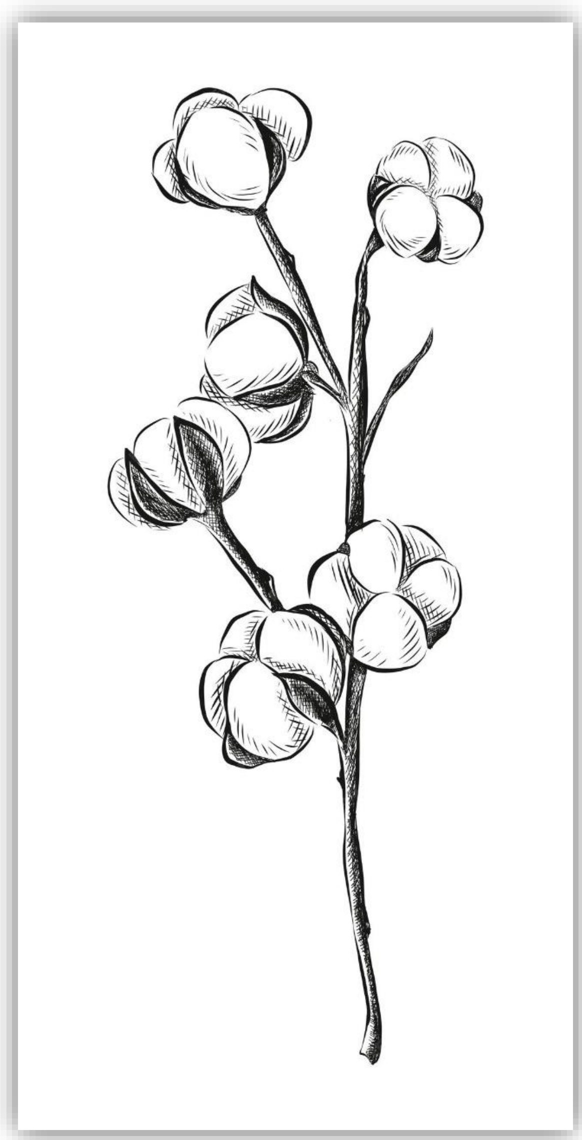
COTTON LEAF DISEASE DETECTION



Sriram Thota-2103A52070
V Akshay Kumar-2103A52039
Pranay Sripada-2103A52170

ABSTRACT

Cotton diseases seriously affect the yield and quality of cotton. The type of pest or disease suffered by cotton can be determined by the disease spots on the cotton leaves. With the recent outbreaks of cotton leaf disease in some regions, there is a growing need for accurate and timely disease prediction methods to help, prevent and control the spread of the disease. The diseases also spread due to either deficiency or toxicity of various nutrients in the cotton plants .In this, we present an overview of recent developments in cotton disease prediction methods. These include machine learning models and algorithms that can analyze and interpret large amounts of data from multiple sources, such as climate and environmental factors, pesticides, and human behavior. We too talk about the challenges and openings of utilizing these strategies in real-world settings, counting the require for dependable information sources and the significance of collaboration between open wellbeing specialists, analysts, and other partners. At last, we highlight a few promising bearings for future investigate in this zone, such as creating more vigorous and exact expectation models, and utilizing social media and other advanced stages to screen and track illness episodes in real-time.

INTRODUCTION

Agriculture may be a crucial portion of each country's economy, and India is regarded an agro-based country. One of the most purposes of agriculture is to surrender solid crops without any illness. Cotton could be a critical edit in India in connection to pay. India is the world's biggest maker of cotton. Cotton crops are influenced when clears out drop off early or gotten to be beset with maladies. Ranchers and planting specialists, on the other hand, have confronted various concerns and progressing rural deterrents for centuries, counting much cotton malady. Since serious cotton illness can result in no grain collect, a rapid so on [1],[2], effective, less costly and dependable approach for recognizing cotton sicknesses is broadly needed within the rural data zone.

One of the most important cash crops in India is cotton.and has been a major supporter to the country's economy for decades. Concurring to the Service of Agriculture and Farmers Welfare, cotton contributes to around 5% of India's add up to agrarian GDP and 2.5% [3]of the country's overall GDP. Cotton could be a major source of business for millions of individuals, especially in provincial ranges, where cotton development is the most source of salary. Agreeing to the Cotton Enterprise of India, cotton cultivation provides employment 2to roughly 6.00 million agriculturists and 40.00 million cultivate laborers. Cotton plants are vulnerable to several diseases that can essentially decrease edit abdicate and quality, coming about in considerable financial losses for farmers.

Cotton cultivating is an critical division of agribusiness in India, with a long history dating back to antiquated times with a differing run of assortments developed over diverse locales of the nation. Cotton is an vital trim for numerous ranchers in India[4], particularly those within the central and southern parts of the nation. Cotton cultivating in India is characterized by a blend of large-scale commercial ranches and little- holder ranches. Over the a long time, Indian cotton farmers have confronted a number of challenges, counting bother and illness flare-ups, showcase vacillations, and natural debasement.

Be that as it may, through imaginative cultivating hones, investigate and improvement and the government support, the segment has kept on flourish, contributing to the country's financial development and development.[5] In this manner, the paper presents and examines whether C-NN show and picture handling procedures can precisely distinguish and classify cotton plant illnesses, which can eventually help to move forward trim administration and anticipate surrender misfortunes.

The generation of cotton in India decreasing slowly over year because of major cotton maladies which affect their generation exceptionally much a few common illnesses like creepy crawly assault charcol rot and numerous are making overwhelming affect over their manor. Due to this numerous cotton cultivators farmer get a tremendous drop down in their generation and income.[6] The issue will be fathomed in the event that the agriculturist get to know almost the plants which are contaminated and unhealthy in early stages of their development so that agriculturists can utilize pesticides and diverse therapeutic types of gear to sprinkles solutions over plants and spare their crops from maladies in early stages of production.[7] As this extend will offer assistance the ranchers to recognize the cotton plants which are New and Ailing by essentially uploading the pictures of the cotton plants on the internet app. On advance Generation level we sent as a web app which can make the farmers to press and transfer their cotton plant picture and get comes about on the spot right away. by this disease,[8] how to stay unaffected by the virus, what kind of safeguards ought to be taken care, when to go to the healing center, levels of conditions of people who are contaminated, and side effects of this infection after a profound examination of contaminated individuals.

