International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

Optimization Accuracy of Cotton Disease Detection using Gradient Boosting Neural Network Techniques

Arwadeep Gupta, Dr. Vivek Rechhariya
Department of Computer Science and Engineering
Lakshmi Narain College of Technology, Bhopal, M.P. India

Abstract:- Cotton is an economically important crop, but its productivity is significantly affected by various leaf diseases that can reduce crop quality and yield. Early and accurate identification of cotton diseases is therefore essential for effective crop management and timely treatment. Traditional disease identification methods rely heavily on manual inspection by experts, which can be time-consuming, subjective, and difficult to apply over large agricultural fields. This study proposes an optimized deep learning-based approach for automatic cotton disease detection using Gradient Boosting Neural Network (GBNN) techniques. The proposed framework utilizes image preprocessing, feature representation, and optimized learning to distinguish between healthy cotton leaves and different disease categories. Gradient boosting is incorporated to improve classification accuracy by iteratively reducing prediction errors and strengthening the model's ability to identify complex disease patterns. The performance of the proposed approach is evaluated using standard classification measures, including accuracy, precision, recall, and F1-score. The optimization strategy aims to enhance the robustness and generalization capability of the model while reducing misclassification among visually similar disease symptoms. The proposed system can provide a reliable and automated solution for early cotton disease diagnosis, supporting farmers and agricultural experts in making timely crop-management decisions. The study demonstrates the potential of optimized Gradient Boosting Neural Network techniques for developing efficient, accurate, and scalable intelligent systems for cotton disease detection and precision agriculture.

Keywords: Cotton Disease (CD), Gradient Boosting Neural Network (GBNN), Accuracy, Recall

I. INTRODUCTION

Cotton is one of the most important commercial crops and plays a significant role in the agricultural economy by providing raw materials for the textile and other allied industries. However, cotton production is considerably affected by various diseases that occur on leaves, stems, and other plant parts. Among these, leaf diseases are particularly important because they can spread rapidly and adversely affect photosynthesis, plant growth, fiber quality, and overall crop yield. Early identification of diseases is therefore essential for implementing appropriate management practices and minimizing economic losses [1].

Traditionally, cotton diseases are identified through visual inspection by farmers or agricultural experts. Although expert-based diagnosis can provide useful results, it is time-consuming, labour-intensive, and influenced by the experience and knowledge of the observer. Moreover, the visual symptoms of different diseases may have similar characteristics, making accurate identification difficult, particularly during the early stages of infection. These limitations create a need for automated and intelligent disease detection systems capable of analysing plant images efficiently and providing consistent classification results.

The rapid development of artificial intelligence, machine learning, and deep learning has created new opportunities for automated plant disease detection. Image-based disease classification methods can extract meaningful patterns from cotton leaf images, such as changes in color, texture, shape, and lesion characteristics. However, the performance of conventional machine learning models may be affected by variations in illumination, image background, disease severity, leaf orientation, and similarities between disease classes. Therefore, developing an optimized learning framework is important for improving the reliability and accuracy of cotton disease classification [2, 3].

Gradient Boosting techniques have demonstrated strong potential for classification problems because they combine multiple weak learners sequentially to construct a more powerful predictive model. By focusing on previously misclassified samples and minimizing prediction errors during successive learning stages, gradient boosting can effectively capture complex relationships within image-derived features. The integration of gradient boosting with neural network-based feature learning can further enhance the capability of the detection system by combining effective feature representation with improved classification performance [4].

In this study, an optimized GBNN approach is proposed for cotton disease detection. The proposed methodology focuses on improving classification accuracy through optimized learning and systematic analysis of cotton leaf images. The model is evaluated using important performance measures such as accuracy, precision, recall, and F1-score to determine its effectiveness in distinguishing healthy and diseased cotton leaves. The objective is to develop an accurate, robust, and automated disease detection framework that can assist farmers and agricultural experts in early diagnosis and timely disease-management decisions. Such an intelligent system can contribute to precision agriculture by reducing dependence on manual inspection and supporting efficient and sustainable cotton production [5, 6].

