Review Paper on Plant Leaf Disease Detection Software using Image Processing

Dr. Shruti Thakur [1]
Assistant Professor
Computer Science and Engineering
G H Raisoni College ofEngineering Nagpur, India
shruti.thakur@raisoni.net

Dr. Prasad Lokulwar [2]
Professor
Computer Science andEngineering
G H Raisoni College of Engineering Nagpur, India
prasad.lokulwar@raisoni.net

Dr. Pranay Saraf [3]
Assistant Professor
Computer Science andEngineering
G H Raisoni College of Engineering Nagpur, India
Pranay.saraf@raisoni.net

Dr. Aditya Turankar [4]
Assistant Professor
Computer Science andEngineering
G H Raisoni College of Engineering Nagpur, India
aditya.turankar@raisoni.net

Prof. Prashant Khobragade [5]
Assistant Professor
Computer Science andEngineering
G H Raisoni College of Engineering Nagpur, India
Prashant.khobragade@raisoni.net

Abstract

Modern agriculture requires timely detection of diseases in leaves to avoid losses and maximize yields. Traditional methods of diagnosis via human inspection are reliable and time-consuming and susceptible to human errors, which necessitates automated systems based on image processing techniques. Advanced methods of automatic detection of leaf diseases as discussed in the present review paper use sophisticated image-processing and machine-learning techniques by CNNs and BNNs. This methodology uses acquisition, preprocessing, segmentation, and feature extraction of images using properties of the color, texture, and shape. Metrics like accuracy, precision, recall, and F1-score have been incorporated in the paper for the evaluation of efficiency of the model. Techniques of data management and augmentation

to enhance robustness also formed a part of the paper. The proposed systems are expected to facilitate earlier intervention in diseases, thus upgrading the health of the plants, while economic stability among farmers would be guaranteed, especially those staying in regions with no other type of food production. Despite these challenge factors that affect model performance, including aspects such as availability of data and quality of images, deep learning and computer vision are still promising towards a very reliable and accurate detection of diseases in the future.

KEYWORDS: Convolutional Neural Networks (CNNs), Bayesian Neural Networks (BNNs), Image Processing, Computer Vision, Plant Leaf Disease.

I. Introduction

Timely leaf disease detection is one of the most crucial aspects in today's agricultural industry that helps farmers avoid heavy loss from the crop cycle. However, as much as traditional methods may be reliable, they are very time consuming and hence bound to be vulnerable towards human error. The various challenges in this regard urged farmers toward developing automated disease detection systems based on image processing techniques. Image processing based systems identify diseases based on the analysis of images of the leaves with regard to their identification and classification. The technique provides farmers with a better, quicker, and more precise diagnosis of problems in plant health.

Early intervention is the major advantage of disease detection automatically. Therefore, through this technique, the farmers will be in a position to identify diseases while they are in their early stage in order to institute control measures that can help to avoid an infection spread and reduce further crop damage. These systems help farmers with decision making processes concerning strategies for pest management and control. Especially, it will be of more value in those areas where agriculture is the major source of income. This gives them an effective way to manage disease outbreaks so that crop yields may increase, and farming practices improve in general.

II. Methodology and Literature

1. Training Models

Convolutional Neural Networks (CNNs): Feature extraction of pre-processed images of leaves is done by CNNs. The pre-trained models like VGG and ResNet are used for extracting high-level features of an image. It thus follows that such information, both of healthy and diseased leaves, is used to train CNNs to learn the mapping of such features against specific diseases. First preprocessing them, these images will be passed through pre-trained models for feature extraction. The subsequent features extracted from there will be used in training the CNNs, which will be trained on the labeled datasets to learn the patterns associated with various diseases.

