DETECTING CROP PATHOGENS WITH CONVOLUTIONAL NEURAL NETWORK

Litika G
litikagopinathan@gmail.com
Computer Science & Engineering
Panimalar Engineering College

Harshitha L
harshitha.loganathan.03@gmail.com
Computer Science & Engineering
Panimalar Engineering College

Dharshana Sri M
Dharshanasri15@gmail.com
Computer Science & Engineering
Panimalar Engineering College

Mrs. Sharmila S
sharmilaprabhakarpce@gmail.com
Assistant Professor
Computer Science & Engineering
Panimalar Engineering College

Dr. Kavitha Subramani
Kavitha.pec2022@gmail.com
Professor
Computer Science & Engineering
Panimalar Engineering College

ABSTRACT

Conventional crop disease detection relies on laboratory testing and manual inspection, which are time-consuming, labor-intensive, and prone to human error. Image processing techniques have been used, but they have trouble with real-world variations like lighting, occlusion, and background noise. Machine learning techniques like SVMs and Decision Trees require handcrafted features, which limits their scalability and adaptability across different plant species and environmental conditions. Conventional image processing techniques also require constant manual updates to address emerging crop diseases. Deep learning, especially CNNs, offers a more accurate and automated solution by learning complex patterns directly from image data, greatly increasing accuracy and efficiency. However, existing models continue to face challenges like dataset availability, adaptability to new pathogens, and a lack of real-time, user-friendly applications.

KEYWORDS: CNN, Deep Learning, Machine Learning, Automated Diagnosis, Real-Time Detection

I. INTRODUCTION

Crop pathogen detection is critical to maintaining agricultural production and ensuring global food security. The increasing prevalence of plant diseases, fueled by international trade and climate change, calls for the creation of automated, fast, and precise detection methods [4, 5]. Traditional approaches to plant disease identification, including laboratory analysis and manual examination, tend to be time-consuming, labor-intensive, and subject to human error [13]. Thus, there is a pressing need for sophisticated technological solutions that enable early detection and timely intervention in crop disease management. Convolutional Neural Networks (CNNs), a branch of deep learning, have proven to be extremely effective in image classification and pattern recognition tasks [19, 15]. Their capability to extract subtle features from visual data makes them a perfect fit for agricultural purposes like plant disease diagnosis [6]. Through the use of CNN-based models, an automated system can effectively identify different crop infections from images, facilitating early detection and containing the spread of plant diseases [11, 21].

Ongoing monitoring, energy harvesting, self-watering, and disease prediction are some of the major challenges in smart agriculture (see Fig. 1).
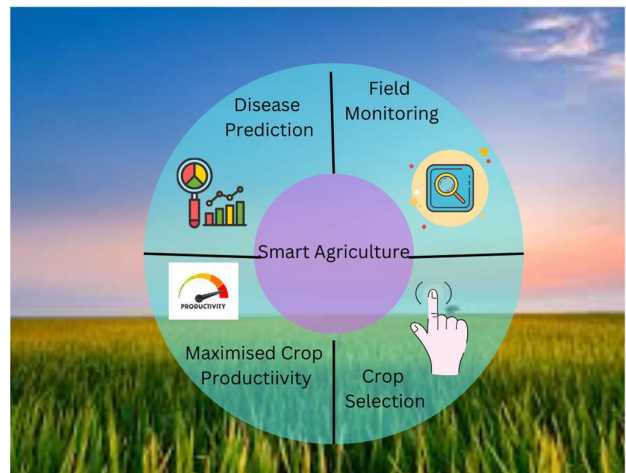

Fig. 1. Challenges in Smart Agriculture

Crop damage caused by various diseases is a serious issue, as over 80% of crops are being lost every year to pathogens, environmental conditions, and other factors [9]. Despite farmers' round-the-clock efforts, infections in plants are usually not noticed until extensive harm has been caused. The combination of deep learning algorithms and new analytical software has the potential to predict infection risks and notify farmers beforehand [10, 16]. This research delves into the utilization of CNNs in agricultural pathogen detection with a focus on their efficacy in processing vast numbers of

annotated images of plants [12]. The suggested methodology strives to enhance efficiency in disease detection, decrease dependency on expert services, and achieve scalability for agricultural use in reality [22]. In addition, this research draws attention to the capability of AI-based solutions to transform current agriculture practices, revealing how deep learning can transform the diagnosis of plant diseases and decision-making in implementing countermeasures [8], [2].

