Hybrid Deep Learning Technique for Leaf Disease Detection System

Pratham Kumar, Adarsh Maurya, Siddhant Singh Rawat
ad0143maurya@gmail.com prathamathurti6398755408@gmail.com
siddhantrawat297@gmail.com

Abstract. Plants play a vital part in providing nourishment in all inclusive. Different natural components cause plant illnesses, resulting in significant generation misfortunes. However manually identifying of plant diseases may be an expensive and time-consuming process. Embracing improved innovations in a machine learning and deep learning can help to resolve these challenges by allowing early detection of plant illness. The later advancements in the use of ML and Deep Learning procedures for the recognizable proof of plant illnesses are investigated in this paper. The study addresses the distributions from 2016 to 2023, and the tests covered in this article show that, using these techniques can be effective to increase the accuracy and skill of plant disease detection. This article also address the challenges and constraints, which are associated with using machine learning and deep learning for plant disease detection ,such as problem with information accessibility, imaging quality issues, and differentiating between both healthy and sick plants. All papers are examined in detail in terms of ML and deep learning data sets, architectures, efficiency metrics, so that we can combine and make the deep learning model which have the accurate accuracy of leaf disease. We hope that this study will be a valuable resource for researchers who work on the detection of plant diseases and insect pests.

Keywords: Plant Diseases, Machine Learning, Deep Learning, Imaging Quality.

1 Introduction

The two third of total population in India are employed on agriculture as India is an agricultural nation. Fruit and vegetable crops are available to farmers in a broad variety. Support from technology can enhance the cultivation. In plants, pathogens may cause disease in any kind of environment [1]. Given that infections are often observed on a plant's leaves, fruits, and stems, early disease detection is crucial to the productive production of crops. Plant diseases are often caused by pathogens, microorganisms, fungus, bacteria, viruses, etc. Plant diseases can occasionally be attributed to unfavorable environmental factors, such as poor soil and water quality [2].

Many methods exist for identifying various plant diseases in their early stages. The traditional approach of detecting plant diseases involves using just visual inspection, which is ineffective when dealing with huge crops. Plant disease diagnosis is accurate, efficient, and requires less time when using the digital image processing approach. Using this method reduces the need for pesticides and saves time, labor, and effort[3]. For the precise diagnosis of plant diseases, many authors provide various methods using digital image processing. Numerous researchers have come up with an extensive variety of image processing algorithms. This study reviews the various image processing techniques for identifying and classifying leaf diseases.

From 2016 to 2023, a number of research studies made contributions to the field of foliar disease detection with deep learning and machine learning methods. These materials were gathered and refined to offer a summary of the difficulties and methods involved in identifying foliar diseases. In order to categorize and identify several plant illnesses, including fig leaf disease, this research concentrated on creating a unique deep learning model, a hybrid random forest multiclass SVM framework, and sophisticated image processing approaches.

Promising outcomes have been observed in the identification and categorization of foliar diseases through the application of ML and DL approaches research has also brought attention to the advantages and disadvantages of these methods. The significance of early illness identification and the necessity for more study to create generalizable models and release more datasets have also been highlighted by this contribution.

Fig. 1. Some of the plant diseases from New Plant Leaf Disease dataset.

2 Leaf Detection

Plants that have infections on their leaves are said to have leaf diseases. Numerous things, such as infections, environmental stressors, and dietary deficits, can lead to them. A variety of symptoms, including as discoloration, spotting, wilting, and defoliation, can be caused by leaf diseases. Plants that suffer from leaf diseases may experience defoliation, lower yields, or even die.

The Reasons behind Leaf Disease is Pathogens, or organisms that can infect plants with illness, are the most frequent causes of leaf diseases. These pathogens may be viruses, bacteria, nematodes, or fungi. The most prevalent kind of plant pathogens are fungi. Many leaf diseases, including as powdery mildew, leaf blight, and leaf spot diseases, can be brought on by them. Leaf diseases can also be caused by bacteria. Water-soaked patches on the leaves are a common symptom of bacterial leaf infections. Viruses can cause significant leaf diseases, although they are less frequent than bacteria and fungi. The emergence of mosaics or leaf stunting are common symptoms of viral leaf infections. Nematodes are microscopic, soil-dwelling organisms that resemble worms. By feeding on plant roots, certain nematodes can spread diseases that affect leaves.

