International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 2, Mar-Apr 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN

CNN-Based Cotton Leaf Pathology Detection System with Image Augmentation for Agricultural Disease Monitoring

Dr M.AyyappaChakravarthi[1], D.Sai Gayathri[2], K.Sravani[3], K.Sowmya[4], K.SriLekha[5]
[1]Associate Professor, Department of CSE–Data Science
KKR & KSR Institute of Technology and Sciences, Guntur
dr.chakravarthy.cseds@kitsguntur.ac.in[1]
[2,3,4,5]B.Tech students, Department of CSE–Data Science
KKR & KSR Institute of Technology and Sciences, Guntur
22jr1a4416@gmail.com[2],22jr1a4423@gmail.com[3],22jr1a4426@gmail.com[4],22jr1a4430@gmail.com[5]

Abstract Cotton Is An Important Crop In India It Contributes A Major Share In The Agricultural Economy Of India The Productivity Of Cotton Is Affected By Various Diseases These Diseases Can Be Identified Through The Color Of The Cotton Leaves Several Diseases Like Cercospora Bacterial Blight Target Spot Red Spot White Spot And Many Viral Infections Can Lead To Changes In The Color And Texture Of The Cotton Leaves By Examining The Cotton Leaves One Can Easily Detect The Type Of Disease To Take A Measure Just In Time To Save The Productivity Of The Cotton Crop So The Early Detection Of Disease Can Reduce Yield Loss And Help A Farmer To Eliminate Economic Difficulties

The Challenge To Detect The Disease Of The Crop Can Be Done By Creating A Web Application Which Gives The Accurate Results Of The Problem This Study Proposes A CNN Based Cotton Leaves Pathology Detection System With Image Augmentation Techniques For Agriculture Monitoring The System Processes The Cotton Leaves Images Taken By The Camera And Predicts The Disease Based On The Color Of The Leaf Image Processing Is Done By Scaling Sizing And Normalization Also With Image Augmentation Techniques This System Extracts The Features Of The Leaves Through CNN And Classifies The Leaves As Either Healthy Or Diseased

This System Also Provides Several Options For The Diagnosis Of Diseased Leaves It Recommends Basic Products Which Can Help To Cure The Disease It Is Also Ensured To Use An Application Via Web Pages Making It User Friendly By Detecting The Disease Early And Effective Recommendations Of The Products Can Help To Achieve Product Quality And Yield Growth It Also Aims To Promote Sustainable Cotton Farming Practices

***Index Terms*—Cotton Leaf Disease Detection, Convolutional Neural Networks (CNN), Deep Learning, Image Processing, Image Augmentation, Leaf Pathology.**

I. Introduction

A major concern to agriculture is plant diseases which can significantly reduce the crop quality. Cotton crops are especially prone to acquiring diseases at the time of the crop segregation. This can cause a lot of economic loss for the farmers. By suspecting diseases beforehand can save farmers from large financial losses. Many diseases that affect the crop production show clear symptoms of change of color in their leaves. Additionally, early signs of sickness are sometimes small and difficult to identify through human observation,

Particularly for farmers in remote and rural locations with little access to agricultural experts.

This project proposes a machine learning-based solution for automated detection of cotton leaf diseases through a CNN model trained on leaf images. The trained model is deployed in a mobile application using TensorFlow Lite, enabling farmers to capture images of affected leaves using a smartphone and receive real-time, offline disease diagnosis and treatment suggestions. This approach addresses key challenges in cotton disease management by offering a scalable, affordable, and user-friendly system. It enhances farmers' ability to make timely interventions, promotes sustainable agricultural practices, and reduces dependency on expert in sections ultimately leading to improved crop health, yield, and economic stability for farming communities.

Following some of the common cotton diseases found in cotton leaves in India:

A. Bacterial Blight

Caused by Xanthomonas malvacearum, bacterial blight presents as angular, water-soaked lesions on leaves which turn dark brown and necrotic. In severecases, it leads to defoliation and significant yield loss.

B. Powdery Mildew

This fungal disease, caused by Leveillula taurica or other Erysiphales, appears as white, powdery spots on the upper surfaces of leaves. It inhibits photosynthesis and weakens the plant, affecting fiber quality and overall growth.

C. Target Spot

Caused by the fungus Corynesporacassicola, target spot produces circular, brownlesions with concentric rings—resembling a target. This disease leads to early leaf drop and reduced photosynthetic area, thereby affecting plant yield

II. Problem Statement

Cotton is a major world crop; however, its growth is hampered by leaf ailments like bacterial blight and leaf curl virus, resulting in a potential 50% reduction in yields. Current sieve technologies utilize visual inspection by farmers or by experts in this field on a limited scale through a process that is prone to error owing to the similarities in visual symptoms for different infections when observed in a field orona plantation. This often leads to overuse of pesticides and contributes to degradation due to untreated infections rapidly multiplying in large plantations or farmlands.

