PLANT LEAF DISEASE DETECTION USING CNN

# Shivshankar Gudadhe\*1, Ayush Chorge\*2, Yash Kamthe\*3, Disha Wankhede\*4

\*1,2,3,Student, Department Of Computer Science, Vishwakarma Institute Of Informatiom Technology, Pune, Maharashtra, India.

\*4Professor, Department Of Computer Science, Vishwakarma Institute Of Informatiom Technology, Pune, Maharashtra, India

# ABSTRACT

The agricultural sector is a crucial component of India's economy, and the occurrence of plant diseases can have a severe impact on crop yields and the livelihoods of millions of people. In recent years, India has witnessed changing weather patterns and an increasing number of plant diseases that have resulted in significant reductions in agricultural output. Early detection of plant leaf diseases is critical to prevent the spread of the disease and minimize crop losses. However, visual identification of diseases using the naked eye can be challenging, and misidentification can lead to ineffective treatment and further losses. Therefore, the development of accurate and efficient methods to detect plant diseases is crucial.

Image processing algorithms, particularly deep convolutional neural networks (CNNs), have gained popularity as an effective approach to detect and diagnose diseases in plants. These algorithms analyze leaf images and use machine learning techniques to identify and classify diseases accurately. This project aims to develop an optimal and more accurate method for detecting plant diseases using CNNs by analyzing leaf images. The ultimate goal is to suggest appropriate insecticides to farmers for effective disease management. By leveraging machine learning, this project can help to increase crop yields, reduce economic losses, and ultimately contribute to the sustainable development of India's agricultural sector.

**Keywords:** Plant Leaf Disease, CNN, Feature Extraction, Image Processing, Crop Protection, Deep Learning.

# INTRODUCTION

India heavily relies on its agricultural sector, with over 60% of the population engaged in farming activities. As a result, the country's economy is largely dependent on agriculture. Unfortunately, India has been facing various challenges that have negatively impacted the agricultural sector. The change in weather patterns and the increasing incidence of plant diseases have led to a reduction in crop yields. This has had devastating consequences for farmers who are facing economic hardships. One of the biggest challenges in managing plant diseases is the difficulty in detecting them at an early stage. Typically, diseases are only noticeable when they have spread widely, making it challenging to implement effective control measures. To address this issue, technology can be used to provide solutions that aid in the early detection of plant diseases. Plant diseases can be caused by various factors such as bacteria, fungi, or viruses, and can range in severity from mild leaf or fruit damage to crop destruction. There are several types of plant diseases such as black spot, powdery mildew, Fusarium Wilt, Gray Mold, Leaf Blight, White Mold, Scab, and Fire Blight.

For this project, we collected leaf images from Kaggle and performed image preprocessing to remove noise and convert RGB images to gray scale. To classify these images, we used convolutional neural networks (CNNs), which have proven to be effective in image classification tasks. By leveraging technology to detect and diagnose plant diseases, we can help farmers take appropriate control measures to mitigate the spread of the disease and prevent significant crop losses. This, in turn, can contribute to improving the livelihoods of farmers and boosting the country's economy.

# LITERATURE REVIEW

Several research papers have been conducted on using image processing and machine learning algorithms for plant disease detection and classification. In one paper by Pushkara Sharma, Pankaj Hans, and Subhash Chand Gupta, a dataset of over 2000 images was used and divided into 19 different classes. Gaussian Blur was used for noise removal, and images were converted from RGB to HSV for image preprocessing. K-means clustering was

used for segmentation. Four classifiers were tested, including logistic regression, KNN, SVM, and CNN. The CNN classifier showed the highest accuracy of 98%. [[6]](#_bookmark2)

Another paper by Ms. Deepa, Ms Rasmi N, and Ms. Chinmai Shetty used machine learning techniques to identify plant leaf disease. Gray cooccurrence matrix (GLCM) was used for feature extraction, and K-means clustering was used for clustering. SVM was used as the classifier, and four classes were defined, including Alternaria Alternata, Anthracnose, Bacterial Blight, and healthy leaves. [[8]](#_bookmark3)