II. NEURAL NETWORK

A NN is an artificial intelligence technique inspired by the structure and functioning of the human brain. It consists of interconnected processing units called neurons, which are generally organized into an input layer, one or more hidden layers, and an output layer. Each connection between neurons is assigned a weight, and the network learns appropriate weights and biases during training by minimizing the difference between predicted and actual outputs.

For cotton disease detection, a neural network can be trained using cotton leaf images or extracted image features to identify patterns associated with healthy and diseased leaves. The input layer receives the relevant image features, while the hidden layers perform successive transformations to learn complex relationships between the features. The output layer produces the final disease classification [7, 8]. During training, optimization algorithms such as gradient descent are used to update the network parameters and reduce classification error.

Neural networks are particularly useful for plant disease detection because disease symptoms can vary in terms of color, texture, shape, lesion distribution, and severity. The ability of neural networks to learn nonlinear and complex feature relationships makes them suitable for distinguishing between visually similar disease classes. When combined with Gradient Boosting techniques, neural-network-based feature learning can be further optimized to improve classification accuracy, robustness, and generalization. Thus, a neural network provides an important foundation for developing an automated and intelligent cotton disease detection system [9].

III. COTTON IMAGE AND DISEASE

Cotton leaf images provide important visual information for the identification and classification of diseases affecting cotton plants. The images generally contain healthy and diseased leaves exhibiting different symptoms such as leaf_curl, fusarium_wilt, bacterial bright and healthy image. These visual characteristics can be analyzed using image processing and neural network techniques to automatically distinguish healthy plants from different disease categories.

For an automated cotton disease detection system, the image dataset is generally divided into training, validation, and testing sets. During preprocessing, images may be resized and normalized to obtain a consistent input format. Data augmentation techniques such as rotation, flipping, cropping, and brightness adjustment can also be applied to increase dataset diversity and improve model generalization. The processed images are then provided to the neural network for feature learning and disease classification.

Figure 1: Leaf_curl disease

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

Figure 2: Bacterial bright

Figure 3: Fussarium_wilt

Figure 4: Healthy Image

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

IV. RESEARCH METHODOLOGY

The proposed research methodology focuses on developing an automated and optimized system for cotton disease detection using Gradient Boosting Neural Network (GBNN) techniques. The complete methodology consists of several sequential stages, including image dataset collection, preprocessing, feature extraction, model development, optimization, training, and performance evaluation. The overall objective is to accurately distinguish healthy cotton leaves from different disease categories based on their visual characteristics.

1. Image Dataset Collection

The first stage involves collecting a dataset of cotton leaf images representing both healthy and diseased conditions. The dataset contains images belonging to different cotton disease categories. Each image is assigned an appropriate class label so that the neural network can learn the distinguishing characteristics of each disease. The dataset should contain sufficient variations in leaf orientation, disease severity, illumination, background, and image quality to improve the robustness of the proposed model.

2. Image Preprocessing

The collected images are preprocessed before being provided to the learning model. Image preprocessing includes resizing the images to a uniform dimension, normalization of pixel values, and removal of unwanted noise or irrelevant information. Data augmentation techniques such as rotation, horizontal and vertical flipping, scaling, cropping, and brightness variation can be applied to increase the diversity of training samples. This process helps reduce overfitting and improves the model's ability to classify images obtained under different conditions.

3. Feature Extraction

After preprocessing, important visual characteristics are extracted from the cotton leaf images. These characteristics include color, texture, shape, edges, lesion patterns, and spatial distribution of infected regions. Neural-network-based feature learning can automatically identify high-level representations from the input images, eliminating the need for completely manual feature engineering. The extracted features provide the basis for distinguishing healthy leaves from different disease categories.

4. Neural Network Model Development

A neural network model is developed to learn the relationship between the extracted image features and their corresponding disease classes. The network consists of an input layer, hidden layers, and an output layer. During training, the model adjusts its weights and biases to minimize the difference between the actual disease labels and the predicted classes. Appropriate activation and optimization functions are used to improve the learning process.

5. Gradient Boosting Optimization

Gradient Boosting is incorporated to improve the classification capability of the proposed framework. The technique builds a sequence of predictive learners, where each successive learner focuses on reducing the errors generated by the previous learners. The optimized model therefore becomes more capable of identifying complex and nonlinear relationships among cotton disease features. Hyperparameters such as learning rate, number of estimators, tree depth, and regularization parameters can be optimized to achieve improved classification performance.

