

Transfer Learning Based Automated Classification Of Cotton Leaf Diseases Using Field Collected Images From The Western Vidarbha Region: A Comparative Study

Mr. Rameshwar S. Devhade¹, Dr. Ranjana D. Raut²

¹Assistant Professor, Department of Electronics, Mahatma Basweshwar Mahavidyalaya, Latur (M.S.),
Latur, Maharashtra, India

²Professor, Department of Applied Electronics, Sant Gadge Baba Amravati University, Amravati (M.S.),
Amravati, Maharashtra, India
doi.org/10.64643/IJIRT209781

Abstract—Cotton is an economically important crop, but its production is significantly affected by various diseases and pests that can reduce crop quality and yield. Manual identification of cotton leaf diseases is time-consuming and depends on expert knowledge of the disease. This study presents an automated cotton leaf disease classification approach using transfer learning and field-collected images from the Western Vidarbha region of India. Initially, 9,361 cotton leaf images were collected using a mobile phone under field conditions. A total of 9,361 cotton leaf images were initially collected using a mobile phone camera under agricultural-field conditions. After systematic duplicate checking and data cleaning, 9,180 unique images were selected. The dataset consisted of six classes: Aphids, Army_worm, Bacterial Blight, Healthy, Powdery Mildew, and Target spot. The cleaned dataset was divided using an 80:20 stratified split, resulting in 7,341 training images and 1,839 validation images. All images were resized to 224 × 224 pixels in size. Four transfer learning architectures, namely MobileNetV2, EfficientNetB0, ResNet50, and DenseNet121, were trained and evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis. Among the evaluated models, ResNet50 achieved the best performance, with an accuracy of 97.99%, followed by DenseNet121 (96.68%) and MobileNetV2 (94.24%). EfficientNetB0 achieved an accuracy of 23.55% under the present experimental configuration. The findings indicate that ResNet50 provides a promising foundation for the automated identification of cotton diseases. Future work will focus on improving model generalization using diverse field conditions and developing lightweight, explainable models for practical mobile-based applications.

Index Terms—Cotton leaf disease, Western Vidarbha, Transfer learning, Deep learning, ResNet50,

DenseNet121, MobileNetV2, EfficientNetB0, Field-collected images, Disease classification.

I. INTRODUCTION

Cotton is one of the most important commercial crops and plays a significant role in the agricultural economy and the textile industry. However, cotton production is affected by various diseases and pest-related disorders that can reduce crop health, productivity, and overall yields. Diseases such as bacterial blight, powdery mildew, and target spot, along with pest-related conditions such as aphid and armyworm infestation, can produce visible symptoms on cotton leaves. Therefore, the early and accurate identification of these conditions is important for timely crop management and reduction of potential yield losses [1], [3], [5]. Traditionally, the identification of cotton leaf diseases and pest-related symptoms has largely depended on visual inspection by farmers, agricultural experts, or plant pathologists. Although expert-based diagnosis can be effective, it may be time-consuming and difficult to perform consistently over large agricultural regions. The visual similarity between the symptoms of different diseases can further increase the difficulty of accurate identification. These challenges create a need for automated image-based approaches that can assist in the rapid and consistent identification of cotton leaf diseases [1], [2]. Recent developments in artificial intelligence, particularly deep learning and convolutional neural networks (CNNs), have provided effective approaches for image-based plant-disease classification. Deep learning models can automatically

