Plant Disease Detection using AI & ML

Saumya Singh, Shreya Singh, Sweta Singh

Student, Student, Student
Computer Science and Allied Branches
Galgotias College of Engineering and Technology (AKTU), Greater Noida, Uttar Pradesh
saumya.21gcebai023@galgotiascollege.edu
shreya.21gcebai047@galgotiascollege.edu
sweta.21gcebai014@galgotiascollege.edu

***Abstract*—Agriculture holds a crucial role in India due to its growing population and increasing demand for food. Consequently, improving crop yields has become essential. One major factor contributing to dip in crop yields is the presence of diseases. These issues can be mitigated through effective detection of plant diseases. Techniques of Machine Learning are particularly useful for this purpose, as they leverage data-driven insights and provide advanced solutions for identifying plant diseases.**

Machine learning-based methods have proven effective in detecting diseases due to their ability to deliver superior results for specific tasks. This review focuses on various techniques used for plant disease detection, incorporating artificial intelligence (AI) through machine learning. These advancements have been applied across multiple fields, leading to significant progress in machine learning and computer vision.

The study includes a comparative analysis of techniques of machine learning, evaluating their effectiveness and usage based on various research papers.

***Keywords*: Artificial Intelligence (AI), Machine Learning (ML), Plant Disease Detection, Predictive Analytics, Crop Health Monitoring, Image Processing.**

I. INTRODUCTION

Advancements in technologies namely IoT (Internet of Things), AI (Artificial Intelligence), and unmanned aerial vehicles (UAVs) have been combined to provide robust support to agriculture, particularly in detecting plant leaf diseases. These innovations enable accurate disease identification and reporting, ensuring timely intervention. Despite these technological strides, modern society shows a declining interest in farming, as many individuals, especially the younger generation, migrate to urban areas seeking safer and more comfortable lives. This migration stems from the numerous challenges and uncertainties faced by farmers in agricultural fields.

The challenge of protecting plants from diseases is intricately linked to climate change and its impact on agriculture. Studies show that fluctuations in climate can affect the development stages and growth rates of pathogens, potentially modifying host resistance and leading to notable shifts in host-pathogen dynamics. Additionally, the global movement of plant diseases has increased, enabling the spread of diseases to regions where they were previously unknown. This creates difficulties, particularly in areas lacking local expertise and resources to manage emerging diseases effectively.

The growing complexity of these issues underscores the importance of leveraging modern technologies like IoT, AI, and UAVs to enhance agricultural resilience and ensure effective disease management strategies. These advancements not only help to mitigate the agriculture from the impact of climate change but also address the challenges posed by the rapid globalization of diseases of plants.

The indiscriminate use of pesticides may contribute to long- term resistance in pathogens, significantly reducing the effectiveness of combating plant diseases. Precision farming, a cornerstone of modern agriculture, emphasizes the need for accurate and timely disease identification to ensure resources are efficiently utilized, minimize waste, and promote healthier crop production. Tackling pathogen resistance and reducing the impact of climate change on agriculture are essential objectives in this regard.

Timely disease detection, including early prevention, has become increasingly vital in the face of environmental challenges. Plant diseases can be identified through various methods, though some diseases exhibit no visible symptoms or are detected too late, requiring advanced diagnostic techniques. Many diseases, however, show visible symptoms, making visual inspection by a skilled professional a key method for identifying illnesses.

Plant pathologists rely on their expertise to recognize characteristic symptoms accurately, which is essential for precise diagnostics. Changes in symptom presentation due to disease can lead to misdiagnosis, particularly for amateurs or hobbyists who lack professional experience. Consequently, refined observational skills and specialized knowledge are necessary to overcome these challenges and ensure reliable disease management in agriculture.
Automated systems that analyze visible symptoms to detect plant health problems can be highly beneficial for both beginner gardeners and experienced experts. These systems act as reliable tools for disease diagnosis, reducing dependency on manual observation and expertise.

Advancements in computer vision have opened new possibilities for precise plant protection and the development of innovative applications in precision agriculture. Techniques such as digital image processing, including color detection and thresholding, have been widely employed for detecting and classifying plant diseases.

