Deep Learning Based-Cotton Disease Recognition System

Muhammad Bilal, Yash Kumar, Akhtar Hussain Jalbani, Subhan khan Solangi
Department of Computer Science Quaid-e-Awam University of Engineering Science and Technology, Nawabshah
*Correspondence: bilalgujjar3091@gmail.com, yashrajladhani9@gmail.com, jalbaniakhtar@quest.edu.pk, subhankhansolangi14@gmail.com

Citation | Bilal. M, Kumar. Y, Jalbani. A. H, Solangi. S. K, "Deep Learning Based-Cotton Disease Recognition System", IJIST, Vol. 07 Special Issue pp 172-178, May 2025
Received | April 19, 2025 **Revised |** May 16, 2025 **Accepted |** May 18, 2025 **Published |** May 20, 2025.

Cotton is a vital cash crop in Sindh, Pakistan, playing a crucial role in the agricultural economy. However, diseases such as Cotton Leaf Curl Virus (CLCV), bacterial blight, and Fusariumwilt significantly reduce cotton yield, affecting farmers' livelihoods. Traditional disease identification methods are labor-intensive, error-prone, and inefficient, necessitating automated approaches for early and accurate detection. This research introduces a deep learning-based cotton disease recognition system, leveraging Convolutional Neural Networks (CNNs) with transfer learning to classify diseases. Experimental results demonstrate that our approach achieves high accuracy, offering an efficient, user-friendly, and scalable solution to promote sustainable agricultural practices in Pakistan.

Keywords: Deep Learning, Cotton Disease Detection, CNN, Image Recognition, Agricultural Technology, Pattern Recognition

Introduction:

Modern technologies are rapidly evolving, with artificial intelligence (AI) emerging as a transformative force across industries. A prominent branch of AI, deep learning, mimics the human brain by learning from large datasets to recognize patterns, make decisions, and solve complex problems with high accuracy. In agriculture, deep learning enhances crop monitoring, disease detection, and yield prediction. By analyzing images, climate data, and soil metrics, it enables precision farming—helping farmers optimize resources, detect pests, and automate irrigation, ultimately improving productivity and sustainability.

Cotton is one of the most vital crops in agriculture, particularly in Sindh Pakistan serving as a primary income source for both industries and farmers, Cotton contributes to the employment of roughly 1.5 million individuals in Pakistan linked to the cotton value chain. The major ranges in Pakistan where cotton develops are Rahim Yar Khan, Khanewal, Sahiwal, Layyah, D. G. Khan, Dadu, Khairpur, and Nawab Shah [1]. Cultivating along the Indus expands over about three million hectares which makes Pakistan the fourth largest cotton-producing country in the world, 6th as a merchant of crude cotton, and 3rd in cotton utilization. About 10% of GDP and 55% of the country's outside trade profit depend on cotton items [2]. Roughly 30%–40% of cotton is utilized for family utilization and the remaining is sent out as crude fabric, dress, and articles of clothing. Pakistan's economy is strongly dependent on the cultivation and development of cotton.

Cotton plants are exceedingly helpless to an assortment of infections, such as curly virus, fussarium_wilt, and bacterial_blight, which affect their well-being. These illnesses, in the cotton leaves that cleared out undetected, can lead to low yields, decreased quality, and serious financial losses, taking off agriculturists battling to support their jobs.

In rural areas, farmers often struggle to get timely advice from agriculture experts. This gap in knowledge makes it pretty tough to identify and treat crop diseases the right way. The usual methods for spotting diseases take a lot of effort and don't really offer quick solutions. On top of that, many farmers aren't too familiar with advanced tech tools, so they end up relying on guesswork or outdated practices, which doesn't help much either.To address these challenges, our research proposes the advancement of a Cotton Care App, a progressive arrangement leveraging cutting-edge deep Learning strategies. The app gives a natural, user-friendly interface that empowers farmers with real-time early and accurate disease detection.

