

Crop Disease Detection and Classification using Deep Learning

Gowthul Alam M M

Department of Computer Science
and Engineering Jain Deemed to be
University, Bengaluru.

G Nikhith Reddy

Department of Computer Science
and Engineering, Jain Deemed to
be University, Bengaluru.

K S Chakradhar Reddy

Department of Computer Science
and Engineering, Jain Deemed to
be University, Bengaluru.

K Sai Vishnu

Department of Computer Science
and Engineering, Jain Deemed to
be University, Bengaluru.

K Vighnesh

Department of Computer Science
and Engineering, Jain Deemed to
be University, Bengaluru.

Abstract— Crop health is essential to both sustainable agriculture and food security. Accurate crop disease classification and early detection are essential for prompt intervention and lowering yield losses. Conventional disease identification techniques are frequently labour-intensive, manual, and prone to human error. This project uses Convolutional Neural Networks (CNNs), a type of deep learning technique, to automatically identify and categorise crop diseases from photos of leaves. The model is trained using a sizable dataset of photos of both healthy and diseased crops, allowing it to recognise complex patterns and characteristics linked to a range of plant diseases. The suggested system provides farmers and agricultural specialists with a scalable and effective solution by exhibiting high accuracy in differentiating between various disease categories and healthy crops. Real-time disease diagnostics can be delivered in the field by incorporating this technology into web-based or mobile applications, giving users access to prompt treatment recommendations and insights. By utilising AI and deep learning, this strategy has the potential to transform crop management and boost agricultural output.

Keywords: deep learning, agricultural production, crop disease, and crop health.

I. INTRODUCTION

Agriculture continues to be a critical pillar of the global economy, playing a particularly vital role in developing and rural countries where a significant portion of the population depends on farming for their livelihood. In such regions, farming is not only a source of income but also a means of sustenance. Among the many factors that influence agricultural success, crop health stands out as a key determinant of both productivity and food security. Healthy crops lead to better yields, improved quality of produce, and stable incomes for farmers. Conversely, any compromise in crop health particularly due to diseases—can lead to massive losses, both economically and in terms of food availability.

Crop diseases, often caused by fungi, bacteria, or viruses, spread rapidly and can devastate large areas of farmland in a short time if not managed properly. Traditional methods of disease identification rely on manual inspection by agricultural experts, which involves visually analyzing symptoms on leaves, stems, and fruits. While effective to an extent, these methods come with several limitations: they are labor-intensive, time-consuming, require the physical presence of experts, and are prone to human error or subjective judgment. Moreover, in remote or under-resourced areas, access to such expertise is often limited or unavailable.

With the advancement of technology and the growing availability of agricultural data, there is an urgent need—and a growing opportunity—to automate and enhance the process of crop disease detection. This is where artificial intelligence (AI), and more specifically deep learning, can play a transformative role. AI-powered solutions can process large volumes of visual data, learn from patterns, and deliver accurate results in real-time, offering immense benefits to the agricultural sector.

This project explores the application of Convolutional Neural Networks (CNNs)—a class of deep learning algorithms highly effective in image recognition tasks—for the purpose of crop disease identification. The system is trained on a comprehensive dataset containing thousands of images of both healthy and diseased plant leaves, allowing the model to learn minute features and visual cues that distinguish various plant diseases. These features, often difficult for the human eye to consistently detect, are used by the model to accurately classify the type of disease present in a leaf sample.

The primary objective of this research is to equip farmers and agricultural stakeholders with a smart, scalable, and easy-to-use tool that enables early disease detection and better decision-making. By deploying this model through a web-based or mobile application, users can simply upload an image or video of the affected plant, receive an instant diagnosis, and access treatment recommendations on the spot. This capability not only helps in mitigating crop losses, but also promotes more sustainable farming practices by encouraging timely interventions and reducing unnecessary use of pesticides.

II. LITERATURE REVIEW

Researchers have started looking at artificial intelligence approaches, particularly deep learning, to increase the precision and effectiveness of agricultural disease detection because of the growing threat posed by plant diseases and the shortcomings of conventional diagnosis methods. Better crop management and increased agricultural output are made possible by these technology developments, which provide scalable and trustworthy substitutes for prompt diagnosis.