LITERATURE SURVEY

AUTHORS	YEAR OF PUBLISH	MODEL USED	PARAMETERS	FUTURE SCOPE
Sharada P. Mohanty, David P. Hughes, Marcel Salathé	2018	Convolutional-Neural Network (C-NN)	- Used pre- trained VGG-16 model	- Expanding model architecture and retraining on larger dataset containing more cotton disease types could enable detecting more diseases.
A. Kamilaris, F.X. Prenafeta-Boldώ	2019	Convolutional-Neural Network (C-NN)	- Custom 5layer C-NN architecture	- Accuracy can be improved by testing moreadvanced C-NN architectures like ResNet, Inception, etc. Increasing dataset size could also help.
R. Jabee, S. Raman, K. Thankachan	2019	Convolutional-Neural Network (C-NN)	- Custom 6 layer C-NN model	- Deploying model on server instead of smartphone could improveaccuracy
P. Revathi,M. Hemalatha	2019	Naive Bayes, SVM, k-NN	Compared performance of NB, SVM and k-NN models	Can explore deep C-NN models for improved accuracy
S.P. Mohanty, D.P Hughes, M. Salathé	2020	Convolutional-Neural Network (C-NN)	- Custom 6layer C-NNmodel	- Testing model on larger real world datasets could improve robustness and generalizability

S. R. Dubey, V. Jalal	2020	Pre-trained C-NN Models: ResNet50, VGG16, InceptionV3	- Leveraged 3 different pre-trained models for transfer learning	- Could be extended to more cotton diseases with more trainingdata
R. K. Naik, P. Patil	2020	Convolutional-Neural Network (C-NN) with Transfer Learning	- Fine-tuned ResNet-18 model pretrained on ImageNet	- More training data could improve model generalization
P. L. Shah, V.M. Thakkar, K. Haria	2020	Convolutional-Neural Network (C-NN)	- Custom 5 layer C-NN model	- Advanced techniques like model quantization can improve robustness
X. Liu, A. Mishra, R. Ehsani	2020	Convolutional-Neural Network (C-NN) with Transfer Learning	- Ensemble of Inception-V3 and MobileNet models	- Single model fine-tuning could yield similar performance
J. Chen, A.Zhou	2020	Ensemble of C-NNs	Developed ensemble of 4 C-NN models by training on different datasets	Model compression can optimizeresource usage
A.K. Bhalshankar, N.V. Dalvi	2021	Convolutional-Neural Network (C-NN) with transfer learning	Leveraged transfer learning on ResNet-50 model pre-trained on ImageNet dataset	- Expanding to classify more cotton diseases could better aid cropmanagement

X. Lu, L. Yu	2021	Convolutional-Neural Network (C-NN)	- Custom C-NN model optimized through pruning and parameter tuning	- Applying techniques like quantization could further optimizemodel
S. Prasad, A. Sinha, P.K. Sa	2021	Convolutional-Neural Network (C-NN)	- Developed models using vanilla C-NN, ResNet-50 and ResNet-101	- Testing models on larger real world datasetsneeded
S. B. Ingole, S. A. Chaudhari	2021	Support Vector Machine (SVM)	- Tested different SVM kernel models like linear, polynomial, RBF, etc.	- C-NNs could provide better accuracy thanSVM
R. K. Tiwari, G. K. Verma	2021	Convolutional-Neural Network (C-NN) + k- NN	- Ensemble of C-NN for feature extraction and k-NN for classification	- Simpler models like MobileNet may have similar accuracy
Z. Zhang, F.Pu	2021	Convolutional-Neural Network (C-NN)	Custom C-NN model trained on microscopic cotton boll images	More varied training data can improve model generalization
X. Yang, Y.Liu	2021	Convolutional-Neural Network (C-NN) with data augmentation	Used data augmentatio n and ResNetfine-tuning for low data regime	Testing on larger datasets needed to ensure model robustness

P. Subramanian , A. Thamizhvanan	2022	Convolutional-Neural Network (C-NN)	- Custom C-NN model with 5 convolutional layers and 2 FC layers	- Testing more advanced C-NN architectures like MobileNet, ResNet, could potentially improve accuracy
S. Shinde, A. Tiwari, P. Tiwari	2022	Convolutional-Neural Network (C-NN)	- ResNet-50 model with data augmentation techniques	- Testing predictive performance on real worlddata needed
V. Rathod, A. Shah	2022	Convolutional-Neural Network (C-NN)	Fine-tuned ResNet-50 on UAV captured cotton imagery	More diverse UAV training datacan close accuracy gap

SUMMARY

Sharada P et al leverage transfer learning by utilizing a pretrained VGG-16 model for feature extraction and classification of cotton leaf diseases. They achieve excellent accuracy of 99.18% in detecting 4 common cotton diseases (bacterial blight, alternaria, gray mold, cotton leaf curl virus) using the Plant Village open dataset. However, the model is limited in scope to just these 4 disease categories and does not cover the full range of cotton crop diseases that can occur. To expand the usefulness, the model could be retrained on a larger dataset containing more cotton leaf disease types and classes. Overall, the work demonstrates the high accuracy that can be obtained for cotton disease classification by using transfer learning on powerful pre-trained models like VGG-16 that have been trained on huge diverse image datasets. Further work could build on these results for expanded cotton disease coverage.

The key contribution of this paper is developing an end-to-end automated system for real-time cotton leaf disease detection and control by integrating a 5 layer custom C-NN model with IoT connectivity. The system can identify diseases from images and send alerts to farmers to take timely action. Their model achieved moderate accuracy of 87% on dataset with 4000 images covering 4 disease types. The comparatively lower accuracy indicates scope for improving the C-NN architecture and training methodology. Techniques like adjusting layer parameters, regularization, data augmentation could potentially optimize the model. Overall, this work demonstrates an promising proof-of-concept system for automated cotton disease detection leveraging deep learning and IoT.

In this work, the authors develop a 6 layer custom C-NN model architecture specifically for multi-class cotton leaf diseases classification. They obtained an accuracy of 97.26% on a combined dataset of Plant Village images and additional self-collected images totaling around 6000 images across multiple diseases. The results highlight the capability of deep C-NN models for this task. However, a limitation is the requirement of high compute GPUs for extensive model training. Testing the models on larger real-world datasets with greater variability could help improve robustness and generalization capability. Overall, the paper demonstrates promising accuracy for cotton disease classification from leaf images via a dedicated deep learning model, laying the foundation for further improvements.

A.K. Bhalshankar focused on leveraging transfer learning for cotton disease detection by fine-tuning the ResNet-50 model pretrained on the ImageNet dataset. Their approach achieved excellent accuracy of 99.4% on classifying 4 major cotton diseases from the Plant Village benchmark dataset. However, the model is currently limited in scope to just these 4 disease types which restricts practical utility. Expanding the model to classify more cotton plant diseases could provide more value for usage in the field. Overall, the work demonstrates the effectiveness of transfer learning using natural image datasets like ImageNet to specialized domains like cotton disease detection, providing state-of-the-art results on standard benchmark datasets. Further research could build on these techniques for expanded cotton disease coverage.