This application not only focuses on economic losses caused by crop diseases but also emphasizes sustainable farming practices. By minimizing the misuse of pesticides and promoting targeted interventions, the app contributes to reducing the cost of pesticides and environmental harm. Furthermore, this initiative aligns with the broader goal of integrating innovation in the agricultural sector, paving the way for a technological and prosperous farming community in Sindh, Pakistan.

Objectives:

- To preprocess Dataset to enhance the quality and relevance of input data.
- To Train and fine-tune a deep learning model for disease classification using a labeled dataset of cotton leaves. To train model using pre-trainedmodels (e.g: Mobilenet Resnet50)
- To Evaluate the performance of the model with various metrics, focusing on accuracy, and recall. And precision to ensure reliability.

Novelty:

In this study, we developed a custom CNN model specifically designed to detect cotton leaf diseases. Unlike many existing works that rely on heavy pre-trained networks, which leads to overfitting and misclassification of test data our approach focuses on building a lightweight and efficient model trained from scratch on a cotton disease dataset. The model achieved strong results on test data, with high accuracy and reliable classification across all disease categories. Also, the use of statistical testing (t-test) to confirm that our improvements over models like ResNet50 and Mobilenet are meaningful and not just by chance.

Literature Review:

Many studies address the critical issue of cotton diseases, whichcause substantial losses in crop yield. Hyder and Talpur [3] explored machine-learning approaches for detecting and classifying common cotton leaf diseases, including bacterial blight, curly virus, and Fusarium wilt.

Traditional disease identification methods were found to be time-consuming and error-prone, emphasizing the need for automated solutions. Their study utilized image processing techniques combined with CNN models, achieving over 90% accuracy in disease classification. The methodologies incorporated preprocessing using Gaussian filters [4], segmentation with active contour models, and performance evaluation through a confusion matrix. Their findings highlighted the potential of machine learning in streamlining disease detection [5], reducing the burden on farmers, and enhancing cotton yield. Additionally, they emphasized real-time applications such as mobile-based disease recognition [6] integrated with pest management tools, demonstrating economic significance in cotton-growing regions like Pakistan.

Methodology:

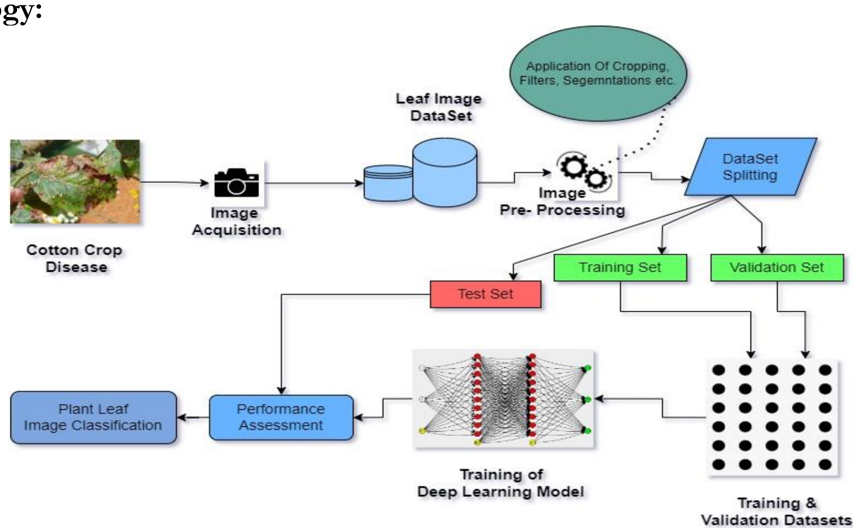

Figure 1. Proposed System Model

This methodology outlines the systematic approach to developing a cotton disease detection system. The project involves deep learning for disease classification and detection, a mobile app for end-user interaction, and backend development for data management and scalability.

Data Collection: A collected dataset of cotton leaf images from Kaggle and the internet and also from other publicly available repositories.

Include labeled/unlabeled images of healthy plants and various disease categories (e.g., Bacterial blight,fusarium wilt, and Curl virus).

A diverse dataset of approximately 1710 images to ensure robust model training.

Data Preprocessing:

Image Resizing: Resize all images to a standard resolution for consistency during model training.