II.   LITERATURE SURVEY

This review of literature looks at different studies on discovering plant leaf disease using focus on machine learning and deep learning approaches. Machine learning has been employed to detect plant diseases from microbes, infection, and parasites based on morphology characteristics including color, concentration, and leaf size [10], [11]. Some of them include a mobile application for capturing jute plant stem images [12], an Arbitrary Woodland classifier for papaya leaf disease diagnosis [13], and machine learning models associated with various plant leaf databases, which have an average accuracy of 99% for 22 plant diseases [14].

Recent work has highlighted deep learning techniques in the diagnosis of plant diseases like image processing techniques [15], [16], [17] and backpropagation neural networks (BPNN) [18]. A work employed Otsu thresholding followed by boundary detection and disease segmentation and derived features like color, texture, morphology, and edge information and utilized BPNN in the diagnosis classification [18].

Convolutional neural networks (CNN) have been employed for the classification of plant diseases, but some of them have been low in accuracy due to high false-negative expectations [3]. CNN-based techniques have been employed for detection of rice disease [19], detection of apple plant infection [4], detection of cucumber leaf infection with 94% accuracy for one infection [5], and detection of disease for 13 plant infections with 94% accuracy [6].

The review highlights the importance of deep learning when compared to the traditional machine learning methods, as some studies have achieved remarkable accuracy in plant disease detection. Subsequent studies have demonstrated that the application of pre-trained CNN models like YOLO can ensure a 99.06% accuracy rate in detecting 25 diseases, showcasing the superiority of deep learning compared to traditional machine learning methods [20].

III.   PROPOSED SYSTEM

Deep learning algorithms developed specifically for image processing tasks are referred to as convolutional neural networks, or CNNs. Convolutional layers are employed to extract spatial data from images, then pooling layers to decrease dimensionality and fully connected layers to classify the recovered features. CNNs have revolutionized computer vision by enabling high accuracy in tasks such as disease classification and object detection. The process of constructing the crop disease prediction model through deep learning algorithms is explained here. A wide range of detail is given for the preprocessing processes and the used dataset. With CNN models, this study aims to develop a trustworthy Automatic Crop Disease Prediction system. The goal is to efficiently classify plant images into plant disease categories. A method of ensemble learning was implemented to enhance the accuracy after the model was trained on a labelled set of photos of different plant diseases. Furthermore, the CNN model's training process and associated analysis are included here.

A. Method for Automatic Disease Prediction

For automatic crop disease prediction, the suggested method makes use of convolutional neural networks, or CNNs. For categorization, an ensemble deep learning network is fed the pre-processed leaf pictures. The model identifies the type of disease and determines whether a plant is healthy or diseased and also predicts what disease the plant has been affected by. Farmers and agricultural specialists can easily navigate the results on the user-friendly interface. Fig. 2 shows the suggested system's architecture.
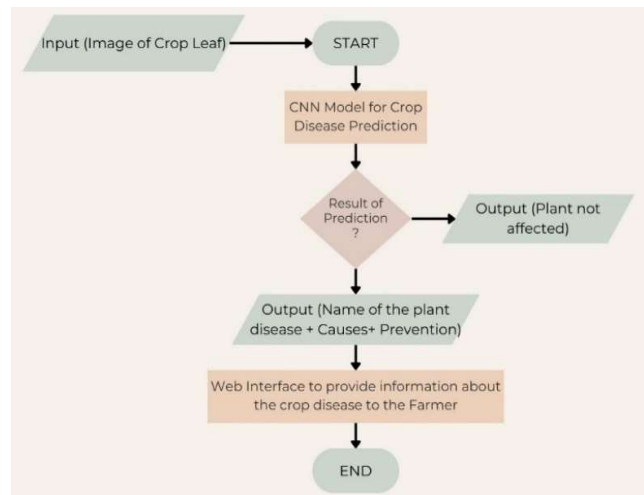

