IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Pepper Bell Plant Disease Detection

[1]Rajeev Kumar, [2]Adarsh Devendra, [3] Ms. Ruchi Sharma
School of Computer Science and Engineering,
Galgotias University, Greater Noida, Uttar Pradesh, India

Abstract: Diseases of pepper bell plants pose a serious threat to food safety and global agriculture. The timely and exact identification of these diseases is essential for their effective control and crop protection. This research looks into using ML and image processing learning techniques to identify plant diseases. The study utilizes various types of images of diseased plant leaves and healthy leaves in high resolution. The images were collected from an actual farm under a wide collection of crops. The model can organize images into the right category to help the farmer and agriculturalists identify this disease affecting the plant. First results show high efficiency and accuracy in diagnosing disease, with the ability to identify the disease an organism is affected with. Moreover, the research examines the usability of this technology on platforms that are easy to use like mobile apps and websites. Results of this study provide an innovative answer to the problems associated with early detection of disease in crops. Through ML and Image Processing we can help farmers to tackle plant diseases and minimize crop losses. Moreover, it will also help improve global food security. We study how ML and Image Processing can help farmers detect diseases at an early stage.

Keywords: convolutional neural network, image processing, pepper bell plants, disease detection, ML, agriculture.

I. Introduction

The early diagnosis of plant diseases using ML and image processing is a novel technique for modern agriculture. Crop disease detection and management using machine learning and image processing is a breakthrough technology for modern farming. Diseases can be detected automatically, quickly, and accurately by the use of the Visual symptoms shown by plant leaves through this innovation. It uses data analysis and computer vision techniques to do this. It could make the controlling of diseases a revolution. Thus, higher agricultural production, less resource consumption, and more sustainable agricultural practices will be achieved by extracting useful features from plant images and training machine learning models. This article gives all the details about the current status of research, their methodology, challenges and applications in this rapidly developing field, with special emphasis on the role played by machine learning and image processing for food security of the globe. This research paper aims to
study the existing machine learning and image processing algorithms for the identification of plant diseases. We'll examine the techniques, algorithms, and datasets that have been applied in this area while highlighting the main challenges and future directions. By knowing their strengths and weaknesses, we can aid in the ongoing attempts to use these technologies to the fullest in agricultural practices.

II.LITERATURE REVIEW

[1] Aakanksha Rastogi, Ritika Arora, and Shanu Sharma released a study titled "Leaf Disease Detection and Grading" that used
fuzzy logic and computer vision technology. To separate the defective ear, K-means clustering is done and textural information is extracted using GLCM and severity is graded using fuzzy logic. We utilized the artificial neural network (ANN) as a classifier which helps to detect the extent of damage of ill leaf. [2] Godliver Owomugisha, John A. Quinn, Ernest Mwebaze, and James Lwasa's paper, "Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease," was recommended. They convert the extracted colour histograms from RGB to HSV and RGB to L*a*b. There are five shape characteristics; peak components create the maximum tree; and categorization is done using area under curve analysis. The decision tree classifiers included options like Decision Tree, Naive Bayes, Random Forest, and Extremely Random Tree. The
Randomized trees provide real-time data, the application is flexible, and the yield score is very high in seven classifiers.[3] The authors looked at the capacity to identify plant disease using characteristics. of colour and texture. They used data of 110 RGB images to check their algorithms. To classify the attributes that were obtained it included the GLCM characteristics the mean and the standard deviation of the image mixed with the Gabor filter and the.mean and standard deviation of the RGB and YCbCr channels. A classifier utilizing a support vector machine was employed for classification. According to the scientists, it would be possible to identify healthy leaves using GCLM traits. They obtained the best accuracy of 83.34% by using all the features and even though colour features and Gabor filter features are considered the best for detection of leaf spots and anthracnose affected leaves respectively.

