Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1295

An Integrated Diagnostic Model for Identifying and Diagnosing Groundnut Leaf Diseases

Sanusi Dahiru Yerima* and Yusuf Musa Malgwi

Department of Computer Science, Faculty of Computing, Modibbo Adama University, Yola, Nigeria

Corresponding Author: sanusidahiru1145@gmail.com

ABSTRACT

Groundnut (Arachis hypogaea L.), also known as peanut, serves as a significant source of edible oil and protein, making it essential for the agricultural economy and food security. However, groundnut plants are susceptible to various diseases, notably leaf diseases. Early and accurate identification of these diseases is crucial for implementing timely management strategies to mitigate losses. In recent years, advancements in deep learning especially convolutional neural networks (CNNs), have revolutionized image recognition tasks, including plant disease identification. This study employed the use of ResNet50 and VGG16 Convolutional Neural Networks (CNNs) with Bayesian Optimization (BO). The dataset comprises both infected and uninfected groundnut leaf images, categorized into six folders based on their status. Performance evaluation demonstrates that ResNet50 and VGG16 model's ability to accurately identify groundnut leaves diseases. Results indicate that the hybrid models ResNet50 and VGG16 effectively captures and distinguishes between infected and uninfected groundnut leaves disease with a high degree of accuracy of 98.72% and 98.22% respectively.

Keywords: Image Processing Convolutional Neural Networks (CNN's), Residual Network (ResNet50), Visual Geometry Group(VGG16), Bayesian Optimization (BO), Computer Vision.

INTRODUCTION

Detecting plant diseases using image processing technique has significant role in agricultural sector,timely detection of epidemics could enhance disease management, potentially mitigation reduce crop loss and control unnecessary expenses (Revati *et al*., 2023). Image processing and deep learning methods such as Convolutional Neural Networks(CNNs) have shown promising results in differentiating between healthy and disease plants in their early stages (Fereshteh *et al*.,2023). In recent times, Agriculture has provided an important contribution to the global economy as the human population continues to increase and urbanization leads to gradual reduction in cultivated land, the agricultural system faces increasingly pressure Hongkun *et al*. (2019).

There is a growing need for efficient and secure agricultural methods to meet the rising demand for food production, to address challenges in large crop production, advanced sensing and driving technologies, along with efficient information and communication technologies are being employed for use in this aspect (Imran Qureshi,2024). These advancements are essential for speeding up the increase in agricultural productivity with greater precision thus promoting the growth of high-quality and high-yield agricultural crops. Over the past decade computer vision inspection system have emerged as a vital tools in agricultural operations, expert and intelligent system based on computer vision algorithms are now common ground in agricultural production management (Aichen *et al*., 2019).

Plant leave disease detection plays an important role in human-computer interaction and AI-driven applications (Zekai *et al*., 2025). Despite many researcher's have focused on improving disease detection using machine learning and deep learning several challenges are yet to be addressed, variations in groundnut disease identification, lighting conditions, and image quality, performance evaluation of the optimizer in relation to other optimizers that are more frequently used (Vijai *et al*. ,2020).In addition there is a need to develop a hybrid CNNs architecture(ResNte50 and VGG16) that can preprocess groundnut leave disease images, extract meaningful features and accurately identify and classify various groundnut diseases.

The aim of this research is to develop a diagnostic system for groundnut disease using Convolutional Neural Network CNN((Resnet50 and VG16) and Bayesian Optimization, objectives are to identify and collect the dataset for the study and preprocess from Mendeley Data and evaluate the performance of the developed model.

Dataset is essential and integral part for the development and evaluation of groundnut leaf disease(Kosalairaman and Nirmala 2025). The dataset is useful to train and validate deep learning and machine learning algorithms for groundnut leaf disease classification and identification (Aishwarya and Padmanabha ,2023). The data.mendely.com introduced by Aishwarya and Padmanabha (2023) which will help in disease detection in groundnut plant that improve model performance and generalization capabilities.