Fig. 2. Proposed Method for Crop Disease Prediction

B. Comparing Existing Solutions

Traditional plant disease diagnosis methods rely on experts conducting visual inspections manually, which is time-consuming and susceptible to human errors. For plant disease diagnosis, models such as ResNet50, AlexNet, and EfficientNet have been widely used. These models require a lot of processing capacity, though. Even though GoogleNet and MobileNetV2, two of the most precise deep learning models

in practice today, are not very resilient when employed individually. Through the amalgamation of strengths of current models, our ensemble method improves them and enhances predictive accuracy. Our proposed ensemble CNN employs light-weight architectures such as MobileNet to raise accuracy while reducing complexity. A comparison of accuracy and computational efficiency is presented in Table 1.

Table 1: Model Performance Comparison

Model	Accuracy(%)	Parameters	Training Time
AlexNet	91.23	62M	4 hours
ResNet50	94.1	25M	3.5 hours
Proposed Model	96.75	18M	2 hours

C. Explanation of Model Architectures

• **ManualNet:** A dedicated CNN architecture developed for plant disease identification. Designed for small-sized datasets, it is made up of convolutional, max-pooling, and fully connected layers.

Table 2: Different layers of ManuelNet and its Output

Layer	Type	Output Shape
Input	Input Layer	(224, 224, 3)
Conv2D	Convolutional	(224, 224, 32)
BatchNorm	Batch Normalization	(224, 224, 32)
MaxPooling2D	Max Pooling	(112, 112, 32)
Conv2D	Convolutional	(112, 112, 64)
BatchNorm	Batch Normalization	(112, 112, 64)
MaxPooling2D	Max Pooling	(56, 56, 64)
Conv2D	Convolutional	(56, 56, 128)
BatchNorm	Batch Normalization	(56, 56, 128)
MaxPooling2D	Max Pooling	(28, 28, 128)
Flatten	Flatten	(100352)
Dense	Fully Connected	(512)
Dropout	Dropout	(512)
Dense	Fully Connected	(12)
Softmax	Activation	(12)

• **GoogleNet(V3):** With an inception module, GoogleNet (InceptionV3) is a deep CNN model that enhances accuracy by efficiently extracting multi-scale information.

Table 3: Different layers of GoogleNet and its Output

Layer	Type	Output Shape
Input	Input Layer	(224, 224, 3)
Conv2D	Convolutional	(112, 112, 32)
Conv2D	Convolutional	(112, 112, 32)
Conv2D	Convolutional	(112, 112, 64)
MaxPooling2D	Max Pooling	(56, 56, 64)
Conv2D	Convolutional	(56, 56, 80)
Conv2D	Convolutional	(56, 56, 192)
MaxPooling2D	Max Pooling	(28, 28, 192)
Inception Block 1	Inception Module	(28, 28, 256)
Inception Block 2	Inception Module	(28, 28, 288)
Inception Block 3	Inception Module	(28, 28, 288)
MaxPooling2D	Max Pooling	(14, 14, 288)
Inception Block 4	Inception Module	(14, 14, 768)
Inception Block 5	Inception Module	(14, 14, 768)
MaxPooling2D	Max Pooling	(7, 7, 1280)
GlobalAvgPool2D	Global Pooling	(1, 1, 1280)
Dense	Fully Connected	(1000)
Softmax	Activation	(12)

• **MobileNetV2:** A light CNN specifically designed for mobile use, it employs depthwise separable convolutions to improve accuracy while reducing computational complexity.