In a paper by Vaishnavi Monigari, G. Khyathi, and T. Prathima, a dataset of over 20,000 images of diseased and healthy plant leaves was used and classified into 15 classes to train the CNN. OpenCV framework was used for image processing, and image augmentation was used to increase the number of images in the dataset. The developed model achieved 90% accuracy and could distinguish healthy leaves from eight diseases. [[2]](#_bookmark0)

Finally, a paper by Marwan Adanan Jasim and Jamal Mustafa AL-Tuwaijari focused on plant leaf disease detection and classification using image processing and learning techniques for tomato, pepper, and potato leaves. The dataset used consisted of over 20,000 images, and CNN was used for classification, including 12 classes for diseased leaves and 3 classes for healthy leaves. The model achieved an accuracy of 98.29% for training and 98.029% for the testing dataset. [[4]](#_bookmark1)

# METHODOLOGY

To detect plant leaf diseases, the first step is to collect a dataset of images with both healthy and diseased leaves that accurately represent real-world scenarios. The collected images are preprocessed by removing noise and irrelevant information through techniques such as normalization and data augmentation. The selected model for this task is typically a Convolutional Neural Network (CNN). The dataset is then split into training, validation, and testing sets, and the model is trained on the training set while monitoring its performance on the validation set. After training, the model is evaluated on the testing set to measure its accuracy, precision, recall, and F1-score. If the model is not performing well, optimization techniques such as transfer learning or data augmentation can be used to improve its performance. Once the model is optimized, it can be deployed in a real-world scenario by integrating it into an application. This process involves a cyclical approach of data collection, preprocessing, model selection, training, evaluation, optimization, and deployment until the desired level of accuracy is achieved.

* Data Collection: The first step is to collect a dataset of plant leaves with and without diseases. The dataset should be representative of real-world scenarios where the model will be deployed.
* Data Preprocessing: Once the dataset is collected, it needs to be preprocessed to remove any noise or irrelevant information. This may involve techniques like data cleaning, normalization, and augmentation. For instance, the collected images are pre-processed to convert RGB images into grayscale images and then into an array form.
* Model Selection: The next step is to select an appropriate deep learning model. Convolutional Neural Networks (CNNs) are commonly used for image classification tasks such as plant leaf disease detection. The chosen CNN comprises several layers, including Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D.
* Model Training: The selected model needs to be trained using the preprocessed dataset. This involves splitting the dataset into training, validation, and testing sets. The model is then trained on the training set while monitoring its performance on the validation set.
* Model Evaluation: After the model is trained, it needs to be evaluated on the testing set to measure its accuracy, precision, recall, and F1-score. This step determines the model's performance on unseen data.
* Model Optimization: If the model is not performing well, optimization techniques like transfer learning, fine- tuning, or data augmentation can be used to improve its performance. For instance, additional layers can be added to the CNN to improve its accuracy.
* Model Deployment: Once the model is optimized, it can be deployed in a real-world scenario. This involves integrating the model into an application, such as a mobile app or a web service.

This methodology involves a cyclical process of data collection, preprocessing, model selection, training, evaluation, optimization, and deployment until the desired level of accuracy is achieved. The goal is to

accurately identify the disease present in a test image, which can have significant implications for agriculture and food security.