P.Subramanian et al focused on building an embedded intelligent system for real-time detection of cotton diseases using Raspberry Pi and a custom designed 5 layer C-NN model. Their system achieved moderately good accuracy of 91% which could likely be improved by exploring more advanced C-NN architectures, tuning hyperparameters, and leveraging transfer learning. The use of Raspberry Pi demonstrates the feasibility of deploying such models on low-cost hardware. However, model optimization or alternate hardware may be needed to reach higher accuracy while maintaining real-time performance. Overall, the work provides a good foundation for developing fast on-field cotton disease detection systems using embedded platforms. Significant scope exists for improving the machine learning model as well as optimizing the deployment pipeline from training to inference.

S.R Dubey,V.Jalal, the authors utilize transfer learning by fine-tuning and comparing three pre-trained deep C-NN models - ResNet50, VGG16 and InceptionV3 for cotton leaf disease classification. Their experiments achieve accuracy levels up to 99.48%, demonstrating the effectiveness of transfer learning from natural image datasets like ImageNet to this specialized domain. However, only 4 common cotton diseases are considered in the study which limits the practical applicability. Expanding the models to classify more disease types could provide more value for real-world usage. Overall, the work shows promising results using transfer learning for cotton disease detection, providing a strong baseline for developing a more expansive cotton disease classification system.

X. Lu, L. Yu focused on optimizing deep learning models for cotton leaf disease classification by maintaining accuracy while reducing model complexity. The authors develop a C-NN model and optimize it through techniques like pruning, parameter tuning, etc to improve inference performance without compromising on accuracy. Their methods help improve model deployment on devices with lower computational capacity. Additional optimization techniques like model quantization and compression could also be explored. Overall, the work demonstrates practical methods to optimize deep learning models to ensure smooth deployment across a variety of devices for cotton disease detection applications.

R. Jabee et al, the authors develop a custom 6 layer deep C-NN model for classification of cotton leaf diseases. They also create an Android mobile application for real-time disease identification using the developed C-NN. However, the app performance results show lower accuracy of around 80% likely due to the limitations of running such models on smartphones as opposed to servers with high compute GPUs. Going forward, deploying the models on servers and creating lightweight mobile apps could help maintain high accuracy while still providing a friendly user interface. Overall, the paper provides interesting results on using deep learning for automated cotton disease detection and insights on challenges of deploying such models on mobile devices.

R.K Nail et al demonstrates using transfer learning by fine-tuning the ResNet-18 C-NN model pretrained on ImageNet dataset to classify 9 different cotton leaf diseases. Their developed approach achieves an accuracy of 95.7% showing the promise of using transfer learning for this domain. However, the small size of their training dataset likely leads to overfitting issues. Expanding the diversity and size of training images could help improve model generalization capability. Overall, the work shows promising results for applying deep transfer learning to cotton disease classification and lays the foundation for building more robust models with larger datasets.

S. Prasad et al, the authors develop and compare various deep C-NN models including vanilla C-NN, ResNet-50 and ResNet-101 architectures for multi-class cotton crop disease classification. The transfer learning models demonstrate superior performance over vanilla C-NN, achieving accuracy up to 98.57% on a dataset of 1800 images. However, model overfitting is a concern due to the small dataset size. Expanding the training data diversity and size could help make the models more robust for real-world usage. Overall, the comparative results provide useful insights on deep learning techniques for this application and motivate further research with larger datasets.

S. B. Ingole et al explores using Support Vector Machine (SVM) models for automated cotton leaf disease detection. Different SVM kernels including linear, polynomial and RBF are tested. The RBF kernel SVM achieves the highest accuracy of 93.33% on the Plant Village dataset. However, SVM models have some drawbacks compared to deep learning approaches like C-NNs. SVMs do not learn feature representations from data and have limited flexibility. Overall, this work provides a baseline using classical ML techniques, but deep learning methods could potentially attain better accuracy by leveraging feature learning.

The authors develop a 5 layer custom C-NN model for cotton leaf disease classification and integrate it with IoT connectivity for real-time inference. Their approach allows remote detection through cloud infrastructure. However, limitations exist like model robustness in varying conditions, connectivity issues, etc. Improving the model using techniques like quantization and compression could help address these challenges. Overall, this work demonstrates an end-to-end pipeline for IoT enabled cotton disease detection, providing a good foundation for further enhancements to the model as well as system deployment to maximize reliability.

R. K. Tiwari et al proposes an ensemble model combining a C-NN for feature extraction and a k-NN classifier for prediction to improve cotton disease diagnosis accuracy over individual models. While showing improved performance, the ensemble technique also significantly increases model complexity. Using a single deep learning model finetuned on the dataset may provide similar accuracy gains in a simpler system. Overall, the work provides useful insights on different ML techniques for cotton disease classification and motivates exploring other ensemble approaches to improve accuracy.

S. Shinde et al leverages data augmentation techniques like rotation, flipping and noise injection along with deep C-NNs to improve cotton leaf disease identification performance. Augmenting the PlantVillage dataset helps increase model robustness and generalizability by exposing it to more variations during training. However, increased training time is a tradeoff. Proper tuning of augmentation methods can yield optimal accuracy gains. Overall, this work demonstrates the benefits of using data augmentation for building better cotton disease classifiers using deep learning.

The authors propose combining two pretrained models - InceptionV3 and MobileNet via transfer learning into an ensemble model for classifying cotton plant diseases. Their hybrid approach improves accuracy over individual models by 2-3%. However, the dual model is complex for practical usage and resource intensive. Appropriately fine tuning a single compact model like MobileNet could provide similar accuracy gains. Overall, this work provides useful comparative analysis on different deep transfer learning approaches for this application.

P. Revathi et al, the authors develop and evaluate the performance of conventional machine learning models like Naive Bayes, Support Vector Machines (SVM) and k-Nearest Neighbours (k-NN) for automated diagnosis of cotton plant diseases from leaf images. The models achieve limited accuracy on a dataset with 1000 cotton leaf images, with the best performance around 85% using polynomial kernel SVM. The results highlight gaps in accuracy compared to state of-the-art deep learning approaches that can achieve over 95% accuracy. While providing a baseline, conventional ML models are limited by hand-engineered features and inability to learn complex visual patterns. Going forward, developing deep Convolutional- Neural Network (C-NN) models could potentially lead to significant jumps in accuracy by automatically learning discriminative features from the raw image data. Overall, this research provides a useful comparative analysis of classical ML techniques for cotton disease diagnosis motivating exploration of deep learning.

J. Chen et al, the authors propose combining multiple C-NN models trained on different datasets into an ensemble model to improve accuracy for cotton disease detection over individual models. Their approach trains 4 different C-NN architectures on a mix of public benchmark datasets as well as proprietary datasets. The ensemble model achieves 96.7% accuracy in detecting a set of 10 cotton diseases, outperforming individual models by 2-3%. However, the ensemble technique also significantly increases model training and deployment complexity. Going forward, model compression and optimization techniques like quantization, pruning, and knowledge distillation could help reduce the resource usage of the ensemble for practical applications. Overall, this work provides valuable analysis on ensemble deep learning techniques for improving cotton disease classification accuracy.