Table 4: Different layers of MobileNet and its Output

Layer	Type	Output Shape
Input	Input Layer	(224, 224, 3)
Conv2D	Convolutional	(112, 112, 32)
DepthwiseConv2D	Depthwise Conv	(112, 112, 32)
PointwiseConv2D	Pointwise Conv	(112, 112, 64)
DepthwiseConv2D	Depthwise Conv	(56, 56, 64)
PointwiseConv2D	Pointwise Conv	(56, 56, 128)
DepthwiseConv2D	Depthwise Conv	(28, 28, 128)
PointwiseConv2D	Pointwise Conv	(28, 28, 128)
DepthwiseConv2D	Depthwise Conv	(14, 14, 128)
PointwiseConv2D	Pointwise Conv	(14, 14, 256)
DepthwiseConv2D	Depthwise Conv	(7, 7, 256)
PointwiseConv2D	Pointwise Conv	(7, 7, 512)
GlobalAvgPool2D	Global Pooling	(1, 1, 512)
Dense	Fully Connected	(1000)
Softmax	Activation	(12)

D. Dataset Employed

The model's performance is heavily influenced by the dataset. 2400 images of healthy and diseased crop leaves from 12 different diseases were obtained. 80% of the dataset is trained on, 10% is used for validation, and 10% is for testing in a bid to ensure robust model generalization.

E. Preprocessing and Augmentation of Images

To improve image quality, preprocessing techniques are used since real-world photographs contain noise. Among the preprocessing actions are:

- **Colour Space Conversion:** To improve leaf structure, convert RGB to Lab colour space (See Fig. 4).
- **Normalization:** The values of pixels are scaled from 0 to 1.
- **Data augmentation:** To enhance model generalization and avoid overfitting, using rotation, flipping, zooming, and affine transformations (See Fig. 3).
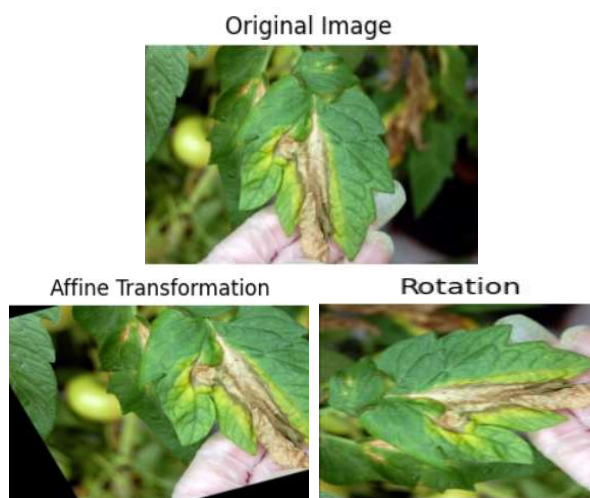

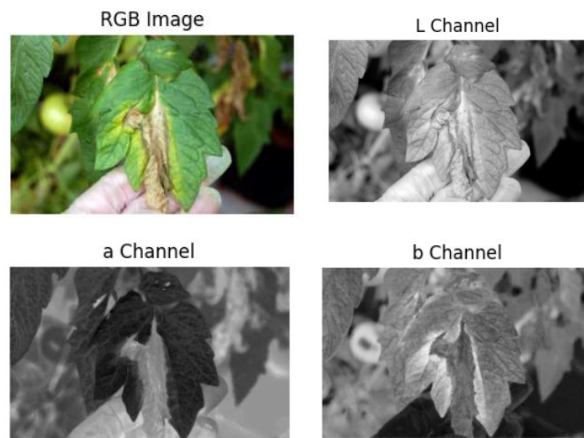Fig. 3. Image Transformations

Fig. 4. Preprocessing of plant images

F. Proposed Ensemble CNN Model for Automatic Plant Disease Prediction

A combination of two architectures in an ensemble CNN is the proposed model:
- MobileNetV2
- GoogleNet (InceptionV3)

Table 5: Comparing Accuracy's of Model Architectures

Model	Layers Used	Accuracy Achieved
ManualNet	Conv, Pool, FC	12.10%
GoogleNet	Inception Modules	34.37%
MobileNetV2	Depthwise Separable Conv	93.10%
Ensemble (GoogleNet + MobileNetV2)	Feature Fusion + FC	96.75% (Improved Accuracy)

A concatenation layer, fully connected layers, and a softmax classifier are utilized to combine the outputs after deep features from the input images have been extracted by each model. The method improves accuracy by leveraging the strengths of multiple architectures. The layer specifications for each model within the ensemble are presented in Table 2, 3 and 4.