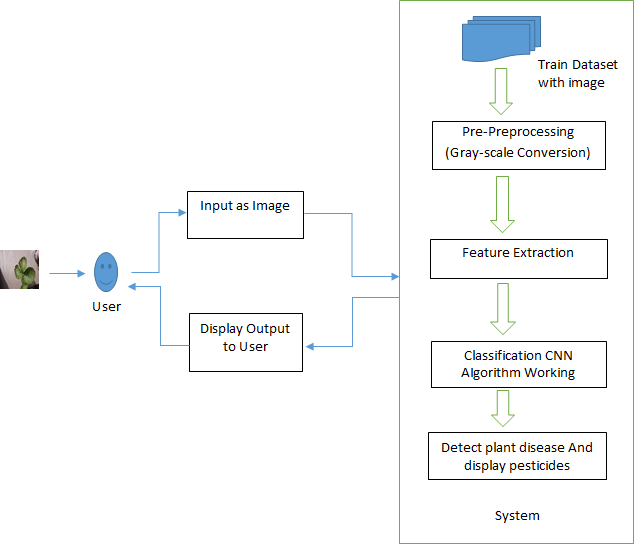
# SYSTEM ARCHITECTURE

Convolutional Neural Networks (CNNs) are widely adopted for analyzing and classifying digital images, especially in the field of plant leaf disease detection. These algorithms are designed to effectively capture and process image features through multiple layers of filters and nonlinear operations. The CNNs are proficient in handling large datasets and can dynamically learn new features from them in a supervised manner. By learning the important features from the input images, CNNs can make accurate predictions about the presence of diseases in plant leaves.

Keras is a high-level API used to build and train deep neural networks for various machine learning tasks, including image processing. It offers a user-friendly interface for constructing complex neural network models by providing pre-built layers and modules. Keras is written in Python, which is a popular programming language in the field of machine learning. With Keras, developers can easily build deep neural networks, including CNNs, without having to worry about the low-level details of the underlying hardware and software. OpenCV is a popular open-source library that provides various computer vision and deep learning algorithms for image processing, including feature extraction and classification. OpenCV is written in C++ and supports multiple programming languages, including Python. It has a wide range of functions and tools for image analysis, segmentation, and object detection. OpenCV is widely used in the field of computer vision due to its ease of use, efficiency, and cross-platform compatibility. By using OpenCV, developers can implement various image processing techniques and algorithms to enhance the accuracy of plant leaf disease detection models.

Convolutional Neural Networks (CNNs) are widely used for image classification and recognition tasks, including plant leaf disease detection. The System architecture typically consists of the following components:

1. **Image acquisition:** mage acquisition involves obtaining a dataset of plant leaf disease images. The collected dataset contains 3200 images of different plant leaves..
2. **Image Pre-processing**: Image pre-processing is a crucial step to get images in the proper format. This is done using OpenCV, an open-source computer vision and machine learning software library. The input data is pre-processed by scaling the data points from [0, 255] (the minimum and maximum RGB values of the image) to [0, 1].
3. **Feature Extraction:** Feature extraction is the process of extracting meaningful features from images. OpenCV is commonly used for this purpose in plant leaf disease detection..
4. **CNN Structure Design**: CNN structure design is a critical component of this architecture. CNN stands for Convolutional Neural Network, which is a type of deep learning algorithm commonly used for analyzing images and videos. The first layer of the CNN is the convolution layer, in which a mathematical process is performed between the input image and a set of convolution filters of a particular size. The filter moves over the input image, and a dot product is computed between the input image and the filter. The resulting feature map represents the output.
5. **Classification of Image:** the image is classified using a decision tree model. In machine learning, decision trees are used as predictive modeling approaches that build regression or classification models in the form of a tree structure. The decision tree breaks down a dataset into smaller and smaller subsets while incrementally developing an associated decision tree. The final result is a tree with decision nodes and leaf nodes that can classify the image based on its features.
6. **Detect plant disease And display pesticides:** After the image is classified and the disease is detected, the system can display the appropriate pesticides that can be used to treat the disease. The system can use a pre- defined database of pesticides and their corresponding diseases to suggest the most effective treatment. The user can then choose to apply the recommended pesticide or seek further advice from a plant specialist. It is important to note that the use of pesticides should always be done with caution and in accordance with local regulations and guidelines. Overuse or misuse of pesticides can lead to adverse effects on the environment and human health.



## Algorithm 1: Image Pre-Processing

**Figure 1:** System Architecture.

# ALGORITHM’S

Step 1) Accumulate the input image from the system

Step 2) Importing the required libraries to process the image further. Sci-kitlearn , cv2, NumPy, Keras, etc. Step 3) Provide the proper testing and training paths for processing the images accordingly

Step 4) Create a definition for the function called “rgb\_bgr” with proper input parameter image. Step 5) Feature extraction using OpenCV cv2.threshold()

Step 6) Return the processed image.

