



Prediction of Plant Leaf Health by Image Contour Segmentation and LSTM Model

Arpit Pethe Scholar, Arjun Rajput Assistant Prof., Dr. Sanjay Sharma,
Dr. Surabhi Karsoliya

Department of CSE, Technocrats Institute of Technology College, Bhopal MP, India.

Abstract- Agriculture is the backbone of the Indian economy and contributes significantly to the GDP. Crop diseases, particularly those affecting leaves, lead to a decline in both the quality and quantity of agricultural produce. Traditional methods like expert diagnosis and pathogen analysis depend on skilled professionals and may be time-consuming. These approaches are also prone to human errors, affecting the accuracy of disease identification and management. This paper has proposed Plant Leaf Health Prediction Model (PLHPM) leaf health prediction model that segment input image and extract features for learning. Image segmentation was done by active contour method, while histogram features were extracted from the image. Extracted histogram features were used for the LSTM model training. Experiment was done on real dataset images of potato. Result shows that proposed Plant Leaf Health Prediction Model (PLHPM) has increased the detection precision and accuracy in less execution time.

Keywords- Image processing, Leaf disease detection.

I. INTRODUCTION

Plant diseases pose a significant global challenge for food industry. Essential crops like wheat, maize, rice, etc., which support a large portion of the world's population, are particularly vulnerable. The presence of various plant diseases often leads to substantial reductions in yield. These diseases, caused by pathogens such as fungi, bacteria, viruses, and other environmental stressors, have a substantial impact on agricultural output. Estimates suggest that plant diseases are responsible for reducing the yields of major food and cash crops by approximately 20% to 40%, leading to severe economic and social ramifications. The financial burden of these losses is staggering, with global agricultural losses due to plant diseases valued at approximately \$220 billion annually.

Given that agricultural productivity is directly linked to food availability, the need for effective disease prevention and management strategies is more pressing than ever. Without adequate control measures, plant diseases can lead to large-scale crop failures, exacerbating food scarcity, driving up food prices, and negatively impacting farmers' livelihoods. Additionally, the improper use of pesticides and other chemical treatments—often resulting from misdiagnosed plant diseases—can further deteriorate soil health, water quality, and ecological balance.

One of the most effective ways to mitigate the impact of plant diseases is through early and accurate detection. Detecting plant diseases primarily involves recognizing visual symptoms that appear on plant leaves, stems, or fruits. These symptoms may manifest as discoloration, spots, wilting, or lesions, each indicating a different type of infection or stress. However, manual disease detection, which relies on human expertise, is often time-consuming, prone to errors, and impractical for large-scale agricultural monitoring. A misdiagnosis can lead to incorrect disease management decisions, such as the selection



of an ineffective chemical treatment, which could not only fail to control the disease but also result in lower crop yields and potential health hazards. Thus, there is a critical need for an efficient and automated plant disease detection system capable of identifying infections at an early stage, thereby enabling timely interventions that can protect entire harvests.

With advancements in artificial intelligence and machine learning, particularly deep learning, numerous automated techniques for plant disease diagnosis have emerged. Deep learning-based models leverage vast datasets of plant images to learn patterns associated with different diseases, enabling highly accurate classification of infections. These models have demonstrated significant potential in improving the speed, accuracy, and scalability of plant disease detection systems. By integrating deep learning techniques into agricultural monitoring, farmers and researchers can make more informed decisions regarding disease control strategies, optimizing both productivity and sustainability.

This research focuses on developing an advanced system for detecting and classifying plant leaf diseases using deep learning techniques. The primary objective is to explore how deep learning algorithms can effectively analyze images of plant leaves and accurately classify various types of diseases. The study aims to improve the precision, efficiency, and adaptability of plant disease detection systems, ensuring that farmers and agricultural professionals can swiftly and accurately diagnose infections and implement appropriate management measures.