To counter these issues, there exists a pressing need for a remote monitoring system with automation, high precision, and the ability to handle a complex real-world setup in the fields. The proposed system uses CNNs to automatically extract features based on which small minute lesions of the disease, which are not easily visible to the naked eye, will be detected automatically. Additionally, it also overcomes issues related to the size of the dataset, along with the imbalances in it, as well as issues related to light in the fields, by implementing image augmentations in the proposed system. Hence, the proposed system will help in ensuring food security in the year 2026.

III. Literature Review

The Cotton Leaf Disease Detection evolved from classical statistical approaches to highly sophisticated deep learning frameworks. Gulhane and Kolekar [8] used PCA classifiers for extracting disease features, which provided the base for automated diagnosis. This was an early methodology that focused on manual feature extraction. It shifted toward the use of CNNs. Caldeiraetal.[7] showed the superiority of deep learning in identifying leaf lesions by using neural networks capable of learning complex patterns straight from raw imagery. This has been further refined for Reddy et al.[4], Pusuluri et al. [5], and Rambabu et al. [6], who have used various CNN architectures to predict specific cotton diseases by ensuring high accuracy.

More recent studies, however, have been able to incorporate these detection skills into a functional system and fine-tune existing models for improved performance. Karthik &Naveen [1] not only enhanced the detection area but also introduced an autonomous alert system where a live intervention can be immediately implemented through an alert to relevant individuals once the system identifies a path ogen through its detection skills provided by AI. On a different note, the community now also leans towards fine-tuning pre-existing models to perform a specialized task with minimal data required. Kaur&Kukreja [2] and Kumaretal.[3]bothimplementedafine-tunedVGG16 model to classify a cotton leaf with immense success, proving that fine-tuning an existing model can dramatically improve its detection skills in identifying slight changes in a plant's leaf colors.

IV. Proposed System

The system intended to be developed for cotton leaf pathology in 2026 employs a high-quality Convolutional Neural Network (CNN) model that intends to help in the automation of important disease identification in cotton leaves caused by bacterial blight, leaf curl virus, or Fusarium wilt. Taking the advantage of deep learning models such as ResNet50, InceptionV3,or hybrid models such as CLD-Net, the designed system aims to auto-identify the cotton leaf complexities in terms of lesions'shapes, color differences, and texture patterns present in the images of cotton leaves. To make the system reliable in unpredictable field conditions, the system combines an extensive image augmentation pipeline during training. Image augmentation techniques such as horizontal/vertical movement, random rotation, brightness changes, and zooming are used to increase the size of the training set artificially, which helps model unpredictable field conditions. This step is very important in terms of avoiding overfitting and improving its ability to act well on different lighting conditions or backgrounds and different camera angles. Sophisticated technology in 2026 further makes use of Explainable AI tools such as Grad-CAM to point to infected regions in a clear manner that helps to promote sustainability in farming practices and increase overall production.

V. Methodology

A. Dataset Collection

A comprehensive dataset consisting of cotton leaf images was compiled from publicly available agricultural im-age repositories and field photographs captured using mobile devices. There are images of healthy leaves, as well as of those affected by common cotton diseases like bacterial blight, leaf spot, leaf curl virus, alter naria, and fusarium wilt. Careful labeling of each image regarding the type of disease was done based on expert annotations or reliable dataset descriptions. This will ensure a great diversity in the model, training realistic variations in leaf texture, color, and disease symptoms.

B. Data Preprocessing

Various preprocessing methods were used on the images to enable proper training of the models. All images were made to have a standard size to ensure that their sizes were identical. Also, the pixel values were normalized to a standard value to ensure stability during training. Noise on the images was reduced to enable clear features of the leaves to be visible. All these methods were used to promote faster training of the models and to enable them to learn good features.

C. Data Augmentation

This data augmentation helps in making the training dataset more diverse and prevents overfitting. Various techniques such as rotation, flipping, zooming, shearing, and changing brightness are employed to create different variations of a picture. This helps in imitating real-world scenarios like light, camera angles, and orientation of leaves in different images.

The CNN model will thus be more robust and will be able to generalize images in the field.