There are some Environmental stressors which can also cause leaf diseases. Drought can cause leaves to wilt and turn brown. Excess moisture can create an environment that is favorable for the growth of pathogens. Extreme temperatures can damage plant cells, making them more susceptible to infection. Nutrient deficiencies can cause leaves to turn yellow or purple.

Depending on the disease's underlying aetiology, leaf illnesses can have a variety of symptoms. Leaves may become discoloured, turning yellow, brown, or purple. Spots on spotting leaves might become elevated, asymmetrical, or round. Drooping and wilting leaves are possible. Defoliation can cause leaves to drop off the plant too soon.

To identify and classify leaf diseases, a variety of image processing techniques have been employed, however, nearly all of these methods have a number of drawbacks, which are outlined below:

- One type of obstacle that significantly lowers efficacy and efficiency in image processing is noise. Lighting effects and technological equipment provide this kind of disruption. The prediction process may be hampered by the existence of noise in a leaf picture, such as pulse and Gaussian noise.
- One of the main obstacles in this experiment is defocused photographs, which are typically taken by electronic gadgets. Defocused pictures become blurry, making this type of image impossible to see properly.
- The majority of academics often use datasets that are spread across several repositories. For the aim of their experiments, exceptionally few of them take pictures of foliar out from rural areas, but clicking photos by themselves brings new challenges.
- Along with many other difficult challenges, these researchers must cope with varying weather, photograph direction, the separation between electrical equipment and sick leaves, and other occlusion-based photos.
- Using a variety of technological equipment, such as digital cameras and cell phones, one may acquire several photos for varied combinations. Another crucial duty is choosing the proper electronic equipment. Due to the inability of mobile phones to focus things accurately, blurry photos are the main problem. Comparably, a digital camera can take pictures of diseased leaves, but there are additional limitations to the device's ability to do so, such as direction and distance from the subject.
- Another challenging aspect for the research study is choosing a huge picture. Even if every detail in a big image may have a significant influence, blurring remains a challenge throughout the process.
- Drone cameras are used in several studies to take pictures of diseased leaves in diverse agricultural settings. Short flight times and the weather have an impact on this camera's fundamental demits. Due to these drawbacks, researchers have started using drone cameras and have produced a number of categorization results depending on the configurations of various models. Drone camera photos have not been gathered for this text.

Despite these challenges, researchers have made significant progress in developing methods to detect leaf diseases in plants. A literature review on plant foliar disease detection systems highlights the importance of accurate identification and timely diagnosis of plant diseases for crop protection. This literature review offers a comprehensive outline of the current status of the plant leaves, such as methods, issues, progress, and progress in this field, and is involved in the detection of plant diseases and involved in precision agriculture provide valuable insight to those involved.

3 Literature Review

Various researchers have conducted literature reviews on plant leaf disease detection systems. This review covers various aspects of plant disease detection, including data acquisition, preprocessing techniques, data augmentation, feature extraction methods, automated disease detection systems, and analytical techniques to improve image quality . This study also highlights the following challenges with existing approaches: For example, the influence of environmental factors, the difficulty of identifying unhealthy leaf parts, and the limitations of some disease detection models.

In this phase of our finding, we will also highlight a few of the articles mentioned below and used as references to demonstrate the varied methodologies used by various researchers in this subject. Allallah[4] introduces a novel approach for automatic identification of tomato leaf diseases, utilizing three compact CNN and transfer learning to extract deep features efficiently. The integration of features from these CNNs and a hybrid feature selection method results in a comprehensive yet lowerdimensional feature set. Six classifiers are employed, with K-nearest neighbor and SVM achieving remarkably high accuracies of 99.92% and 99.90%, respectively, using only 22 and 24 features.

