RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Issue 3: 428-435 (March 2026)

Generation and Utilization of an Augmented Cashew Leaf Dataset for Disease Analysis using Transfer Learning

Sumit Dey[1], Raj Kumar Goel[1], Arvind Kumar Chaurasiya[2], Rubul Kumar Bania[3]

10.18805/IJARe.A-6479

ABSTRACT

Background: The scarcity of publicly available datasets for the detection of cashew leaf disease (DCLD) (*Anacardium occidentale* L.). This study represents the development and evaluation of cashew leaf disease dataset of three districts. Here, Dataset_1 with 1274 original images and Dataset_2 with 2548 augmented images to increase its diversity. Images were capture using smartphone with natural field conditions, background, various resolutions, angles, Pre-processing and data augmentation techniques were applied to enhance model performance.

Methods: Multiple deep learning models, including DenseNet121, EfficientNetB0, restNet50 and Vision Transformer were trained using transfer learning to classify four cashew leaf condition: Anthracnose, Healthy, Leaves Spot and Powdery Mildew.

Result: DenseNet121 achieved high class wise precision and recall, with overall accuracy between 96% to 97%. EfficientNetB0 improved upon this result, attaining overall accuracy between 96% and 98% with F-1 score up to 97%. RestNet50 and ViTB16 achieved the highest performance, with over accuracy to 99.93% respectively.

Key words: Anthracnose, DenseNet121, EfficientNetB0, Leaves spot, Powdery mildew, Vision transformer.

INTRODUCTION

The cashew (*Anacardium occidentale* L.) has become one of the most commercially important cash crops in the tropics of India. It's economically and nutritionally rich and even through its exports oriented crop. However, peoples are significantly engaged with production to processing and also provides livelihood to millions of farmers. India has become one of the world's largest producers of cashew traditionally. Further, cultivation occurs mainly in the states of Kerala, Maharashtra, Odisha, Andhra Pradesh, Jharkhand, West Bengal and the north-eastern states, including Meghalaya (Sarath *et al.*, 2025; Srinivasarao *et al.*, 2022).

A few diseases like Anthracnose, Powdery Milde and Leaf Spot can cause serious damages to a plant. In addition to these three diseases maintaining healthy leaves can also be useful in baseline monitoring. Dead spots could result in leaf dropping or wilting. Anthracnose leads to dead spots with dropping of leaves and powdery mildew appears as white powdery fungal growth on the upper surface of leaf which weakens plant vigor. Further, leaf spot leads to scattering of necrotic patches. If, it is not done early detection and treatment effectively, this could lead to rapid spreading of diseases causing yield losses in cashew farms and deteriorate the quality of nuts (Adeniyi and Asogwa, 2023; Bisht and Roy, 2025; Mehta *et al.*, 2025).

Traditional disease detection in cashew plantations relies primarily on manual inspection by farmers, agronomist and plant pathologists. While effective to some extent, manual diagnosis is often time-consuming, subjective and prone to errors. Further more, the overlapping visual symptoms of multiple diseases make differentiation challenging for non-experts. Large-scale monitoring of plantations, particularly in diverse agro-climatic regions like Meghalaya, is nearly impossible through human observation alone. These challenges emphasize the need for automated, scalable and reliable diagnostic systems for DCLD (Sandhi *et al.*, 2025; Singh *et al*., 2023).

Finding sick plants in cashew farms has usually been done by people looking at them, like farmers, farm experts and plant doctors. Although this works somewhat, checking by hand usually effort-demanding, depends on who is looking and can have mistaken. Also, because different sicknesses can look similar, it can be hard for people who aren't experts to tell them apart. Watching over big farms, especially in different growing areas such as Meghalaya, is falmost impossible just by having people look. These problems highlight how important it is to have automatic,

[1]Department of Computer Application, North-Eastern Hill University, Tura-794 002, Meghalaya, India.
[2]Department of Horticulture, North-Eastern Hill University, Tura-794 002, Meghalaya, India.
[3]Department of Computer Science, Birangana Sati Sadhani Rajyik Vishwavidyalaya, Golaghat-785 621, Assam, India.