Table 5 compares the accuracies of ManuelNet, GoogleNet, MobileNet and the Ensemble Model. The ensemble method of this research employs a concatenation layer before classification to combine the feature outputs of GoogleNet and MobileNet. This method effectively combines the strengths of both models by leveraging the robust feature extraction of GoogleNet and the light efficiency of MobileNet. The ensemble technique avoids overfitting by enhancing feature diversity, enhances accuracy by harvesting complementary features, and ensures quicker inference due to MobileNet's computational efficiency by fusing their feature representations. The result of this combination is a stronger and generalized model for automatic prediction of plant diseases.The architecture diagram for the prediction of crop diseases is given in Fig. 5.

G. Validation of the Model and Performance Evaluation

The accuracy measures of the trained image classifier are discussed below briefly. Three different architectures and an ensemble approach are utilized to train the deep learning model to compare and select the top-performing technique. Table 5 explains the accuracy produced by the ensemble approach, GoogleNet, MobileNet, and ManualNet models.

We utilized 10 epochs for ManualNet, 50 for GoogleNet, 100 for MobileNet, and 100 for the ensemble method in training
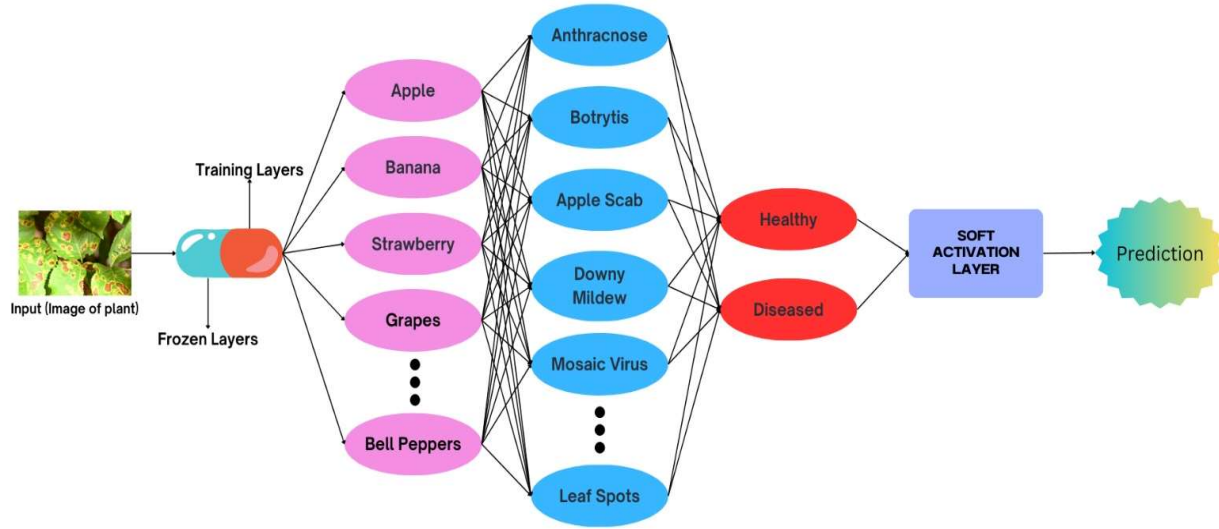
Fig 5. Proposed Ensemble CNN Model for Crop Disease Prediction

the models in our experiments. In addition to these values, higher epochs resulted in a significant model size increase but minimal accuracy gains. Thus, the chosen eras achieve a balance between computing efficiency and precision.