Rest of paper is organized into few more sections. Section 2 presents a comprehensive review of other researchers' work. In Section 3, we discuss the functionality and architectural design of the proposed deep learning model. Section 4 provides an in-depth evaluation of the model's components through extensive experimentation, analyzing its performance in comparison to existing approaches. Finally, Section 5 concludes the work outcomes with future direction.

II. RELATED WORK

E. Moupojou et al. [8] introduced an ensemble-based framework for identifying plant diseases using images captured directly from the field. Their method incorporates the Segment Anything Model (SAM) to identify objects within the images, followed by image processing steps that enhance the segmentation results. To separate leaf regions from the background, they employed a Fully Convolutional Data Description (FCDD) model. The disease classification task is then handled by a model trained on the PlantVillage dataset, ensuring precise identification of plant ailments.

Pawar et al. [9] underscored the critical role of early disease diagnosis in minimizing crop yield losses. They proposed a systematic image processing workflow that includes stages such as image acquisition, preprocessing, segmentation, feature extraction, and classification. Automating these steps allows for rapid and dependable diagnosis, empowering farmers to take swift and effective action.

Aithal et al. [10] addressed the importance of plant disease recognition in the context of Indian agriculture. They proposed an automated system where farmers can upload images of infected leaves and receive instant analysis along with symptom descriptions and treatment advice. Their model, trained on data from crops such as grape, potato, and tomato using the PlantVillage dataset, achieved high accuracy, indicating its practical utility for improving farming outcomes.

Reddy et al. [11] focused on machine learning-based methods for detecting plant diseases, particularly using Support Vector Machine (SVM) and Random Forest algorithms. Their findings suggest that such approaches offer a scalable and effective alternative to traditional manual inspection methods, thereby increasing efficiency and reducing the labor involved.



In [12], researchers conducted a comparative study of SVM, Multi-Class SVM, and a hybrid model combining Random Forest and Decision Tree classifiers. While all approaches successfully identified plant diseases, the ensemble model exhibited higher confidence levels, suggesting greater reliability for practical applications. The study emphasized the importance of confidence scores in supporting decisions for effective disease control.

In [13], a two-dimensional Convolutional Neural Network (2D CNN) was proposed to classify diseases in tomato and cotton crops. The model featured a shallow architecture, offering computational efficiency without compromising accuracy. Evaluation metrics such as confusion matrices, ROC curves, and AUC values indicated that the model outperformed various transfer learning approaches, validating its effectiveness in practical disease detection scenarios.

III. PROPOSED METHODOLOGY

This section outlines the proposed Plant Leaf Health Prediction Model (PLHPM). Figure 1 illustrates the overall flow of the training process. Each block in the figure is explained in detail in the following subsections. The model is developed using the Image Potato Dataset (IPD). The system is divided into two primary modules. The first module performs feature extraction from leaf images. The second module involves training a mathematical model using extracted features. Together, these modules enable accurate prediction of plant leaf health.

PI <-Plant_Image_Preprocessing(IPD)

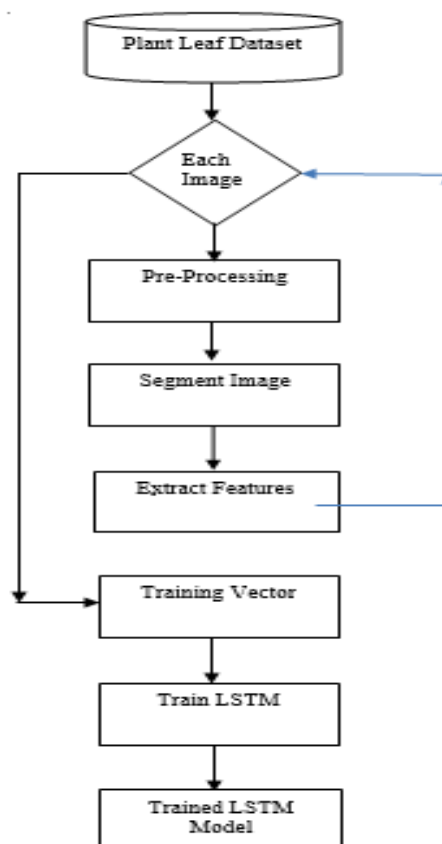


Fig. 1 Block diagram of proposed PLHPM model.