Corresponding Author: Raj Kumar Goel, Department of Computer Application, North-Eastern Hill University, Tura-794 002, Meghalaya, India. Email: raj9921@yahoo.com
ORCID: https://orcid.org/0000-0001-9684-060X

How to cite this article: Dey, S., Goel, R.K., Chaurasiya, A.K. and Bania, R.K. (2026). Generation and Utilization of an Augmented Cashew Leaf Dataset for Disease Analysis using Transfer Learning. *Indian Journal of Agricultural Research.* **60(3):** 428-435. doi: 10.18805/IJARe.A-6479.

Submitted: 04-11-2025 **Accepted:** 13-01-2026 **Online:** 22-01-2026

easy-to-use and trustworthy systems for finding diseases in cashew leaves (Demilie, 2024).

Artificial intelligence (AI), machine learning (ML) and deep learning (DL) have provided transformative solutions to such agricultural challenges. DL models, exclusively Convolutional Neural Networks (CNNs), have proved leading edge performance in plant disease classification tasks using leaf images (Mohanty *et al.*, 2016; Goel and Vishnoi, 2025). These models can automatically extract hierarchical features, enabling them to distinguish between subtle disease symptoms with high accuracy. Similarly, Vision Transformers (ViTs) have emerged as promising architectures for agricultural image classification. Furthermore, the use of transfer learning with pretrained models such as EfficientNet, DenseNet, ResNet and Inception has enhanced performance, allowing models to generalize effectively (Dogra *et al.,* 2023; Pakruddin and Hemavathy, 2025; Dinkar *et al*., 2025).

Data augmentation techniques are widely used to artificially expand training datasets and reduce overfitting. These techniques help ensure that models are more robust to changes in lighting, orientation and leaf shape (Wang *et al.,* 2025). However, one of the biggest challenges in developing reliable AI-based disease detection systems in agriculture is the availability of domain-specific and region-specific datasets. While benchmark datasets exist for crops such as rice, corn, tomato and potato (Alomar, 2023; Rastogi and Singh, 2024; Topi *et al*., 2025).

To fill this gap, we present a dataset focused on the analysis of cashew leaf diseases for computational purposes. Three Garo Hils districts of Meghalaya, India were selected due to their large cashew plantations and prevalence of various plant diseases. It contains a large number of images taken in these region using smartphones. This dataset is a valued resource for scholars, helping them develop advanced algorithmic models.

The proposed classification and identification of foliar diseases, as well as the datasets used, are described in Section 2. A description of developed methodology is detailed in Section 3, while results and discussion is given in Section 4, followed by conclusions and future scope.

MATERIALS AND METHODS

Three Garo Hills districts of Meghalaya, India, were selected due to their large cashew plantations and the prevalence of various plant diseases. Fieldwork lasted several months, (20 months in 2023-25) during which various leaf samples were collected under different environmental conditions and at different growth stages. The research investigation was conducted in the Department of computer application, North-Eastern Hill University Tura Campus, Meghalaya. Fig 1 showing the district's on the map of Meghalaya, India from where the image data of Cashew Leaf collected.

To obtain high-quality images suitable for DL/ML, a smartphone with a 64-megapixel camera was used. This choice provided a good balance between portability and image resolution, facilitating the capture of detailed leaf characteristics necessary for accurate disease classification. The resulting images had a resolution of 4640 × 2610 pixels and 72 dpi (Fig 2).

Table 1 summarizes recent related studies using various plant disease datasets. It highlights the attributes, number of samples and classes, as well as the dataset creator. The specification table presents the dataset's key characteristics and metadata, including the subject area, data type and data collection process. It also provides details about the data format, storage method and source location (Table 2).

The dataset setup workflow, depicting the sequential procedure followed for image acquisition, preprocessing, annotation and dataset structuring. Each phase ensures data integrity, uniformity and suitability for subsequent model training and performance evaluation (Fig 3).

Causes of the cashew leaf diseases

Some varieties are susceptible to several leaf diseases that reduce photosynthetic capacity and fruit yield. The main causes of these diseases are fungal pathogens, bacterial infections and environmental circumstances. Fungal pathogens are the greatest culprits, which attack leaves under various conditions like high humidity and warm temperatures in addition, abiotic stress factors such as prolonged leaf wetness, soil moisture imbalance and wounds caused by pruning or insect feeding create entry



Fig 1: Leaves images of cashew plants collected from the 3 districts of Meghalaya.



Fig 2: GPS photos of data collection method.