MATERIALS AND METHODS

Data Description

The dataset for this study were identified and collected from Mendeley Data, It contain both healthy and disease leaves images.and it was chosen for this study due to its comprehensive collection of label groundnut leave diseases images. The images are categorized into Six main classes, representing the health status of the groundnut leaves (Table 1).

Table 1: Dataset of Groundnut leaf.

Category	Number of Images
Leaf	1871
Early leaf spot	1731
Late leaf spot	1896
Nutrition deficiency	1665
Rust	1774
Early Rust	1474
Total	**10,361**

Data Preprocessing

The image of groundnut leaves have been divided into six separate groups based on their condition. These groups consist of 1871 healthy leaf images, 1731 early leaf spot images, 1896 late leaf spot images, 1665 nutrition deficiency images, 1724 rust images and 1474 early rust images. Out of which 2380 images have been utilize for training the dataset while 818 have been use for testing purposes. The images were first resized to suit the model input and convolutional layers. A [224, 224, 3] images sizes have been used for both models. Next, features have been extracted by the Resnet50 and VGG16 CNN models. A total of 18000 images have been use.

CNN Models Training and Evaluation

The suggested groundnut leaf dataset was trained the CNN using hyperparameter acquired through Bayesian optimization. The dataset was sourced from Ashwaraya and Pandamanabha 2023) which contains 1871 leaf,1731 Early leaf spot,1896 late leaf spot, 1665 nutrition deficiency, 1774 rust and 1474 early rust, the total number data in the dataset is 10361 number of images, image categorized into six classes. For the validation of the groundnut leaf disease, only images of infected leaves were selected from the dataset. Image from the training have been employed to train a model, which have been selected to classify images from the test set. The validation dataset was used to monitor the models performance during training and evaluate its performance on the testing dataset in other to obtain unbiased performance metrics.

Performance Evaluation Metrics

The optimize CNN-Bayesian model for groundnut leaf disease detection have been assessed using the following metrics: Accuracy, Classification error, Precision and Recall.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$\text{Classification error} = \; = \frac{FP+FN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+TN} \quad (4)$$

$$\text{F1Score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}*\text{Recall}} \quad (5)$$

RESULTS AND DISCUSSION

Convolutional Neural Networks (CNNs) are given in this chapter. Specifically, two pre-trained architectures ResNet50 and VGG16 were used to perform transfer learning for disease classification tasks. Bayesian optimization was used to optimize hyperparameters such as learning rate, batch size, and dropout rate for enhancing model performance. The training and test dataset employed was obtained from the public data repository on Mendeley Data, where there are labelled images of groundnut leaves infected by various diseases. Python Programming environment along with TesnsorFlow and matplot libraries for integrating advanced CNN models such as ResNet50 and VGG16 with Bayesian Optimization for fine-tuning hyperparameters were used for the implementation.

To leverage transfer learning, we used ResNet50 and VGG16 as feature extractors. The models were pre-trained on the ImageNet database and used without the final (top) classification layers. The output of each model's convolutional base was retained so that meaningful image features related to groundnut leaf patterns could be learned. Table 2 shows some features extracted from images by the deep learning models (Resnet50 and VGG16) and were used in training the neural network model for identifying groundnut leaf diseases. 1000 features were extracted for each of the 1800 image used with 300 images from each category

Feature Extraction with Pretrained CNNs

Extracted Features for Each Class Using ResNet50 and VGG16

Table 2: Extracted Features per class using ResNet50 and VGG16

Class	ResNet50 - Feature 1	ResNet50 - Feature 2	ResNet50 - Feature 3	VGG16 - Feature 1	VGG16 - Feature 2	VGG16 - Feature 3
Healthy Leaf	0.563	-0.234	1.223	0.432	-0.567	0.312
Early Leaf Spot	-0.324	1.153	-0.782	0.231	1.112	-0.621
Late Leaf Spot	0.951	-0.527	0.738	0.856	-0.432	0.478
Nutrition Deficiency	0.032	-1.112	0.894	0.145	0.233	-0.367
Rust	1.432	0.547	-0.965	1.034	0.125	-0.512
Early Rust	-0.234	1.089	0.615	0.321	-0.213	0.845

Explanation of table 2

***ResNet50 Features**:* The first three feature columns are the values from the extracted feature vector with the ResNet50 model for each class.