III.PROBLEM STATEMENT

A thrilling and sophisticated activity that calls for expertise and proficiency in a variety of fields is the design and construction of an automatic plant disease detection from a given image tool. Above all other problems, conventional disease detection techniques are not reliable. These methods usually require special expertise and good eyesight, so they can't be used widely in agriculture. Also, they can't detect diseases early on when remedying diseases is possible. The aim of is to design and implement a mechanized system that is reliable and can detect numerous plant diseases quickly and accurately from visual cues so as to solve this issue. Successfully implementing this will allow human inspectors to handle less work and give early treatment, thus enhancing agricultural productivity. This consists of training models on multiple datasets that cover a wide range of conditions and species in order to cement the robustness and adaptability of the system. The study further seeks to make this technology user-friendly for end users, such as farmers and agricultural experts by designing a friendly interface. Scalability is another important matter; the use of new technologies such as drone-based or remote sensing technology may enable improvement of the system's capacity to monitor widespread diseases. This research basically tackles the urgent problem of identification of plant disease and proposes a game-changing solution based on machine learning and image processing. The purpose of this study is to enhance the performance, accuracy, and functionality of plant disease detection.

IV.METHODOLOGY

To check whether a leaf is diseased or healthy, we need to follow steps. It refers to training classifiers, feature extraction, classification, and preprocessing. To avoid too much noise interference, every image size is reduced to a single normalized value. After that features are taken from a previously processed image. Basic steps to identify the disease in plants

PRIOR TO PROCESSING

The disease detection system converts the received image into a 2D RGB matrix for analysis. It can detect image types as leaf and classify the received input image into the predominant colour on the basis of size of integers in the matrix. After that, each photo is resized to the same size.

SEGMENTATION

The input image has the information needed for processing. This is however the most serious problem as the affected area could be anywhere in the image. The K-means technique divides the image into smaller patches, which are studied on an individual basis. It is then utilized to find the area that is damaged. Only the damaged areas are selected for the purpose of finding a Cure. It saves damaged areas for later examination when it finds them. The K-means iterative method divides a dataset into K non-overlapping sub-groups.

FEATURE EXTRACTION

Detecting a disease is the most important step in the entire process. Transforming raw visual data into numerical features is known as feature extraction. When less material is needed to transport the data and can handle it without the loss of any, the outcomes are

better. original data set's information. Visual elements such as color and intensity are modified to highlight the hidden features of the disorders. These derived features will assist in the faster and accurate identification of disease relative to a validation database for foliar disease.

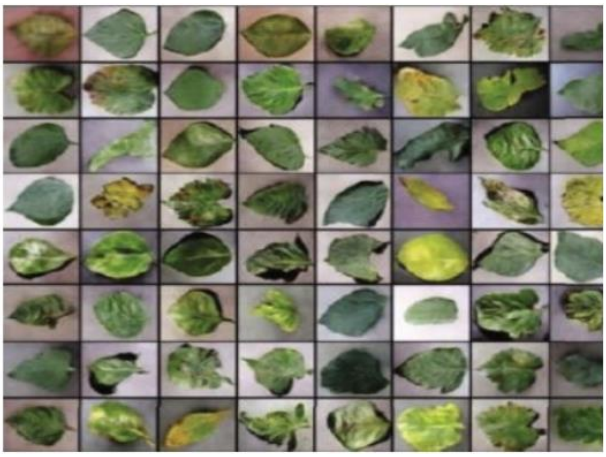
CLASSIFICATION

To classify and diagnose plant diseases is to apply ML and image processing techniques to building a model. This model will help in discriminating between healthy plants and those which are diseased. The data was preprocessed to standardise image sizes and improve the quality and diversity of images before accurate labelling is done to differentiate between healthy and unhealthy plants frequently mentioning the disease categories. To facilitate an objective analysis, the dataset is split into sets that will train, validate and test the model. Feature extraction, often by convolutional neural networks (CNNs), allows for automatic learning of relevant features for the classification of diseases.

To further improve performance predictions, there are phases of model selection, training and evaluation and post-processing with demand for recall also, accuracy and precision.

V.DATASET

The dataset consisted of high-resolution images of pepper bell plant leaves. Each image was labeled as either healthy or diseased, specifying the type of disease. Key features included: Healthy Leaves: Uniform green color without spots or lesions. Diseased Leaves: Visible symptoms such as discoloration, spots, or mold.