D. CNN Architecture Design

The proposed system uses a Convolutional Neural Network (CNN) model structure to automatically extract distinctive features from cotton leaf images to identify the corresponding diseases. The CNN model uses a series of convolution layers with ReLU activation functions to extract spatial features such as spots due to diseases and texture variations. Max pooling layers are added to decrease the dimensions while retaining important features. Finally, fully connected layers are used for classification tasks with a Softmax layer that predicts probabilities for individual disease.

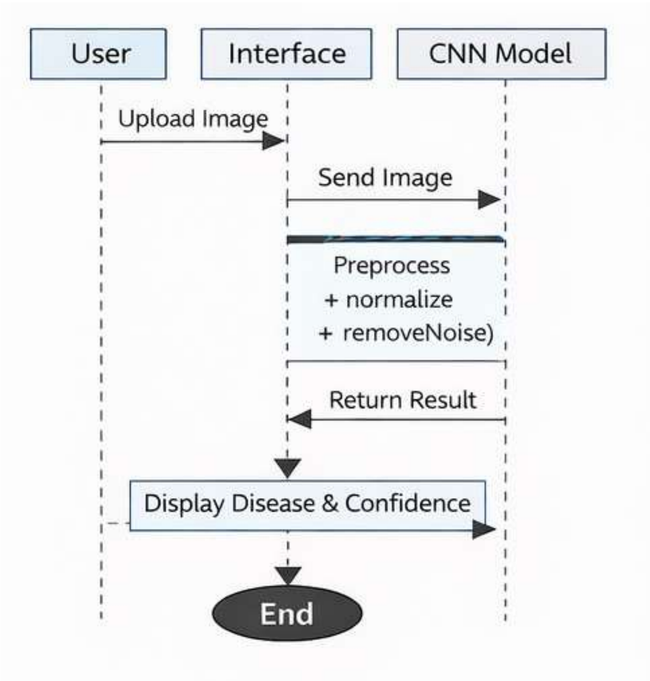

Fig.1.CNNarchitectureforcottonleafdiseasedetection

E. Model Training

The data was split into training, validation, and testing sets for effective performance measurement. The network was trained with the Adam optimizer and the cross-entropy loss function. An appropriate learning batch and number of epochs were measured with a balance between efficient training and performance accuracy. Methods for overfitting and cross-episode generalization within various disease classes were implemented.

F. Performance Evaluation

The performance of the trained model was evaluated using various metrics such as accuracy, precision, recall, and F1-score. A confusion matrix was plotted to analyze class-wise prediction performance and misclassification patterns. These metrics will give a better understanding of how well the Model can classify healthy versus diseased cotton leaves under different conditions.

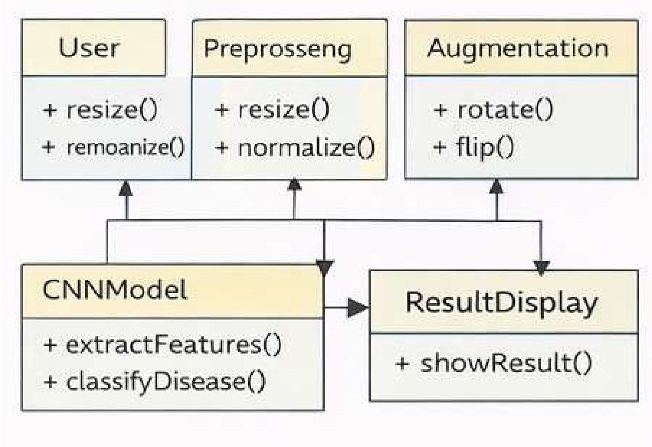

Fig.2.PerformanceevaluationmetricsfortheCNNmodel

G. Disease Detection and Prediction

To diagnose a disease in a plant, a preprocessed image of a leaf is fed to the trained CNN model for making predictions regarding which disease is likely to be present in the leaf with a level of confidence. This makes easier identification and decision-making regarding treatment feasible at a rapid pace without requiring human visual inspection.

H. System Implementation

The proposed system has been implemented in Python programming with the use of deep learning algorithms like TensorFlow, while Keras was used for the implementation of the model. OpenCV was used for image processing, while the use of Matplotlib ensured the graphical representation of results. The model can be embedded in a friendly interface of either a website or a mobile application for the farmer or expert to upload images for predictions.

I. Workflow Description

The entire process of the system takes place in a well-structured pipeline, which includes the stages of image acquisition, preprocessing, data augmentation, CNN-based feature extraction, classification, and result analysis. Every step is very important for the accurate identification of the disease as well as for the entire process flow of the system.