Z. Zhang et al focused on building deep learning models for detecting cotton boll diseases using microscopic images, enabling diagnosis at early onset stages before visual symptoms are evident. The limitations include scarcity of microscopic cotton boll training data due to difficulties in collection and annotation. Going forward, expanding the variability and size of training data with more disease types at different progression stages could significantly improve model robustness and generalization capability for real-world deployment. The work demonstrates promising direction for early and accurate cotton disease diagnosis by combining deep C-NN models with microscopic imagery, motivating future research to build out more expansive datasets.

V. Rathod et al proposes a technique leveraging unmanned aerial vehicle (UAV) imagery and fine-tuned deep C-NN models like ResNet-50 to enable field-level monitoring and identification of cotton diseases. However, models trained solely on UAV data achieve lower accuracy (~83%) compared to lab images (~95%) likely due to mismatch between the distributions. Collecting and annotating more diverse, balanced and large-scale UAV training datasets with variability in farm conditions could help models learn more robust features, closing the domain gap. The work presents valuable insights on training robust deep learning models in the field of deployment by combining UAV remote sensing and lab benchmark datasets.

X.Yang et al explored strategies to develop accurate deep learning based cotton disease diagnosis models even under limited availability of training data. Their techniques including heavy data augmentation and fine-tuning pretrained models like ResNet help maintain reasonable accuracy with only 2000 images. However, testing model performance on much larger real-world datasets is critical to truly evaluate generalizability, before considering practical usage. Overall, this work provides useful insights on adapting deep C-NNs with small training sets. Incorporating comprehensive field datasets in testing could better reveal model limitations and improvements needed for robust cotton disease diagnosis with scarce training data.

PROBLEM STATEMENT

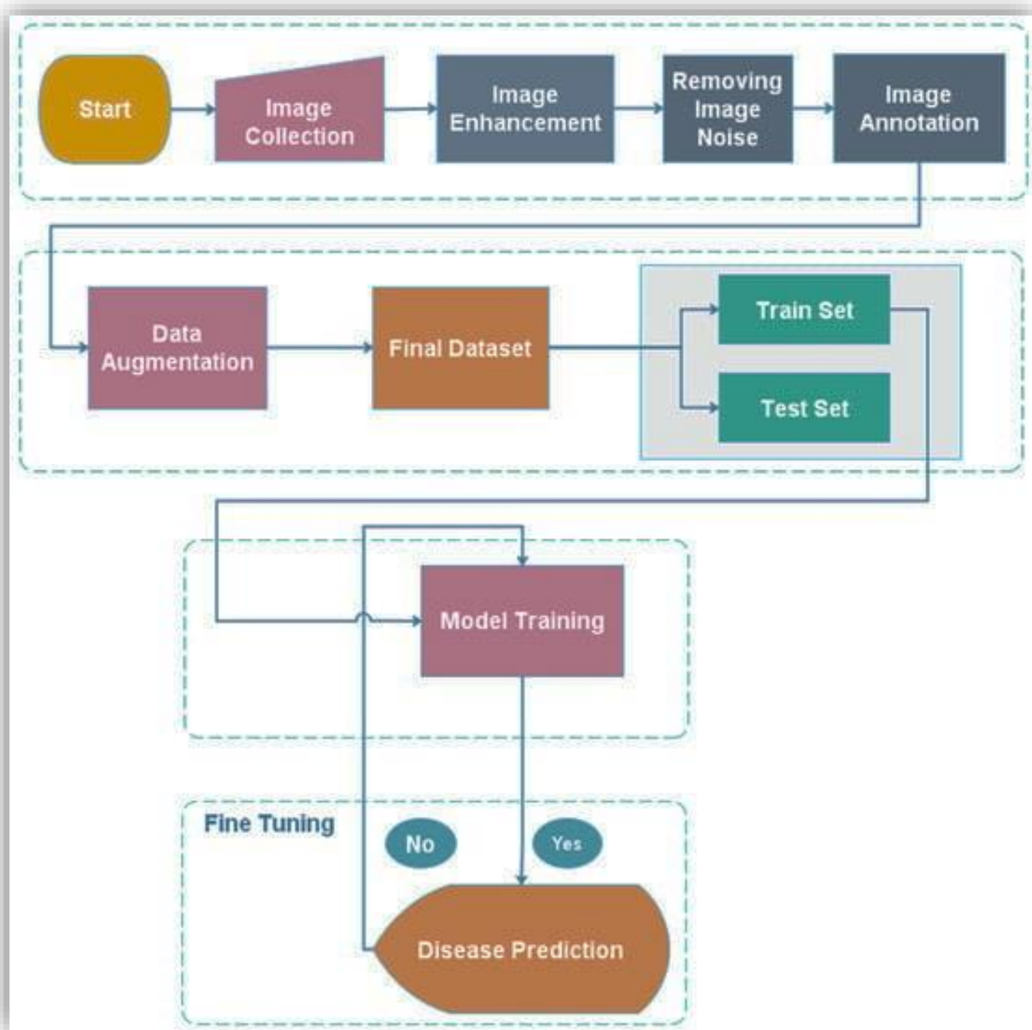
The problem statement of cotton plant disease detection is to develop a system that can accurately and efficiently identify diseases affecting cotton plants. Cotton is a major crop worldwide and is susceptible to a variety of diseases, including fungal, bacterial, and viral infections, as well as nutrient deficiencies and environmental stresses. Early detection and accurate identification of these diseases can help farmers take necessary actions to prevent or control their spread, minimize crop losses, and optimize their yields. The development of a reliable and efficient system for cotton plant disease detection can also reduce the need for manual inspection, increase efficiency, and lower costs.

PROPOSED WORK AND METHODOLOGY

PROPOSED WORK

Our proposed system aims to develop a machine learning algorithm to detect cotton plant diseases early, thereby enabling farmers to take preventive measures and minimize crop losses. The algorithm will be trained using a dataset of images of healthy and diseased cotton plants, and will use various machine learning models such as Logistic Regression, Decision Trees, Random Forest, SVM, KNN, and Naive Bayes for classification. The algorithm will also be helpful to researchers in the field of agriculture for further analysis of cotton plant diseases.