***VGG16 Features**:* The last three columns are the values from the VGG16 model.

All feature columns correspond to the output of the GlobalAveragePooling2D layer after the images have been subjected to ResNet50 or VGG16, having reduced the high-dimensional output to an acceptable level.

t-SNE Feature Visualization

To better understand how well the ResNet50 and VGG16 feature extraction works, t-SNE was applied to map high-dimensional feature representations to 6D space. The model was optimized using Bayesian optimization with optimization search space of estimators(50,200), max dept(5,30) with Random state=42 .This provides a clean visualization of how well the CNN's have discriminated the different disease classes along the extracted features. Point Explanation: the closer the points are, the more similar their feature vectors are in the original 6D space, however the farther the point are, the more different they are.

It can be noted from Figure 1, the t-SNE plots ResNet50 achieves a clearer separation among some of the classes, notably between Early Leaf Spot and Late Leaf Spot. However, from Figure 2 ,VGG16 tends to have some overlap between Two Early Leaf Spots and Late Leaf Spot, suggesting that VGG16's extracted features may not be as discriminative as those obtained by ResNet50 for these classes.

Confusion Matrix for ResNet50

Figure 3 demonstrate the confusion matrix for ResNet50 model evaluation, shows the accurate and inaccurate classification counts across all the disease from the six classess.All classes have very high correct predictions for the values on the diagonal, ranges from 294–298.Nutritious Deficiency correct classification sample with minor errors followed by Rust and Early Rust achieved 296. The model performed well at identifying Nutrition Deficiency and Rust (>95% accuracy). Minor Confusions from Matrix: A few Healthy Leaf images were misclassified as Early Leaf Spot or Late Leaf Spot. Late Leaf Spot was wrongly confused with Early Leaf Spot 6 times. Two(2) misclassifications in Nutrition Deficiency.