Deep learning models were used for image-based plant disease identification in a research by Mohanty et al. [1]. They developed a Convolutional Neural Network (CNN) model with an accuracy of 99.35% using a dataset of 54,306 photos with 26 illnesses and 14 crop species. The findings demonstrated that CNNs perform noticeably better than conventional image processing and machine learning techniques. The study did point out that more optimisation could be necessary for real-world deployment in order to manage noisy and less-than-ideal picture inputs.

CNNs were used in a deep learning-based method for automated plant disease detection by Sladojevic et al. [2]. They used a proprietary dataset of photos of both healthy and sick leaves to train their algorithm. The model showed promise for real-time illness detection and attained excellent classification accuracy. In order to improve model generalisation, the researchers stressed the need of using appropriate picture preprocessing and augmentation strategies.

Ferentinos [3] looked at the application of deep learning models to the diagnosis of plant diseases in greenhouse settings. CNN architectures like AlexNet and GoogLeNet, which were trained on a variety of tomato plant disease datasets, were used in the study. The top-performing model achieved 99.53% classification accuracy. Although the study stressed the necessity of model testing in field settings, it came to the conclusion that CNNs might be an effective tool for diagnosing plant diseases in controlled settings.

For the categorisation of plant diseases, Too et al.

[4] compared many CNN architectures, including as VGG16, ResNet50, and MobileNet. Their research centred on striking a balance between computing efficiency and precision. MobileNet provided a lightweight substitute appropriate for mobile applications, which is advantageous for real-time diagnosis in field settings, even if VGG16 and ResNet50 shown good accuracy.

In order to create a smart agricultural solution, Picon et al. [5] investigated the integration of CNNs with Internet of Things (IoT) equipment. Their solution enabled farmers to use drones or mobile cameras to take pictures of leaves, which a CNN model would subsequently process to identify diseases. The study emphasised how crucial connection and latency are when implementing such systems in isolated agricultural regions.

III. METHODOLOGY

This research makes use of deep learning methods, namely Convolutional Neural Networks (CNNs), to create a crop disease detection system that is both accurate and efficient. Data collection, preprocessing, model selection and training, assessment, and deployment are the steps in the technique. Every stage is intended to guarantee the system's resilience and usefulness in real time.

3.1 Dataset Classification

Collecting a thorough and varied dataset is an essential first step in developing a reliable and effective crop disease detection system. Images of both healthy and damaged crop leaves make up the dataset, which was chosen based on the literature and the introduction's goals. These photos are gathered from a variety of public sources, including real-time field photography, research databases like PlantVillage, and agricultural picture collections. The emphasis is especially on crops like rice, potatoes, and tomatoes that are common in Indian agriculture. This guarantees that the technology can identify common plant illnesses that Indian farmers commonly face and is in line with real-world conditions.

3.2 Data Preprocessing

Preprocessing becomes crucial when the dataset is gathered in order to get it ready for model training. Preprocessing involves common practices like scaling all photos to the same size for uniformity and normalising pixel values to increase training efficiency, as various studies have shown. Rotation, flipping, and zooming are examples of data augmentation techniques used to improve the dataset and avoid overfitting. By learning from a wide range of image orientations and illumination circumstances, these phases aid in the model's improved generalisation. Furthermore, every image is appropriately labelled according to the crop type and the particular illness it displays, guaranteeing the model receives precise and understandable input for categorisation.

3.3 Model Architecture

Convolutional Neural Networks (CNNs), a potent deep learning architecture ideal for image categorisation applications, are used in this study. Because CNNs can automatically learn visual characteristics from data without human interaction, the introduction highlights their usage. CNNs are selected for their ability to successfully capture intricate patterns in pictures, and they are motivated by recent developments in deep learning that have been emphasised in the literature. By examining minute and complex visual clues found on the leaves, the model is intended to differentiate between different plant diseases. Lightweight models like MobileNet or richer architectures like ResNet may also be taken into consideration for testing, depending on the accuracy objectives and computing needs.

3.4 Model Training

The dataset is divided into training, validation, and testing sets for efficient learning. In order for the CNN model to understand the relationship between image attributes and illness classifications, it must be fed labelled pictures throughout the training phase. To modify the model's weights during training, optimisation methods like the Adam optimiser are frequently employed. The model converges effectively and overfitting is avoided with the use of early halting and learning rate modification. The objective is to create a model that, even when shown photos taken in varied settings, can reliably distinguish between many disease kinds and healthy leaves.