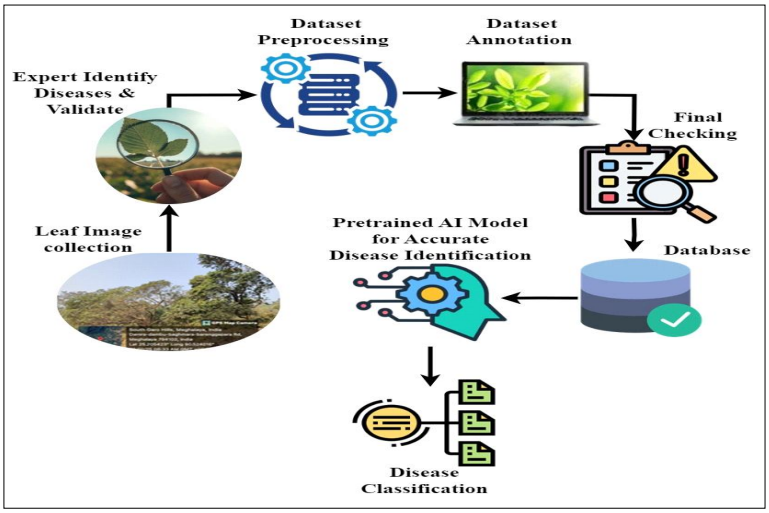


Fig 3: Dataset setup workflow.

Table 1: Recent related work on various dataset.

Year	Attributes	No. of samples	No. of classes	Dataset created
2022	High resolution images	1835	30	Mendeley dataset (Divyasree and Sheelarani, 2022).
2022	Plain background (PB)	-	4	Author-generated (Kaur and Singh, 2022).
2023	Uniform illumination, PB	300	18	Author-generated (Thanikkal *et al.*, 2023)
2022	PB	2515	40	Author-generated (Roopashree *et al.*, 2022)
2022	PB	1500	50	Author-generated (Rao *et al.,* 2022)
2023	PB	1300	10	Author-generated (Thanikkal *et al.*, 2023)
2023	Plain and non varying intensity levels/varying background	6900/5900	80/40	Author-generated (Pushpa and Rani, 2023)
2024	Different background	2159	5	Author-generated (Mustofa *et al.*, 2024)

Table 2: Specification table.

Subjects	Computing and agricultural field.
Particular subject area	Computer vision, image processing and DCLD.
Types of data	Image file with 1274 images.
Data capture process	In this study, no fixtures were used to capture images; instead, we considered using real smartphone-based image capture scenarios to photograph plant leaves, the photos were taken in different weather conditions, at different locational distances, lighti ng conditions of different seasons.
Data format	JPEG file.
Details of data collection	The images captured for the dataset present real-world challenges for plant disease identification and recognition. Only cashew leaves species were photographed from different projection angles using a fixed-lens smartphone camera, illustrating differences in leaf color, multiple resolutions and different distances. The dimension and resolution of collected images is 4640 × 2610 pixels and 72 dpi.
Data source location	Three districts namely Garo Hills in Meghalaya, India, were selected due to their extensive cashew plantations and the prevalence of various plant diseases. The fieldwork spanned several months, during which leaf samples were collected under varying environmental conditions and at various growth stages.
Data accessibility	Data can be access from digital repository.
Any article related to the dataset	None of the article published on this dataset.
Data validation	With the help of a plant pathologist/ professionals/ expert, the images were subsequently classified into 4 categories namely, Anthracnose, Leaves spot, Powdery mildew and Healthy leaf. Manual annotation of images was done to record disease types and their spatial locations; moreover, the annotations images were saved in the database for usage with various Artificial Intelligence models.

points for pathogens and increase infection (Akinwale and Esan, 2021). Here, four categories of cashew leaf diseases are as described below.

Powdery mildew

It is widespread and unhelpful diseases affecting cashew caused by *Erysiphe quercicola* (formerly *Oidium anacardii*). It manifests as white type, powdery fungal growth on the upper portion of leaves, young shoots and inflorescences. Infected tissues become diluted and make infections lead to leaf curling, yellowing, early defoliation and suppression of new growth. When the pathogen attacks inflorescences, it causes flower abortion, resulting in poor fruit and nut set (Smith *et al.*, 1995).

Anthracnose

It is caused by *Colletotrichum gloeosporioides* and devastating foliar syndrome of cashew. It typically begins as small, water-soaked spots with sunken on young or mature leaves, which later expand into irregular brown to black necrotic lesions. As the infection advances, the lesions may coalesce, leading to extensive leaf blight and premature defoliation (Monteiro *et al.*, 2022).