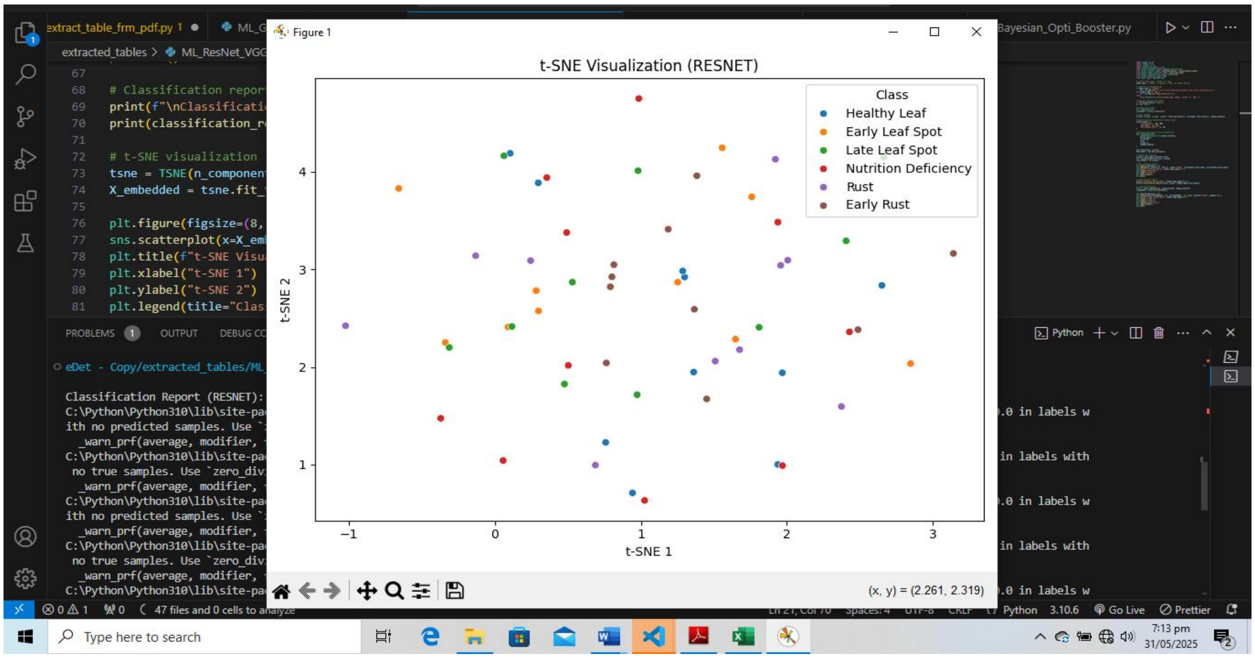

Figure 1: ResNet50 Visualisation.

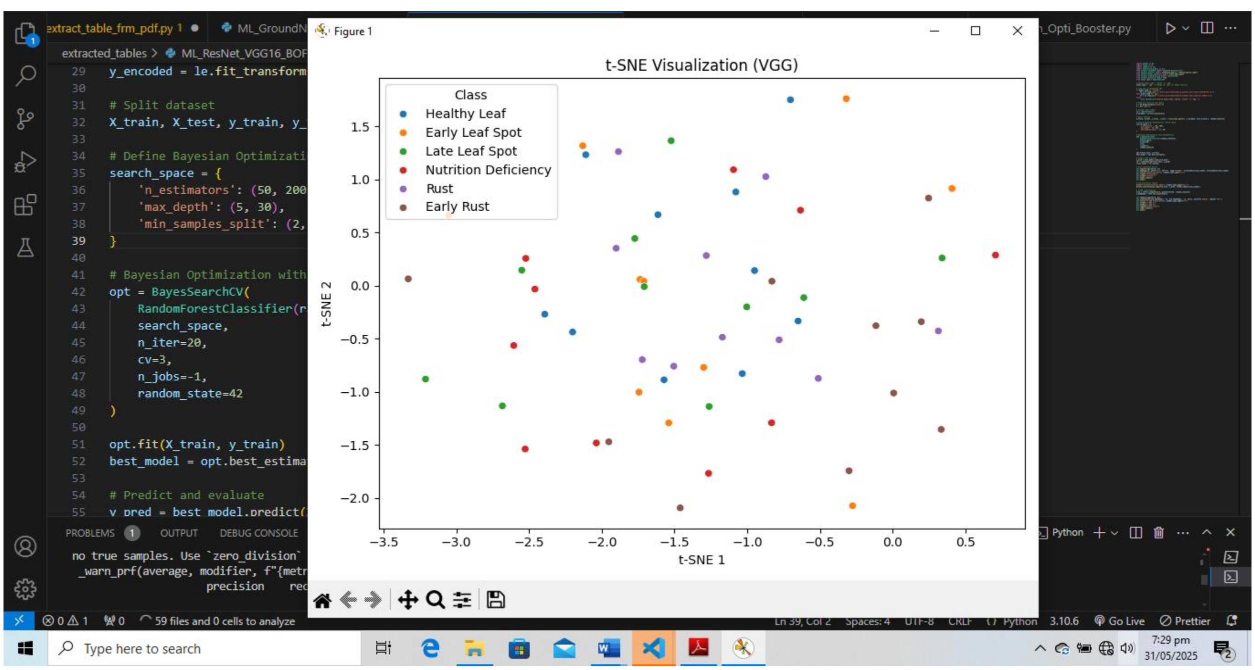

Figure 2: VGG16 Visualisation.

Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041
E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1295

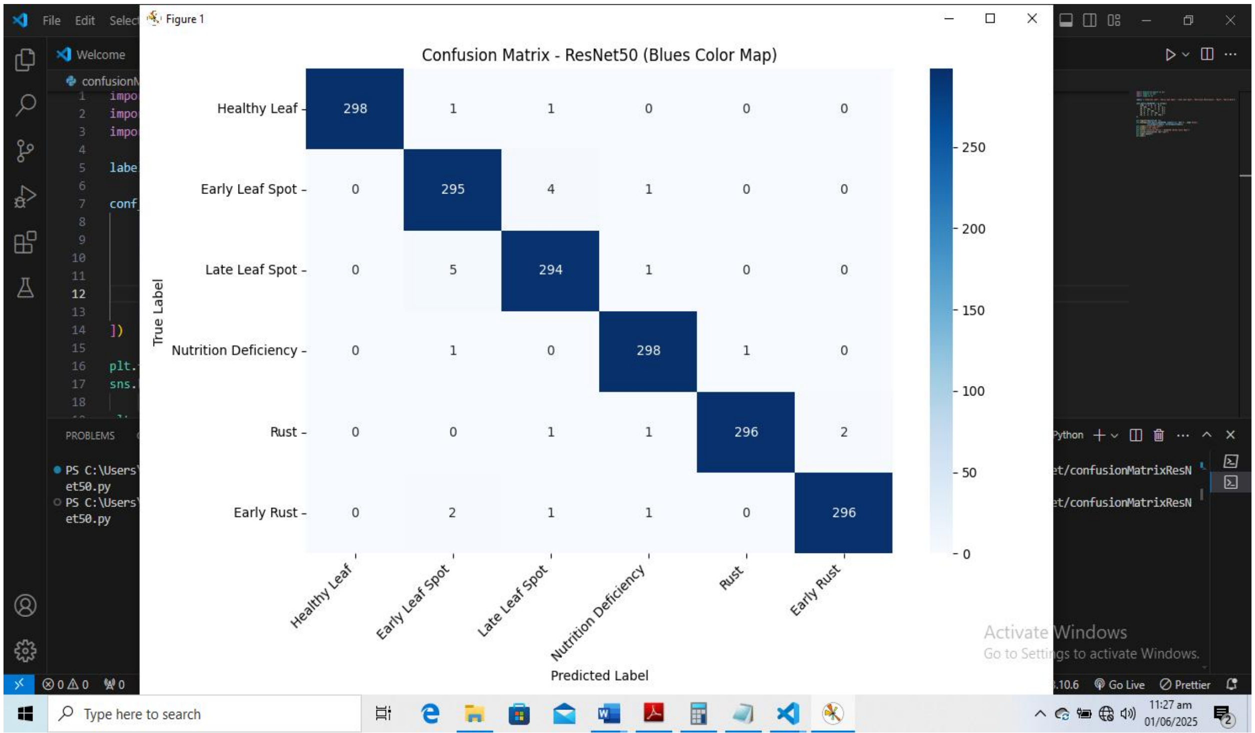

Figure 3: ResNet50 Confusion Matrix

Table 3: ResNet50 Evaluation Metrics per Class

Class	Precision	Recall	F1 Score
Healthy Leaf	1.0000	0.9933	0.9966
Early Leaf Spot	0.9547	0.9833	0.9687
Late Leaf Spot	0.9768	0.9800	0.9784
Nutrition Deficiency	0.9900	0.9933	0.9916
Rust	0.9966	0.9899	0.9932
Early Rust	0.9933	0.9866	0.9899

Evaluation Metrics for ResNet50 Result by Macro Averages

Macro = unweighted average across all classes
Accuracy 0.9872 ≈ 98.72%
Classification Error 0.0266
Precision 0.9852
Recall 0.9877
F1-Score 0.9864

Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041
E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1295

VGG16 Confusion Matrix

Figure 3 demonstrate the confusion matrix for ResNet50 model evaluation which shows the accurate and inaccurate classification counts across all the disease from the six classess.All classes have very high correct predictions for the values on the diagonal, ranges from 292–298. Nutritious Deficiency and Rust achieved 298 correct classification sample with minor errors. The model performed well at identifying Nutrition Deficiency and Rust (>94% accuracy).

