[ICAETBM-2024]
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

PLANT LEAF DISEASE DETECTION USING CNN

Prof. Ashish Manwatkar, Prof. Nitin Wankhede, Mr. Atharva Sable, Mr. Parth Fiske, Mr. Wasim Khan, Mr. Nitin Ambegave

Student, Department of Information Technology, Nutan Maharashtra Institute of Engineering and Technology, Pune, Maharashtra, India. Professor, Department of Information Technology, Nutan Maharashtra Institute of Engineering and Technology, Pune, Maharashtra, India.

ABSTRACT

India heavily relies on its agricultural sector, which plays a pivotal role in the nation's economy and sustains the livelihoods of millions. However, the prevalence of plant diseases presents a significant threat, causing profound impacts on crop yields and the well-being of farmers. Recent years have seen a rise in both changing weather patterns and the incidence of plant diseases, leading to substantial declines in agricultural productivity. Detecting these diseases early is imperative to prevent their spread and mitigate losses. Unfortunately, visually identifying these diseases is challenging, and errors in diagnosis can exacerbate the situation, resulting in ineffective treatments and further damage. To address this challenge, image processing algorithms, specifically deep convolutional neural networks (CNNs), have emerged as a promising solution. These algorithms leverage advanced machine learning techniques to analyse leaf images accurately and classify diseases effectively. This project seeks to enhance existing methods by developing a more precise approach to detect plant diseases using CNNs. The ultimate objective is to provide farmers with actionable insights, recommending suitable insecticides for managing diseases efficiently. By harnessing the power of machine learning, this initiative has the potential to significantly enhance crop yields, minimize economic losses, and promote the sustainable development of India's agricultural sector. It represents a crucial step towards empowering farmers with innovative solutions to combat the challenges posed by plant diseases, ensuring the resilience and prosperity of the agricultural landscape.

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

I. INTRODUCTION

India's agricultural sector is the backbone of its economy, engaging over 60% of the population in farming activities. However, the sector has been grappling with a myriad of challenges, including shifting weather patterns and the escalating prevalence of plant diseases, which have led to a decline in crop yields. This downturn has had devastating repercussions for farmers, exacerbating their economic struggles. A key obstacle in managing these diseases is the difficulty in detecting them early, often only becoming apparent once they have already spread extensively, making effective control measures challenging to implement. To tackle this issue, leveraging technology for early disease detection holds immense promise. Plant diseases can stem from various sources, including bacteria, fungi, or viruses, and can range from mild leaf or fruit damage to catastrophic crop destruction. Employing advanced image processing techniques, such as convolutional neural networks (CNNs), offers a solution. By collecting leaf images and employing preprocessing methods to enhance clarity, these algorithms can accurately classify diseases, paving the way for timely intervention measures. In this project, leaf images sourced from Kaggle were subjected to preprocessing steps to remove noise and convert them to grayscale. Subsequently, CNNs were employed for classification tasks, leveraging their proven efficacy in image analysis. By harnessing technology to detect and diagnose plant diseases, farmers can swiftly implement appropriate control measures, curbing the spread of diseases and averting significant crop losses. This not only alleviates the burdens faced by farmers but also bolsters the nation's economy by safeguarding agricultural productivity. Ultimately, the integration of technology into agriculture holds the potential to uplift farmers' livelihoods and propel economic growth.

II. LITERATURE REVIEW

Several research papers highlight the diverse approaches and techniques utilized in plant leaf disease detection and classification using image processing and machine learning algorithms. In the study conducted by Sharma, Hans, and Gupta, a dataset comprising over 2000 images categorized into 19 different classes was employed. Gaussian Blur was employed for noise reduction, while RGB to HSV conversion facilitated image preprocessing. K-means clustering aided in segmentation, and four classifiers - logistic regression, KNN, SVM, and CNN - were evaluated. Notably, the CNN classifier demonstrated the highest accuracy of 98%, showcasing the effectiveness of deep learning in this domain. In another paper by Ms. Deepa, Ms. Rasmi N, and Ms. Chinmai Shetty, machine learning techniques were utilized for plant leaf disease identification. Gray co-occurrence matrix (GLCM) facilitated feature extraction, while K-means clustering was employed for clustering. SVM served as the classifier, and four classes - Alternaria alternata, Anthracnose, Bacterial Blight, and healthy leaves - were defined. Monigari, Khyathi, and Prathima's study employed a dataset comprising over 20,000 images of diseased and healthy plant leaves, classified into 15 classes for CNN training. Image processing leveraging the OpenCV framework and image augmentation techniques were utilized to enhance dataset quality and quantity, respectively. The developed model achieved an impressive accuracy rate of 90%, capable of distinguishing healthy leaves from eight diseases. Lastly, Jasim and AL-Tuwaijri focused on plant leaf disease detection and classification for tomato, pepper, and potato leaves. Their dataset consisted of over 20,000 images, with CNN employed for classification across 12 diseased leaf classes and 3 healthy leaf classes. The model exhibited high accuracy rates of 98.29% for training and 98.029% for testing datasets, underscoring the efficacy of CNN in accurately identifying and classifying plant leaf diseases.