Plant Image Pre-Processing: Input potato plant leaf image is preprocessed by transforming image into square matrix. As plant image have background that is not the part of study and act as noise. In order to identify background image section segmentation is applied in image. So let's IPD each image pass in pre-processing function which gives an processed image PI as an output.

Active contour segmentation:

The active contour model, originally introduced by Osher and Sethian, has evolved into a widely adopted mathematical and computational approach used across various domains such as image processing, fluid dynamics, pattern recognition, and computer vision. In the context of image segmentation, this method offers several advantages over traditional techniques, particularly in handling complex shape variations and dynamic topological changes. The proposed approach in this work incorporates the active contour method with the following key steps:

- A randomly initialized matrix, matching the dimensions of the input image, is first generated and integrated with the image to aid in initializing the contour. Although natural contours may already exist within the image, this step enhances the effectiveness of contour generation.
- The next phase involves locating contour positions and generating contour outlines, which form the basis for initial image segmentation.
- Once the contours are established, the segmentation process is refined by assessing the proximity of each pixel to the segmented regions. Pixels with a negative distance value are classified as part of the segment, while those with a positive distance are considered outside the segment.
- Finally, the segmented areas are continuously updated by analyzing surrounding pixel intensities. The update function is designed to accommodate shape deformations, allowing the contour to adapt flexibly. The active contour model dynamically adjusts to either merge or split regions based on evolving pixel characteristics, leading to more accurate and flexible segmentation outcomes.

SI<-Active_Contour_Segmentation(PI)

Feature Extraction Module A histogram is an essential feature used in image analysis to capture the distribution of pixel intensities within an image. In the case of unhealthy plant images, the color of the leaves gradually changes from green to yellowish and then to brown, indicating the presence of disease or stress. The classification of such images heavily relies on color-based features, as these variations in hue and intensity serve as key indicators of plant health [18]. To effectively analyze these color changes, the foreground region of the image, which was segmented in the previous step, is used for extracting histogram feature values. The histogram represents the frequency distribution of pixel intensities in the segmented region, allowing for a quantitative assessment of color variations. In this model, sixteen bins are utilized for histogram calculation, ensuring a detailed and structured representation of color variations across different intensity levels. This approach enhances the accuracy of classification by providing meaningful color-based information for disease detection.

Feature Extraction Module A histogram is an essential feature used in image analysis to capture the distribution of pixel intensities within an image. In the case of unhealthy plant images, the color of the leaves gradually changes from green to yellowish and then to brown, indicating the presence of disease or stress. The classification of such images heavily relies on color-based features, as these variations in hue and intensity serve as key indicators of plant health [18]. To effectively analyze these color changes, the foreground region of the image, which was segmented in the previous step, is used for extracting histogram feature values. The histogram represents the frequency distribution of pixel intensities in the segmented region, allowing for a quantitative assessment of color variations. In this model, sixteen bins are utilized for histogram calculation, ensuring a detailed and structured representation of color variations across different intensity levels. This approach enhances the accuracy of classification by providing meaningful color-based information for disease detection.



