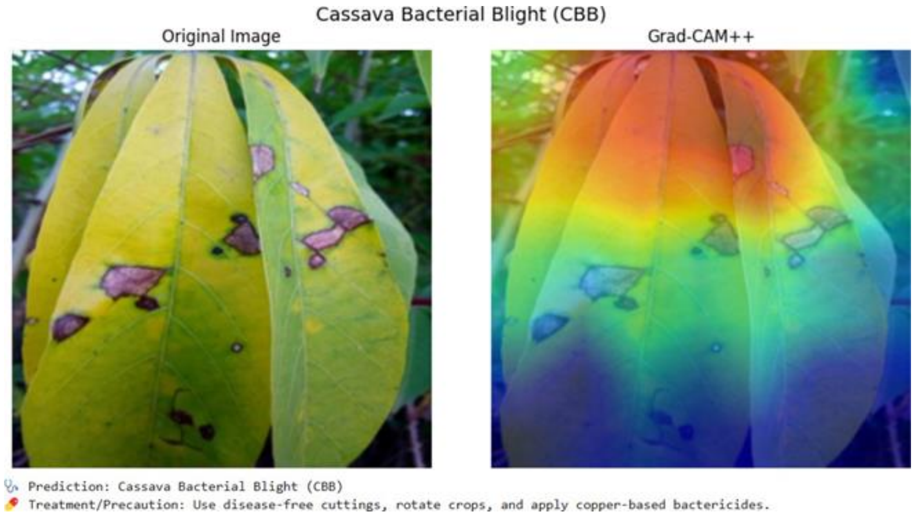

Figure 10: Grad-CAM++ for diseased leaf

Among the models compared, the distilled model was the best compromise on accuracy and efficiency. While the teacher model provided slightly greater precision, its size and latency presented deployment challenges. The student model, though small, had lost some abstraction capability (Figure 11).

Figure 11: Grad-CAM++ for healthy leaf

The distilled model was built on the teacher's understanding and yet maintained efficiency, with good test performance. Explainability outputs validated the quantitative results. Both SHAP and Grad-CAM++ confirm that the models are targeting biologically significant features in the image. This visual consistency enhanced the classifier's robustness, opening the door to practical deployment in agricultural field applications.

6. Discussion

The results of this research underscore the potential of employing deep learning and explainable artificial intelligence to identify cassava leaf diseases with high precision, even in limited setups. The comparative assessment of various models reveals that performance need not be highly complex, particularly when methods such as knowledge distillation are employed to transfer representational ability from small to large models. This section provides a critical evaluation of the most important results, their applicability in real-world usage, and the future potential of AI-based plant disease diagnosis systems.

6.1. Model Efficiency vs. Performance

One of the main aims of the study was to maintain high classification performance while keeping computational overhead small. The student model, a customised CNN optimised for size, achieved a test accuracy of 74%, comparable to or slightly better than that of the VGG19-based teacher model. This performance is impressive given the much smaller number of parameters and lower inference time required by the student network. The light structure of the student architecture makes it amenable to deployment on mobile devices or in edge environments, e.g., in-field crop diagnostics or UAV integration for aerial crop imaging. Knowledge distillation also enhanced the system by leveraging the teacher model's soft-label supervision and transferring it to a distilled student model constructed with DenseNet201. The procedure maintained high-level feature abstractions while ensuring efficiency. The distilled model achieved an identical test accuracy of 74%, while precision and recall were similar to those of the teacher. These findings indicate that strategic model compression with KD does not necessarily sacrifice performance and can be a key approach for developing deployable solutions without excessive computing resources.

6.2. Significance of Explainability

In agriculture, where stakeholders range from researchers to farmers and policy-makers, explainability is essential. Reliable predictions are not enough without an understanding of the AI model's decision-making process. This research fulfilled this gap by combining two complementary explainability methods: SHAP (SHapley Additive Explanations) and Grad-CAM++. SHAP provided high-resolution information about which areas of a cassava leaf image contributed to the final classification. By assigning weights to input pixels, SHAP plots identified necrotic lesions, discolouration patterns, and structural anomalies as the most important features used by the model. Such local interpretability is extremely beneficial in agronomic studies and extension services, where domain experts need verification of computerised evaluations. Grad-CAM++, run on the PyTorch version of the student model, produced visual heatmaps of attention that pinpointed spatial regions of the image most relevant to the predicted class. These attention maps showed the strongest correlation with known disease symptoms and provided a type of class-specific explanation consistent with the biological nature of cassava diseases. Notably, this enabled researchers not only to verify that the model learned relevant patterns but also to provide greater transparency for end users. The inclusion of these explainability tools reflects a broader movement in AI ethics and accountability, ensuring that AI models in sensitive domains remain interpretable, trustworthy, and diagnostically sound.