Leaf spot

Leaf spot disease is a common foliar problem in cashew plantations, primarily caused by *Pestalotiopsis* species, with recent reports identifying *Neopestalotiopsis clavispora* as an emerging pathogen. Initial symptoms typically appear as small, circular to irregular brown lesions scattered across the leaf surface. As the disease progresses, these lesions enlarge and may coalesce, forming extensive necrotic patches (Manjunatha *et al.*, 2023).

Table 3: Distribution before and after augmentation.

Cashew leaf diseaseclass	Original images dataset_1	After augmentation dataset_2
Healthy	375	637
Powdery mildew	222	637
Leaf spot	275	637
Anthracnose	402	637
Total	1274	2548

Healthy

These leaves observed appear completely healthy. It has a smooth texture, consistent light green color with lusture and no visible signs of damage or infection. The veins are clearly defined and the leaf is intact without any deformities or discoloration. Healthy leaves such as this one are a positive sign of good plant health and effective crop management (Akinwale and Esan, 2021). Fig 4 demonstrate each class images of the datasets.

Methodology

Data augmentation not only addresses the class imbalance problem but also advances the model's generalization capability by enabling it to adapt to different variations of the same leaf state. Therefore, the augmented dataset provides a more solid foundation for deep learning classification, reduces overfitting and improves the model's strength in real-world scenarios. Table 3 shows the distribution before and after augmentation.

Data augmentation

To balance the class distribution and increasing dataset diversity data augmentation using the Keras Image Data Generator was performed. The augmentation approaches are as:

- Rotation up to 20°.
- Shifting width and height by 10%.
- Shearing up to 10%.
- Scaling within a 20% range.
- Horizontal flipping.

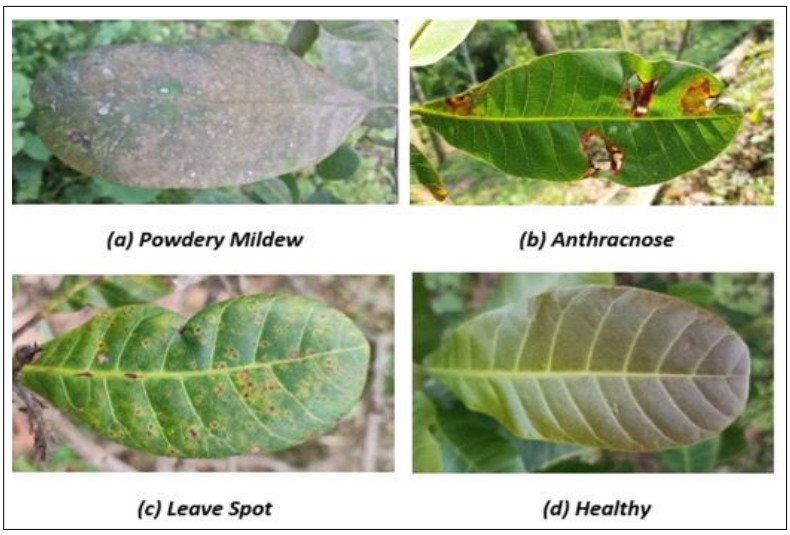


Fig 4: Demonstrate each class images of our dataset.

- Filling empty pixels using nearest neighbors after the transformation.

These transformations produce synthetic yet realistic variations of the original leaf images and enrich the dataset with new samples that reflect real-world variations.

Balancing the dataset

The exact number of images per class was generated to 637, resulting in a balanced dataset covering all four classes.

Dataset annotation by expert

The collected leaves data were classified into four categories manually by agricultural professional according to their respective diseases: “Leave-spot”, “Anthracnose”, “Powdery Mildew” and “Healthy”. This classification was performed in collaboration with agricultural professionals and is based on theoretical/practical evidence gathered prior to implementing the data collection methodology. Dataset annotation process is shown in Fig 5. The complete classification criteria are outlined below.

This pipeline study includes of three phases. In the initial phase, data gathering and data pre-processing are performed. In the second phase, four different DTL based models are used to extract various features from the input images during the training process (Fig 6). Finally, in the third phase, the classification task is performed using a fully connected layer. Those four models were selected because they represent a balanced combination of efficiency (EfficientNet-B0), feature richness (DenseNet-121), reliability (ResNet-50) and innovation (ViT-B16). Together, they can provide a comprehensive set of transfer learning approaches for evaluating the cashew nuts datasets.