6. Model Training and Testing

The prepared dataset is divided into training, validation, and testing subsets. The training data are used to learn the model parameters, while the validation data are used for monitoring the learning process and selecting suitable hyperparameters. Finally, the unseen testing dataset is

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

used to evaluate the generalization capability of the proposed model. The training process is repeated for suitable epochs or iterations until the model achieves stable performance without excessive overfitting.

7. Performance Evaluation

The effectiveness of the proposed cotton disease detection model is evaluated using standard classification metrics. Accuracy measures the overall percentage of correctly classified images, while precision indicates the proportion of correctly predicted positive samples. Recall measures the ability of the model to correctly identify disease samples, and the F1-score provides a balanced measure of precision and recall. A confusion matrix can additionally be used to analyze class-wise classification performance and identify disease categories that are frequently misclassified.

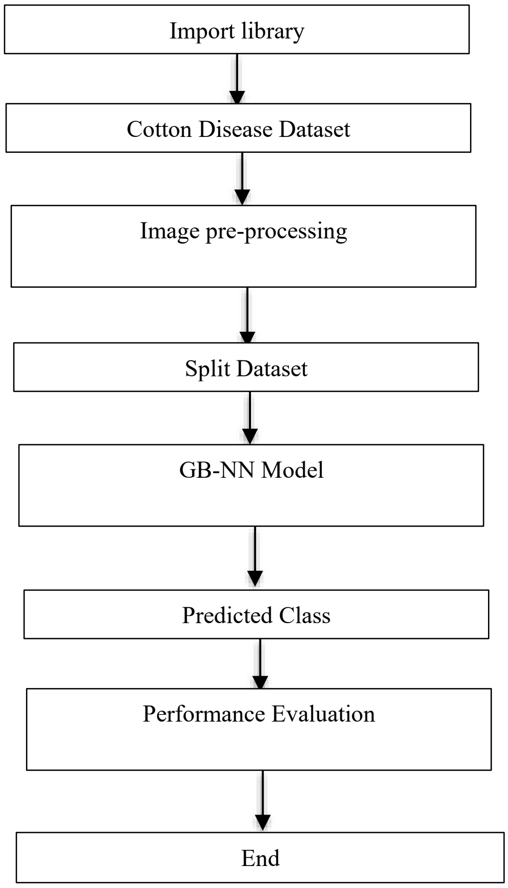

Figure 5: Flow Chart

V. SIMULATION RESULT

The performance of the proposed model was evaluated using accuracy, precision, recall, and F1-score. Accuracy represents the overall proportion of correctly classified images, while precision indicates the correctness of positive disease predictions. Recall measures the ability of the model to identify actual disease samples, and the F1-score provides a balanced assessment of precision and recall. The obtained results can be presented using a comparative performance graph to clearly demonstrate the effectiveness of the proposed GBNN model.

Table I. Recall Results

Cotton Disease		**Recall**
Healthy Image		97.18%
Cotton Disease	Cotton leaf curl	96.88%
	Bacterial bright	97.77%
	Fussarium_wilt	98.33%