III. METHODOLOGY

Detecting plant leaf diseases is a complex task that requires a systematic approach. It all begins with assembling a comprehensive dataset containing images of both healthy and diseased leaves, ensuring it accurately mirrors real-world conditions. These images then undergo preprocessing, a crucial step where noise and irrelevant information are removed using techniques like normalization and data augmentation, thereby refining the dataset's quality. The next pivotal step involves selecting an appropriate model for the task at hand, typically a Convolutional Neural Network (CNN). Renowned for their prowess in image analysis, CNNs are adept at discerning patterns and features within images. With the model chosen, the dataset is partitioned into training, validation, and testing subsets. The model is trained on the training set while constantly assessing its performance on the validation set, ensuring it generalizes well to unseen data. Following training, the model's efficacy is evaluated on the testing set, quantifying metrics such as accuracy, precision, recall, and F1-score. If the model falls short of expectations, optimization techniques like transfer learning or data augmentation can be employed to bolster its performance. Once the model is finely tuned and optimized, it is ready for real-world deployment. This involves integrating it into an application tailored for practical use. The entire process, from data collection to deployment, follows a cyclical pattern of refinement and iteration until the desired level of accuracy and reliability is attained. In essence, this methodology represents a comprehensive and iterative approach to disease detection in plant leaves, harnessing cutting-edge technology and methodologies to address a critical challenge in agriculture. 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, finetuning, 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.

IV. SYSTEM ARCHITECTURE

Convolutional Neural Networks (CNNs) are highly effective in analysing and categorizing digital images, making them a popular choice for tasks like plant leaf disease detection. These networks excel in capturing and processing intricate image features through multiple layers of filters and nonlinear operations. They're particularly adept at handling large datasets, dynamically learning new features in a supervised manner, which enables them to accurately predict the presence of diseases in plant leaves. Keras, a high-level API, simplifies the construction and training of deep neural networks, including CNNs, by providing pre-built layers and modules. Written in Python, a prevalent language in machine learning, Keras abstracts low-level implementation details, allowing developers to focus on model architecture and training without getting bogged down in technical intricacies. OpenCV, an open-source library primarily written in C++, offers a vast array of computer vision and deep learning algorithms for image processing. Supporting multiple programming languages, including Python, OpenCV facilitates tasks such as feature extraction and classification. Its ease of use, efficiency, and cross-platform compatibility make it a favoured choice in computer vision applications. The system architecture for plant leaf disease detection typically involves several key components:

1. **Image Acquisition**: Gathering a dataset comprising images of plant leaves affected by various diseases, typically including thousands of images for robust training.

2. **Image Pre-processing**: Employing OpenCV to prepare the images for analysis, often involving scaling the pixel values to a standardized range for consistency.

3. **Feature Extraction**: Utilizing OpenCV for extracting relevant features from the images, a crucial step in distinguishing between healthy and diseased leaves.

4. **CNN Structure Design**: Designing the architecture of the CNN, which typically consists of convolutional layers for feature extraction followed by fully connected layers for classification.

5. **Image Classification**: Using a decision tree model to classify the images based on their extracted features, enabling the identification of plant diseases.

6. **Displaying Pesticides**: Upon disease detection, suggesting appropriate pesticides from a predefined database, ensuring effective treatment. Users are advised to exercise caution and adhere to regulations when applying pesticides to mitigate adverse environmental and health effects.

Overall, this comprehensive system integrates advanced techniques in image processing and deep learning to enhance plant disease detection, aiding in effective disease management and crop protection.