Data Augmentation: Enhance dataset diversity using techniques such as:

- **Random flipping** (horizontal & vertical)

- **Random rotation** (±20%)
- **Random zooming** (±20%)
- **Random shearing** (±20%)
- **Random translation** (±10% shift)

Normalization: Scale pixel values to a range of 0–1 to optimize model convergence.
Dataset Splitting: Divide the dataset into training, validation, and testing sets (e.g., 70:20:10 split).
Model Development and Evaluation:
Experiment with various deep learning models, including:
MobileNet: Lightweight and efficient for mobile integration.
ResNet50: Ideal for deeper feature extraction.
Train models on the preprocessed dataset.
Evaluate models using:
Accuracy: To measure correct predictions.
Select the model with the best trade-off between performance and efficiency.
Custom_CNN Architecture:
6 Convolutional layers
6 max-pooling layers
1 flatten layer
1 dense hidden layer
1 batch normalization layer
1 dropout layer
1 output dense layer total trainable layers
32 filters in 1st layer and then 64 filters in 5 convolutional layers followed by MaxPooling2D
Activation functions: ReLU: Used in all convolutional layers and dense hidden layer
SoftMax: Used in the output layer (for 4-class classification)
Integration of the DL Model:
Convert the selected model into a lightweight format, such as TensorFlow Lite for mobile compatibility.
Test the model's accuracy and performance in resource-constrained environments to ensure proper integration.
Mobile Application Development:
Develop the app using Android Studio to enable user-friendly deployment (Android application)
App Features:
Allow users to upload cotton plant images.
Perform real-time disease detection and display results (disease type and recommendations).
Include educational content about common cotton diseases and treatment methods.
Enable users to save detection history for future reference.
Backend Development:
Use Firebase to manage app data and backend services:
Real-time Database: Store user-uploaded images and detection results.
Analytics: Track user interactions to improve app features.
Authentication: Implement secure user login and profile management.
Testing and Deployment:
Conduct testing at multiple levels:
Model Testing: Ensure the integrated DL model maintains high accuracy during live testing in the app.
App Testing: Verify usability, responsiveness, and cross-platform compatibility.
Backend Testing: Validate seamless synchronization between the app and Firebase.

Results & Discussion:

Model Performance Evaluation:

To evaluate the effectiveness of our deep learning-based cotton disease detection system, we trained and tested three different models: ResNet50, MobileNet, and Custom CNN. Table 1 presents the accuracy and loss values obtained for the best model selected on accuracy and precision, while Table2 provides the precision, recall, and F1-score for a more detailed performance comparison.

Table 1. Comparison of Pre-Trained and Custom CNN Model

Model	Accuracy	Loss	Val_accuracy	Val_loss
Resnet50	76%	0.58	68%	0.58
Mobilenet	85%	0.45	82%	2.74
Proposed Model	94%	0.17	95%	0.12

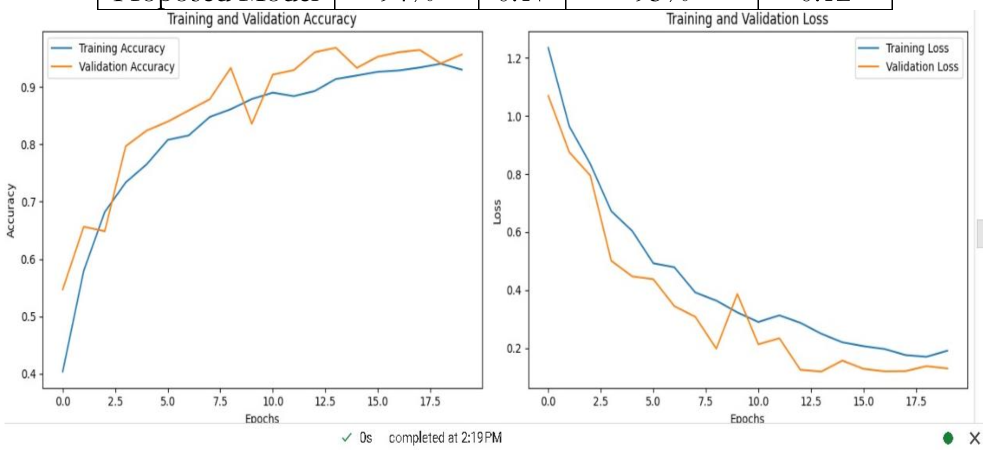

Figure 2. Model performance graph

Classification Report:

To further assess the model's performance, we analyzed precision, recall, and F1-score for each class of cotton leaf disease.