Input: SI, PI

Output: PIF // Plant Image Feature

1. Loop 1:m // Number of pixels in x axis's
2. Loop 1:n // Number of pixels in y axis's
3. If SI Not Equal to 0 // Selected Segment of Image
4. Hist_image • PI[m,n];
5. EndIf
6. EndLoop
7. EndLoop
8. PIF • Estimate_Histogram(Hist_image, b)

LSTM Model The learning process of an LSTM model for plant leaf health prediction begins with data preprocessing, where time-series data from environmental sensors or image features is normalized and structured into sequences. During forward propagation, the LSTM processes input sequences using forget, input, and output gates to retain relevant temporal information. The model then computes a loss function (e.g., categorical cross-entropy for classification) to measure prediction accuracy. Using backpropagation through time (BPTT), gradients are computed, and weights are updated using optimizers like Adam. Gradient clipping helps prevent exploding gradients. Over multiple epochs, the model learns patterns and improves accuracy. Evaluation metrics such as accuracy, precision, and confusion matrix assess performance. Once trained, the model predicts leaf health status from new data sequences.

Proposed CLHPM Algorithm

Input: IPD

Output: PHP-LSTM // Plant Health Prediction by LSTM

1. Loop 1:p // IPD have p number of images
2. PI • Plant_Image_Preprocessing(IPD)
3. SI • Active_Contour_Segmentation(PI)
4. PIF • Extract_Features(SI, PI)
5. Train_Vector • H
6. EndLoop
7. Train_Data • [Train_Vector, DO] // DO: Desired Output
8. LSTM • Initialize(Train_Data)
9. PHP-LSTM • Train(LSTM, Train_Data)

Above algorithms takes plant dataset as input and extract features form image to collect training vector and train the LSTM model.

IV. EXPERIMENT AND RESULT

The experiment was conducted on a computer system equipped with an Intel i3 processor and a hardware configuration consisting of 4 gigabytes (GB) of RAM. This setup provided the necessary computational resources for processing and analyzing the images used in the study. The MATLAB software environment was utilized for the development and implementation of the proposed model, as it offers a wide range of tools and functionalities for image processing, machine learning, and data analysis. The experimental work was carried out using a real dataset consisting of crop leaf images, ensuring that the results accurately reflect practical agricultural scenarios. The dataset contained various samples of diseased and healthy leaves, allowing for a comprehensive evaluation of the model's



performance. A detailed description of the dataset, including the number of images, types of plant species, and disease categories, is presented in Table 1 [14].

Table 1 Sample Image

Healthy Image	
Early Blight	
Late Blight	

Results

Table 2 Precision value of leaf class prediction models.

Testing Image Set	Previous Model [13]	PLHPM
25x3	0.8163	0.9388
50x3	0.76	0.92
75x3	0.9467	0.96
100x3	0.8788	0.9495
120x3	0.912	0.952

Table 2 shows plant leaf class prediction precision values for different testing image sets. It was found that use of active contour for image segmentation has increases the work efficiency of learning. Proposed PLHPM model has increases the precision value by 8.6117% as compared to [13].



Table 3 Recall value of leaf class prediction models.

Testing Image Set	Previous Model [13]	PLHPM
25x3	1	1
50x3	1	1
75x3	1	1
100x3	1	1
120x3	1	1

Table 3 shows that both model has successfully identify all set of infected images in different testing dataset sizes. As all images are correctly classify hence recall values are 1.

Table 4 F-Measure value of leaf class prediction models.

Testing Image Set	Previous Model [13]	PLHPM
25x3	0.8636	0.9583
50x3	0.8989	0.9684
75x3	0.9726	0.9796
100x3	0.935	0.974
120x3	0.954	0.9754

Table 4 shows f-measure values of plant leaf health class prediction of comparing models. Proposed PLHPM modle has uses the histogram feature form the segmented image and it was found that most of noise were remain in other segmented region. This increases the learning of the trained model.

Table 5 Accuracy value of leaf class prediction models.

Testing Image Set	Previous Model [13]	PLHPM
25x3	92	97.33
50x3	93.88	97.96
75x3	98.23	98.67
100x3	96.15	98.4
120x3	97.03	98.38

Table 5 shows plant leaf class prediction accuracy percentage values for different testing image sets. It was found that use of active contour for image segmentation has increases the work accuracy efficiency of learning. Proposed PLHPM model has increases the precision value by 2.69% as compared to [13].