Minor Confusions from Matrix: A few Healthy Leaf images were misclassified as Early Leaf Spot or Late Leaf Spot. Late Leaf Spot was wrongly confused with Early Leaf Spot 5 times. 2 misclassifications in Nutrition Deficiency.

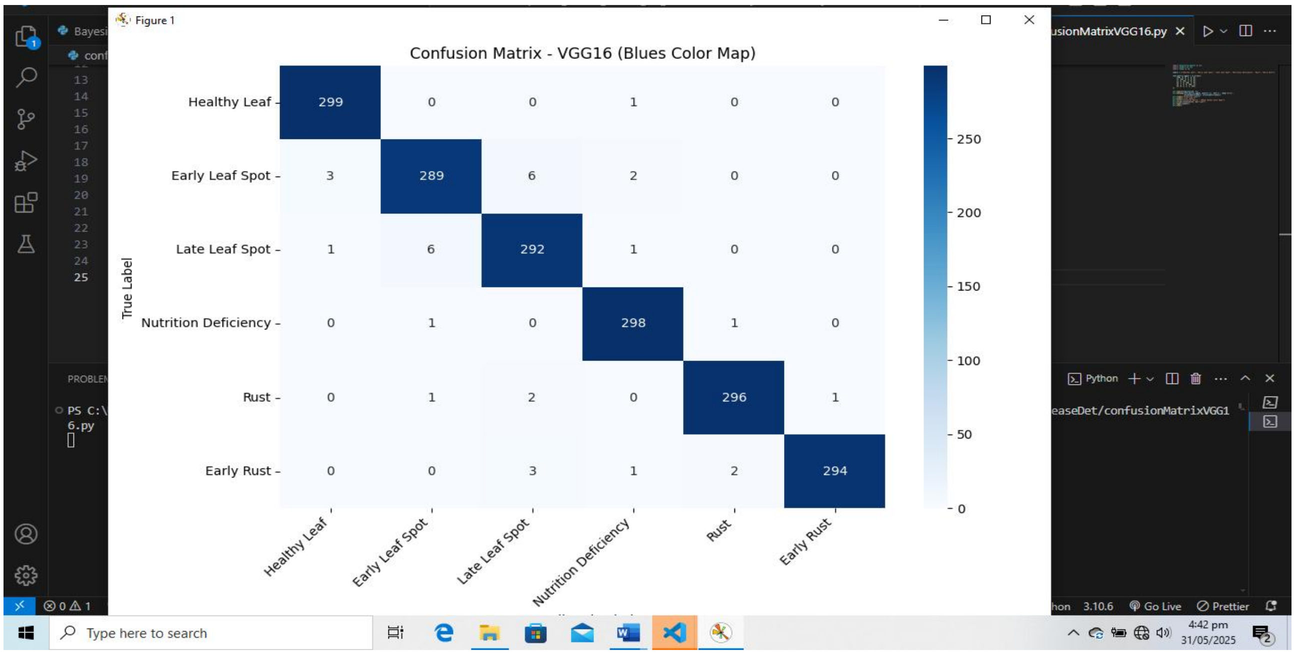

Figure 4: 5 VGG16 Confusion Matrix.

Table 4: VGG16 Evaluation Metrics Per Class

Class	Precision	Recall	F1 Score
Healthy Leaf	0.9868	0.9970	0.9919
Early Leaf Spot	0.9731	0.9633	0.9682
Late Leaf Spot	0.9637	0.9733	0.9685
Nutrition Deficiency	0.9835	0.9933	0.9884
Rust	0.9899	0.9866	0.9882
Early Rust	0.9966	0.9800	0.9882

Evaluation Metrics for VGG16 Result

Accuracy 0.9822 ≈98.22%
Classification Error 0.0766
Precision 0.9823
Recall 0.9823
F1-Score 0.9822

Performance Comparison

T able 5 Displays the models comparison with DenseNet-169 accuracy.