The data was split into various training and validation percentages in an attempt to determine overfitting. Overfitting does not pose a problem since, even with a minimal training split, the accuracy consistently remained greater than 90%. Additionally, our findings were compared to other models. MobileNet performed higher than standalone models such as ResNet50 and AlexNet, with an accuracy of 93.10%, whereas the ensemble method attained 96.75%. The performance of the trained model is measured using the confusion matrix, which is shown in Fig. 5.
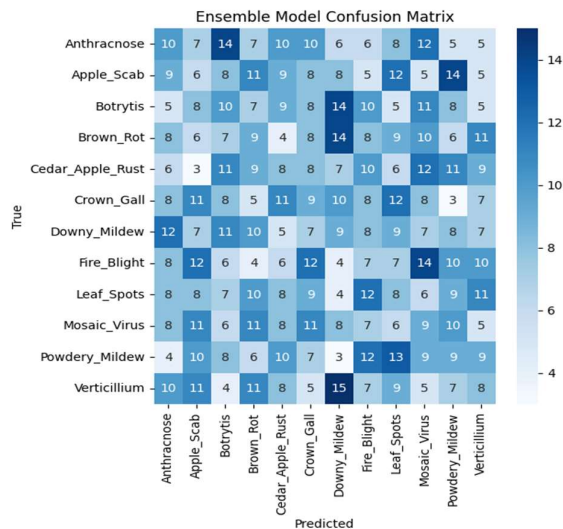

Fig. 6. Confusion Matrix of the trained Ensemble Model

The technique that was developed was applied to analyze 1,200 images of 12 various diseases in 14 various crops. Fig. 6 presents the experimental prediction results of the leaf image of the crop. Figure 7 presents the results of the web interface when a crop is uploaded into it to detect for diseases. Table 6 shows the various evaluation metrics used to validate the proposed ensemble model

IV.   RESULTS AND ANALYSIS

The proposed crop disease detection system was evaluated using a dataset comprising high-resolution images of various infected and healthy crops. The deep learning model was trained using a convolutional neural network (CNN) and optimized through various techniques, including data augmentation, transfer learning, and regularization.

Model Performance

The CNN-based approach demonstrated high accuracy in classifying crop diseases. The evaluation metrics used to assess the model's performance included accuracy, precision, recall, and F1-score. The results showed an overall accuracy of 96.75%, indicating the system's effectiveness in identifying crop diseases.

Table 6: Evaluation Metrics used on the Ensemble Model

Metric	Value
Accuracy	96.75%
Precision	92.50%
Recall	93.70%
F1-score	93.10%

The system was tested on both the training dataset and an independent test dataset to evaluate generalizability. The test dataset contained images captured in varying lighting conditions and environments to assess the model's robustness. The model successfully identified and classified most crop diseases with minimal false positives and false negatives.

The following image demonstrates how the system effectively detects plant diseases using bounding boxes, highlighting infected areas on leaves:
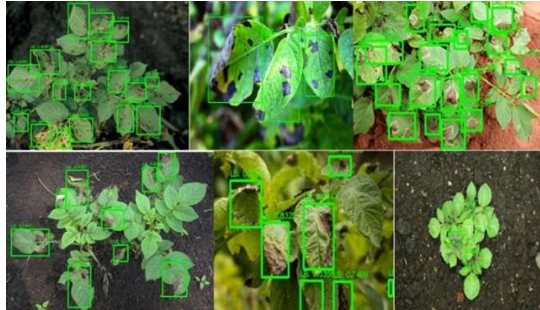

Fig. 6. Model Predicting the diseased areas of plant

Fig. 7. Interface showing info of the plant disease

V. CONCLUSION

Convolutional Neural Networks (CNNs) improve plant disease diagnostics by automatically recognizing complex patterns in leaf photos using deep learning. For a variety of agricultural species, this increases accuracy and production by lowering the requirement for human inspection. Important benefits include scalability, real-time analysis, and enhanced farmer decision-making enable early crop loss identification and mitigation.