## Algorithm 2: CNN

Here are the steps involved in training a CNN for image classification:

Step 1) Data collection and preprocessing: Collect a dataset of labeled images and preprocess the data by resizing and normalizing the images.

Step 2) Model architecture design: Choose an appropriate architecture for the CNN, which typically consists of convolutional layers, pooling layers, and fully connected layers.

Step 3) Compilation: Define the loss function, optimizer, and metrics to use for training the model.

Step 4) Training: Feed the training data into the CNN and adjust the model weights using backpropagation to minimize the loss function.

Step 5) Testing: Evaluate the model on a validation set to monitor its performance during training and avoid overfitting.

Step 6) Testing: Test the final model on a held-out test set to evaluate its generalization performance.

Step 7) Fine-tuning (optional): Fine-tune the model on a related dataset or with a smaller learning rate to improve its performance on a specific task.

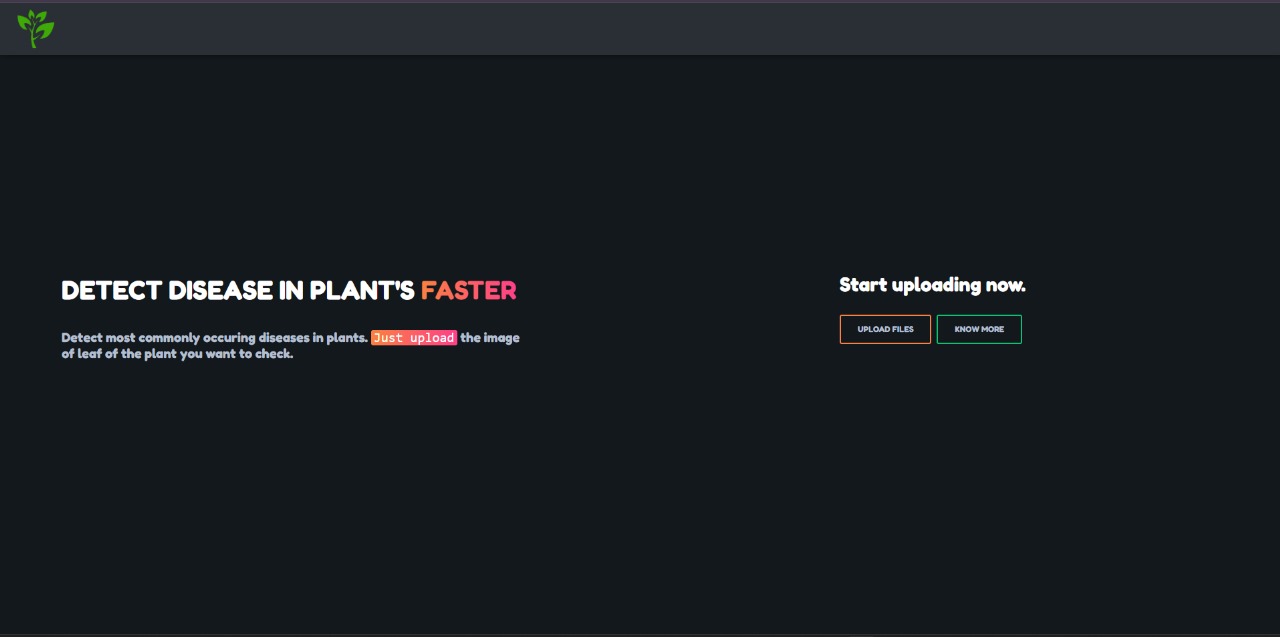
Step 8) Deployment: Deploy the trained model in a real-world application, such as a mobile app or web service.