Developing a dataset is a vital step in diagnosing plant diseases through image analysis. In this research, images of healthy and infected leaves were collected directly from agricultural fields. These visuals, accompanied by relevant tabular data, capture the physical changes that occur in leaves due to disease. The dataset was carefully organized to ensure clarity and ease of understanding. It reflects a range of leaf colors and textures, often indicating the presence of specific infections. Figure 2 presents a representative sample from the Plant Village dataset, showcasing the visual impact of various plant diseases.

Table 1: Disease recognition and detection based on images (year-wise from 2019 to 2021).

Paper	Year	Database used	Image used	Accuracy
Atila et al. [1]	2021	PlantVillage	61,486	0.984
Reddy et al. [2]	2021	PlantVillage	54,305	0.984
Sembiring et at. [3]	2021	PlantVillage	17,641	0.972
Adit et al. [4]	2020	PlantVillage	76,000	0.980
Guo et al. [5]	2020	PlantVillage	61,486	0.992
Khan et al. [6]	2020	PlantVillage	7,733	0.978
Saleem et al. [7]	2020	PlantVillage	61,486	0.730
Goncharov et al.[8]	2019	PlantVillage	54,306	0.870
Lin et al. [9]	2019	Powdery mildew data set	20	0.961
Nagasubramanian et al. [10]	2019	Hyperspectral image data	111	0.957
Ozguven et al. [11]	2019	Sugar beet data set	155	0.955
Ramcharan et al. [12]	2019	Cassava disease data set	720	0.790
Stewart et al. [13]	2019	Aerial image data set	3,000	0.780
Turkoglo et al. [14]	2019	Turkey plant data set	1,965	0.956

VI. ALGORITHM DESCRIPTION

A CNN is a type of deep learning model that is typically used for visual data analysis. For detecting symptoms of disease, CNN application is used on plant leaf images. The layers in Models have specialized intelligently designed for the implementation of tasks. The convolutional layers use filters that can be trained to understand the important underlying patterns of the input plant leaf images like edges or texture or any other individual identity part of the disease.

The convolution process can be mathematically described as.

$$S(i, j) = (I * K)(i, j) = \sum m \sum n \, I(i - m, j - n) \, K(m, n).$$

Here, S(i,j) is the output feature map, I is the input image and K is the convolution kernel. This helps the network in learning finer details between diseased and healthy leaf images.

A max pooling layer is a layer that reduces the size of the image by discarding the pigmented pixels. The max-pooling operation takes the highest value from a set area.

$$M(i, j) = \max\{I(i-1, j-1), I(i-1, j), I(i, j-1), I(i, j)\}.$$

These layers help decrease computation because they reduce the amount of data. The fully connected layers at the end of the CNN take the features from the previous layers and make a prediction. Activation functions (ReLU) are added to the model to enhance the model's ability to identify data relationships. Reusing and modifying a pre-trained model on a comparatively smaller dataset is possible because of transfer learning. This approach significantly improves performance even when only a limited amount of labeled plant data is available. The use of CNN in this context not only enhances accuracy but also provides a strong foundation for international agriculture and food security. Increased crop yields, decreased economic losses, and improved food security are all results of effective disease scalable and effective agricultural disease monitoring systems.

VII. RESULTS

The dataset table provides a summary of several recent studies that employed machine learning and image processing to identify plant diseases. It includes information such as the datasets used, number of images, sources, and the accuracy achieved.

The results indicate that models trained on the Plant Village dataset consistently reached high accuracy—often exceeding 97%. This reflects the effectiveness of the proposed methods. However, some variability in results exists, likely due to differences in the size and quality of datasets, and the specific techniques used.

These findings highlight the growing potential of automated plant disease detection. Such innovations are essential for modern agriculture, improving crop management and enhancing food security. It's crucial for researchers to assess both the strengths and limitations of existing studies to identify future opportunities.

In our project, we used a CNN model that takes a leaf image as input, analyzes it, and outputs the disease name and relevant recommendations. The trained model achieved impressive accuracy rates: 98.9% on the test set and 98.7% on the validation set. This means that the model has learnt significant visual information and is highly reliable for predicting disease.