Javed[5] examined a hybrid deep learning-based framework, comprising Guava Infected Patches Modified MobileNetV2 and U-Net (GIP-MU-NET), Guava Leaf Segmentation Model (GLSM), and Guava Multiple Leaf Diseases Detection (GMLDD) models, is proposed for real-time detection of multiple diseases from a single guava leaf.

Gulzar[7] evaluates the effectiveness of five deep learning models (AlexNet, VGG16, InceptionV3, MobileNetV3, and EfficientNet) for classifying sunflower diseases using image data. EfficientNetB3 demonstrated the highest precision, recall, F1-score, and accuracy (0.979), outperforming other models like AlexNet (0.865), VGG16 (0.965), InceptionV3 (0.954), and MobileNetV3 (0.969)

Rege[8] focuses on early plant disease detection for food security, employing ML and DL techniques on leaf images. The proposed model, based on EfficientNetB3 and B5 architectures, along with the CIVE masking method. The ensemble average model achieves an impressive 99.78% accuracy,

Abisha[9] introduces an innovative method for identifying and classifying infected brinjal leaves employing DCNN,RBFNN. The study collected 1100 images of brinjal leaf diseases caused by five different species and 400 images of healthy leaves from Indian agricultural farms. The fusion of DCNN achieved a mean accuracy of 93.30%, outperforming RBFNN, which achieved 76.70% accuracy without fusion and 87% with fusion

Malik[10] addresses the importance of preserving plants in agriculture for economic growth and proposes a hybrid model for early recognition and classification of sunflower diseases. Focusing on Alternaria leaf blight, Phoma blight, downy mildew, and Verticillium wilt, the study utilizes VGG-16 and MobileNet, both transfer learning models, and employs the stacking ensemble learning approach to create a hybrid model. The dataset, comprising 329 sunflower images across five categories, is built by the author with Google Images.

Table 1. Literature Survey

S.no	**Author**	**Technology**	**Advantages**	**Disadvantages**
1	Abisha,M Mutawa, et al. (2023) [4]	RBFNN, CNN	This technique uses to lower the noise and boost the standard of the picture through image advancement with accuracy 97.3%.	In this data collection of images is less.
2	Ullah , Najah Alsubaie, et al. (2023) [5]	EfficientNetB3 used for various image classification tasks, MobileNet for its efficiency and speed	Reduce the computation cost and improve the recognition performance with accuracy 99.92%.	Not focused on designing structured sparse model for large leaf image database.

3	Gulzar, Ünal, et al. (2023) [6]	CNN(ALexNet, VGG16, InceptionV3 and MobileNetV3)	It illustrates the promise of DL in earlier illness finding and classification of sunflower disease with accuracy 97.9%.	Highly complex algorithm are used and time consuming process.
4	Rege, Ranaware, et al. (2023) [7]	EfficientNetB3, EfficientNetB5, CIVE masking method	This plant disease detection system can contribute to the development of automated systems for agricultural applications with accuracy 98.68%.	There is no development how to handle the gray scale images.
5	Omneya Attallah (2023) [8]	CNN, SVM, KNN, NB, LDA	It employing a hybrid choice of features strategy to selects and generates a complete feature with a set of smaller dimensions with accuracy 97.3%.	Difficult to predict K-Value.
6	Malik, Vaidya, et al. (2022) [9]	VGG-16, MobileNet	High accuracy of 89.2% in the recognition of sunflower leaf disease.	It only uphold the sunflower leaf diseases application.
7	Altalak, Ammaduddin et al. (2022) [10]	CNN, SVM	Researchers suggested a hybrid model that makes uses of DL methods to identify plant leaf diseases early on.	Highly complex algorithm are used and time consuming process.
8	IlaydaYa g , AytaçAltan et al. (2022) [11]	FPA, CNN, SVM	The suggested model classifying the specific plant leaf diseases in real time with high accuracy of 93.97%.	It does not perform very well and dataset is very small.
9	Sumitra Nuanmeesri (2021)[12]	VGG1,SVM classifier	This is a fusion of multiple method for classifying diseases on rose leaves with an accuracy of 92.26%.	It is not suitable for large datasets.
10	Shubham Raje (2021) [13]	Random Forest, CNN, KNN	Best result for the detection of diseased leaves is 60%.	Different initial partitions can result in different final clusters.