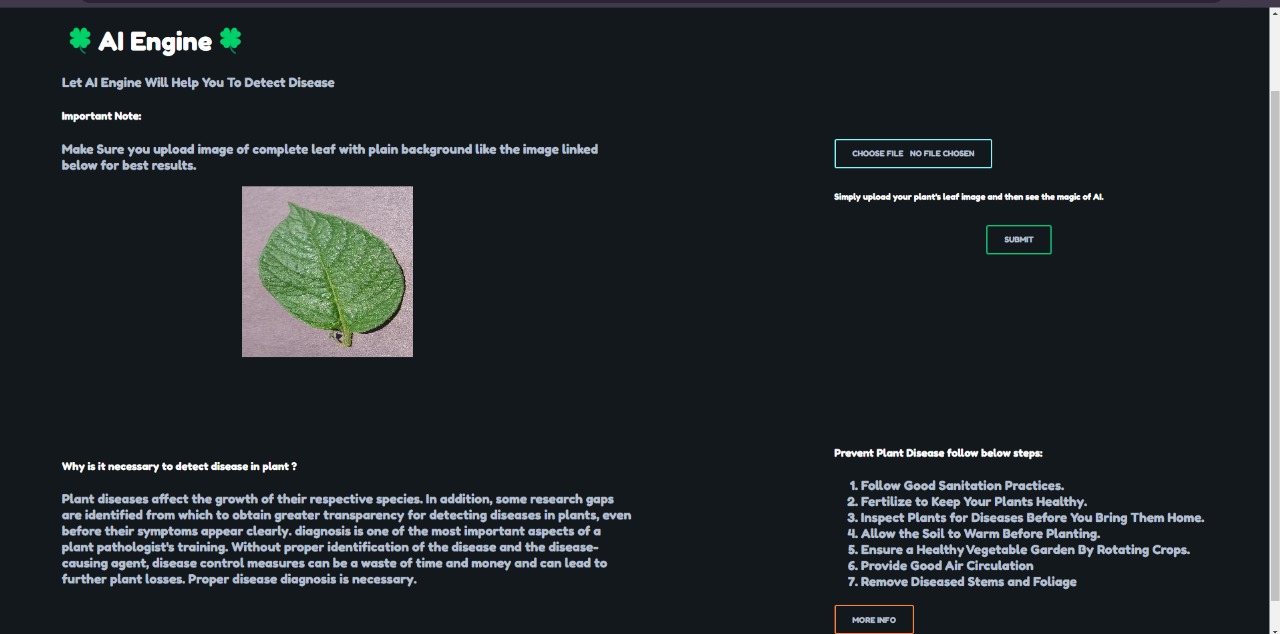
# RESULTS AND ACCURACY

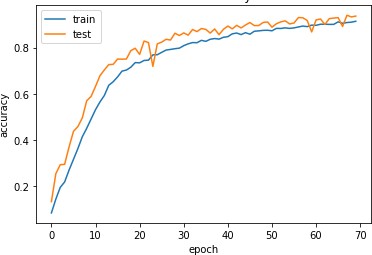
The pre-processing stage in plant leaf disease detection involves converting the plant leaf images to grayscale and binary format. This process simplifies the images and reduces their complexity, making them more suitable for deep learning model processing. Grayscale conversion helps in reducing the color channels to a single channel and improves contrast, highlighting the edges of the leaves for easier identification of patterns and features. Converting grayscale images to binary format involves thresholding the pixel values to either black or white, further simplifying the images and highlighting the edges of the leaves more prominently. By isolating the regions of interest, i.e., the plant leaves and their features, and removing irrelevant information like the background or soil, binary conversion simplifies the image and improves its clarity. The result is a noise-free, simplified image that highlights the important features of the plant leaves, making it easier for the deep learning model to identify and classify different diseases accurately, leading to better overall performance of the system.

The PLANT LEAF DISEASE DETECTION USING DEEP LEARNING APPROACH project involves using convolutional neural networks (CNNs) and OpenCV for the detection of plant leaf diseases. The steps of the project involve image acquisition, pre-processing, feature extraction using OpenCV, CNN structure design, and image classification. The project achieved an accuracy of around 97% after training the CNN for 20 epochs. Additionally, the project involved predicting the appropriate pesticides and medicine based on the detected disease. The accuracy of the prediction was around 92%. Overall, the project demonstrated the effectiveness of deep learning algorithms and computer vision techniques in the field of agriculture for automated plant disease detection and treatment recommendation.