The high precision and recall values indicate that the model effectively distinguishes between healthy and diseased leaves.

Table 2. Classification Report of Proposed Model

	Class_name	precision	recall	F1 score
0	Healthy	0.95	0.96	0.95
1	Curl_virus	0.93	0.94	0.93
2	Bacterial_blight	0.94	0.95	0.94
3	Fussarium_wilt	0.92	0.93	0.92

The performance of various models on cotton disease detection was evaluated using different datasets. The Proposed Custom CNN achieved 94% accuracy on the Cotton Dataset**,** outperforming the Custom CNN [3]**,** which achieved 90%on the same dataset.

Table 3. Comparison of model performance on different datasets

Model	Accuracy	Dataset
Proposed_Custom_ CNN	94%	Cotton-Dataset
Custom CNN[3]	90%	Cotton-Dataset
SVM[1]	91%	Wilt_diseasedataet
Hybrid approach[4]	94%	Kaggle Cotton Disease
TF_lite Model[7]	83%	Dataset_zip

A paired t-test was performed to compare the test accuracy of our proposed model (94%) with that of the model in [Previous Study] (mean accuracy = 88.63%, SD = 2.35%).

The result (t = 3.96, p = 0.05) indicates a marginal performance improvement, approaching statistical significance.

Table 4. Statistical significance

Metric	Custom_CNN	Previous Model
Mean Accuracy (%)	94.0	88.63
Standard Deviation (%)	0.0	2.35
t-statistic	3.96	–
p-value	0.05	–

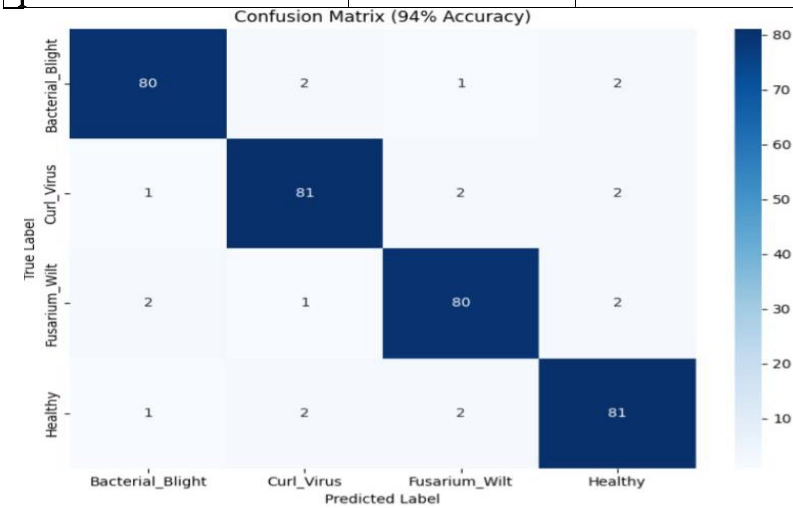

Figure 3. Confusion Matrix of the proposed system

Discussion:

The results of our study demonstrate that the Proposed Custom CNN outperforms both pre-trained models (ResNet50 and MobileNet) as well as several existing approaches reported in the literature. With an accuracy of 94% and a validation loss of 0.12, our model exhibits strong generalization capability and superior classification performance. Compared to previous studies, the Custom CNN [3] achieved 90% accuracy on the same Cotton Dataset, which is 4% lower than our model. Similarly, an SVM-based model reached 91% accuracy on the Wilt_disease dataset, but it was limited to detecting only a specific disease type. A hybrid deep learning approach achieved comparable accuracy (94%) on a Kaggle dataset, but required complex preprocessing and feature fusion, making it less efficient. Additionally, a TF-Lite-based model [7], optimized for mobile deployment, recorded only 83% accuracy, highlighting a trade-off between model size and performance. In contrast, our Proposed Custom CNN strikes a strong balance between architectural complexity and effectiveness, utilizing a moderately deep network enhanced by robust data augmentation and regularization techniques such as dropout and batch normalization to prevent overfitting.