In recent years, deep learning techniques, especially convolutional neural networks (CNNs), have gained significant popularity for detecting plant diseases. As an advanced area of machine learning, deep learning has achieved remarkable results in fields such as computer vision, pharmaceuticals, and bioinformatics. Its advantage lies in the ability to process raw data directly without the need for handcrafted features, making it a powerful tool for plant disease diagnostics.

Two primary factors results in the success of deep learning in both academic and industrial applications. First, the availability of vast amounts of data generated daily enables the development of robust models. Second, advancements in Graphics Processing Units (GPUs) provide the computational power necessary to train deep models efficiently through parallel processing. These capabilities have significantly improved the accuracy and scalability of automated plant disease detection systems.

II. LITERATURE REVIEW

Table 1 Table for Literature Review

Author	**Title**	**Techniques used**	**Accuracy**
[1] A. B. atil and S. D. Khirade	Detection of Plant Diseases Using Image Processing	Back Propagation neural network (BPNN) and Digital image processing techniques	87%
[2] V. S. Rajpurohit and S. S. Sannakki	Classificatio n of Pomegranate Diseases.	Neural network classifier	97.30%
[3] R.V. Kshirsagar and P. R. Rothe	Identificatio n of Cotton Leaf Diseases	Back Propagation Neural Network Classifier	85.52%
[4] Medha Wyawahare and Shiroop Madiwalar	A comparative study of Plant disease identificatio n	Support vector machine classifier	83.34%
[5]Deepsikh a, Garima Shrestha, M. Das and N. Dey	Detection of Plant Diseases Using CNN	Convolutiona l neural network	88.80%
[6] D. Ward, P. Moghadam, E. Goan, P. Sikka, S. Jayawardena And E. Hernandez	Detection of Plant Diseases using Hyperspectral Imaging.	k-means clustering algorithm	93%
[7] Nazki et al.	Plant disease recognition.	Generative Adversarial Network and Deep CNN	86.1%
[8] Uday Pratap Singh, Sukirty Jain, Siddharth Singh Chouhan, and Sanjeev Jain	Classificatio n of Mango Leaves Infected by Anthracnose Disease	Multilayer convolution al neural network (MCNN)	97.13%
[9] Vijai Singh	Disease detection of Sunflower leaf	Particle Swarm Optimizatio n Algorithm.	98%
[10] Sumita Mishra, Diksha Rajpal and Rishabh Sachan	Disease recognition of Corn plants.	Deep Convolution Neural Network	88.46%
[11] Parul Sharma, Wiqas Ghai and Yash Paul Singh Berwal	Disease Detection in plants.	Convolution Neural Network	98.6%
[12] Qiong Ren, Hai Han and Hui Cheng.	Plant Disease Detection	Random Forest	85%
[13] Manjunath Badiger, Sachin CN shetty, Varuna kumara and Sudhir poojary	Leaf and Skin disease detection.	K-means algorithm and SVM classifier	96.3%
[14] Aryan Garg	Plant leaf Image Classification.	Resnet-50 Deep Learning Model	76.229%
[15] Srdjan Sladojevic, Andras Anderla, Marko Arsenovic, Dubravko Culibrk and Darko Stefanovic	Detection of Plant Diseases by Leaf Image Classification	Deep Convolution Neural Network	96.3%

III. PROPOSED WORK

Traditional farming often involves labour-intensive activities such as manually collecting data, managing adverse weather conditions, and applying pesticides to control diseases. These practices can pose significant risks to farmers, particularly in drought-prone regions. Considering the difficulties encountered in traditional farming, there is a pressing need for predictive, data-driven solutions to assist farmers in addressing agricultural challenges in real time more efficiently.

To address these challenges, we propose a method utilizing a Decision Tree Classifier to predict diseases in Cotton plants. This approach uses variables specifically temperature, moisture of soil and other environmental factors to identify potential threats to crops. By leveraging predictive analytics, farmers can gain valuable insights to make timely and informed decisions, reducing risks and improving crop health. This approach has the potential to revolutionize traditional farming practices by offering dependable, data-driven support tailored specifically to the needs of cotton cultivation.