11	Kartikeyan, Gyanesh et al. (2021) [14]	SVM, ANN, KNN	It was shown that classification approaches were frequently employed for the verification and detection of illnesses in leaf images. With an accuracy of 95.12%	Difficult to predict K-Value.
12	Zhang et al. (2021) [15]	VGG-16	DL approaches may identify plant leaf diseases with an accuracy of 85%, provided that there is enough data available for training.	It does not perform very well, and this model has poor robustness.
13	Sathvick M., et al. (2020) [16]	CNN, SVM, RNN	Developed the application that will increase the farmers' access to and utility of this technology with accuracy of 98%.	Different initial partitions can result in different final clusters.
14	Nagaraju, P. Chawla et al. (2020) [17]	CNN, SVM	Detecting plant illness by enhancing system achievement with accuracy of 98.5%.	It is not suitable for large datasets.
15	A. Dixit, S. Nema et al. (2018) [18]	SVM , ANN	It covers different techniques for identifying wheat illness. We will employ different algorithms to get findings that are more accurate.	The variable number of neurons that is usually much less than the number of training points.

4 Problem Identification

The process of automatically recognizing and classifying plant diseases through the analysis of leaf images is known as automatic leaf detection. This technology lowers the cost of manual labour and increases the effectiveness of foliar detection. There are several approaches for automatic leaf detection, including the use CNN, single class classifiers, and DL techniques. These methods train a model on a dataset of leaf images and utilize them to classify new images as healthy or diseased.

Major problems encountered with various disease detection systems in accordance with literature review include:

1. **Image Segmentation:** Accurate segmentation of leaf images is important for disease detection, but can be difficult due to the different shapes, sizes and colour of leaves.
2. **Limited Data Set:** The availability of limited data sets can affect the accuracy of the detection system. A larger dataset helps improve the performance of the system.

3. **Overfitting:** When a model is trained excessively well on its training data set, overfitting takes place, resulting in poor performance on new data. Regularization techniques help prevent overfitting.
4. **Hardware Limitations**: Using DL models for disease detection requires large amounts of computational resources, which can be difficult on low-power devices such as smartphones.

5 Objective of Research

To resolve these problems, some possible solution are:

1. **Image augmentation:** Applying operations like rotation, scaling, and flipping to pre-existing images can be used as image augmentation techniques to produce more training data and enhance the efficiency of your segmentation models.
2. **Transfer learning:** For improved model performance on limited data, use transfer learning by starting a new model on a smaller dataset using an existing pre-trained model.
3. **Regularization:** Apply regularization techniques to improve model generalization and avoid overfitting, such as weight decay and dropout.
4. **Model optimization:** Reduce the size and complexity of your model using model optimization techniques like quantization and pruning to make it appropriate for low-power devices.

6 Proposed Methodology

Hybrid models can improve the performance of existing leaf detection systems by integrating multiple techniques such as machine learning and deep learning. By combining the best features of both methods, the model is able to identify foliar diseases with greater adaptability and accuracy. Hybrid models also have the ability to combine the features extracted from one model with the classification capabilities of another, creating a more complete and efficient disease detection system. A hybrid strategy that combines several models and algorithms may enhance generalization and flexibility when applied to different data sets, hence enhancing leaf detection systems' overall performance.

The proposed plant leaf disease detection method uses a hybrid model for training and validation on the publicly available New Plant Diseases Dataset.

Data Collection: The research paper use a freely accessible Leaf image called New Plant Diseases Dataset on Kaggle for the execution of Leaf images. There are 87000 RGB images representing healthy and diseased plant leaves divided into 38 different categories. The “New Plant Diseases Dataset” used in this study, includes colour casts, poor contrast and hazy leaf images.

Data Preprocessing: Clean and preprocess datasets, resize images, normalize pixel values, and augment data using various methods like brightness alteration, mirroring, and rotation to increase model robustness. This can be achieved using libraries such as OpenCV, NumPy, and TensorFlow in Python.