learn discriminative visual features from leaf images and classify them into different categories. Transfer learning has further improved the practicality of deep learning for agricultural image analysis by allowing models pretrained on large image datasets to adapt to specific crop-disease classification tasks. Several studies have investigated CNN and transfer-learning approaches for cotton disease identification using architectures such as ResNet, DenseNet, MobileNet, EfficientNet, Xception, and Inception [2], [4], [5], [6]. Existing research has demonstrated promising classification performance; however, substantial differences exist among studies in terms of the dataset source, number of disease classes, image acquisition conditions, dataset size, preprocessing procedures, and evaluation protocols. Several studies have used publicly available datasets or relatively limited disease categories, whereas field-based datasets have been developed with different class compositions and experimental settings [3], [4], [5]. These differences make direct comparisons of model performance across studies difficult. Recent studies have also explored the comparative evaluation of multiple deep learning architectures, demonstrating that model performance can vary considerably depending on the dataset and experimental configuration [5], [6], [7]. To address these limitations, the present study developed a self-collected cotton leaf image dataset obtained directly from agricultural field conditions using mobile-device image acquisition. The dataset contains six classes: Aphids, Army_worm, Bacterial_Blight, Healthy, Powdery_Mildew, and Target_spot. A systematic data-cleaning procedure was performed to identify duplicate images, resulting in 181 duplicate relationships. After cleaning, the dataset contained 9,180 clean images. The dataset was subsequently divided using an 80:20 stratified split, resulting in 7,341 training and 1,839 validation images. All images were resized to 224×224 pixels for training and evaluation of the model. A comparative experimental framework was established using four transfer-learning architectures: MobileNetV2, EfficientNetB0, ResNet50, and DenseNet121. The models were evaluated using the same validation dataset and common classification metrics, including accuracy, precision, recall, F1-score, class-wise performance, and confusion matrix analysis. Among the evaluated architectures, ResNet50 achieved the highest validation accuracy of 97.99%, followed by DenseNet121 (96.68 %) and MobileNetV2 (94.24 %).

EfficientNetB0 achieved an accuracy of 23.55% under the present experimental configuration. These results demonstrate the importance of evaluating multiple architectures on a specific field-collected dataset rather than assuming that a particular architecture will provide the best performance across various datasets. Therefore, the comparative evaluation provides an empirical basis for selecting an appropriate transfer-learning architecture for the developed cotton leaf disease dataset.

II. LITERATURE REVIEW

Deep- and transfer-learning techniques have been increasingly applied to automate cotton leaf disease classification because of their ability to learn discriminative visual features directly from leaf images. Existing studies have investigated conventional convolutional neural networks (CNNs), pretrained transfer learning models, field-collected datasets, and more recently, hybrid CNN–Transformer approaches. However, substantial differences remain in terms of dataset sources, disease categories, image acquisition conditions, preprocessing procedures, and evaluation protocols.

Azath and Bruck (2021), investigated cotton leaf disease classification using a CNN-based approach. They used a cotton leaf disease dataset containing 3,117 images across four classes and reported an accuracy of 96.4%. This study demonstrated the effectiveness of CNNs in automated cotton disease classification. However, this study considered a relatively limited number of disease categories compared with the six-class classification problem addressed in the present study.

Islam et al. (2023), investigated transfer learning for cotton leaf disease prediction using VGG16, VGG19, InceptionV3, and Xception. Their dataset contained 2,310 images, and the reported accuracies were 90.22%, 96.74%, 97.83%, and 98.70% for VGG16, VGG19, InceptionV3, and Xception, respectively. Xception achieved the highest accuracy of 98.70%. This study also demonstrated the practical use of the selected model through a web-based smart application. However, the authors identified issues related to class imbalance and the adaptability of the model to novel disease cases, indicating the need for further evaluation of robustness.

Anwar et al. (2023), evaluated several deep transfer learning models for automated cotton plant disease detection. Their dataset contained 2,654 images, and models, including DenseNet169 and ResNet50V2, were investigated. DenseNet169 and ResNet50V2 achieved accuracy levels of approximately 96% under the experimental conditions. This study demonstrated the benefits of transfer learning and the influence of image resolution on model performance. However, the dataset and disease categories were different from those considered in the present study, limiting the direct comparison of the reported performance with our experimental results.

Hyder and Talpur (2024), investigated cotton leaf disease detection using machine- and deep-learning approaches. Their study evaluated models, including CNN, VGG16, and ResNet50, on a cotton disease dataset. They reported strong performances for several models, demonstrating the applicability of deep learning to cotton disease recognition. However, their study was based on a four-class dataset containing approximately 1,710 images, which differs from the six-class field-collected dataset used in the present study.