CONCLUSION

In this study, authors show that image processing and machine learning can be employed to diagnose plant disease effectively. Several key takeaways emerge from the study.

1. CNNs are highly effective for automating plant disease recognition, even across different species and environments.
2. Data quality is critical—well-structured and pre-processed datasets significantly improve model performance.
3. Collaboration with plant pathology experts ensures the model stays relevant and adapts to emerging diseases.
4. User-friendly interfaces make this technology accessible to farmers and agricultural professionals.

This study helps with better disease control and agricultural sustainability by connecting powerful AI techniques to the feasible and practical use in the field. This helps to increase crop yields, reduce losses, and enhance food security across the globe. The project presents an effective way to find plant diseases and also enables new inventions in precision farming and other sustainable agriculture practices.

Importance of research for project health and environment. Management is essential as the world is becoming complicated with added problems in agriculture. Overall, the research paper "Detection of Plant Disease Using Machine Learning and Image Processing" contributes significantly to agriculture. It offers an innovative solution to a common agricultural problem and can remodel how we maintain and protect our crops through the use of machine learning and image processing. This study opens the path for future developments in disease detection, sustainable agriculture, and international food security, in addition to serving as a monument to the promise of technology in agriculture.

REFERENCES

1. Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using Efficient Net deep learning model," Ecological Informatics, vol. 61, article 101182, 2021.

2. S. R. G. Reddy, G. P. S. Varma, and R. L. Davuluri, "Optimized convolutional neural network model for plant species identification from leaf images using computer vision," International Journal of Speech Technology, 2021

3. A. Sembiring, Y. Away, F. Arnia, and R. Muharar, "Development of concise convolutional neural network for tomato plant disease classification based on leaf images," Journal of Physics: Conference Series, vol. 1845, article 012009, 2021.

4. V. V. Adit, C. V. Rubesh, S. S. Bharathi, G. Santhiya, and R. Anuradha, "A Comparison of Deep Learning Algorithms for Plant Disease Classification," in Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies, Lecture Notes in Electrical Engineering, vol. 643, pp. 153–161, Springer, Singapore, 2020.

5. Y. Guo, J. Zhang, C. Yin et al., "Plant disease identification based on deep learning algorithm in smart farming," Discrete Dynamics in Nature and Society, vol. 2020, Article ID 2479172, 11 pages, 2020.

6. M. A. Khan, T. Akram, M. Sharif, and T. Saba, 2019.

7. M. H. Saleem, S. Khanchi, J. Potgieter, and K. M. Arif, "Image based plant disease identification by deep learning meta-architectures," Plants, vol. 9, no. 11, article 1451, 2020.

8. Goncharov, G. Ososkov, A. Nechaevskiy, and I. Nestsiarenia, "Disease detection on the plant leaves by deep learning," in Selected Papers from the XX International Conference on Neuroinformatics, in Advances in Neural Computation, Machine Learning, and Cognitive Research II, pp. 151– 159, Moscow, Russia, 2019.

9. K. Lin, L. Gong, Y. Huang, C. Liu, and J. Pan, " Deep learning based segmentation and quantification of cucumber powdery mildew using convolutional neural network," Frontiers in Plant Science, vol. 10, p. 155, 2019.

10. K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathysubramanian, "Plant disease
identification using explainable 3D deep learning on hyperspectral images," Plant Methods, vol. 15, no. 1, p. 98, 2019.

11. M. M. Ozguven and K. Adem, “Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms,” Physica A: Statistical Mechanics and its Applications, vol. 535, article 122537, 2019.

12. A. Ramcharan, P. McCloskey, K. Baranowski et al., “A mobile based deep learning model for cassava disease diagnosis,” Frontiers in Plant Science, vol. 10, p. 272, 2019.

13. E. L. Stewart, T. Wiesner-Hanks, N. Kaczmar et al., “Quantitative phenotyping of Northern Leaf Blight in UAV images using deep learning,” Remote Sensing, vol. 11, no. 19, article 2209, 2019.

14. M. Türkoğlu and D. Hanbay, “Plant disease and pest detection using deep learning-based features,” Turkish Journal of Electrical Engineering and Computer Sciences, vol. 27, no. 3, pp. 1636–1651.