Feature Extraction: Various feature extraction techniques are used including colour, texture, and shape-based features, transfer learning, and segmentation. Colour, texture, and shape-based features are widely used to differentiate between healthy and diseased plants. Transfer learning is used to leverage knowledge from large datasets, reducing the need for extensive feature engineering. Segmentation is used to segment diseased components from leaves using techniques like K-means clustering or custom segmentation models.

Classifier: Classification and detection of plant leaf diseases are also accomplished through the use of DL technology. It analyze and categorizes leaf infections utilizing MATLAB picture processing based on their physical appearance.

Hybrid Deep Learning Model: To increase the accuracy and efficiency of disease detection, a hybrid DL model that combines the advantages of several DL algorithms has been proposed for foliar diseases. Hybrid models typically include the use of CNN and various other DL techniques such as conditional random fields (CRFs) and convolutional auto encoders to detect and classify foliar diseases.

Disease Classification: In image processing technology, classification is the most difficult task. Precisely forecasting the value of a given discrete class variable utilizing a vector of characteristics or predictors is the primary objective of classification. When detecting plant diseases, images are classified based on whether they are infected or not.

Feature Selection: Feature selection is an important step in detecting foliar diseases. Various studies have proposed different methods for feature selection, such as the Salp-Swarm algorithm, feature selection based on DL, and on gray-level co-occurrence matrix (GLCM). These methods aim to identify the most relevant features from leaf images, such as: Colour and texture features to improve disease detection accuracy.

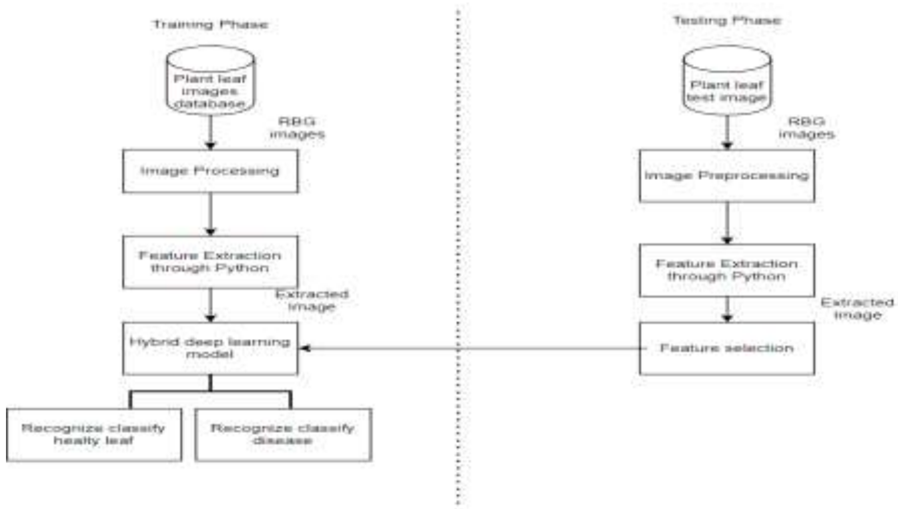

Fig. 2 Flow Diagram for proposed Methodology.

Conclusion And Future Work

This overview discusses various image processing techniques for the diagnosis of plant disease. This survey consists of four basic steps: preprocessing, feature extraction, classification, segmentation. Nevertheless, the majority of this field's research is restricted to lab-based investigations and mostly depends on gathered images of pests and plant diseases. It is crucial to collect photos from different plant growth levels, seasons, and demesne, in order to improve the model's of robustness and generalization. Early detection and management of plant disease and pests is critical for preventing and managing their transmission and growth; consequently, combining atmospheric and plant health information, such as moisture and temperature, is required for effective diagnosis and forecasting. To realize the full benefits of this innovation, agriculture and plant security specialists must collaborate, combining their expertise and proficiency with DL algorithm models, and incorporating the results into farming tools. Based on articles from 2016 to 2023, this report examines current advances in utilizing and DL methods for plant disease spotting. Every step is compared with respect to its performance, advantages, disadvantages, and techniques. Our outcomes show that DL model are extremely good at identify and categories diverse leaf diseases. After conducting the survey, we have come to the conclusion that by pre-processing methods we can contribute to higher segmentation accuracy. In the many papers there are no constraints on data acquisition.