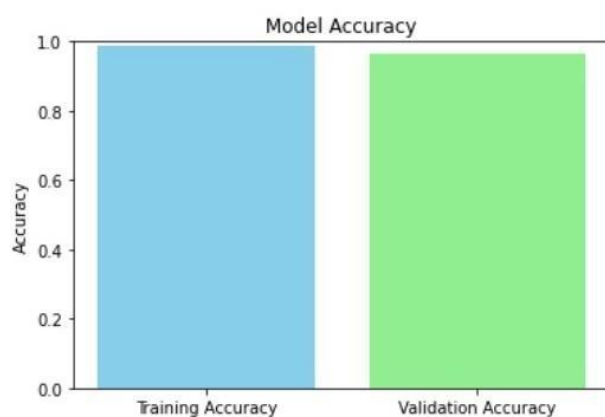


Fig 1.1 Model Accuracy

3.5 Application Deployment

Following training and evaluation, the model is incorporated into a useful application intended for farmers and agricultural specialists to utilise in the real world. According to the introduction, this system is meant to be accessed by mobile or online platforms, allowing for real-time illness diagnosis right in the field. When a user uploads a picture or video of a sick leaf, the app rapidly analyses the data and offers a diagnosis and some basic treatment recommendations. By enabling farmers to make well-informed decisions more rapidly, this technology eventually enhances crop health and agricultural output. In order to ensure acceptance even in rural or resource-constrained locations, the program strives to be lightweight and user-friendly.

IV. IMPLEMENTATION

The first step in putting this concept into action is gathering a large picture dataset of both healthy and sick plant leaves. Images of crops including rice, potatoes, and tomatoes were collected from publically accessible sources like the Plant Village dataset and field-level picture archives since the focus is on crops that are important to Indian agriculture. To make sure the model can distinguish between crops that are afflicted and those that are not, the dataset was organised into many classes that represent distinct crop illnesses in addition to a class for healthy leaves. To preserve consistency and enhance model performance, the photos were preprocessed after the data was collected. This required implementing data augmentation methods including rotation, zooming, brightness modifications, and horizontal flipping, as well as scaling all photos to a consistent resolution, usually 224x224 pixels, and normalising pixel values to fall within a predetermined range. These additions broaden the variety of datasets and improve the model's ability to generalise when tested in real-world scenarios. The foundation for supervised learning was established by labelling each image with the name of the crop and the associated illness.

Using Convolutional Neural Networks (CNNs) to classify images is the main focus of this study. CNNs were chosen for their capacity to identify minute visual patterns in plant leaves based on a review of related research and current developments. In order to extract spatial characteristics, the architecture consists of many convolutional layers, pooling layers, and finally, fully linked dense layers for classification. Transfer learning may be utilised to shorten training times while preserving high accuracy with pre-trained models like ResNet50 or MobileNet, depending on performance needs. The gathered dataset was used to refine these models. An 80-10-10 split of the training, validation, and test datasets was used to train the model. The Adam optimiser and categorical cross-entropy as the loss function were used for training. The training process was accelerated by the use of GPU acceleration using platforms such as Google Colab. To prevent overfitting and guarantee effective model convergence, strategies including early halting and learning rate scheduling were used. To guarantee stability and dependability, the model's performance on the validation set was regularly assessed during training.

Performance measures including accuracy, precision, recall, and F1-score were computed in order to assess the trained model. A thorough grasp of the model's categorisation skills across several illness categories was made possible by these measurements. In order to visualise the distribution of accurate and inaccurate predictions and aid identify classes with a higher rate of misclassifications, a confusion matrix was also created. The finished trained model was included into an online application that was made to be user-friendly for agricultural officers and farmers.

Users may take pictures of crop leaves with their phones or submit images to the interface. When an image is received, the program uses the trained model to interpret it and shows the anticipated illness name along with basic recommendations for prevention or therapy. While HTML, CSS, and JavaScript were used to create the frontend, Python frameworks like Flask or Fast API were used to develop the backend. With the use of this technology, field-level disease identification may be done in real time, allowing for prompt treatments that can lower output losses and promote sustainable farming methods.