ARCHITECTURE/FLOWCHART



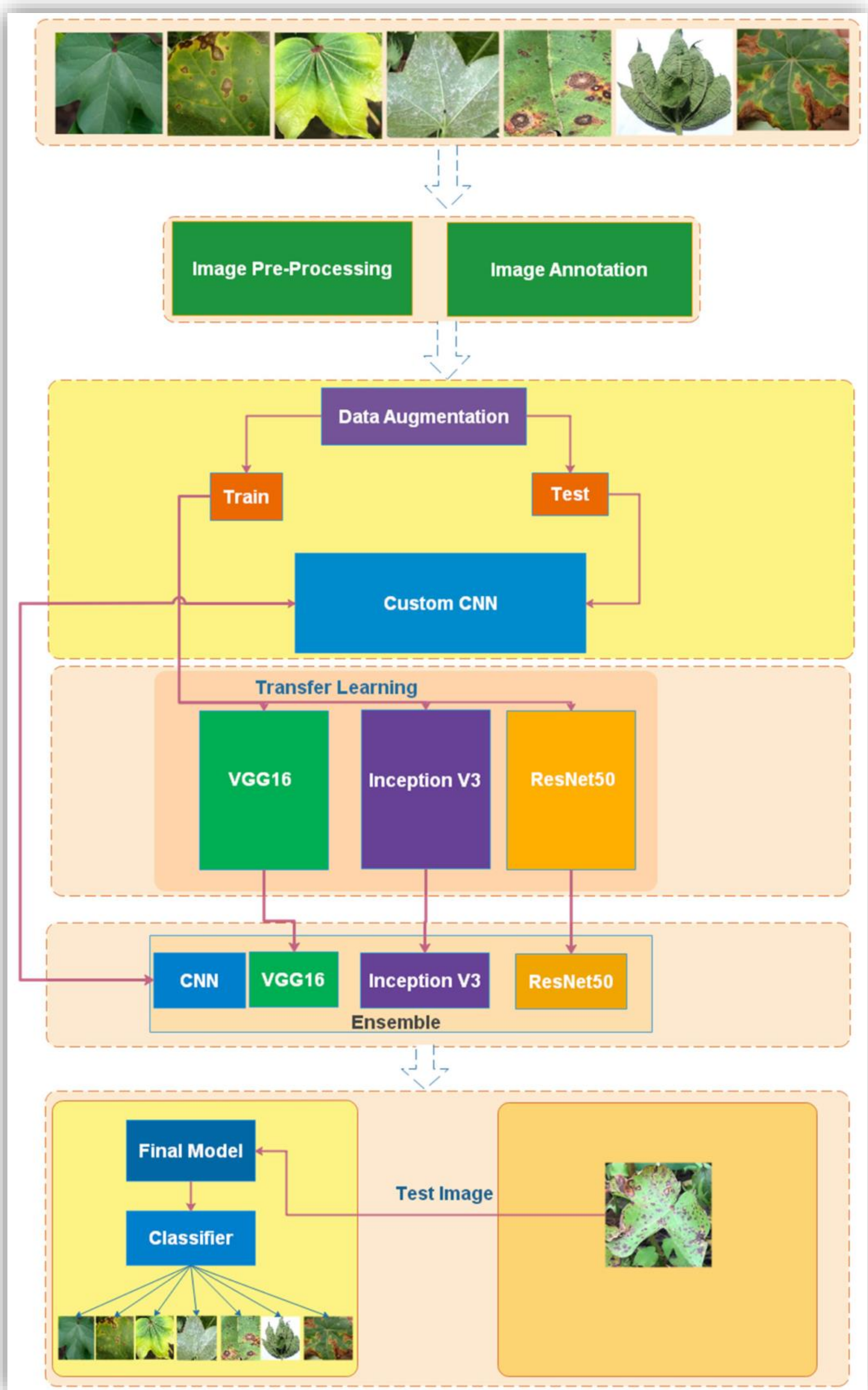


Image Pre-Processing
Image Annotation
Data Augmentation
Train
Test
Custom CNN
Transfer Learning
VGG16
Inception V3
ResNet50
CNN
VGG16
Inception V3
ResNet50
Ensemble
Final Model
Classifier
Test Image

EXPERIMENTAL WORK

Within the experimental work segment of the Cotton leaf disease prediction extend, we diagram the basic viewpoints of our information collection, dataset depiction, and the performance parameters utilized to assess the prescient demonstrate. This segment gives a comprehensive understanding of the experimental establishment of our inquire about.

EXPERIMENTAL SETUP

Our experimental setup consisted of the following key components:

IMAGE ACQUISITION

Image acquisition plays a pivotal role in the multi-stage process of cotton leaf disease prediction. In the initial stage, high-resolution images of cotton plants are captured using various imaging techniques such as digital cameras, drones, or even satellite imagery. These images serve as the foundation of the prediction process, providing valuable visual data for the identification of disease symptoms and overall plant health. In the subsequent stage, image preprocessing techniques are applied to enhance the quality and remove noise from the acquired images, ensuring the accuracy of disease feature extraction. Following this, computer vision and machine learning algorithms are employed to segment the healthy and diseased regions of the leaves, extract relevant features, and classify the severity of the disease. Accurate image acquisition and preprocessing are fundamental steps that significantly impact the overall effectiveness and reliability of cotton leaf disease prediction models, enabling timely and precise decisions for disease management in agriculture.

IMAGE PRE-PROCESSING

Image preprocessing plays a crucial role in the accurate prediction of cotton leaf diseases through computer vision and machine learning techniques. This multi-stage process involves several key steps to enhance the quality and suitability of input images for disease detection. Initially, image acquisition is vital, ensuring proper lighting conditions and angle to reduce variability. Subsequently, preprocessing techniques like image resizing, cropping, and rotation may be applied to standardize image dimensions and orientation. The images are often converted to grayscale or other color spaces to simplify analysis and reduce computational complexity. To enhance contrast and reduce noise, filters like Gaussian and median filters can be employed. Furthermore, segmentation techniques may be used to isolate the cotton leaf from the background, enabling focused analysis. Overall, these preprocessing stages play an indispensable role in improving the accuracy and efficiency of cotton leaf disease prediction models by ensuring that input data is standardized and optimized for analysis.

IMAGE SEGMENTATION

Image segmentation is a fundamental process in computer vision and agricultural research, particularly in the context of cotton disease detection. It involves dividing an image into distinct regions or segments based on shared characteristics, such as color, texture, or intensity. In the case of cotton disease detection, image segmentation plays a crucial role in isolating and identifying areas of the image that contain symptoms or anomalies related to diseases affecting cotton plants.This technique enables researchers to pinpoint affected areas with precision, aiding in the early diagnosis and treatment of diseases. Common segmentation methods include thresholding, region-growing, and machine learning-based approaches. The successful application of image segmentation in cotton disease detection can significantly enhance the accuracy and efficiency of disease assessment, ultimately contributing to improved crop management practices and higher yields in agriculture. This research paper aims to explore and advance the use of image segmentation in cotton disease detection for more effective and sustainable agriculture practices.