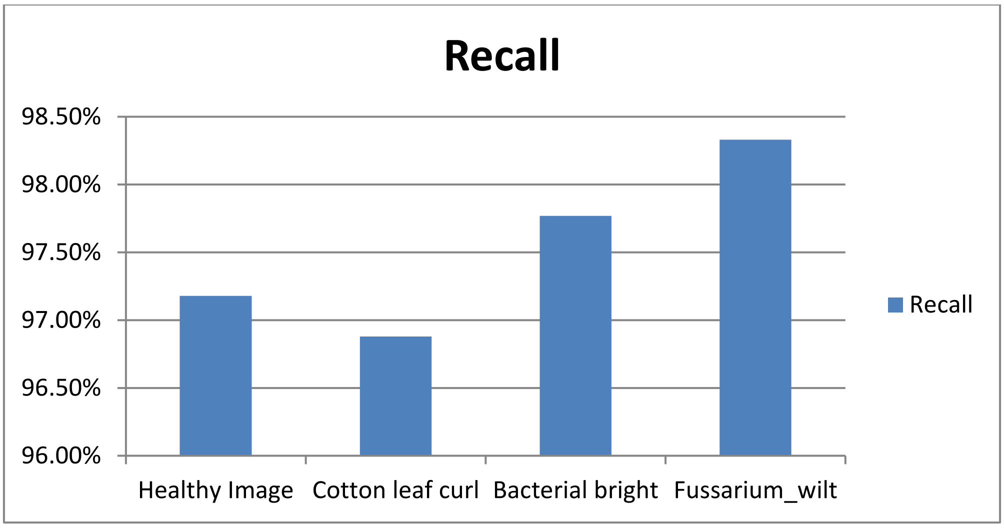

Figure 6: Graph of Recall

The simulation results indicate that the proposed model performs effectively in identifying different cotton leaf disease categories. The model achieved a recall of 97.18% for healthy images, indicating that most healthy cotton leaves were correctly recognized. For cotton leaf curl, the recall was 96.88%, demonstrating a high ability to detect this disease. The model achieved a higher recall of 97.77% for bacterial blight, showing reliable identification of bacterial infection symptoms. The highest recall of 98.33% was obtained for Fusarium wilt, indicating that this disease category was detected most effectively by the proposed approach. Overall, the recall values ranging from 96.88% to 98.33% demonstrate that the proposed model has strong disease-identification capability and can effectively distinguish between healthy cotton leaves and different disease conditions.

Table II. Precision Results

Cotton Disease		Precision
Healthy Image		100%
Cotton Disease	Cotton leaf curl	100%
	Bacterial bright	100%
	Fussarium_wilt	100%

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

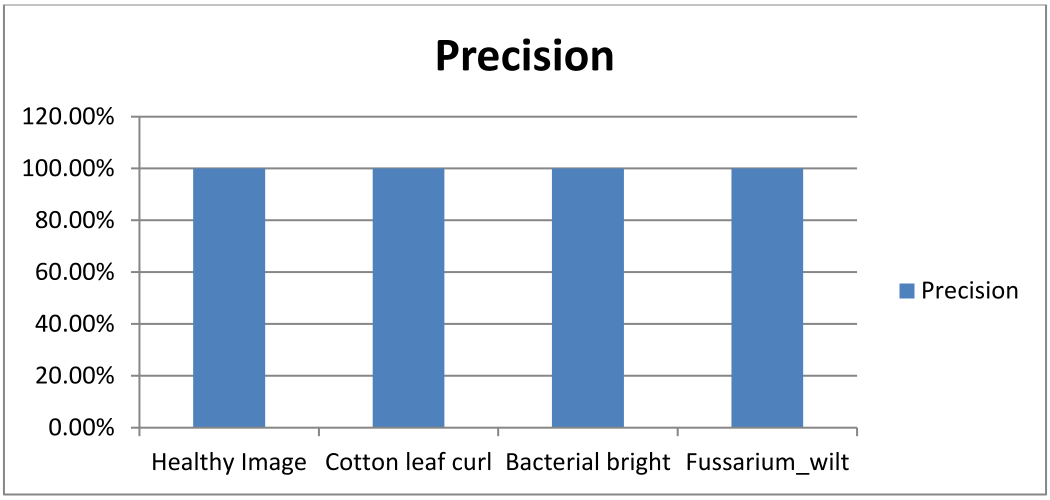

Figure 7: Graph of Precision

Table IIII: Accuracy Results

Cotton Disease		**Accuracy**
Healthy Image		97.18%
Cotton Disease	Cotton leaf curl	96.88%
	Bacterial bright	97.77%
	Fussarium_wilt	98.33%