Ramya and Kalimuthu (2024), evaluated InceptionV3 for cotton leaf disease classification using a dataset of 1,710 images belonging to four classes. The model achieved an accuracy of approximately 97%. Although this study demonstrated the effectiveness of transfer learning for cotton disease recognition, its smaller four-class dataset provided a different experimental setting from the six-class classification framework used in the present study.

Bishshash et al. (2024), developed a field-based cotton leaf disease dataset using field surveys conducted between October 2023 and January 2024. The dataset contained 2,137 original images and 7,000 augmented images, representing multiple cotton disease, pest, and stress categories. InceptionV3 achieved an overall accuracy of 96.03%. This study is particularly relevant to the present research because it demonstrates the value of field-acquired cotton leaf images under varying environmental conditions. Nevertheless, its dataset composition, augmentation strategy, and class structure differ from those of the self-collected six-class dataset used in the present study.

Singh et al. (2024), and other recent studies have further investigated transfer learning architectures for cotton plant disease classification. A comprehensive

comparative analysis reported the use of architectures, including EfficientNet, Xception, ResNet, Inception, VGG, DenseNet, and MobileNet. This study demonstrated that different architectures can produce substantially different performance levels depending on the dataset and experimental configuration. For example, EfficientNetB3 achieved a very high performance in the reported experiments. This evidence highlights the importance of evaluating model architectures within specific datasets and experimental settings rather than assuming that one architecture will consistently outperform the others.

Nandini and Deepa (2025), conducted a comparative analysis of MobileNetV2 and ResNet50 for cotton leaf disease classification. Their study used 1,450 images belonging to four categories: bacterial blight, cotton leaf curl virus, fusarium wilt, and healthy leaves. This study is particularly relevant to the present work because MobileNetV2 and ResNet50 were also evaluated in our experimental framework. However, the smaller dataset and four-class problem differ from our six-class-field dataset.

III. MATERIAL AND METHODS

The overall working experimental setup of the proposed cotton leaf disease classification system is illustrated in Fig. I. The workflow includes field data collection, data cleaning, image preprocessing, dataset preparation, transfer learning model selection, model training, comparative analysis, and performance evaluation. This systematic workflow ensures a standardized and reproducible approach for automated classification of cotton leaf diseases.

A. Field Data Collection

The cotton leaf image dataset was collected directly from agricultural fields using a mobile phone camera under natural-field conditions. The collected images captured real-world variations in leaf appearance, background, illumination, orientation, and disease symptoms. Initially, the dataset consisted of 9,361 images, which were categorized into six classes: Aphids, Army_worm, Bacterial_Blight, Healthy, Powdery_Mildew, and Target_spot. A data-cleaning process was performed to identify and remove the duplicate images. After duplicate removal, 9,180 unique images were retained and used in the experimental analysis.

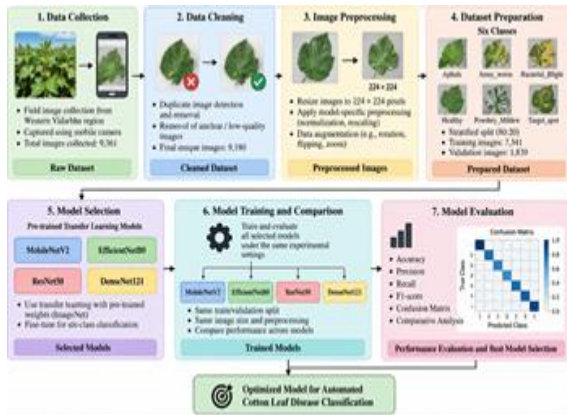


Fig. 1: Working experimental setup for automated cotton leaf disease classification using field-collected images from the Western Vidarbha region.