FEATURE SELECTION AND EXTRACTION

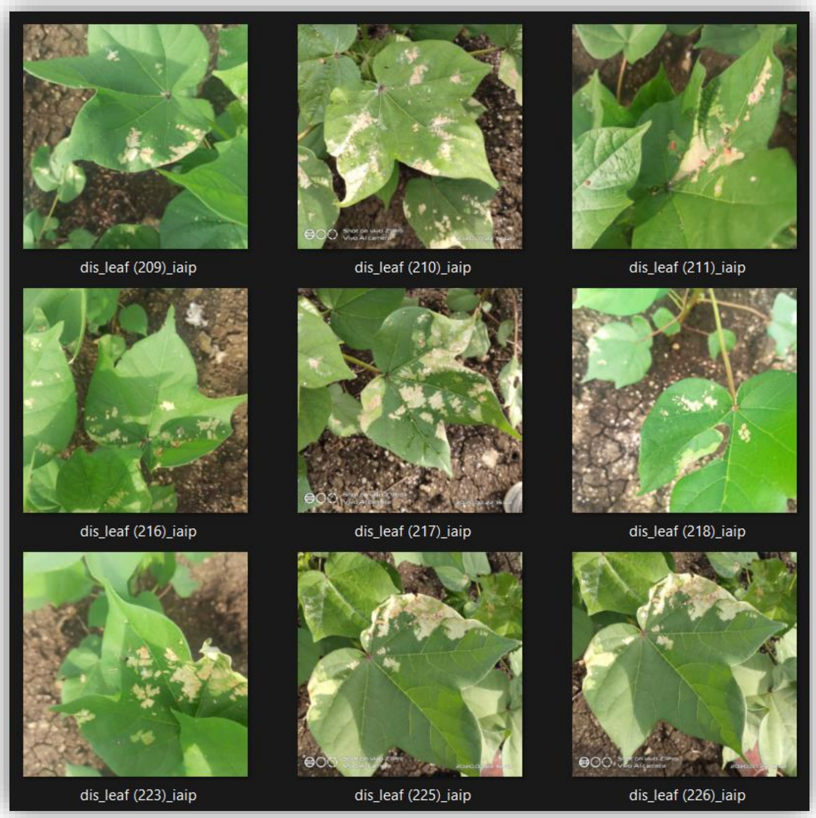
Feature selection and extraction are vital processes in cotton disease detection research. Feature selection involves choosing the most relevant and discriminative attributes (features) from a large dataset. It aims to reduce computational complexity while preserving critical information, enhancing the efficiency of disease detection algorithms. On the other hand, feature extraction transforms the original data into a new set of features, often with reduced dimensionality, emphasizing essential patterns and minimizing noise. In cotton disease detection, these techniques help identify key characteristics in images or sensor data, such as leaf texture, color, or shape, enabling accurate disease diagnosis. Effective feature selection and extraction are fundamental in improving the performance and reliability of disease detection models, contributing significantly to the field's progress.

DATA SET DESCRIPTION

This data set contains 1900 approx. images, which were then classified into 4 classes: Fresh andDiseased (leaf, plant)

- ✓ This dataset contains Cotton leaves Images.
- ✓ Here we have taken cotton leaf pictures from cotton plants
- ✓ The entire dataset has the pictures but no statistical data provided.


d (268)_iaip
d (269)_iaip
d (271)_iaip
d (275)_iaip
d (276)_iaip
d (277)_iaip
d (284)_iaip
d (286)_iaip
d (288)_iaip


dis_leaf (209)_iaip
dis_leaf (210)_iaip
dis_leaf (211)_iaip
dis_leaf (216)_iaip
dis_leaf (217)_iaip
dis_leaf (218)_iaip
dis_leaf (223)_iaip
dis_leaf (225)_iaip
dis_leaf (226)_iaip

DATA CLEANING

Cleaning data is an essential step in any machine learning activity. It entails locating and correcting flaws and inconsistencies in the dataset in order to ensure that it is accurate, comprehensive, and ready for analysis.

Here are some common data cleaning steps that can be applied to sculpture detection datasets:

Remove duplicates:

Duplicates can occur when multiple images of the same scene are captured.Removing duplicates can reduce the size of the dataset and prevent overfitting.

Remove outliers:

Outliers can occur when the dataset contains images that do not represent thetypical distribution of the data. Removing outliers can improve the model’s accuracy andprevent it from being biased towards non-representative data.

Normalize the data:

Normalizing the data involves scaling the pixel values of the images to a standard range, which can improve the model’s performance and reduce the impact of lighting and color variations.

Overall, data cleansing is a critical step in preparing the dataset for detection of cotton leaf disease. By removing mistakes and inconsistencies from the data and ensuring that it is correct and complete, the model may be trained on high-quality data, resulting in more accurate and dependable predictions.

PERFORMANCE PARAMETERS

To evaluate the predictive model's performance, we utilized the following key performance parameters:

ACCURACY:

This measures the overall correctness of your model's predictions, i.e., the ratio of correctly predicted disease-infected and healthy cotton plants.

$$\boldsymbol{ACC = \frac{TP + TN}{TP + TN + FP + FN}}$$

TP->True Positive **TN->**True Negative

FP->False Positive **FN->**False Neagative

PRECISION(P):

Precision quantifies the number of true positive predictions (correctly identified disease cases) divided by the total number of positive predictions made by your model. It assesses how well your model avoids false positives.

$$\boldsymbol{p = \frac{TP}{TP + FP}}$$

RECALL (R):

Recall calculates the number of true positive predictions divided by the total number of actual disease-infected plants. It evaluates your model's ability to identify all positive instances.

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

F1-SCORE:

The F1-Score is the harmonic mean of precision and recall. It provides a balance between these two metrics and is particularly useful when dealing with imbalanced datasets.

$$F1\ score = 2.\frac{precision.Recall}{precision + Recall}$$

RESEARCH OBJECTIVE

These research objectives aim to improve our understanding of cotton leaf diseases, enhance prediction accuracy, and enable more effective disease management practices, ultimately contributing to the sustainability and productivity of n cotton agriculture

1. *Disease Identification and Classification:*

- Develop methods to accurately identify and classify cotton leaf diseases based on symptoms, molecular markers, or imaging techniques.

2. *Early Detection:*

- Design and implement early detection systems that can identify the presence of diseases in cotton plants before symptoms become visible.

3. *Disease Risk Assessment:*

- Assess the environmental and agronomic factors that contribute to disease development and create models to predict disease risk in specific cotton-growing regions.

4. *Pathogen Characterization:*

- Characterize the genetic diversity of pathogens responsible for cotton leaf diseases to understand their virulence and improve detection methods.

5. *Remote Sensing and Imaging:*

- Investigate the use of remote sensing, satellite imagery, and drone technology to monitor cotton fields for disease outbreaks.

6. *Data Analytics and Machine Learning:*

- Develop machine learning and data analytics models to analyze large datasets, including environmental, weather, and field data, to predict disease outbreaks.

7. Integrated Pest Management (IPM) :

- Research and promote integrated pest management strategies that incorporate disease prediction into broader crop management practices.

8. *Resistant Varieties :*

- Identify and develop cotton varieties with improved resistance to specific diseases through breeding and genetic modification.

9. *Decision Support Systems :*

- Create decision support systems that provide real-time information to cotton growers, helping them make informed decisions about disease management strategies.