This approach focuses on utilizing plant imaging combined with machine learning techniques to identify diseases plants. Classification methods, including **Naive Bayes (NB)** and **Convolutional Neural Networks (CNN)**, are employed and evaluated to determine the most accurate model for disease prediction. The objective of this study is to identify the optimal model that provides superior accuracy, offering a dependable solution for effective disease detection and enhancing crop health management.

A.CNN

Convolutional Neural Networks (CNNs) represent a specialized deep learning model designed for processing structured grid-like data such as images with a complex network structure designed to perform convolution operations. These operations involve applying filters to input data (such as images) to take out important features, for instance textures, edges and shapes. The network is composed of several layers, such as convolutional layers, pooling layers, and fully connected layers, which collaborate to identify visual patterns and generate predictions.

CNNs are especially effective for image-related tasks, such as image recognition, classification, and segmentation, making them widely used in applications like object detection, facial recognition, and plant disease detection. The convolutional layers in CNNs enable the model to automatically learn spatial hierarchies of features, improving its ability to identify intricate patterns in data.

Below are references that highlight the application of Convolutional Neural Networks (CNNs) in detecting plant diseases:

1. **Mohanty, S. P., Hughes, D. P., & Salathé, M**. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, 1419.

-This study investigates the application of Deep Learning techniques for analyzing and solving complex problems in the given context, particularly CNNs, for detecting plant diseases from images. It demonstrates how CNNs can be trained to categorize images of plants and identify disease symptoms.

2. **Ferentinos, K. P.** (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311-318.

-This paper presents the implementations of deep learning methods, including CNNs, in detecting diseases of plants. It compares the performance of CNN models with traditional machine learning approaches and highlights the effectiveness of CNNs in achieving high classification accuracy.

3. **Suresh, P., & Muneeswaran, K.** (2020). Automatic plant disease detection using deep learning algorithms. Journal of Ambient Intelligence and Humanized Computing, 11(6), 2281- 2291.

-This research discusses various techniques of deep learning including CNNs, for automatic detection of plant diseases. The authors review different models and datasets used in the field and propose CNN-based solutions for precise and proficient plant disease identification.

4. **Zhang, C. & Zhang, Y.** (2019). Plant disease detection based on convolutional neural network and image processing. In 2019 IEEE 3rd International Conference on Computer and Communications (ICCC), 978-983.

-This paper explores the integration of CNNs and image processing techniques for detecting plant diseases. It examines the structure of CNN models and their ability to accurately classify plant diseases using images captured under different conditions.

5. **Ubbens, J. R., & Ji, Z.** (2019). Plant disease detection using deep learning: A review. Artificial Intelligence in Agriculture, 1, 10-18.

-A comprehensive review of deep learning applications in detection of plant diseases. The paper covers various deep learning models, including CNNs, and their applications to identify and categorise plant diseases using image datasets.

6. **Nath, S., & Bhattacharya, M.** (2020). Deep learning- based plant disease detection and classification: A survey. Journal of Artificial Intelligence in Agriculture, 2, 50-63.

-This paper surveys the applications of deep learning, particularly CNNs, in detection of plant diseases. It discusses various architectures and techniques for accurate identification and categorization of plant diseases from images.

These references explore the adoption of CNNs in plant disease detection and provide detailed insights into how deep learning can improve agricultural practices.

B. NAÏVE BAYES

The Naive Bayes algorithm is an efficient and straightforward machine learning technique grounded in **Bayes' Theorem**. It operates under the assumption that features are independent of each other, given the class label, which is often not the case in real-world scenarios but still works well in practice for many applications.

Here are several studies that have applied the Naive Bayes algorithm in detection of plant diseases using artificial intelligence and machine learning:

1. **Plant disease identification using Machine Learning: A hybrid approach combining Naïve Bayes and Decision Tree Algorithm.**

-Summary: This study addresses plant disease identification by proposing a hybrid model that combines Naïve Bayes with a decision tree algorithm. The hybrid approach aims to enhance classification accuracy by leveraging the strengths of both algorithms.

2. A **Framework for Plant leaf Detection** Using K-Nearest Neighbour (KNN) Algorithm and Accuracy comparison with Naive Bayes (NB) Algorithm.