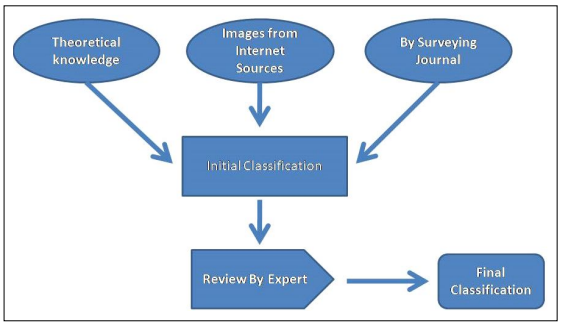


Fig 5: Dataset annotation process.

The accuracy of a deep CNN model is largely depended on quality of the dataset used for training. To ensure reliable performance, a through data cleaning process is carried out after data collection. This process involves removing any faulty or irrelevant images from the dataset. Furthermore, all images are resized to a uniform dimension of 224×224 pixel. Which help reduce computational complexity during training and enhance overall performance of the models. Some of the details of the applied DTL models are given below.

EfficientNet-B0

EfficientNet-B0, developed using a novel compound scaling method that uniformly balanced a network’s depth, width and resolution in a systematic and efficient manner. Unlike traditional scaling methods that arbitrarily increase these dimensions, EfficientNet implements a balanced scaling strategy that improves accuracy at low computational costs. EfficientNet-B0, being a baseline version, provides an excellent balance between model size, computational efficiency and accuracy, making it highly deployment.

DenseNet-121

DenseNet-121 is a deep CNN that introduces the concept of dense connectivity. Here. Its every layer is receiving feature

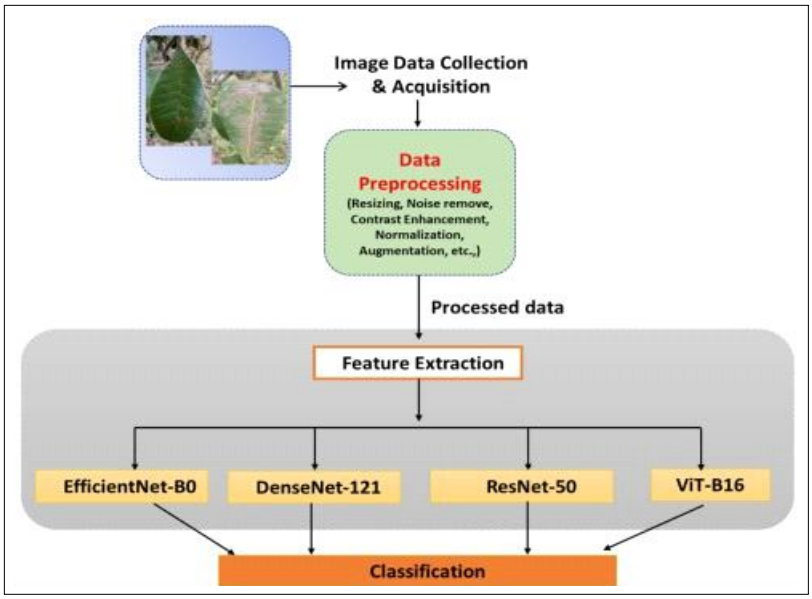


Fig 6: Workflow diagram for DCLD.

maps from its all-preceding layers and sends its feature maps to all subsequent layers. The 121-layer version strikes a balance between efficiency and depth, achieving high performance with low redundancy and is particularly effective in tasks requiring rich feature representation.

ResNet-50

ResNet-50 is a widely adopted CNN architecture that presented the concept of residual learning through shortcut (or skip) connections. The "50" in ResNet-50 refers to its 50 convolutional layers, making it a moderately deep model that balances computational cost and accuracy. ResNet-50 has become a benchmark model in computer vision, demonstrating strong generalization across diverse image classification and feature extraction tasks.

Vision transformer (ViT-B16)

ViT-B16 represents a significant departure from CNNs by adopting the transformer architecture, originally designed for NLP, to computer vision tasks. This model split the input image into fixed-size patches (in this case, 16×16. It is offering superior performance on large-scale datasets. The ViT-B16 variant, with its baseline configuration and 16-pixel patch size, strikes a stability between programmatic efficacy and correctness in vision tasks.