Table 6 Execution Time value of leaf class prediction models.

Testing Image Set	Previous Model [13]	PLHPM
25x3	104.797	13.2308
50x3	204.9166	20.7261
75x3	504.537	73.501
100x3	436.157	43.1298
120x3	523.1209	70.0782



Execution time for the image class prediction models were shown in table 6. It was found that use of histogram feature extraction time was less hence prediction time of model is also less in all set of experimental data.

V. CONCLUSION

Plant leaf is very help to identify the health and its production, hence analysis and learning more about it is very useful. This paper has proposed model that uses active contour based segmentation of image. Segmented region of image help to extract feature from effective portion of image. Histogram feature from the image train the LSTM model for learning. It was found that use of histogram reduces the prediction time and increases the detection accuracy. Experiment was done on potato plant leaves taken from local agriculture land. Result shows that proposed PLHPM has increased the precision by 8.61% and accuracy by 2.69% as compared to existing model in [13]. In future scholars can develop similar for tree leaf health detection such as mango, apple, etc.

REFERENCES

1. Jones, R.A. Global plant virus disease pandemics and epidemics. *Plants* 2021, 10, 233.
2. Liliane, T.N.; Charles, M.S. Factors affecting yield of crops. *Agron.-Clim. Chang. Food Secur.* 2020, 9.
3. Savary, S.; Willocquet, L. Modeling the impact of crop diseases on global food security. *Annu. Rev. Phytopathol.* 2020, 58, 313–341.
4. United Nations Children's Fund (UNICEF). *The State of Food Security and Nutrition in the World 2020: Transforming Food Systems for Affordable Healthy Diets.*
5. J. K. D. Treceñe, "Brassicaceae leaf disease detection using image segmentation technique", *Proc. IEEE EUROCON 19th Int. Conf. Smart Technol.*, pp. 30-34, Jul. 2021.
6. M. S. Selvi, S. J. Rani, A. Nisha and K. Sneha, "Tomato leaf disease detection using image processing", *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, pp. 367-371, Jul. 2022.
7. H. Nagar and R. S. Sharma, "Pest detection on leaf using image processing", *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, pp. 1-5, Jan. 2021.
8. E. Moupojou, F. Retraint, H. Tapamo, M. Nkenlifack, C. Kacpah and A. Tagne, "Segment Anything Model and Fully Convolutional Data Description for Plant Multi-Disease Detection on Field Images," in *IEEE Access*, vol. 12, pp. 102592-102605, 2024,
9. Pawar, M. Singh, S. Jadhav, V. Kumbhar, T. Singh and S. K. Shah, "Different crop leaf disease detection using convolutional neural network", *Proc. Int. Conf. Appl. Mach. Intell. Data Anal. (ICAMIDA)*, pp. 966-979, 2023.
10. R. Aithal, R. K. Anil and D. Angmo, "Rural tourism in India: Case studies of resilience during crisis", *Worldwide Hospitality Tourism Themes*, vol. 15, no. 1, pp. 63-73, Jan. 2023.
11. P. Chaitanya Reddy, R. M. S. Chandra, P. Vadiraj, M. Ayyappa Reddy, T. R. Mahesh and G. Sindhu Madhuri, "Detection of plant leaf-based diseases using machine learning approach", *Proc. IEEE Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS)*, pp. 1-4, Dec. 2021.
12. R. Kumar et al., "Hybrid Approach of Cotton Disease Detection for Enhanced Crop Health and Yield," in *IEEE Access*, vol. 12, pp. 132495-132507, 2024.
13. H. I. Peyal et al., "Plant Disease Classifier: Detection of Dual-Crop Diseases Using Lightweight 2D CNN Architecture," in *IEEE Access*, vol. 11, pp. 110627-110643, 2023
14. <https://www.kaggle.com/datasets/vipooooool/new-plant-diseases-dataset>