-Summary: This paper presents an outline for detecting diseases of plant leaf using the KNN algorithm and compares its accuracy with the Naive Bayes algorithm. The study provides insights into the performance differences between these two classifiers in the context of plant disease detection.

3. **Detection of Plant Leaf Diseases Using Naive Bayes ML Algorithm.**

-Summary: This research explores the effectiveness of the Naive Bayes algorithm in identifying plant leaf diseases. The findings suggest that Naive Bayes is an efficient and reliable method for this purpose, contributing to improved crop output and quality.

4. **Prediction of Plant Diseases Using Machine Learning Algorithms**

-Summary: This study examines several ML algorithms, including Naive Bayes, for prediction of plant disease. It discusses the application of Convolutional Neural Network, Decision Trees, Naive Bayes theorem, Artificial Neural Networks algorithms in predicting plant diseases.

5. **Accurate and Efficient Detection of Plant Leaf Diseases Using Support Vector Machine (SVM) and Naive Bayes.**

-Summary: This study assesses the performance of the Naive Bayes classification algorithm for detecting leaf diseases and compares its accuracy with the Support Vector Machine (SVM) algorithm. The research seeks to identify the more efficient classifier for detecting plant leaf diseases.

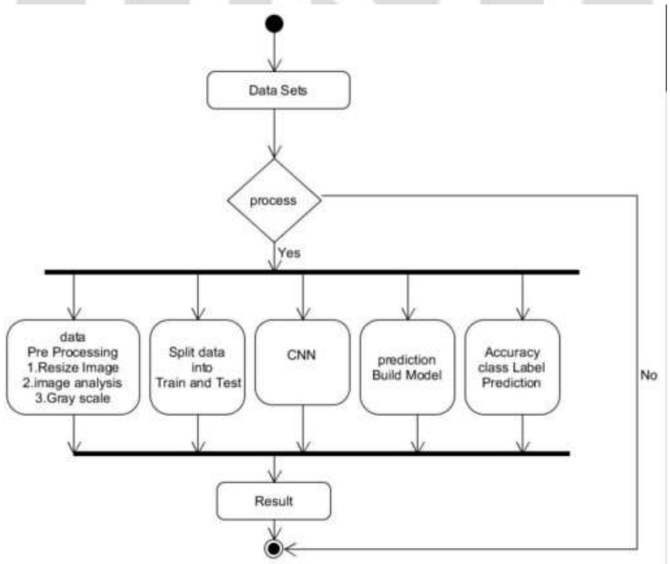

Fig. 1 Activity Diagram for the detection of Plant Diseases

IV. CHALLENGES

A. Dataset size problem

Currently, methods of deep learning methods are extensively applied across several tasks of computer vision, with plant disease and detection of pests emerging as a specialized utilization in agriculture. However, a significant challenge in this field is the scarcity of high-quality datasets for agricultural diseases of plants and pests. Unlike open-source standard libraries such as ImageNet, which contains over 14 million sample images, self-collected datasets for plant disease detection are typically smaller and require substantial effort for data labelling.

In functioning, some diseases of plants have a low occurrence rate and acquiring images for these diseases is costly. This often results in datasets with only a few samples, severely limiting the functionality of deep learning techniques in identifying diseases of plants and pests. Addressing this issue is crucial for advancing the effectiveness of these methods in agriculture.

B.Data amplification, synthesis and generation

Data amplification plays a pivotal role in enhancing the performance of models of Deep Learning, primarily in plant disease and pest detection. An enhanced data augmentation strategy can significantly enhance the model's accuracy and robustness. One of the most widespread techniques for amplifying plant disease and pest image datasets is through image processing operations. These include transformations such as mirroring, rotating, shifting, warping, filtering, and adjusting contrast. These operations generate multiple variations of the original samples, effectively increasing the size and diversity of the dataset.

C.Fine-tuning classical network model and Transfer learning

Transfer learning (TL) is an effective technique that leverages knowledge gained from large, generic datasets and applies it to specialized tasks with smaller datasets. Numerous studies have shown the effectiveness of transfer learning in agricultural applications. For example, **Oppenheim et al.** collected images of infected potatoes, varying in size, colour, and shape, under natural lighting, and classified them by fine-tuning the VGG network. The results showed that transfer learning and fine- tuning were successful in tailoring the model to this particular task.