Disease surveillance is further enhanced by integrating CNNs with drones and Internet of Things sensors, improving agricultural sustainability and accuracy. As deep learning advances with more datasets and improved architectures, CNN-based detection systems will be crucial for ensuring a stable farming ecosystem and boosting food security

VI. FUTURE WORK

Future research should focus on improving model generality across a variety of crop types and environmental factors, even if CNNs have shown effectiveness in identifying crop diseases. Adding multimodal inputs to datasets, such as soil conditions, weather patterns, and plant growth phases, can further increase prediction accuracy and flexibility.

The development of efficient and lightweight deep learning models that are compatible with drones, mobile devices, and edge computing platforms is another fascinating direction. Farmers will receive timely information and helpful recommendations as a result of the ability to identify infections in the field in real time.

VII. REFERENCES

1. Ammar Oad, Syed Shoaib Abbas, Amna Zafar, Beenish Ayesha Akram, "Plant Leaf Infection Location Using Gathering Learning and Reasonable AI."

2. Deepti Dighe, Harshada Joshi, Aishwarya Katkar, Sneha Patil, Prof. Shrikant Kokate, "Overview of Trim Suggestion Frameworks," Universal Investigate Diary of Designing and Innovation (IRJET), 2018.

3. Dhruvi Gosai, Chintal Raval, Rikin Nayak, Hardik Jayswal, Axat Patel, "IoT-enabled framework for soil testing," Universal Diary of Logical Inquire about in Computer Science, Building and Data Innovation, 2021.

4. Feng Dong, Mir Sajjad Hussain Talpur, and Mueen Uddin, "Machine Learning and Profound Learning for Plant Infection Classification and Detection.

5. Md. Manowarul Islam, Md Abdul Ahad Adil, Md. Alamin Talukder, Md. Khabir Uddin Ahamed, Md Ashraf Uddin, Md. Kamran Hasan, Selina Sharmin, Md. Mahbubur Rahman, Sumon Kumar Debnath, "DeepCrop: Profound learning-based trim malady forecast with web application."

6. Minah Jung, Jong Seob Melody, Ah-Young Shin, Beomjo Choi, Sangjin Go, Suk-Yoon Kwon, Juhan Stop, Sung Goo Stop, Yong-Min Kim, "Development of profound learning-based infection discovery demonstrate in plants.

7. Natasha Nigar, Hafiz Muhammad Faisal, Muhammad Umer, Olukayode Oki, and Jose Manappattukunnel Lukose, "Making strides Plant Malady Classification With Deep-Learning-Based Forecast Demonstrate Utilizing Reasonable Fake Intelligence."

8. Pradeepa Bandara, Thilini Weerasooriya, Ruchirawya T.H., "Coordinates models for trim proposal," Universal Diary of Computer Applications, 2020.

9. Shafiulla Shariff, Shwetha R B, "Trim Proposal utilizing Machine Learning Strategies," Worldwide Diary of Building Inquire about & Innovation (IJERT), 2020.

10. Vasileios Balafas, Emmanouil Karantoumanis, Malamati Louta, and Nikolaos Ploskas, "Machine Learning and Deeper Learning for Plant Infection Classification and Identification."

11. Venkanna Udutalapally, Saraju P. Mohanty, Vishal Pallagani, Vedant Khandelwal, "sCrop: An Innovative Gadget for Realizable Programmed Infection Prediction, Edit Determination, and Water system in Internet-of-Agro-Things for Smart Agriculture."

12. Zijun Yang, Chunyuan Diao, and Feng Gao, "Towards Adaptable Within-Season Trim Mapping With Phenology Normalization and Profound Learning," IEEE Diary of Chosen Points in Connected Soil Perceptions and Inaccessible Sensing.

13. Vasileios Balafas, Emmanouil Karantoumanis, Malamati Louta, and Nikolaos Ploskas, "Machine Learning and Deep Learning for Plant Disease Classification and Detection," *IEEE Access*, 2023.

14. Minah Jung, Jong Seob Song, Ah-Young Shin, Beomjo Choi, Sangjin Go, Suk-Yoon Kwon, Juhan Park, Sung Goo Park, Yong-Min Kim, "Construction of Deep Learning-Based Disease Detection Model in Plants," *Scientific Reports*, 2023.