Table 5: Models(ResNet50 and VGG16) Comprison with DenseNet-169

Models	Classification	Confusion Matrix	Accuracy
Densenet-169	5 Classes	Not clear as result of misclassifications across Nutrition Deficiency and Rust Leaf,Late Leaf spot and Rust as Early Rust	99.83%
ResNet50/VGG16	6 Classes	The Models demonstrated highly level of accurracy in classifying Nutrition deficiency 99.3%, Rust and Early Rust 98.7% Respectively	98.72% and 98.22% Respectively

DISCUSSION

The objective is to increase the efficacy of identification and classification of groundnut leaf diseases with the help of deep learning models, ResNet50 and VGG16 convolutional neural network models with Bayesian optimization method.The chosen architecture has established a new standard, with state-of-the-art accuracy levels of 98.72% and 98.22% for ResNet50 and VGG16 respectively.

CONCLUSION

The performance evaluation of the RESNET50 and VGG16 models for disease classification demonstrates excellent performance, where both models have large True Positive Rates (TPR) and small False Negative Rates (FNR) for different disease classes. RESNET50 outperforms VGG16 all the time, demonstrating its efficiency in identifying significant features for disease Identification. Both models excel in early disease detection, particularly in distinguishing between Nutrition Deficiency, Early Rust, and Rust. There are performance gaps among certain classes, where RESNET50 has higher accuracy for Early Leaf Spot, late Leaf Spot and Early Rust. Generally, the outcomes show the promise of deep learning models in plant disease. diagnosis, RESNET50 being one of the Favorite choices due to its better performance.

REFERENCES

Aichen W., Wen .Z & Xinhua W. (2019, February 05). A Review on Weed Detection Using
Ground-Based Machine Vision and Image Processing Techniques. *Computer and Electronics in Agriculture*, pp. Vol 158 pp 226-240. https://www.sciencedirect.com/science/article/pii//s0168169918317150()

Aishwarya M.P. & Padmanabha R. (2023). *Dataset of Groundnut Plant Leaf for Classification and Detection*. Retrieved August 23, 2023, from Journal Home Page: www.elsevier.com/locate/dib

Fereshteh S., Hamed T.G., Seyedmojtaba S, Seyedal .H, Samuel M. and Fartash V. (2023). *Application of Image Processing and Transfer Learning for*

Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1295

Detection of Rust Disease. India: Science Report 23-5233. https://doi.org/10.1038/s41598-023-31942-9

Hongkun T., Tianhai .W, Yadong .L, Xi .Q, & Yanzhou .L. (2019). *Computer Vision Technology in Agricultural Automation- A Review.* Guangzhou, China: Information Processing in Agriculture. www.elsevier.com/locate/inpa

Imran Qureshi (2024). Integrating Few-Shot Learning and Multimodal Image Enhancement in GNut: A Novel Approach to Groundnut Leaf Disease Detection *https://www.mdpi.com/journal/computers.https://doi.org/10.3390/computers13120306*

Journal of Harbin Engineering University, 44 (5).

Revati R.N.,Sawant S.D., Ingle P.M. and More V.(2023). Detection of Early Leaf Spot of Groundnut Using Neural Network. *International Research Journal of Engineering and Technology*, 10, 2395-0072.

Kosalairaman T. and Nirmala A.(2025). Prediction of Groundnut Leaf Disease Detection and Classification Using Augmented Capsule Neural Network. *DOI: https://doi.org/10.22399/ijcesen.771.Dr NGP Arts and Science college*

Vijai .S, Namita .S, and Shikha .S.(2020). A Review of Imaging Techniques for Plant Disease Detection . *Journal of Artificial Intelligent in Agriculture,* 229-242.

.

Zekai Lv, Shangbin Yang, Shichuang Ma, and Hui Zhang (2025).Efficient Deployment of Peanut Leaf Disease Detection Models on AI Devices. *Multidisciplinary Digital Publishing Institue (MDPI). https://doi.org/10.3390/ agriculture15030332.*