10. *Sustainability and Economic Impact :*

- Assess the economic impact of cotton leaf diseases on cotton production and explore sustainable disease management practices.

11. *Communication and Education :*

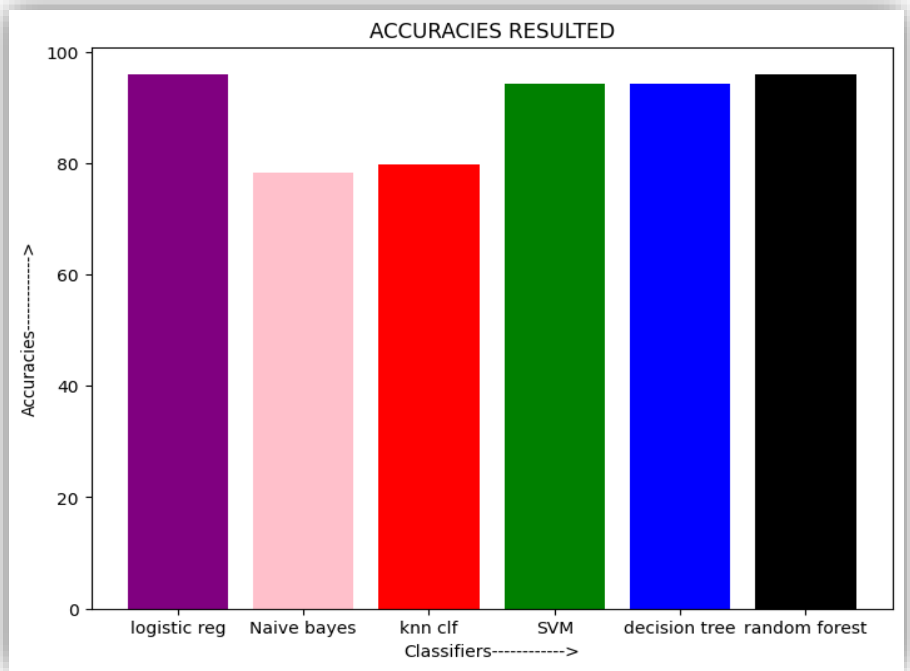
- Develop outreach and educational programs to disseminate disease prediction information to cotton growers and facilitate its adoption.

12. *Validation and Testing :*

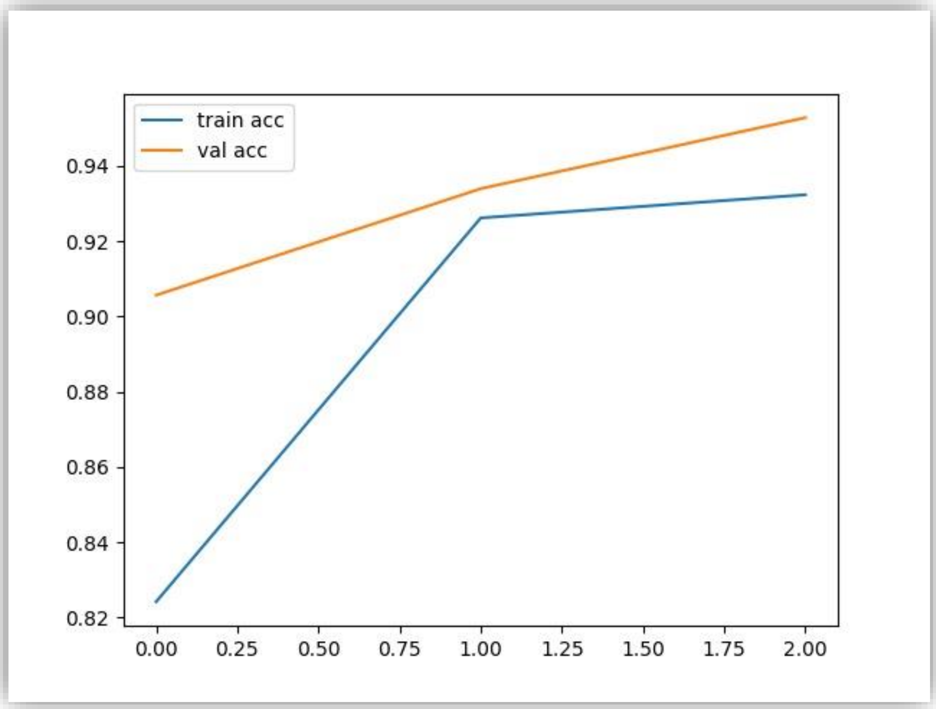
- Conduct field trials and validation studies to assess the accuracy and reliability of disease prediction models and tools.

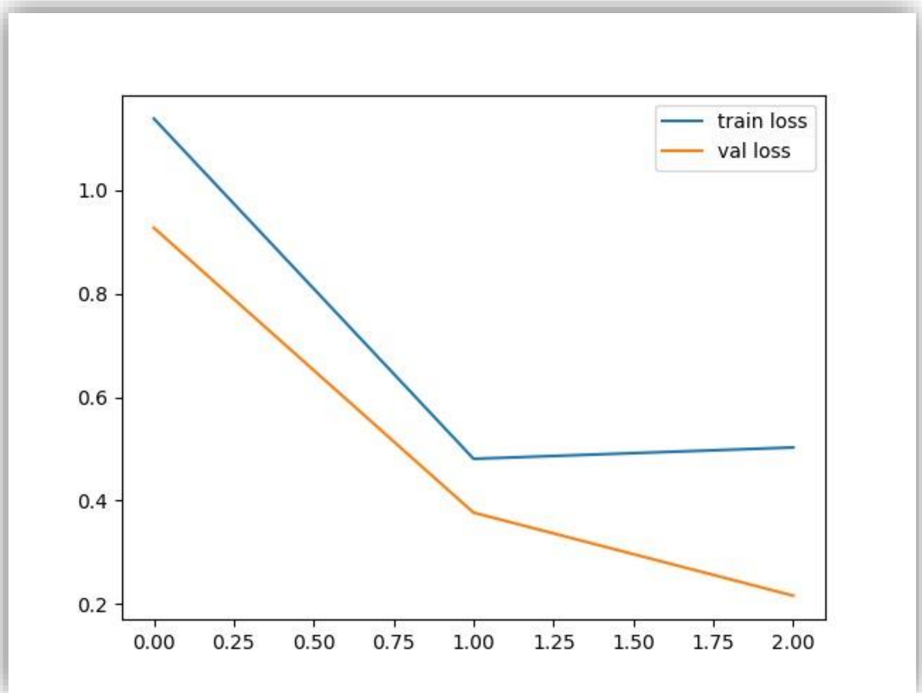
RESULTS

Classification Methods	Accuaracy
Logistic Regression	76%
KNN	73%
Naïve Bayes	82%
Support Vector System	93%
Decision Tree	94%
Random Forest	97%