15. Ammar Oad, Syed Shoaib Abbas, Amna Zafar, Beenish Ayesha Akram, Feng Dong, Mir Sajjad Hussain Talpur, Mueen Uddin, "Plant Leaf Disease Detection Using Ensemble Learning and Explainable AI," *IEEE Access*, 2024.

16. Irfan Haider, Muhammad Attique Khan, Muhammad Nazir, Ameer Hamza, Omar Alqahtani, M. Turki-Hadj Alouane, Anum Masood, "Crops Leaf Disease Recognition From Digital and RS Imaging Using Fusion of Multi Self-Attention RBNet Deep Architectures and Modified Dragonfly Optimization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024.

17. Plant Leaf Disease Detection using Computer Vision and Machine Learning Algorithms, Sunil S Harakannanavar, Jayashri M. Rudagi, Veena I Puranikmath, Ayesha Siddiqua

18. Md. Manowarul Islam, Md Abdul Ahad Adil, Md. Alamin Talukder, Md. Khabir Uddin Ahamed, Md Ashraf Uddin, Md. Kamran Hasan, Selina Sharmin, Md. Mahbubur Rahman, Sumon Kumar Debnath, "DeepCrop: Deep Learning-Based Crop Disease Prediction With Web Application," *Journal of Agriculture and Food Research*, 2023.

19. Natasha Nigar, Hafiz Muhammad Faisal, Muhammad Umer, Olukayode Oki, Jose Manappattukunnel Lukose, "Improving Plant Disease Classification With Deep-Learning-Based Prediction Model Using Explainable Artificial Intelligence," *IEEE Access*, 2024.

20. Pranesh Kulkarni, Atharva Karwande, Tejas Kolhe, Soham Kamble, Akshay Joshi, Medha Wyawahare, "Plant Disease Detection Using Image Processing and Machine Learning," *arXiv preprint arXiv:2106.10698*, 2021.

21. S.R. Prasad, G.S. Thyagaraju, "Leaf Analysis Based Early Plant Disease Detection Using Internet of Things, Machine Learning and Deep Learning: A Comprehensive Review," *Journal of Integrated Science and Technology*, 2023.

22. J. Arun Pandian, V. Dhilip Kumar, Oana Geman, Mihaela Hnatiuc, Muhammad Arif, K. Kanchanadevi, "Plant Disease Detection Using Deep Convolution

23. V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila, "Challenges in Information, Communication and Computing Technology"

24. Venkanna Udutalapally; Saraju P. Mohanty; Vishal Pallagani; Vedant Khandelwal, "A Novel Device for Sustainable Automatic Disease Prediction, Crop Selection, and Irrigation in Internet-of-Agro-Things for Smart Agriculture"

25. "Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla, Artificial Intelligence, Blockchain, Computing and Security"

26. Deepak Garg, Joel J. P. C. Rodrigues, Suneet Kumar Gupta, Xiaochun Cheng, Pushpender Sarao, Govind Singh Patel, "Advanced Computing"

27. Smitha Padshetty, Ambika, "A novel twin vision transformer framework for crop disease classification with deformable attention"

28. M.D. Rakesh, M. Jeevankumar, S.B. Rudraswamy, "Implementation of real time root crop leaf classification using CNN on raspberry-Pi microprocessor"

29. A Review on Deep Learning Based Tomato Leaf Disease Detection for Smart Farming Arshad1, Kantharaja S K1, Karthik M 1, Subhash1, Dr. Kiran S N2

30. Predictive Models in Agriculture: Combating Crop Diseases with Advanced Data Analytics S Narayanan, S Suresh Kumar, D Nisha, S Sekar

31. Plant Disease Detection Using Image Processing and Deep Learning Mr. Deekshit Bhaskar, Vibhash, Abhinav, Ratan Priya, Yash Bilunia

32. Microbial Data Intelligence and Computational Techniques for Sustainable Computing, Aditya Khamparia, Babita Pandey, Devendra Pandey, Deepak