B. Dataset Splitting

The cleaned dataset was divided into training and validation subsets using an 80:20 stratified split ratio. Stratification was employed to maintain approximately the same class distribution in both the subsets. A fixed random seed of 42 was used to make the dataset partitions reproducible. The resulting dataset consisted of 7,341 training and 1,839 validation images. The validation dataset was kept separate from the training data and used for model performance evaluation. The detailed class-wise distribution is presented in Table I.

Table I: Stratified Training and Validation Distribution

Class	Total	Training (80%)	Validation (20%)
Aphids	1,531	1,224	307
Army_worm	1,535	1,228	307
Bacterial_Blight	1,530	1,224	306
Healthy	1,531	1,224	307
Powdery_Mildew	1,531	1,224	307
Target_spot	1,522	1,217	305
Total	9,180	7,341	1,839

C. Image PreProcessing

All input images were resized to a fixed spatial resolution of 224×224 pixels before being supplied to deep learning models. A batch size of 32 was used for the model training and evaluation. Because the four architectures used different input preprocessing conventions, model-specific preprocessing was applied. The MobileNetV2 and EfficientNetB0 inputs were scaled by dividing the pixel values by 255.0. The

ResNet50 inputs were processed using the standard ResNet50 preprocessing function provided by TensorFlow/Keras. Similarly, the DenseNet121 inputs were processed using the standard DenseNet preprocessing function. The same model-specific preprocessing procedure was applied during the final validation evaluation to ensure consistency between training and evaluation.

D. Transfer Learning Models

To investigate the effectiveness of different convolutional neural network architectures for cotton leaf disease classification, four pretrained deep learning models were selected: MobileNetV2, EfficientNetB0, ResNet50, and DenseNet121. These architectures were selected because they represent different CNN design approaches and provide a useful basis for comparative evaluation using a common dataset and experimental protocol.

- MobileNetV2:** MobileNetV2 is a lightweight convolutional neural network architecture designed for computational efficiency. Its relatively compact architecture makes it suitable for applications in which computational resources are limited. In this study, MobileNetV2 was evaluated as a lightweight baseline for cotton-leaf disease classification.
- EfficientNetB0:** EfficientNetB0 is a computationally efficient CNN architecture that uses compound scaling to balance the network depth, width, and input resolution. It was included to evaluate whether an efficient modern CNN architecture could provide competitive performance on the collected dataset.
- ResNet50:** ResNet50 is a 50-layer residual convolutional neural network that uses residual connections to facilitate the training of deeper networks. It was selected because residual learning has demonstrated strong performance in image classification tasks and provides an important benchmark for evaluating cotton disease classification.
- DenseNet121:** DenseNet121 employs dense connectivity between layers, allowing the feature information to be propagated through multiple layers of the network. This architecture was included to investigate whether dense feature reuse could improve the recognition of visually similar symptoms of cotton leaf diseases.

E. Model Evaluation

After training, all four models were evaluated using the same 1,839-image validation set. Model-specific preprocessing was applied during the evaluation to ensure that the input representation was consistent with the preprocessing used for each architecture. The classification performance was evaluated using four commonly used metrics:

- a) Accuracy: Accuracy measures the proportion of correctly classified images among all evaluated images. For the multiclass problem, accuracy represents the total proportion of validation images assigned to the correct class.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- b) Precision: Measures the proportion of correctly predicted positive samples among all samples predicted to be a particular class.

$$Precision = \frac{TP}{TP + FP}$$

- c) Recall: Recall measures the proportion of correctly identified samples of a particular class among all actual samples that belong to that class.

$$Recall = \frac{TP}{TP + FN}$$

- d) F1-Score: The F1-score provides a harmonic mean of precision and recall. For the six-class classification problem, macro- and weighted-average metrics were also considered to provide an overall assessment of the classification performance across classes.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

IV. RESULT AND DISCUSSION

The experimental results demonstrated that ResNet50 achieved the best overall classification performance among the evaluated deep learning models. As presented in Table II, ResNet50 obtained an accuracy of 97.99%, precision of 98.01%, recall of 97.99%, and F1-score of 97.99%, outperforming DenseNet121, MobileNetV2, and EfficientNetB0.