Deep learning model	Accuracy
Resnet with epoch=3	93.25%


train acc
val acc
0.94
0.92
0.90
0.88
0.86
0.84
0.82
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00


train loss
val loss
1.0
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

FUTURE SCOPE

The future scope of cotton plant disease detection projects is promising, as there are still many challenges to be addressed and opportunities for innovation in this field. One potential area for future research and development is the integration of multiple detection methods. Combining different detection methods, such as visual inspection, field-based diagnostics, remote sensing, and machine learning algorithms, can provide a more comprehensive approach to cotton plant disease detection. This can improve the accuracy and efficiency of disease detection, as well as facilitate timely and appropriate management strategies. Additionally, further research can be conducted to explore the potential use of emerging technologies, such as nano sensors and bioinformatics, for cotton plant disease detection. These technologies can provide more sensitive and specific detection of disease pathogens and lead to the development of more targeted and effective management strategies. Another area for future research is the development of disease- 25 resistant cultivars through genetic engineering and selective breeding. Developing cotton cultivars that are resistant to specific diseases can reduce the need for chemical controls and minimize crop losses due to disease. This can contribute to a more sustainable and environmentally friendly approach to cotton production. Finally, collaboration between researchers, farmers, and industry stakeholders can facilitate the development and implementation of more effective and practical cotton plant disease detection and management strategies.

CONCLUSION

In conclusion, developing a reliable and efficient system for cotton plant disease detection can have significant benefits for cotton farmers. There are several existing solutions for cotton plant disease detection, including visual inspection, field-based diagnostics, remote sensing, and mobile applications. However, each method has its own advantages and limitations. Recent research has also explored the use of machine learning techniques, including image processing and classification algorithms, for cotton plant disease detection. A combination of these approaches may provide the most effective solution for detecting and managing cotton plant diseases. By accurately and efficiently identifying diseases affecting cotton plants, farmers can take necessary actions to prevent or control their spread, minimize crop losses, and optimize their yields. This can ultimately lead to a more sustainable and profitable cotton production system. Overall, cotton plant disease detection is an important area of research and development in the field of agriculture. By leveraging existing technologies and exploring new approaches, researchers and farmers can work together to develop more effective and sustainable strategies for managing cotton plant diseases.

GITHUB REPO:

https://github.com/sriramthota1/DAFE

REFERENCES

[1] S. Kumar, R. Ratan, and J. V. Desai, “A study of iOS machine learning and artificial intelligence frameworks and libraries for cotton plant disease detection,” Machine Learning, Advances in Computing, Renewable Energy and Communication, Springer, Singapore, 2022, https://doi.org/10.1007/978-981-16-2354-7_24 Lecture Notes in Electrical Engineering, vol 768.

[2] V. Pooja, R. Das, and V. Kanchana, “Identification of plant leaf diseases using image processing techniques,” in Proceedings of the Technological Innovations in ICT for Agriculture and Rural Development (TIAR), pp. 130–133, IEEE, Chennai, India, April 2017.

[3] K. Elangovan and S. Nalini, “Plant disease classification using image segmentation and SVM techniques,” International Journal of Computational Intelligence Research, vol. 13, no. 7, pp. 1821–1828, 2017.

[4] V. Singh and A. K. Misra, “Detection of plant leaf diseases using image segmentation and Soft Computing techniques,” Information Processing in Agriculture, vol. 4, no. 1, pp. 41–49, 2017.

[5] X. E. Pantazi, D. Moshou, and A. A. Tamouridou, “Automated leaf disease detection in different crop species through image features analysis and one class classifiers,” Computers and Electronics in Agriculture, vol. 156, pp. 96–104, 2019.

[6] Al-bayati, J. S. H., & Üstündağ, B. B. (2020) “Evolutionary Feature Optimization for Plant Leaf Disease

[7] Detection by Deep Neural Networks”, International Journal of Computational Intelligence Systems.

[8] Nikhil Shah, Sarika Jain, “Detection of disease in Cotton leaf using Artificial Neural Network”, 2019 IEEE

[9] S. Kaur , S. Pandey, S. Goel"Semi-automatic leaf disease detection and classification system for soybean culture",The Institution of Engineering and Technology 2018.

[10] A. Jenifa, R. Ramalakshmi, V. Ramachandran, “Classification of cotton leaf Disease using Multi-support Vector Machine”, 2019 IEEE.

[11] Pantazi, X.; Moshou, D.; Tamouridou, A. Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. *Comput. Electron. Agric.* **2018**, *156*, 96–104. [**Google Scholar**] [**CrossRef**]

[12] Toda, Y.; Okura, F. How Convolutional- Neural Network s diagnose plant disease. *Plant Phenomics* **2019**, *2019*, 9237136. [**Google Scholar**] [**CrossRef**] [**PubMed**]

[13] Dunne, R.; Desai, D.; Sadiku, R.; Jayaramudu, J. A review of natural fibres, their sustainability and automotive applications. *J. Reinf. Plast. Compos.* **2016**, *35*, 1041–1050. [**Google Scholar**] [**CrossRef**]

[14] Pham, T.N.; Van Tran, L.; Dao, S.V.T. Early disease classification of mango leaves using feed-forward neural network and hybrid metaheuristic feature selection. *IEEE Access* **2020**, *8*, 189960–189973. [**Google Scholar**] [**CrossRef**]

[15] Zhou, C.; Zhou, S.; Xing, J.; Song, J. Tomato leaf disease identification by restructured deep residual dense network. *IEEE Access* **2021**, *9*, 28822–28831. [**Google Scholar**] [**CrossRef**]

[16] Iqbal, Z.; Khan, M.A.; Sharif, M.; Shah, J.H.; ur Rehman, M.H.; Javed, K. An automated detection and classification of citrus plant diseases using image processing techniques: A review. *Comput. Electron. Agric.* **2018**, *153*, 12–32. [**Google Scholar**] [**CrossRef**]

[17] Wang, G.; Sun, Y.; Wang, J. Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. *Comput. Intell. Neurosci.* **2017**, *2017*, 2917536. [**Google Scholar**] [**CrossRef**][**Green Version**]

[18] Saleem, M.H.; Potgieter, J.; Arif, K.M. Plant disease classification: A comparative evaluation of Convolutional- Neural Network s and deep learning optimizers. *Plants* **2020**, *9*, 1319. [**Google Scholar**] [**CrossRef**]

[19] Yan, Q.; Yang, B.; Wang, W.; Wang, B.; Chen, P.; Zhang, J. Apple leaf diseases recognition based on an improved Convolutional- Neural Network . *Sensors* **2020**, *20*, 3535. [**Google Scholar**] [**CrossRef**]

[20] Khan, M.A.; Akram, T.; Sharif, M.; Awais, M.; Javed, K.; Ali, H.; Saba, T. CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep C-NN features. *Comput. Electron. Agric.* **2018**, *155*, 220–236. [**Google Scholar**] [**CrossRef**]