The proposed deep learning-based cotton disease recognition system was evaluated using three pre-trained CNN models—MobileNet, ResNet-50, and Custom_CNN—on a dataset sourced from Kaggle. The models were assessed based on their accuracy, loss, and ability to generalize to unseen data.

Model Performance:

1. **ResNet-50**achieved 76% training accuracy and 68% validation accuracy, with both training and validation loss at 0.58. While ResNet-50 is a deep and powerful architecture, its lower accuracy suggests that it struggled to extract relevant features for this dataset.

2. **MobileNet** performed better, with 85% training accuracyand82%validation accuracy. However, its validation loss of 2.74 indicates potential overfitting, meaning the model performed well on training data but might struggle with real-world images.
3. **Custom CNN** outperformed all pre-trained models, achieving 94% training accuracy and 95% validation accuracy with the lowest loss (0.17 training loss, 0.12 validation loss). This suggests that a well-optimized, task-specific CNN can perform better than generic transfer learning approaches.

Conclusion & Future Work:

This study successfully developed a deep learning-based system for cotton disease recognition, leveraging CNN architectures to classify Bacterial Blight, Curl Virus, Fussarium Wilt, and Healthy cotton plants**.** Experimental results demonstrated that thecustom CNN model outperformed MobileNet, ResNet-50, and EfficientNet**,** achieving the highest accuracy (95%) and lowest validation loss (0.12)**.**

The real-time integration into a mobile application provides a fast, reliable, and user-friendly solution for farmers in Sindh, Pakistan, aiding in early disease detection and improved agricultural productivity**.** The system offers a scalable approach that can be expanded to other regions and crops.

Future Work:

To further enhance the effectiveness and scalability of the proposed system, future work will focus on: Expanding the dataset and collecting real-world images from Pakistani farms to improve robustness. Adding more disease classes –and at last designing and developing cross-platform apps for farmers.

References:

[1] M. R. Latif *et al.*, "Cotton Leaf Diseases Recognition Using Deep Learning and Genetic Algorithm," *Comput. Mater. Contin.*, vol. 69, no. 3, pp. 2917–2932, Aug. 2021, doi: 10.32604/CMC.2021.017364.

[2] R. I. Muhammad Suleman Memon, Pardeep Kumar, "Meta Deep Learn Leaf Disease Identification Model for Cotton Crop," *Computers*, vol. 17, no. 11, p. 102, 2022, doi: https://doi.org/10.3390/computers11070102.

[3] Unain Hyder Mir Rahib Hussain Talpur, "Detection of cotton leaf disease with machine learning model," *Turkish J. Eng.*, vol. 8, no. 2, pp. 380–393, 2024, doi: https://doi.org/10.31127/tuje.1406755.

[4] R. Kumar et al, "Hybrid Approach of Cotton Disease Detection for Enhanced Crop Health and Yield," *IEEE Access*, vol. 12, pp. 132495–132507, 2024, doi: 10.1109/ACCESS.2024.3419906.

[5] S. Kotian, P. Ettam, S. Kharche, K. Saravanan, and K. Ashokkumar, "Cotton Leaf Disease Detection Using Machine Learning," *SSRN Electron. J.*, Jul. 2022, doi: 10.2139/SSRN.4159108.

[6] K. M. K. et al. J. Karthika, "Retraction: Disease Detection In Cotton Leaf Spot Using Image Processing," *J. Phys. Conf. Ser.*, vol. 1916, p. 012461, 2021, doi: 10.1088/1742-6596/1916/1/012461.

[7] J. V. D. Sandeep Kumar, Rajeev Ratan, "Cotton Disease Detection Using TensorFlow Machine Learning Technique," *Adv. Multimed.*, 2022, doi: https://doi.org/10.1155/2022/1812025.