6.3. Real-World Applications and Deployability

The system has practical implications in real-world agriculture, particularly in rural and low-resource settings where cassava is the primary crop. The student model, especially the distilled version, demonstrates the computational efficiency required for mobile deployment, enabling farmers to diagnose plant disease in the field using smartphone cameras. The compact size of the model guarantees local inference, eliminating the need for continuous internet connectivity, a significant consideration in remote locations. In addition, the system can be scaled up to drone-based disease monitoring, where lightweight, real-time inference is essential for monitoring large cassava plantations. UAVs with onboard embedded systems can load pre-trained student or distilled variants to evaluate field health and create geo-tagged disease maps to facilitate timely intervention and containment. In institutional settings, the system may be integrated with centralised farm management systems and dashboards. Agricultural officers or cooperatives may track disease outbreaks and trends by aggregating data from various users. Such platforms may also facilitate policymaking by determining high-risk areas and optimising resource allocation for disease control interventions.

7. Limitations and Challenges

The system is not without limitations, despite its encouraging results. One of the primary challenges is generalizability. The training set, while rich in diversity and quality, might not fully represent the full range of lighting conditions, image resolutions, and environmental backgrounds encountered in real-world deployment. This difference could result in performance loss in actual deployment environments. Data augmentation strategies during training reduced this effect, but future work can leverage domain adaptation techniques or synthetic data generation. Another limitation is that explainability techniques rely on model robustness. SHAP values and Grad-CAM++ maps are only valid if the underlying model has learned semantically meaningful representations. In under-trained or overfitted models, explanations can be misleading. This requires careful validation of both model performance and explanation quality before deployment in the real world. Additionally, the system currently relies solely on image inputs. Although this makes deployment easier, other metadata, such as geographic location, time of year, and soil conditions, could further enhance prediction accuracy and contextual applicability. Multi-modal inputs are an area for future development.

7.1. Future Directions

The findings of this study open up several paths for further development. One short-term one is the development of existing pipelines through ongoing learning tactics, enabling models to adjust to new disease strains or evolving symptoms over time without being constrained from the beginning. This would make the system more durable and better able to respond to real-world dynamics. Adding Vision Transformer (ViT) architectures might further enhance performance, particularly in detecting faint disease patterns. Although ViTs are computationally costly, their parameter-efficient variants and pruning strategies can potentially enable them to be deployed in mobile settings. Federated learning is another potential avenue, supporting decentralised training on user devices without compromising privacy. This approach would facilitate the building of a worldwide, distributed cassava disease detection network that learns to improve its performance as more users use the system. Finally, adding SHAP and Grad-CAM++ dashboards to the system's visual analytics interface would make it even more useful to agronomists and researchers. These dashboards would not just be diagnostic platforms but could also serve as training sites to instruct farmers on disease symptoms and severity levels.

8. Conclusion and Future Work

This paper proposes a robust, lightweight, and explainable deep learning architecture for cassava leaf disease classification, leveraging knowledge distillation and explainable AI methodologies. The system is based on a teacher-student setup, in which a VGG-19-based teacher shares knowledge with a lightweight custom CNN student. In addition, a DenseNet-based distilled model acts as an optimised link between performance and computational cost. All models were tested on a real cassava disease dataset with five large classes, and performance metrics such as accuracy, precision, and recall were reported for the training and test sets. The outcomes show that the student model, trained via knowledge distillation, performs similarly to the larger teacher model but is much more parameter- and computationally efficient. The distilled model offered the best trade-off between accuracy and generalizability. This supports the practical usability of the system, especially in situations where deployment on mobile or low-resource edge devices is critical. Explainability has been a primary pillar of this work. Using SHAP and Grad-CAM++ visualisations, the system not only produced accurate predictions but also provided understandable insights into its decision-making. These visual explanations are key components in establishing trust among end-users, farmers, and agricultural professionals who depend on transparency when deciding what affects crop management and food security.