By investigating the advantages and disadvantages of various approaches, and providing insightful information for scholars and industry experienced .This study devote to the enhancement of plant illness detection and circumvention.

In the future, the model may be refined by analyzing different dataset leaf samples and extracting key features implementing hybrid approaches.

References

1. Alston, J. M., & Pardey, P. G. (2014). Agriculture in the global economy. Journal of Economic Perspectives, 28(1), 121-146.
2. Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification by deep learning—a review. IEEE Access, 9, 56683-56698.
3. Pantazi, X. E., Moshou, D., & Tamouridou, A. A. (2019). Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. Computers and electronics in agriculture, 156, 96-104.
4. Omneya Attallah, "Tomato Leaf Disease Classification via Compact Convolutional Neural Networks with Transfer Learning and Feature Selection", *Horticulture* 2023, vol,*9* ed.2.
5. Rashid, J., Khan, I., Ali, G., Alturise, F., & Alkhalifah, T. (2023). Real-Time Multiple Guava Leaf Disease Detection from a Single Leaf Using Hybrid Deep Learning Technique. Computers, Materials & Continua, 74(1).
6. Alsubai, S., Dutta, A. K., Alkhayyat, A. H., Jaber, M. M., Abbas, A. H., & Kumar, A. (2023). Hybrid deep learning with improved Salp swarm optimization based multi-class grape disease classification model. Computers and Electrical Engineering, 108, 108733.
7. Gulzar, Y., Ünal, Z., Aktaş, H., & Mir, M. S. (2023). Harnessing the power of transfer learning in sunflower disease detection: A comparative study. Agriculture, 13(8), 1479.
8. Rege, P., Ranaware, P., & Wagh, M. A hybrid ensemble approach for plant disease detection and classification.
9. Abisha, S., Mutawa, A. M., Murugappan, M., & Krishnan, S. (2023). Brinjal leaf diseases detection based on discrete Shearlet transform and Deep Convolutional Neural Network. Plos one, 18(4), e0284021.
10. Malik, A., Vaidya, G., Jagota, V., Eswaran, S., Sirohi, A., Batra, I., ... & Asenso, E. (2022). Design and evaluation of a hybrid technique for detecting sunflower leaf disease using deep learning approach. Journal of Food Quality, 2022, 1-12.
11. Altalak, M., Ammad uddin, M., Alajmi, A., & Rizg, A. (2022). Smart agriculture applications using deep learning technologies: A survey. Applied Sciences, 12(12), 5919.
12. Yağ, İ., & Altan, A. (2022). Artificial intelligence-based robust hybrid algorithm design and implementation for real-time detection of plant diseases in agricultural environments. Biology, 11(12), 1732.
13. Nuanmeesri, S. (2021). A hybrid deep learning and optimized machine learning approach for rose leaf disease classification. Engineering, Technology & Applied Science Research, 11(5), 7678-7683.
14. Raje, S. (2021). Detecting Diseases in Rice Leaf Using Deep Learning and Machine Learning Techniques (Doctoral dissertation, Dublin, National College of Ireland).

15. Kartikeyan, P., & Shrivastava, G. (2021). Review on emerging trends in detection of plant diseases using image processing with machine learning. International Journal of Computer Application, 975, 8887.
16. Mulla, R. A., Pawar, M. E., Banait, S. S., Ajani, S. N., Borawake, M. P., & Hundekari, S. (2023). Design and implementation of deep learning method for disease identification in plant leaf. International Journal on Recent and Innovation Trends in Computing and Communication, 11, 278-285.
17. Ushadevi, G. (2020). A survey on plant disease prediction using machine learning and deep learning techniques. Inteligencia Artificial, 23(65), 136-154.
18. Oppenheim, D., Shani, G., Erlich, O., & Tsror, L. (2019). Using deep learning for image-based potato tuber disease detection. Phytopathology, 109(6), 1083-1087.