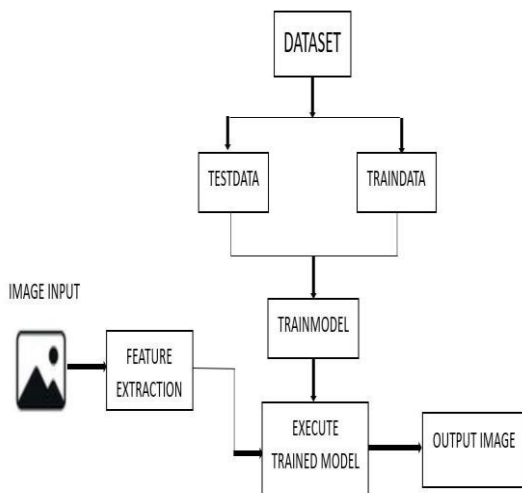


Fig 1.2 System Archite

III. RESULT AND DISCUSSION

Using pictures of both healthy and sick leaves, the CNN-based crop disease classification system demonstrated excellent accuracy in recognising a variety of plant diseases. With high accuracy, recall, and F1-scores, the model successfully differentiated between various illness categories, particularly for visually dissimilar groups. There were very few misclassifications, and they usually happened amongst illnesses with comparable symptoms. Transfer learning and data augmentation improved model resilience and decreased overfitting. Real- time illness detection is now possible thanks to the trained model's effective integration into an intuitive online application. All things considered, the method worked well as a practical, scalable, and effective way to help farmers with early diagnosis and improved crop management.

5.1 Image Output

When the user provides the leaf image as the input to the application, it is then passed on to the model for disease detection. The processed image will be stored in a predefined path of the folder respectively with the plant name and disease name. The result is then displayed as an image where the disease area is marked inside bounding boxes along with the disease name and the confidence score

is displayed adjacent to it as shown in fig 1.3. If the image provided is of a healthy leaf then the name of the leaf is along with confidence score is displayed as shown in fig 1.4



Fig 1.3 Input of Diseased leaf image

Disease Name: Apple__Cedar_apple_rust



Fig 1.4 Output of Disease leaf image with name of disease

VI. CONCLUSION

This study shows how Convolutional Neural Networks (CNNs), a kind of deep learning, may be used to accurately and efficiently identify agricultural diseases. The program was able to learn intricate visual cues and produce accurate predictions by using a sizable collection of photos of both healthy and damaged crop leaves. Farmers may access the trained model by integrating it into a web-based application, which allows for rapid action and real-time illness diagnosis. This method improves crop health monitoring and increases overall production, which not only lessens reliance on manual diagnosis but also promotes sustainable agriculture.

VII. REFERENCE

- [1] S. Maheswaran, S. Sathesh, P. Rithika, I. M. Shafiq, S. Nandita, and R. Gomathi, "Detection and classification of paddy leaf diseases using deep learning (cnn)," in International Conference on Computer, Communication, and Signal Processing, 2022: Springer, pp. 60-74.
- [2] S. Maheswaran, R. Gomathi, S. Sathesh, D. Kumar, G. Murugesan, and P. Duraisamy, "Real-time Implementation of YOLO V5 Based Helmet with Number Plate Recognition and Logging of Rider Data using PyTorch and XAMPP," in 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023: IEEE, pp. 1- 7.
- [3] Shill, Apu, and Md Asifur Rahman. "Plant disease detection based on YOLOv3 and YOLOv4." In 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), pp. 1-6. IEEE, 2021.
- [4] Song, ChauChung, ChihLun Wang, and YiFeng Yang. "Automatic detection and image recognition of precision agriculture for citrus diseases." In 2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 187-190. IEEE, 2020.
- [5] S. Sathesh, S. Maheswaran, T. Karthi, N. Kumar, R. Kumar, and R. Sabarishwaran, "Sensor Based Agribot For Agricultural Field," in 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023: IEEE, pp. 1- 6.
- [6] S. Maheswaran, S. Sathesh, A. Kumar, R. Hariharan, R. Ridhish, and R. Gomathi, "YOLO based Efficient Vigorous Scene Detection And Blurring for Harmful Content Management to Avoid Children's Destruction," in 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), 2022: IEEE, pp. 1063-1073.