Similarly, **Too et al**. assessed various classical networks by fine-tuning and comparing their performance. The experimental findings revealed that the accuracy of DenseNets improved with more iterations, highlighting the potential of transfer learning to boost model performance over time.

Chen et al. utilized transfer learning and fine-tuning to identify rice disease images in challenging background conditions, achieving a remarkable average accuracy of 92.00%. This demonstrated that transfer learning surpasses training models from scratch, emphasizing its capability to enhance both accuracy and efficiency in plant disease detection tasks.

D.Lighting problems

Previous studies on plant disease and pest detection have primarily used images captured in controlled environments, such as boxes of indoor light. While this approach helps minimize the impact of external lighting, thereby simplifying the image processing, it differs significantly from images taken under natural sunlight. Natural light is highly dynamic, and the camera's ability to capture varying light conditions is limited. This limitation can lead to colour distortion in images when the lighting falls outside the camera's optimal range.

Additionally, the angle and distance from which images are collected can vary, causing significant changes in the visible specifications of plant diseases and pests. These variations present substantial challenges for visual recognition algorithms, making it harder to consistently identify and classify plant conditions in real-world scenarios.

E.Detection speed problem

In contrast to traditional methods, deep learning algorithms typically produce superior results in detecting plant diseases and pests. However, they also come with higher computational complexity. To achieve high detection accuracy, deep learning models must thoroughly learn the characteristics of images, which increases computational load. This can slow down detection speed, making it challenging to meet real-time requirements. On the other hand, reducing computational complexity to improve speed can result in insufficient training, leading to false or missed detections.

Hence, it is essential to develop efficient algorithms that strike a balance between detection accuracy and speed. Deep learning-based plant disease and pest detection in agricultural applications generally involves three key stages: data labelling, model training, and model inference. While model accuracy is often prioritized, real-time applications place more emphasis on model inference efficiency. Most current methods focus heavily on recognition accuracy, with less attention given to inference efficiency.

V.CONCLUSION

In this study, automated plant disease detection systems were examined, focusing on methods that integrate convolutional neural networks (CNNs) with phytopathological expertise to extract symptomatic signals. However, the challenge of semantically cataloging datasets with sufficient labelled samples and representative data remains a significant obstacle, particularly given the complexities of real-world agricultural conditions.

Our findings highlight substantial advancements in CNN-based plant disease detection methods. Despite the inherent complexity of datasets, often comprising images captured in real agricultural settings, classical CNN architectures combined with optimization and customization techniques demonstrated notable accuracy improvements. Additionally, there is a growing trend toward developing novel CNN models specifically for recognizing plant diseases.

One promising development is the use of multispectral and hyperspectral imaging, which provides more comprehensive data than traditional RGB imagery. However, our analysis also revealed several gaps: crops like grains, millets, and stone fruits, which hold significant economic and nutritional importance, were often overlooked. Moreover, many trained models failed to generalize effectively to images from diverse datasets and real-world scenarios. Most research efforts have been confined to diagnosing diseases of specific crops or a limited range of diseases, indicating a need for broader and more adaptable solutions.

In summary, while DL-based approaches for plant disease detection have made remarkable progress, challenges such as dataset variability, generalizability, and overlooked crop categories must be addressed to create more robust and practical methodologies.

REFERENCES

[1] S. D. Khirade and A. B. Patil, "Plant Disease Detection Using Image Processing," 2015 International Conference on Computing Communication Control and Automation, 2015, pp. 768-771, doi: 10.1109/ICCUBEA.2015.153.

[2] S. S. Sannakki and V. S. Rajpurohit," Classification of Pomegranate Diseases Based on Back Propagation Neural Network," International Research Journal of Engineering and Technology (IRJET), Vol2 Issue: 02 | May-2015.

[3] P. R. Rothe and R. V. Kshirsagar," Cotton Leaf Disease Identification using Pattern Recognition Techniques", International Conference on Pervasive Computing (ICPC),2015.