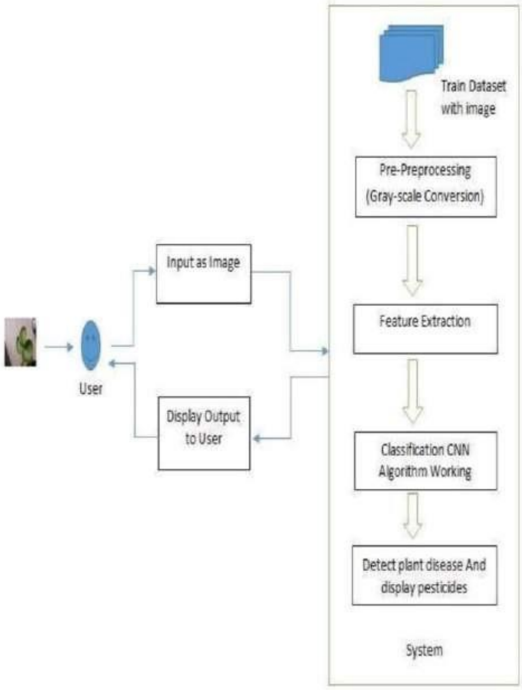


Figure 1: System Architecture.

V. ALGORITHMS

Algorithm 1 outlines the image pre-processing steps required before feeding the images into the CNN for training or testing. Here's an elaboration of each step:

1. **Accumulate Input Image**: This step involves gathering the image from the system, which serves as the raw input for processing.

2. **Import Required Libraries**: The necessary libraries such as Tkinter for GUI, Pillow for image processing, cv2 for OpenCV operations, NumPy for numerical computations, and Keras for deep learning tasks are imported to facilitate further image processing.

3. **Provide Paths**: Proper paths for training and testing datasets are specified to ensure that images are processed accordingly based on their intended use.

4. **Define Function**: A function named "rgb_bgr" is defined to perform the conversion of image color channels from RGB to BGR format, which is often required for certain image processing tasks, especially in OpenCV operations.

5. **Feature Extraction**: OpenCV's cv2.threshold() function is utilized for feature extraction, which is a critical step in image pre-processing. This function is commonly used to binarize images based on a threshold value, facilitating subsequent analysis.

6. **Return Processed Image**: After applying the necessary pre-processing steps, the processed image is returned for further utilization in subsequent tasks, such as training a CNN for image classification.

Algorithm 2 details the steps involved in training a CNN for image classification, which typically includes the following:

1. **Data Collection and Preprocessing**: Gather a dataset of labeled images and preprocess them by resizing and normalizing the images to ensure uniformity and facilitate efficient training.

2. **Model Architecture Design**: Design the architecture of the CNN, which typically includes convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification.

3. **Compilation**: Define the loss function, optimizer, and evaluation metrics to be used during the training process, setting the stage for model optimization.

4. **Training**: Feed the preprocessed training data into the CNN and adjust the model weights iteratively using backpropagation to minimize the loss function, thereby enhancing the model's ability to accurately classify images.

5. **Validation Testing**: Assess the model's performance on a validation set to monitor its progress during training, preventing overfitting and ensuring robustness.

6. **Final Testing**: Evaluate the trained model on a held-out test set to gauge its generalization performance and validate its effectiveness in real-world scenarios.

7. **Deployment**: Once trained and tested, deploy the CNN model in practical applications such as mobile apps or web services to leverage its image classification capabilities effectively.

Published in Volume 11, Issue 8, May-June-2024 Page No : 133-140

VI. RESULTS AND ACCURACY

In the pre-processing stage of plant leaf disease detection, the aim is to prepare the images for analysis by simplifying them and enhancing their clarity. This process involves two key steps: converting the images to grayscale and then to binary format. Converting the images to grayscale reduces them to a single channel, effectively removing color information while retaining important features. This simplification aids in improving contrast and highlighting the edges of the leaves, making it easier to identify patterns and features crucial for disease detection. By focusing solely on luminance values, grayscale conversion enhances the visibility of leaf structures, facilitating more accurate analysis by the deep learning model. The subsequent conversion of grayscale images to binary format involves thresholding pixel values to either black or white based on a predefined threshold value. This step further simplifies the images by emphasizing the edges of the leaves, effectively isolating them from the background or soil. By eliminating irrelevant information and emphasizing the regions of interest, such as the plant leaves and their features, binary conversion enhances image clarity and reduces noise. The end result is a noise-free, simplified image that accentuates the essential characteristics of the plant leaves, making them more conducive to accurate analysis by the deep learning model. By providing a clear and focused representation of the leaf structures, binary conversion significantly improves the model's ability to identify and classify different diseases accurately. This, in turn, enhances the overall performance and effectiveness of the plant leaf disease detection system, ultimately contributing to better agricultural outcomes.