Table II: Performance Comparison of Different Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ResNet50	97.99	98.01	97.99	97.99
DenseNet121	96.68	96.70	96.68	96.69
MobileNetV2	94.24	94.24	94.24	94.22
EfficientNetB0	23.55	16.42	23.55	13.51

The class-wise F1-score comparison presented in Table III further confirms the consistent performance of ResNet50 in all six cotton leaf classes. ResNet50 achieved the highest F1-score for Healthy (99.67%) and Powdery_Mildew (99.67%), followed by Army_worm (98.87%). The model also demonstrated strong classification performance for aphids (96.45%), Bacterial_Blight (96.43%), and target spots (96.84%). Compared with ResNet50, DenseNet121 showed slightly lower F1-scores across the disease classes, particularly for target _spot. MobileNetV2 also provided satisfactory classification performance, with lower F1-scores than the two deeper models, while offering the advantage of reduced computational requirements for resource-constrained deployment.

Table III: Class-Wise Performance Comparison

Disease Class	ResNet 50 F1 (%)	DenseNet1 21 F1 (%)	MobileNet V2 F1 (%)
Aphids	96.45	95.50	92.81
Army_worm	98.87	98.68	96.90
Bacterial_Blight	96.43	95.72	91.87
Healthy	99.67	98.54	96.56
Powdery_Mildew	99.67	97.72	95.73
Target_spot	96.84	93.94	91.42

The confusion matrix analysis further indicated that most classification errors occurred between Target_spot and Bacterial_Blight in the ResNet50 model, whereas the remaining classes were classified with relatively high reliability. This suggests that the visual characteristics of these two diseases may be similar, resulting in occasional misclassification.

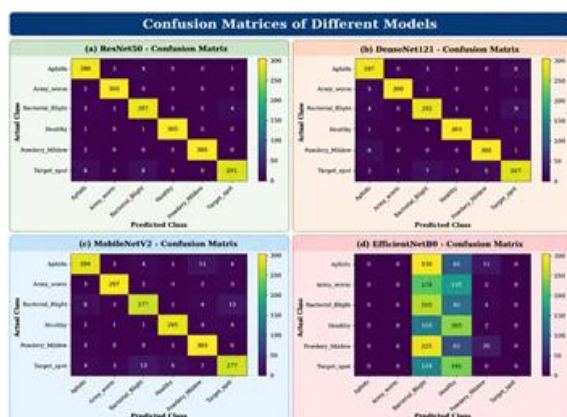


Fig. II. Confusion matrix of different models on the validation dataset for six-class cotton leaf disease classification.

Based on the overall evaluation, class-wise F1-scores, and confusion matrix analysis, ResNet50 was selected as the final model for the proposed cotton disease classification system. Its high and consistent classification performance makes it the most suitable model among the evaluated architectures for reliable cotton-leaf disease identification.

V. DISCUSSION

The comparative evaluation showed that ResNet50 achieved the highest accuracy (97.99%), followed by DenseNet121 (96.68%) and MobileNetV2 (94.24%), whereas EfficientNetB0 achieved 23.55% under the present configuration. These results demonstrate that the model performance varies with the architecture and experimental settings. Under the common experimental protocol adopted in this study, ResNet50 provided the best performance among the four evaluated architectures, achieving an accuracy of 97.99 %. Future work will focus on testing the model using images from diverse locations and environmental conditions, improving generalization, and developing lightweight and explainable models for real-world mobile-based cotton disease identification applications.

ACKNOWLEDGMENT

I, Mr. Rameshwar S. Devhade, sincerely thanks to my research guide Dr. Ranjana D. Raut for her invaluable guidance and continuous support throughout this research work. The author also expresses deep

gratitude to Dr. Sujata N. Kale, Head and Professor of the Department of Applied Electronics, Sant Gadge Baba Amravati University, Amravati, for her encouragement and valuable suggestions, continuous motivation, valuable insights for providing a supportive research environment. Special thanks to farmers, agricultural agencies, and dataset providers for their contributions. Finally, the author acknowledges the constant support of colleagues, friends, and family.