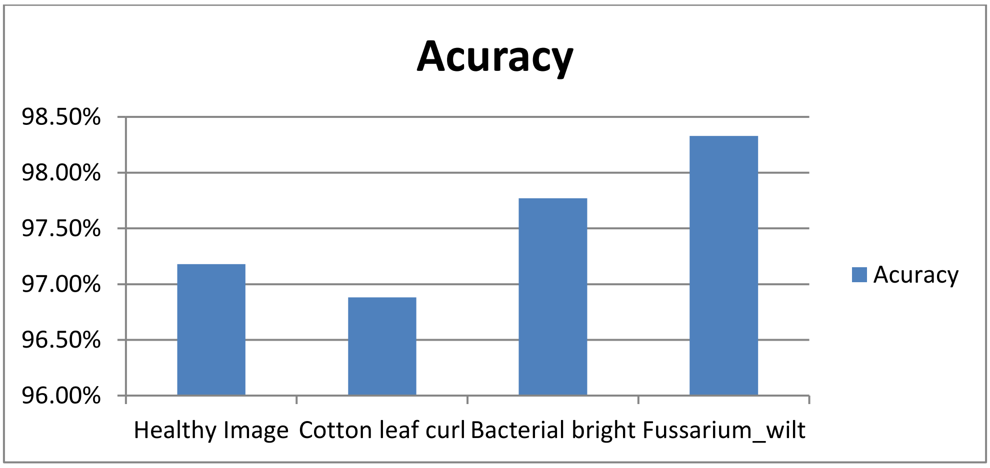

Figure 8: Graph of Accuracy

Table IV: F1-score Results

Cotton Disease		**F1-score**
Healthy Image		98.57%
Cotton Disease	Cotton leaf curl	98.42%
	Bacterial bright	98.87%
	Fussarium_wilt	99.16%

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

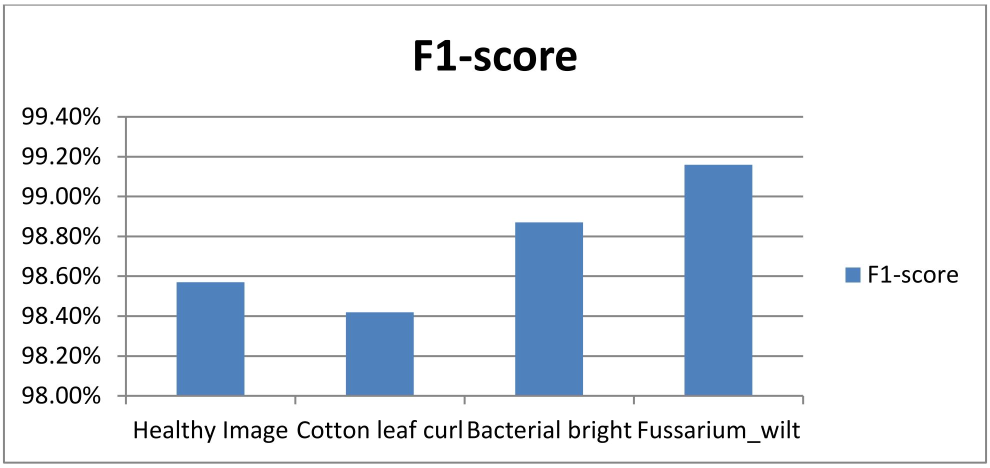

Figure 9: Graph of F1-score

The F1-score results demonstrate the strong and balanced classification performance of the proposed model for different cotton leaf categories. The model achieved an F1-score of 98.57% for healthy images, indicating a highly reliable balance between precision and recall. For cotton leaf curl, an F1-score of 98.42% was obtained, showing effective identification despite the visual similarity of disease symptoms. The model achieved 98.87% for bacterial blight, reflecting highly accurate disease classification. The highest F1-score of 99.16% was obtained for Fusarium wilt, indicating the best overall classification performance among the evaluated categories. Overall, the F1-scores ranging from 98.42% to 99.16% demonstrate that the proposed model provides highly consistent and reliable classification of both healthy and diseased cotton leaf images.

VI. CONCLUSION

The study presented an optimized GBNN approach for the automatic detection and classification of cotton leaf diseases. The proposed method addresses the limitations of conventional manual disease identification, which can be time-consuming, subjective, and dependent on expert knowledge. By utilizing image-based analysis and optimized gradient boosting learning, the proposed framework is capable of identifying important visual patterns associated with different cotton disease conditions.

The integration of gradient boosting with neural network-based learning improves the model's ability to handle complex disease features and reduce classification errors. The performance evaluation using accuracy, precision, recall, and F1-score provides a comprehensive assessment of the proposed model's effectiveness. The approach demonstrates strong potential for achieving reliable and consistent cotton disease detection, particularly in cases where disease symptoms exhibit similar visual characteristics.