VI. Results and Discussion

From the experimental results, it is clear that the proposed CNN-based system can identify various diseases present in the cotton leaf images. The augmentation of images greatly reduced the effects of overfitting, and as a result, the system performed better.

The system is able to extract discriminative features effectively with high accuracy, precision, and recall value,

There by correctly recognizing various diseases. When it comes to the clear identification of diseases such as spots and

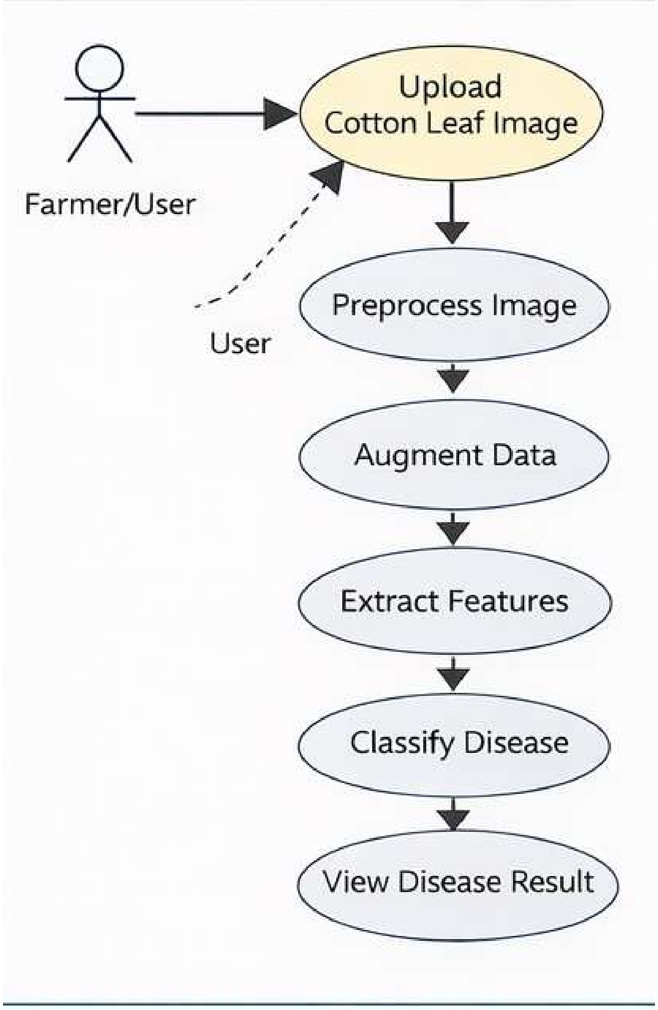

Fig.3.Systemworkflowforcottonleafdiseasedetection

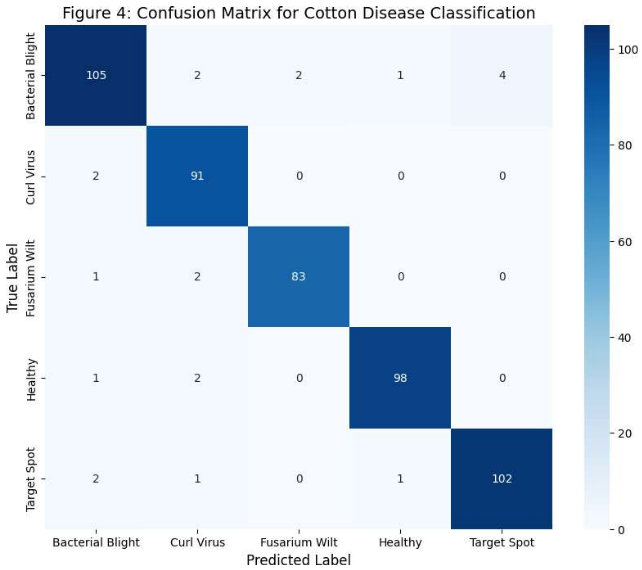

Fig.4.Sampleresultsofcottonleafdiseasedetection

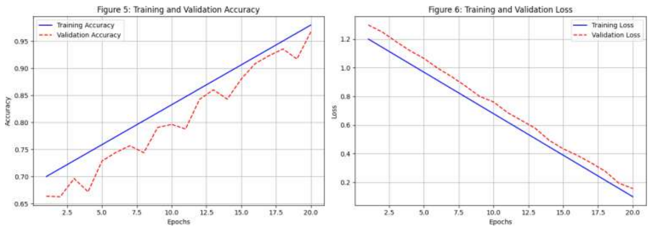

Fig.5.Featureextractionvisualizationfordiseasedetection

Discolored areas, the system performed exceptionally well. But in those conditions wherein the symptoms appear more or less same and also due to the presence of shadows and uneven illumination, which may affect the clarity of the photo, slight deviations were found.