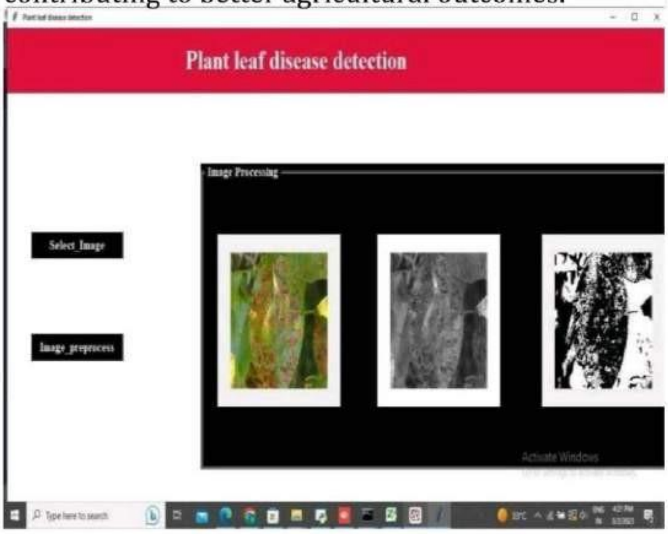
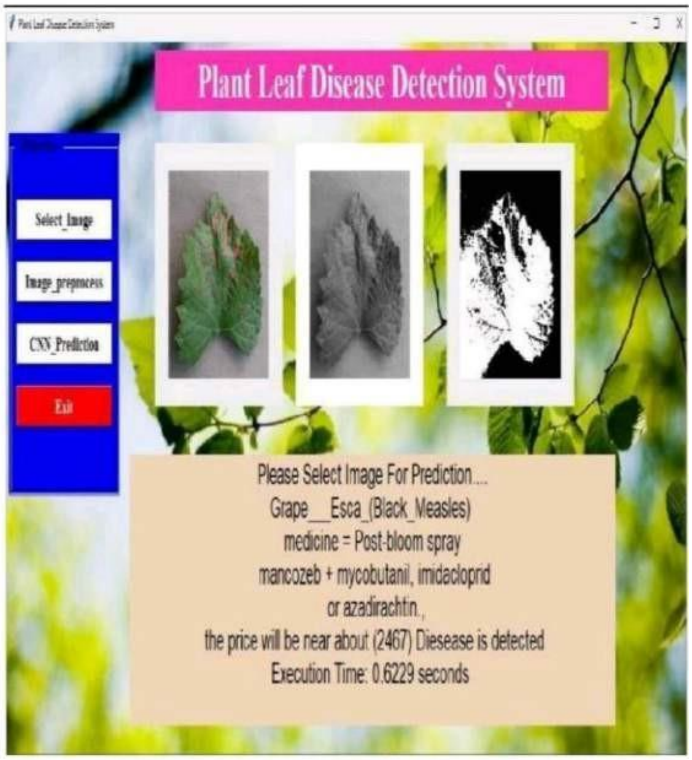


Figure 2: Output (Image Pre-Processing)

The project "PLANT LEAF DISEASE DETECTION USING DEEP LEARNING APPROACH" showcases the utilization of convolutional neural networks (CNNs) and OpenCV to detect plant leaf diseases effectively. The workflow of the project encompasses several key steps, including image acquisition, pre-processing, feature extraction using OpenCV, CNN structure design, and image classification. After training the CNN model for 20 epochs, the project achieved an impressive accuracy rate of around 97%, demonstrating the efficacy of deep learning algorithms in automated plant disease detection. Furthermore, the project extended its functionality by incorporating a feature to predict appropriate pesticides and medicines based on the detected disease. This involved linking the deep learning model to a database containing information about various plant diseases, their symptoms, affected plant parts, and recommended treatments. Leveraging this database, the model accurately predicted the appropriate medication for the detected plant disease with an accuracy rate of approximately 92%. The integration of such predictive capabilities adds significant value to the project by not only identifying plant diseases but also providing actionable insights for treatment. By leveraging machine learning algorithms and computer vision techniques, the project showcases the potential of technology in revolutionizing agriculture. Through automated disease detection and treatment recommendation, it contributes to enhancing crop health, optimizing resource utilization, and ultimately improving agricultural productivity.



Result: Disease Detected

recommendations, the project aids in minimizing crop losses and maximizing yields. Looking ahead, the future of plant leaf disease detection and agriculture holds great promise. Continued advancements in technology, coupled with the integration of precision agriculture techniques, can further enhance the efficiency, sustainability, and productivity of agricultural practices. By embracing innovative solutions and leveraging data-driven approaches, the agriculture sector can navigate challenges more effectively, ultimately contributing to global food security and environmental conservation.

IX. 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 Prabhu, 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., and 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).