Overall, the proposed GBNN-based framework can serve as an effective intelligent tool for early cotton disease diagnosis and precision agriculture. Automated detection can help farmers and agricultural experts identify diseases at an early stage and take appropriate management actions, thereby reducing crop losses and improving productivity. In future work, the proposed system can be further enhanced using larger and more diverse datasets, real-field images, advanced optimization techniques, and lightweight models suitable for mobile or IoT-based agricultural applications.

REFERENCES

[1] R. Kumar *et al.,* “Hybrid Approach of Cotton Disease Detection for Enhanced Crop Health and Yield”, in *IEEE Access*, vol. 12, pp. 132495-132507, 2024.

[2] I. Ahmed, G. Habib and P. K. Yadav, “An approach to identify and classify agricultural crop diseases using machine learning and deep learning techniques”, *Proc. Int. Conf. Emerg. Smart Comput. Informat. (ESCI)*, pp. 1-6, Mar. 2023.

[3] S. Bondre and D. Patil, “Recent advances in agricultural disease image recognition technologies: A review”, *Concurrency Computation Pract. Exper.*, vol. 35, no. 9, Apr. 2023.

[4] M. M. Islam, M. A. Talukder, M. R. A. Sarker, M. A. Uddin, A. Akhter, S. Sharmin, et al., "A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture", *Intell. Syst. Appl.*, vol. 20, Nov. 2023.

[5] R. Mahum, H. Munir, Z.-U.-N. Mughal, M. Awais, F. S. Khan, M. Saqlain, et al., "A novel framework for potato leaf disease detection using an efficient deep learning model", *Hum. Ecol. Risk Assessment Int. J.*, vol. 29, no. 2, pp. 303-326, Feb. 2023.

[6] L. Goel and J. Nagpal, "A systematic review of recent machine learning techniques for plant disease identification and classification", *IETE Tech. Rev.*, vol. 40, no. 3, pp. 423-439, May 2023.

[7] C. Sarkar, D. Gupta, U. Gupta and B. B. Hazarika, "Leaf disease detection using machine learning and deep learning: Review and challenges", *Appl. Soft Comput.*, vol. 145, Sep. 2023.

[8] Y. Yuan, L. Chen, H.Wu, and L. Li, “Advanced agricultural disease image recognition technologies: A review,” Inf. Process. Agricult., vol. 9, no. 1, pp. 48–59, Mar. 2022.

[9] V. Kathole and M. Munot, “Deep learning models for tomato plant disease detection,” in Advanced Machine Intelligence and Signal Processing. Singapore: Springer, pp. 679–686, 2022.

[10] Q. Pan, J.-F. Qiao, R. Wang, H.-L. Yu, C. Wang, K. Taylor, and H.-Y. Pan, ‘‘Intelligent diagnosis of northern corn leaf blight with deep learning model,” J. Integrative Agricult., vol. 21, no. 4, pp. 1094–1105, 2022.

[11] U. Chauhan, “ResTS: Residual deep interpretable architecture for plant disease detection,” Inf. Process. Agricult., vol. 9, no. 2, pp. 212–223, Jun. 2022.

[12] Vallabhajosyula, V. Sistla, and V. K. K. Kolli, ‘‘Transfer learning-based deep ensemble neural network for plant leaf disease detection,” J. Plant Diseases Protection, vol. 129, no. 3, pp. 545–558, Jun. 2022.

[13] Chen, Z. Yuan, S. Chen, and X. Zou, “Plant disease recognition model based on improved YOLOv5,” Agronomy, vol. 12, no. 2, p. 365, Jan. 2022.

[14] Wang, X. He, K. Feng, and H. Zhu, ‘‘Plant disease detection and classification method based on the optimized lightweight YOLOv5 model”, Agriculture, vol. 12, no. 7, p. 931, Jun. 2022.

[15] J. Karthika, M. Santhose and T. Sharan, "Disease detection in cotton leaf spot using image processing", *J. Phys. Conf. Ser.*, vol. 1916, no. 1, 2021.