REFERENCES

- [1] Zekiwo Azath, M., and Bruck, A., "Deep learning-based image processing for cotton leaf disease and pest diagnosis," *Journal of Electrical and Computer Engineering*, vol. 2021, pp. 1–10, Art. no. 9981437, 2021, doi: 10.1155/2021/9981437.
- [2] Islam, M., et al., "A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture," *Smart Agricultural Technology*, 2023, doi: 10.1016/j.atech.2023.100238.
- [3] Anwar, S., Soomro, S. R., Baloch, S. K., Patoli, A. A., and Kolachi, A. R., "Performance analysis of deep transfer learning models for the automated detection of cotton plant diseases," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11561–11567, 2023, doi: 10.48084/etasr.6187.
- [4] Gülmez, B., "A novel deep learning model with the Grey Wolf Optimization algorithm for cotton disease detection," *Journal of Universal Computer Science*, vol. 29, no. 6, pp. 595–626, 2023, doi: 10.3897/jucs.94183.
- [5] Hyder, U., and Talpur, M. R. H., "Detection of cotton leaf disease with machine learning model," *Turkish Journal of Engineering*, vol. 8, no. 2, pp. 380–393, 2024, doi: 10.31127/tuje.1406755.
- [6] Kukadiya, H., Arora, N., Meva, D., and Srivastava, S., "An ensemble deep learning model for automatic classification of cotton leaves diseases," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 3, 2024.
- [7] Bishshash, P., Nirob, A. S., Shikder, H., Sarower, A. H., Bhuiyan, T., and Noori, S. R.

- H., “A comprehensive cotton leaf disease dataset for enhanced detection and classification,” *Data in Brief*, vol. 57, Art. no. 110913, 2024, doi: 10.1016/j.dib.2024.110913.
- [8] Johri, P., Kim, S., Dixit, K., Sharma, P., Kakkar, B., Kumar, Y., Shafi, J., and Ijaz, M. F., “Advanced deep transfer learning techniques for efficient detection of cotton plant diseases,” *Frontiers in Plant Science*, vol. 15, Art. no. 1441117, 2024, doi: 10.3389/fpls.2024.1441117.
- [9] Nandini, R., and Deepa, “Deep learning for cotton leaf disease detection: Comparative analysis of MobileNetV2 and ResNet50,” in *Proceedings of the IEEE 1st International Conference on Recent Trends in Computing and Smart Mobility (RCSM)*, 2025, doi: 10.1109/RCSM67767.2025.11507397.
- [10] Bhavani, R., “Detection of leaf diseases in soybean plant using autoencoder and multinomial logistic regression,” *Legume Research*, Apr. 2025, doi: 10.18805/lr-5461.
- [11] Dixit, S., Srinivasan, K., Kumar, A., Haripriya, A., and Bohre, K., “Classification and recognition of soybean leaf diseases in Madhya Pradesh and Chhattisgarh using deep learning methods,” *Apr. 2023*, pp. 1–6, doi: 10.1109/PCEMS58491.2023.10136030.
- [12] Upadhyay, A., et al., “Deep learning and computer vision in plant disease detection: A comprehensive review of techniques, models, and trends in precision agriculture,” *Artificial Intelligence Review*, vol. 58, no. 3, Jan. 2025, doi: 10.1007/s10462-024-11100-x.
- [13] Dhage, S., and Garg, V. K., “Cotton plant fungal disease classification using deep learning models,” in *Proceedings of the IEEE International Conference on Emerging Trends in Engineering Technology and Science (ICETET-SIP)*, Apr. 2023, pp. 1–5, doi: 10.1109/ICETET-SIP58143.2023.10151608.
- [14] Jing, Q., “A lightweight deep learning model for crop disease detection on mobile devices,” *International Journal of Engineering Science and Information Technology*, vol. 5, no. 3, pp. 636–643, Jul. 2025, doi: 10.52088/ijesty.v5i3.1535.