The accuracy of a deep CNN model is largely depended on quality of the dataset used for training. To ensure reliable performance, a through data cleaning process is carried out after data collection. This process involves removing any faulty or irrelevant images from the dataset. Furthermore, all images are resized to a uniform dimension of 224×224 pixel. Which help reduce computational complexity during training and enhance overall performance of the models.

RESULTS AND DISCUSSION

A custom image dataset of cashew (*Anacardium accidentale* L.) leaf disease was developed from field samples collected from the Gro Hills region of the Meghalaya India. The dataset was systematically classified by us into four distinct classes "Healthy", "Anthracnose", "Powdery_Mildew" and "Leaves_Spot". To the evaluate the robustness of the dataset and the impact of the image enhancement two separate version were prepared Dataset_1 original dataset with original images and Dataset_2 augmentation images. This enhancement increases the diversity, realistically capturing variation in leaf orientation and natural deformation, which is critical for improving model performance as shown in the Table 4.

We employed four DL architecture, like DenseNet121, EfficeineNetB0, RestNet50 and Vision Transformer (VitB16) and perform classification through transfer learning. All model were trained alone for 20 epochs and performance was evaluated through standard metrics including precision, recall, F-1 score and overall accuracy. All model showed significant performance and improvements after applying data augmentation, confirming the positive impact of dataset expansion and diversity. Densent121 improves from 87% to 95% accuracy, reflecting stronger learning stability on the datasets after augmentation. EfficentNetB0 gain from 91% to 97% confirming its efficiency in handling augmented, complex image features. RestNet50 achieved from 94% to 99.87% accuracy demonstrating excellent feature extraction and resilience against overfitting. Vit-B16, a transformer-based model, gained the highest accuracy 99.93% after augmentation, proving its strong capacity to leverage spatial and contextual features in plant leaf images. Below Fig 7 shows the before and after augmentation performance.

Table 4: Dataset_1 (original) and Dataset_2 (augmented), classification into four categories.

Model	Epoch	Dataset class	Precision in %		Recall in %		F1-score in %		Accuracy in %	Accuracy in % augmented
DenseNet121	20	Anthracnose	94	97	78	97	85	97	87	95
		Healthy	90	95	96	99	93	97		
		Leave_spot	74	98	83	92	93	95		
		Powdery_mildew	93	92	87	94	90	93		
EfficentNetB0	20	Anthracnose	97	98	84	97	90	98	91	97
		Healthy	94	95	96	96	95	96		
		Leave_spot	79	98	90	97	84	98		
		Powdery_mildew	93	96	93	98	93	97		
RestNet50	20	Anthracnose	1	95	86	98	93	97	94	99.87
		Healthy	94	92	1	96	97	94		
		Leave_spot	85	95	93	91	89	93		
		Powdery_mildew	1	96	93	90	97	93		
VitB16	20	Anthracnose	1	90	89	98	94	94	91	99.93
		Healthy	90	86	98	94	94	90		
		Leave_spot	83	1	83	74	83	85		
		Powdery_mildew	93	85	93	92	93	88		

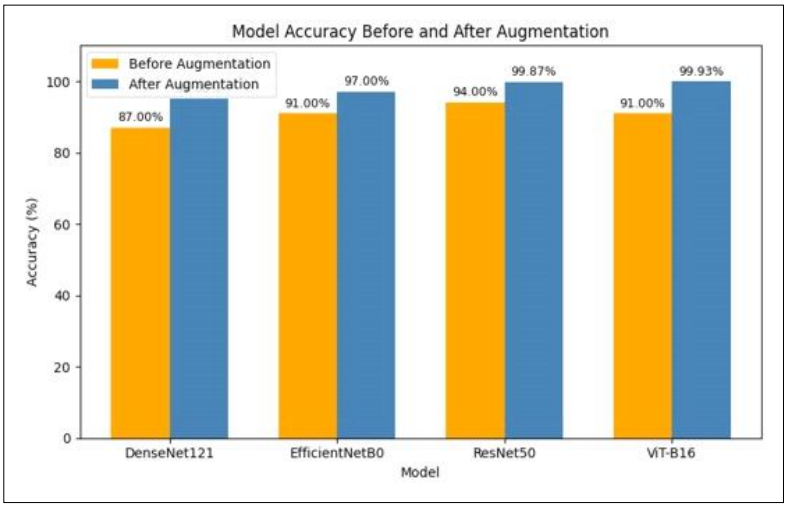


Fig 7: Before and after data augmentation performance.