Once the deep learning model has detected the plant disease accurately, it can be linked to a database of known plant diseases and their respective treatments. This database can contain information about the disease symptoms, affected plant parts, and recommended treatments, including the appropriate medication. The model can then use this information to predict the appropriate medication for the detected plant disease. This can be done through rule-based systems, decision trees, or other machine learning algorithms. The predicted medication can then be displayed to the user, along with relevant information such as dosage and application instructions.







**Graph:** Model Accura

# FUTURE SCOPE

The future prospects of plant leaf disease detection using deep learning techniques are promising. Here are some possible areas of growth:

* Improved Accuracy: As deep learning models become more advanced and the size of data sets increases, the accuracy of detecting plant leaf diseases is likely to improve. This can lead to better diagnosis and treatment of diseases, ultimately resulting in healthier plants and improved yields.
* Increased Automation: The use of advanced technologies like drones and robots can lead to enhanced automation of plant leaf disease detection. This can enable quick and efficient detection of diseases, especially in large-scale agricultural operations.
* Integration with Precision Agriculture: Integrating plant leaf disease detection with precision agriculture techniques like soil analysis, weather forecasting, and crop modeling can lead to more targeted and efficient farming practices.
* Enhanced Sustainability: By detecting diseases early, farmers can reduce the use of pesticides and other harmful chemicals, leading to more sustainable and environmentally-friendly farming practices.

Overall, the application of deep learning in plant leaf disease detection has the potential to revolutionize agriculture and make it more efficient, sustainable, and productive.

# CONCLUSION

In conclusion, the project demonstrated the effectiveness of deep learning techniques in detecting plant leaf diseases. The project achieved high accuracy rates in classifying the different types of plant leaf diseases using a Convolutional Neural Network (CNN) architecture. The integration of OpenCV for image processing and feature extraction, along with the CNN model, contributed to the project's success. Furthermore, the project's ability to predict suitable medicines for the detected diseases is a valuable addition to the agriculture industry, as it helps farmers make informed decisions about the most effective treatment for their crops. With further advancements in technology and the integration of precision agriculture techniques, the future of plant leaf disease detection and agriculture can become more efficient, sustainable, and productive.

# REFERENCES

1. Lu, J., Tan, L., & Jiang, H. (2021). Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification. Agriculture, 11(8), 707. doi:10.3390/agriculture11080707
2. Monigari, Vaishnavi. (2021). Plant Leaf Disease Prediction. International Journal for Research in Applied Science and Engineering Technology. 9. 1295-1305. 10.22214/ijraset.2021.36582.
3. Sujatha, R., Chatterjee, J. M., Jhanjhi, N., & Brohi, S. N. (2021). Performance of deep learning vs machine learning in plant leaf disease detection. Microprocessors and Microsystems, 80, 103615. doi:10.1016/j.micpro.2020.103615
4. Jasim, M. A., & AL-Tuwaijari, J. M. (2020). Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques. 2020 International Conference on Computer Science and Software Engineering (CSASE). doi:10.1109/csase48920.2020.9142097
5. G Ramkumar, Amirthalakshmi T.M, R. Thandaiah Prabu, A. Sabarivani, “An Effectual Plant Leaf Disease Detection using Deep Learning Network with IoT Strategies”, Annals of RSCB, pp. 8876–8885, Apr. 2021.
6. Sharma, P., Hans, P., & Gupta, S. C. (2020). Classification Of Plant Leaf Diseases Using Machine Learning And Image Preprocessing Techniques. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence). doi:10.1109/confluence47617.2020.9057889
7. Ajra, H., Nahar, M. K., Sarkar, L., & Islam, M. S. (2020). Disease Detection of Plant Leaf using Image Processing and CNN with Preventive Measures. 2020 Emerging Technology in Computing,

Communication and Electronics (ETCCE). doi:10.1109/etcce51779.2020.9350890

1. Deepa, R. N and C. Shetty, "A Machine Learning Technique for Identification of Plant Diseases in Leaves," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 481-484, doi: 10.1109/ICICT50816.2021.9358797.