8.1. Future Work

There are several significant ways that this work may be extended. Multi-modal Integration: Adding metadata such as geographic coordinates, weather, and historical disease data can increase model robustness and versatility across different conditions:

- **Disease Class Expansion:** The addition of other cassava diseases and other crops will broaden the system's usefulness across agricultural applications.
- **Deployment on Edge Devices:** Tuning the final student and distilled models for Android and IoT deployments would facilitate real-time prediction in offline or limited-bandwidth environments.
- **Interactive Explainability:** Constructing interactive dashboards that enable users to iterate on predictions, view SHAP/Grad-CAM overlays, and simulate scenarios can facilitate decision-making at scale.
- **Temporal Monitoring:** Including temporal models to observe disease progression over time across a plantation may help enable preventive actions before yield is significantly impacted.
- **Federated Learning:** The application of privacy-preserving, decentralised learning models will enable the system to continuously learn from distributed real-world data without violating individual data ownership.

Lastly, the research depicts an executable, interpretable, and effective pipeline for cassava disease classification, which blends transfer learning prowess with knowledge distillation and interpretability of AI. The platform could be a precursor to widespread adoption in major agricultural nations, helping reduce their reliance on large-scale labour forces and, ultimately, address food sustainability alongside precision agriculture in developing nations.

Acknowledgement: The authors gratefully acknowledge SRM Institute of Science and Technology for its support throughout this research. Its resources and academic guidance contributed significantly to the completion of this work.

Data Availability Statement: The data used in this study are publicly available and can be obtained from the Kaggle Cassava Leaf Disease Classification competition section.

Funding Statement: No funding grant was received from any funding organisation in the for-profit, commercial, or non-for-profit sector for this study.

Conflicts of Interest Statement: On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics and Consent Statement: This study does not include any studies on human participants or animals by the authors. Consent is not applicable.

References

1. A. G. C. Junior, F. S. D'a Silva, and R. Rios, "Deep learning-based transfer learning for classification of cassava disease," *arXiv preprint arXiv:2502.19351*, 2025. Available: https://arxiv.org/pdf/2502.19351 [Accessed by 26/07/2025].
2. D. G. Ayalde, J. C. G. Londoño, A. Q. Mosquera, J. L. L. Melendez, W. Gimode, T. Tran, X. Zhang, and M. G. Selvaraj, "AI-powered detection and quantification of post-harvest physiological deterioration (PPD) in cassava using YOLO foundation models and K-means clustering," *Plant Methods*, vol. 20, no. 1, p. 178, 2024.
3. J. Amara, B. Konig-Ries, and S. Samuel, "Explainability of deep learning-based plant disease classifiers through automated concept identification," *arXiv preprint arXiv:2412.07408*, 2024. Available: https://arxiv.org/pdf/2412.07408 [Accessed by 10/10/2024].
4. P. S. Thakur, P. Khanna, T. Sheorey, and A. Ojha, "PlantXViT: Explainable vision transformer enabled convolution neural network for plant disease identification," *arXiv preprint arXiv:2207.07919*, 2022. Available: https://arxiv.org/abs/2207.07919 [Accessed by 16/10/2024].
5. H. N. Ngugi, A. E. Ezugwu, A. A. Akinyelu, and L. Abualigah, "Revolutionizing crop disease detection with computational deep learning," *Environmental Monitoring and Assessment*, vol. 196, no. 3, p. 302, 2024.
6. M. K. Gohil, A. Bhattacharjee, R. Rana, K. Lal, S. K. Biswan, N. Tiwari, and B. Bhattacharya, "A hybrid technique for plant disease identification and localisation real-time," *arXiv preprint arXiv:2412.19682*, 2024. Available:

file:///C:/Users/moham/Downloads/A_Hybrid_Technique_for_Plant_Disease_Identificatio.pdf [Accessed by 20/10/2024].
7. J. Kalezhi and L. Shumba, “Cassava crop disease prediction and localization using object detection,” *Crop Protection*, vol. 187, no. 1, p. 107001, 2025.
8. A. Upadhyay, N. S. Chandel, K. P. Singh, S. K. Chakraborty, B. M. Nandede, M. Kumar, A. Subeesh, K. Upendar, A. Salem, and A. Elbeltagi, “Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture,” *AI Review*, vol. 58, no. 3, p. 92, 2025.
9. E. C. Nnadozie, O. N. Iloanusi, O. A. Ani, and K. Yu, “Detecting Cassava Plants under Different Field Conditions Using UAV-Based RGB Images and Deep Learning Models,” *Remote Sensing*, vol. 15, no. 9, p. 2322, 2023.
10. C. Bajpai, R. Sahu, and K. J. Naik, “Deep learning model for plant-leaf disease detection in precision agriculture,” *International Journal of Intelligent Systems Technologies and Applications*, vol. 21, no. 1, pp. 72-91, 2023.
11. Y. Li, L. Ma, and N. Sun, “A bilinear transformer interactive neural networks-based approach to fine-grained recognition and protection of plant diseases for gardening design,” *Crop Protection*, vol. 180, no. 6, p. 106660, 2024.
12. T. R. Chhetri, A. Hohenegger, A. Fensel, M. A. Kasali, and A. A. Adekunle, “Towards improving prediction accuracy and user-level explainability using deep learning and knowledge graphs: A study on cassava disease,” *Expert Systems with Applications*, vol. 233, no. 12, p. 120955, 2023.
13. S. P. Khabusi, P. Pheroijam, and S. Kshetrimayum, “Attention-Based Approach for Cassava Leaf Disease Classification in Agriculture,” *2023 International Conference on Energy, Power, Environment, Control, and Computing (ICEPECC)*, Gujrat, Pakistan, 2023.
14. V. Thonduri, B. T. V. Reddy, G. C. Balaji, G. Sasank, and D. Y. Naik, “MobileNetV2 for plant disease detection: a scalable deep learning framework,” *International Journal of Advance Research and Innovative Ideas in Education*, vol. 11, no. 2, pp. 833-838, 2025.
15. P. Pragada, K. G. A. Teja, A. Nandini, H. H. Kumar, and B. Hyma, “Improving Precision Agriculture by Utilizing Resnet152 for Cassava Plant Disease Detection and Enhanced Crop Health Monitoring,” *Archives*, vol. 13, no. 6, pp. 1193-1198, 2023.
16. B. Grieve, S. Duffy, M. M. Dallas, J. T. Ascencio-Ibanez, V. Alonso-Chavez, J. Legg, L. Hanley-Bowdoin, and H. Yin, “Early Detection of Plant Virus Infection Using Multispectral Imaging and Machine Learning,” *Plant Health Cases*, 2024. Available: https://www.cabidigitallibrary.org/doi/10.1079/planthealthcases.2024.0010 [Accessed by 31/10/2024].
17. H. P. Khandagale, S. Patil, V. S. Gavali, S. V. Chavan, P. P. Halkarnikar, and P. A. Meshram, “Design and Implementation of FourCropNet: A CNN-Based System for Efficient Multi-Crop Disease Detection and Management,” *arXiv preprint arXiv:2503.08348*, 2025. Available: https://arxiv.org/abs/2503.08348 [Accessed by 11/07/2025]
18. J. F. Tusubira, B. Akera, S. Nsumba, J. Nakatumba-Nabende, and E. Mwebaze, “Scoring Root Necrosis in Cassava Using Semantic Segmentation,” *arXiv preprint arXiv:2005.03367*, 2020. Available: https://arxiv.org/abs/2005.03367 [Accessed by 07/10/2024]
19. V. Kiruthika, S. Shoba, M. Sendil, K. Nagarajan, and D. Punetha, “Hybrid ensemble - deep transfer model for early cassava leaf disease classification,” *Heliyon*, vol. 10, no. 16, p. e36097, 2024.
20. R. Srivats, A. Sharma, V. Kalyanasundaram, and S. Saravanan, “Explainable AI with SHAP and Grad-CAM: A Novel Approach to Visualizing Cassava Leaf Health using CIVE, ECI Indices, and an Advanced UNet Framework,” *2025 International Conference on Innovation in Computing and Engineering (ICE)*, Greater Noida, India, 2025.