Challenges and limitation

Collecting a diverse dataset of cashew leaves in the filed presents many challenges including geographical barriers: remote plantation is difficult to reach due to poor transportation rugged terrain. Environmental variability: changes in light, shadow and weather can lead to inconsistent image quality. Certain disease only appear in certain season, causes repeated field investigation. Same disease can present different patterns depending on severity, leaf age and stress which complicate accurate in labeling the data. A limitation of the dataset is class imbalance, as healthy leaves were more abundant than infected leaves. Environmental factors such as variations in lighting and background clutter presented additional challenges during the data collection phase.

CONCLUSION

To ensure accurate labeling of disease categories and lesion locations, the dataset was manually annotated with the assistance of domain experts. Each image was carefully annotated with detailed information about the type and condition of disease symptoms and these annotated images were stored in a structured database for seamless integration with various AI models. After applying data augmentation techniques, substantial improvements in model performance were observed, validating the positive impact of dataset expansion and diversity. The consistent performance increase across all models strongly indicates that the prepared cashew leaf dataset is accurately labeled, robust and suitable for ML and DL applications.

Both manual and computational validation confirm that the dataset is stable, reliable and scientifically rigorous. The interdisciplinary approach integrating expertise from plant pathologist/ professionals/ experts and computational modeling further strengthens and credibly reinforces the dataset's robustness and applicability for advanced agricultural research. The DenseNet121 model showed an increase in accuracy from 87% to 95%, while ResNet50 achieved a significant increase from 94% to 99.87%, demonstrating excellent feature extraction efficiency and resilience to overfitting. The Transformer-based model ViT-B16 achieved the highest accuracy of 99.93% after augmentation.

Disclaimers

The authors are responsible for the accuracy and completeness of the information provided, but do not accept any liability for any direct or indirect losses resulting from the use of this content.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Adeniyi, D.O. and Asogwa, E.U. (2023). Dynamics of diseases and insect pests of cashew tree. *Forest Microbiology.* pp. 265-284.

Akinwale, M.G. and Esan, E.B. (2021). Diseases of cashew (*Anacardium occidentale* L.): A review of pathogens, symptoms and management strategies. *Journal of Plant Pathology.* **103(1):** 1-14.

Alomar, K. (2023). Data Augmentation in Classification and Segmentation: Recent reviews and applications. *International Journal of Computer Vision Applications.* **9(2):** 46. https://doi.org/10.3390/jimaging9020046.

Bisht, S. and Roy, S. (2025). Optimizing role assignment for scaling innovations through AI in agricultural frameworks: An effective approach. *Advanced Agrochem.* **4(2):** 106-113.

Demilie, W.B. (2024). Plant disease detection and classification techniques: A comparative study of the performances. *Journal of Big Data.* **11(1):** 5.

Dinkar, S., Jayapriya K., Pallerla, N., Jadhav, D. Kunal and Shet, R.A. (2025). Convolutional neural networks for the intelligent and automated detection of mango leaf disease to enhance crop health management. *Agricultural Science Digest.* **45(6):** 1004-1010. doi: 10.18805/ag.DF-717.

Divyasree, G. and Sheelarani, M. (2022). Ayurvedic leaf identification using deep learning model: VGG16. Available at SSRN 4091254.

Dogra, R., Rani, S., Singh, A., Albahar, M.A., Barrera, A.E. and Alkhayyat, A. (2023). Deep learning model for detection of brown spot rice leaf disease with smart agriculture. *Computers and Electrical Engineering.* **109:** 108659.

Goel, R.K. and Vishnoi, S. (2025). Space-age evolution-remote sensing and IoT for productive and sustainable agricultural landscape. *Sustainable Futures.* **10:** 101280.

Kaur, P.P. and Singh, S. (2022). Random Forest Classifier used for Modelling and Classification of Herbal Plants Considering Different Features using Machine Learning. In Mobile Radio Communications and 5G Networks: Proceedings of Second MRCN 2021. Singapore: Springer Nature Singapore. (pp. 83-94).

Manjunatha, S.V., Nalini, M.S. and Shankara, H.N. (2023). First report of neopestalotiopsis clavispora causing leaf blight of cashew in India. *Plant Disease.* **107(8):** 2148.

