Cassava Disease Classification – Kaggle Competition Report

Team Name: Emmanuel_Tiana

Emmanuel Kirui Barkacha[1] and Rabemanantsoa Andriamianja Tiana[2]
[1]ebarkacha@aimsammi.org
[2] arabemanantsoa@aimsammi.org

Abstract

This project focuses on the challenge of cassava leaf disease classification using deep learning techniques. Given a dataset of labeled cassava leaf images, our goal was to accurately classify them into five categories: four disease types and healthy leaves. We explored various convolutional neural network architectures focusing on transfer learning with pretrained models such as DenseNet121, EfficientNet-B4 and ConvNeXt. Through systematic experimentation involving data augmentation and fine-tuning, we achieved strong performance on the public leaderboard. ConvNeXt emerged as the best-performing model, demonstrating high generalization capabilities.

KeyWords: Deep learning; Convolutional neural networks (CNN); Transfer learning; Pretrained models;

1. Introduction

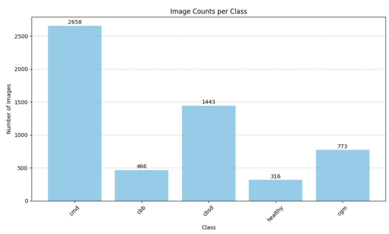
Cassava is important food crop in Sub-Saharan Africa, providing carbohydrates to millions and cultivated by over 80% of smallholder farmers[1] . Despite its resilience, cassava is vulnerable to several diseases which cause significant yield losses. These diseases are difficult to distinguish visually and often require expert inspection. The task involves building a model that can classify cassava leaf images into one of five categories. We explore modern CNN architectures and training techniques to develop an accurate image classification model.

2. Dataset

The dataset consists of 21,367 cassava leaf images, with 9,436 labeled and 12,595 unlabeled samples. Each labeled image falls into one of five classes: four diseases Cassava Brown Streak Disease (CBSD), Cassava Mosaic Disease (CMD), Cassava Bacterial Blight (CBB) and Cassava Green Mite (CGM) or healthy.

(a) CBB **(b)** CBSD **(c)** CGM **(d)** CMD **(e)** Healthy

Figure 1. Sample cassava leaf images for each class: CBB, CBSD, CGM, CMD, and Healthy.



The data was collected in Uganda using smartphones and annotated by agricultural experts at NaCRRI [1].

There is a noticeable class imbalance, with CMD and CBSD having the most samples, while CBB and Healthy are less represented. This dataset was originally used in a 2020–2021 Kaggle competition [2] and reflects real-world image quality and field conditions.

Figure 2. Distribution of training images across cassava leaf classes.

3. Methodology

We used transfer learning with pretrained convolutional neural networks (CNNs) to classify cassava leaf images. All models were trained using PyTorch on a Tesla P100 GPU via Kaggle Notebooks.

3.1 Data Preprocessing

The images were resized to 380×380 pixels to match the input size required by the EfficientNet-B4 architecture[3]. Data normalization was performed using the standard ImageNet mean and standard deviation values. To improve generalization and address class imbalance, we applied data augmentation: random flips, rotations, color jitter, cropping and resizing.

3.2 Model Architectures

We experimented with three CNNs pretrained on ImageNet such as: **DenseNet121**: Efficient and strong baseline, **EfficientNet-B4**[3]: Optimized through compound scaling and **ConvNeXt-Tiny**[4]: Modern CNN inspired by transformer design, achieving the best performance.

3.3 Training Details

Models were fine-tuned using the AdamW optimizer with architecture-specific learning rates and cosine annealing scheduling. We employed cross-entropy loss with label smoothing to improve generalization. The dataset was split into 80% training, 10% validation, and 10% test using stratified sampling to preserve class distributions. Due to resource constraints and the relatively large input image size, we used smaller batch sizes. This allowed us to train for more epochs and improve model convergence. **Metrics:** Accuracy and macro F1-score.

4. Experiments and Results

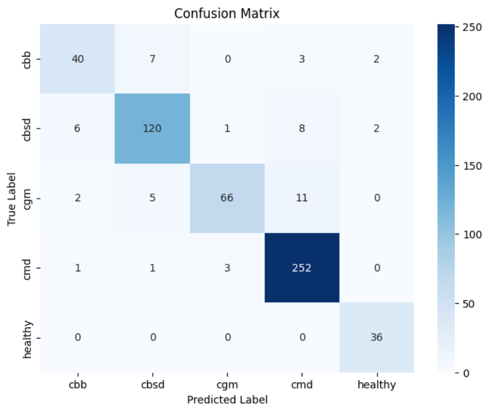
We evaluated all models on the validation and test sets using accuracy and macro F1-score. ConvNeXt-Tiny consistently outperformed other models in both metrics.

4.1 Model Performance

Table 1. Performance of different models on the validation set.

Model	Val Accuracy (%)	Val Macro F1-Score	Local Test Accuracy (%)
DenseNet121	89.93	0.8507	88.87
EfficientNet-B4	91.71	0.8720	90.11
ConvNeXt-Tiny	**91.87**	**0.8882**	**90.81**

4.2 Confusion Matrix



To better understand class-wise performance, we plotted the confusion matrix for the best-performing model (ConvNeXt-Tiny), shown in Figure 3. The model performed particularly well on CMD and CBSD, which had the most training samples. Performance on the less represented classes like CBB and Healthy was slightly lower but still acceptable.

Figure 3. Confusion matrix for ConvNeXt-Tiny on the validation set.

4.3 Leaderboard Performance

Our final submission using ConvNeXt-Tiny achieved **91.456% accuracy** on the public leaderboard, placing us as the top team. Overfitting was avoided by using data augmentation and careful validation practices using f1 score.

5. Conclusion

This project focused on cassava leaf disease classification using deep learning and transfer learning. Among the models tested, ConvNeXt-Tiny demonstrated the best performance, achieving over 91% accuracy and strong generalization to unseen data. We found that image size selection, effective data augmentation and label smoothing significantly improved model robustness, especially for underrepresented classes. Future improvements could include ensembling multiple models, using test-time augmentation (TTA) and using pseudo-labeling to leverage the unlabeled data.

References

[1] E. Mwebaze, T. Gebru, A. Frome, S. Nsumba, and J. Tusubira, "icassava 2019 fine-grained visual categorization challenge," *arXiv preprint arXiv:1908.02900*, 2019. [Online]. Available: https://arxiv.org/abs/1908.02900

[2] Kaggle, "Cassava leaf disease classification," https://www.kaggle.com/competitions/cassava-leaf-disease-classification, 2020, accessed: 2025-08-09.

[3] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *arXiv preprint arXiv:1905.11946*, 2019. [Online]. Available: https://arxiv.org/abs/1905.11946

[4] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 11 976–11 986. [Online]. Available: https://arxiv.org/abs/2201.03545