[4] S. C. Madiwalar and M. V. Wyawahare, "Plant disease identification: A comparative study," 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), 2017, pp.13-18,doi: 10.1109/ICDMAI.2017.8073478

[5] G. Shrestha, Deepsikha, M. Das and N. Dey, "Plant Disease Detection Using CNN," 2020 IEEE Applied Signal Processing Conference (ASPCON), 2020, pp. 109-113, doi: 10.1109/ASPCON49795.2020.9276722.

[6] P. Moghadam, D. Ward, E. Goan, S. Jayawardena, P. Sikka and E. Hernandez, "Plant Disease Detection Using Hyperspectral Imaging," 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2017, pp. 1-8, doi:10.1109/DICTA.2017.8227476.

[7] Haseeb Nazki, Sook Yoon, Alvaro Fuentes, Dong Sun Park "Unsupervised image translation using adversarial networks for improved plant disease recognition" Published by Elsevier B.V,(2020).

[8] Uday Pratap Singh, Siddharth Singh Chouhan, Sukirty Jain, And Sanjeev Jain "Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease" (2019).

[9] Vijai Singh "Sunflower leaf diseases detection using image segmentation based on particle swarm optimization" 2019 Published by Elsevier,(2019).

[10] Sumita Mishra, Rishabh Sachan, Diksha Rajpal "Deep Convolutional Neural Network basedDetection SystemforReal-timeCornPlant DiseaseRecognition" 2020 Published by Elsevier B.V, (2019).

[11] Parul Sharma, Yash Paul Singh Berwal, Wiqas Ghai "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation" open access 2019 Published by Elsevier B.V, (2019).

[12] Qiong Ren , Hui Cheng and Hai Han "Research on machine learning framework based on random forest algorithm" : AIP Conference Proceedings, (2017).

[13] Manjunath Badiger, Varuna kumara,Sachin CN shetty,Sudhir poojary "Leaf and skin disease detection using image processing" Global Transactions Proceedins,(2022).

[14] Aryan Garg"Image Classification Using Resnet-50 Deep Learning Model"Analytics vidya,(2022).

[15] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic "Deep Neural NetworksBasedRecognition ofPlant Diseases by Leaf Image Classification", Volume 2016 Hindawi Publishing Corporation, (2016).

[16] Aakanksha Rastogi, Ritika Arora and Shanu Sharma," Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic" 2nd International Conference on Signal Processing and Integrated Networks (SPIN)2015.

[17] Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa," Automated VisionBased Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease ", Preceding of the 1'st international conference on the use of mobile ICT in Africa ,2014.

[18] S. C. Madiwalar and M. V. Wyawahare, "Plant disease identification: A comparative study," 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), 2017, pp. 13-18, doi: 10.1109/ICDMAI.2017.8073478.

[19] S. D.M., Akhilesh, S. A. Kumar, R. M.G. and P. C., "Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight," 2019 International Conference on Communication and Signal Processing (ICCSP), 2019, pp. 0645- 0649, doi: 10.1109/ICCSP.2019.8698007.

[20] Shanwen Zhang, Subing Zhang,Chuanlei Zhang, Xianfeng Wang, Yun Shi “Cucumber leaf disease identification with global pooling dilated convolutional neural network” Published by Elsevier B.V, (2019).

[21] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).

[22] Mohanty SP, Hughes DP and Salathé M (2016) Using Deep Learning for Image-Based Plant Disease Detection. Front. Plant Sci. 7:1419. doi: 10.3389/fpls.2016.01419.

[23] R. M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," in IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, no. 6, pp. 610- 621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.

[24] Md Nasim Adnan “Improving theRandom Forest Algorithm byRandomly Varying the Size of the Bootstrap Samples”Adnan,(2014).

[25] Niveditha M, Pooja R, Prasad Bhat N, shashank N, “Plant disease detection using machine learning” IEEE (2021).

[26] C.K. Sunil, C.D. Jaidhar, N. Patil, Cardamom plant disease detection approach using EfficientNetV2, in: IEEE Access, vol. 10, 2021, pp. 789–804, https://doi.org/10.1109/ACCESS.2021.3138920, 2021.

[27] G. Wang, Y. Sun, J. Wang, Automatic image-based plant disease severity estimation using deep learning, Comput. Intell. Neurosci. 2017 (2017) 1–8. https://doi.org/10.1155/2017/2917536, 2917536