Mehta, R.A., Kumar, P., Prem, G., Aggarwal, S. and Kumar, R. (2025). AI-powered innovations in agriculture: A systematic review on plant disease detection and classification. *Indian Journal of Agricultural Research.* **59(9):** 1321-1330. doi: 10.18805/IJARe.A-6371.

Mohanty, S.P., Hughes, D.P. and Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science.* **7:** 1419.

Monteiro, F., Santos, F. and Silva, L. (2022). Disease-causing agents in cashew: A review. *Agronomy.* **12(10):** 2553.

Mustofa, S., Ahad, M.T., Emon, Y.R. and Sarker, A. (2024). BD papaya leaf: A dataset of papaya leaf for disease detection, classification and analysis. *Data in Brief.* **57:** 110910.

Pakruddin, B. and Hemavathy, R. (2025). Performance evaluation of deep learning models for multiclass disease detection in pomegranate fruits. *Indian Journal of Agricultural Research.* **59(10):** 1535-1542. doi: 10.18805/IJARe.A-6396.

Pushpa, B.R. and Rani, N.S. (2023). DIMPSAR: Dataset for Indian medicinal plant species analysis and recognition. *Data in Brief.* **49:** 109388.

Rao, R.U., Lahari, M.S., Sri, K.P., Srujana, K.Y. and Yaswanth, D. (2022). Identification of medicinal plants using deep learning. *Int. J. Res. Appl. Sci. Eng. Technol.* **10:** 306-322.

Rastogi, R. and Singh, P. (2024). Fruit disease detection using colour, texture and ANN: A sustainable approach for smart cities. *International Journal of Agriculture Innovation, Technology and Globalisation*. **4(1):** 63-96.

Roopashree, S., Anitha, J., Mahesh, T.R., Kumar, V.V., Viriyasitavat, W. and Kaur, A. (2022). An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach. *Measurement*. **200:** 111484.

Sandhi, A., Kumar, R., Bhardwaj, R., Kumar, D., Rana, A.K., Ajala, O., Deepak, A. and Salau, A.O. (2025). Optimized deep learning framework for pomegranate disease detection using nature-inspired algorithms. *Plant Methods.* **21(1):** 1-36.

Sarath, C.M.C, Sahib, P.R., Achuthan, K. and Chandran, R. (2025). From farm to table: India's transition towards cashew production and sustainable livelihood of smallholder farmers. *International Journal of Agricultural Sustainability*. **23(1):** 2557083.

Singh, M., Aulakh, S.S. and Bimbraw, S.A. (2023). Manmade problems in indian agriculture and their solutions: A review. *Agricultural Reviews*. **44(3):** 343-349. doi: 10.18805/ag.R-2382.

Smith, D.N., King, W.J., Topper, C.P., Boma, F. and Cooper, J.F. (1995). Alternative techniques for the application of sulphur dust to cashew trees for the control of powdery mildew caused by the fungus *Oidium anacardii* in Tanzania. *Crop Protection.* **14(7):** 555-560.

Srinivasarao, C., Jasti, V.N.S.P., Kondru, V.R., Bathineni, V.S.K., Mudigiri, R., Venati, G.V., Priyadarshini, P., Abhilash, P.C. and Chaudhari, S.K. (2022). Land and water conservation technologies for building carbon positive villages in India. *Land Degradation and Development.* **33(3):** 395-412.

Thanikkal, J.G., Dubey, A.K. and Thomas, M.T. (2023). An efficient mobile application for identification of immunity boosting medicinal plants using shape descriptor algorithm. *Wireless Personal Communications.* **131(2):** 1189-1205.

Thanikkal, J.G., Dubey, A.K. and Thomas, M.T. (2023). Deep-morpho algorithm (DMA) for medicinal leaves features extraction. *Multimedia Tools and Applications*. **82(18):** 27905-27925.

Topi, A., Sana, M., Daja, M. and Topi, D. (2025). Applications of machine learning models in near-infrared spectroscopy for small-grain quality control. *Asian Journal of Dairy and Food Research*. **44:** 142-150. doi: 10.18805/ajdfr.DRF-565.

Wang, Z., Wang, P., Liu, K., Wang, P., Fu, Y., Lu, C.T., Aggarwal, C.C., Pei, J. and Zhou, Y. (2025). A comprehensive survey on data augmentation. *IEEE Transactions on Knowledge and Data Engineering*. pp 20.