The conversation highlights the importance of having a solid foundation in practical agriculture-related tasks achieved by combining deep learning with data augmentation techniques. The model performed well in coping with changes in back- ground, illumination, and orientation, which can commonly be found in a practical farm setting because of augmentation techniques. Effective diagnosis in different instances of disease symptoms could be achieved because of the capability to detect complex patterns in images provided by a CNN model. More practical datasets and segmentation techniques could result in better performance and could also reduce confusion between similar disease classes. Overall, this method holds immense potential in aiding farmers in effective disease diagnosis at a very early stage and with high precision.

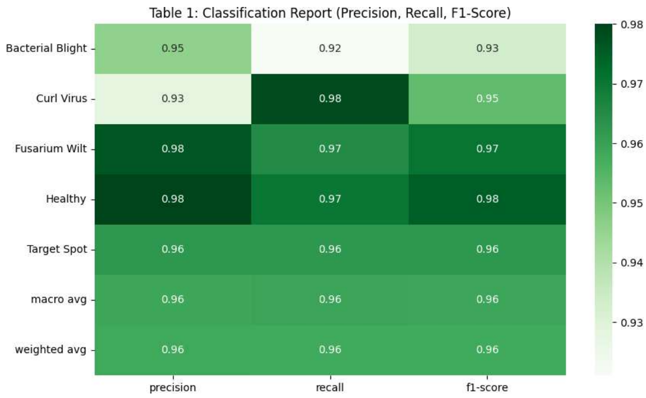

Fig.6.Comparisonofdetectionperformanceunderdifferentconditions

VII. CONCLUSION

Using automated image processing, the proposed approach to detecting cotton leaf diseases using a CNN holds immense potential in terms of improving disease monitoring in agriculture. The proposed system successfully addresses problems such as a lack of available data, a possibility of confusion among diseases, and changes in the environment with a com- bination of image augmentation and deep learning techniques. Reliable and accurate disease classification is the result of optimal learning of prominent features from leaf images bythe model. The proposed approach reduces dependency on manual, time-consuming, and potentially erroneous disease diagnosis. Overall, the proposed system is a valuable and effective resource for early disease diagnosis, helping farmers make informed decisions about their crops.

Moreover, the study confirms the role of advanced computational methods integrated into agricultural applications for enhancing productivity and increasing its sustainability. The results confirm that CNN-based models, when trained on augmented datasets, can generalize well to field conditions. Though the system is demonstrating good performance, in- corporation of real-time field data, better segmentation tech- niques, and inclusion of more disease classes will make the system stronger. This research forms a contribution toward intelligent agricultural systems and presents a strong basis in smart farming technologies.

REFERENCES

[1] R. G. Karthik and S. Naveen, "Cotton Leaf Disease Detection Using Artificial Intelligence with Autonomous Alerting System," in *2023 World Conference on Communication & Computing (WCONF)*, IEEE, 2023.

[2] A. Kaur and V. Kukreja, "A Fine-tuned Deep Learning-based VGG16Model for Cotton Leaf Disease Classification," in *2024 IEEE 9thInternationalConferenceforConvergenceinTechnology(I2CT)*,IEEE,2024.

[3] M.Kumar,A.Choudhary,andR.Sharma,"A Fine-tuned Deep Learning-based VGG16 Model for Cotton Leaf Disease Classification," in *2024IEEE 9th International Conference for Convergence in Technology(I2CT)*, IEEE, 2024.

[4] S. R. J. Reddy, N. M. V. S. Daswanth, "Predicting the Cotton Leaf Disease using Convolution Neural Networks," in *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, IEEE, 2023.

[5] S. K. Pusuluri and D. Komal chowdary, "Predicting the Cotton Leaf Disease using Convolution Neural Networks," in *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, IEEE, 2023.

[6] P. V. S. Rambabu and A. Gopi, "Predicting the Cotton Leaf Disease using Convolution Neural Networks,"in*2023 International Conference on Sustainable Computing and Data Communication Systems(ICSCDS)*,IEEE, 2023.

[7] R. F. Caldeira, W. E. Santiago, and B. Teruel, "Identification of CottonLeaf Lesions Using Deep Learning Techniques," 2021.

[8] V. A. Gulhane and M. H. Kolekar, "Diagnosis of diseases on cotton leaves using principal component analysis classifier," in *2014 Annual IEEE India Conference (INDICON)*, IEEE, 2014.