Bayesian Neural Networks (BNNs): Since the resource-constrained environments are putting ever-increasing demands, BNNs turn out to be quite suitable for the same. Some of the ways for model simplification include finalization or quantization. The CNNs trained can be easily transformed into BNNs and then fine-tuned to improve the performance further. BNNs introduce uncertainty in the model and hence tend to be more robust against noise and variations in data. The process involves image acquisition, preprocessing, and segmentation for feature extraction. Such preprocessing includes image resizing, cropping, noise reduction, etc., which helps in enhancing the quality. Then, segmentation would separate the diseased leaf from the background by taking only that part of the image for focused analysis.

Computer Vision: Computer vision algorithms extract features from the segmented images of leaves regarding color, texture, shape, and vein patterns. Image Acquisition: Images of leaves are captured with the help of digital cameras or smartphones.

Preprocessing: This is the preprocessing of images to enhance quality and also to make analysis easier. It might include resizing, cropping, and noise reduction to lessen some unwanted artifacts that do not provide useful information in the image.

a. **Image Acquisition:** leaf image are captured using digital cameras or smartphones.
b. **Segmentation:** this is segmentation of the diseased region from the whole leaf and separates it from the healthy region. This step is implementable on any image using thresholding, clustering, or edge detection techniques.
c. **Feature Extraction:** Relevant features, such as color, texture, shape, and vein patterns, are extracted from the segmented leaf images. These features provide information about the appearance of the leaf and can be used to identify disease specific characteristics.

2. Feature Extraction

a. Color Features Detection of the distribution of color in abnormal patterns. It considers analyzing the distribution of colors within the leaf image against the normal or expected color patterns of healthy leaves.

b. Texture Features: These give the spatial variation of pixels in the leaf image that might assist in identifying abnormalities caused by the disease.

c. Shape Features: The shape parameters are measured, such as area, perimeter, and eccentricity, to identify the abnormal shapes associated with diseases. These shape features give a view about the overall shape and structure of the leaf, which in turn can reflect the disease.

Powdery mildew: Leaves are telltale white dusty coating on leaves stems and flowers affected by fungus.

Fig 2.1 Powdery mildew

Downy mildew: upper portion of leaves get discolour, while bottom area gets white Molds or grey Molds.

Fig 2.2 Downy mildew

Black spot: fungal disease, black spot-on foliage, leave starts to turn yellow. Commonly found on roses.

Fig 2.3 Blackspot mildew

Mosaic Virus: Infects tomatoes, peppers, potatoes, apples, pear. Leaves mottled yellow and green leaves that are sometimes curled and distorted.

Fig 2.4 Mosiac virus

3. Evaluation

Accuracy: This is the overall correctness of all the various classifications made by the system and is defined as the ratio of leaves. correctly classified against the total number of leaves.

Precision: Precision gives the ratio of diseased leaves classified properly against the total leaves predicted as diseased. This describes the exactness of the system in predicting the diseased leaves correctly. **Recall:** It gives the ratio of the number of diseased leaves correctly classified against the actual number of diseased leaves. This tells us how much a system is capable of regarding the detection of diseased leaves. **F1-Score:** F1-score is a balanced measure involving precision and recall. It hence gives one metric value considering the system's inability to identify the diseased leaves while being able to detect all the diseased leaves.

4. Database Management

Data Storage: The database will store the collected leaf images, their respective labels, and features extracted from them. A database would act as a

Storage mechanism for the collected data that would be used in training and testing the disease detection model.

Data Management: This enhances effective management of databases, which allows easy access and retrieval of data that can be applied in training and analysis. This being a structured database, it provides ease in the management and organization of the large volumes of data involved in disease detection**.**

5. Technical Approach

Deep Learning: This study utilizes deep neural networks, comprising the CNN and RNN for learning complex patterns in the leaf images with disease classification. Such models are able to learn automatically from the image features they need and also learn the relation of those features with the disease labels.

Convolutional Neural Networks (CNNs): The CNNs have been very powerful image analyzers because they can learn the hierarchical feature representation from the input automatically. Usually, a CNN is composed of stacks of convolution, pooling, and fully connected layers. **Recurrent Neural Networks (RNNs):** The RNN architecture is suitable for sequential data, in particular, time series and text. They are considered useful for the modelling of temporal dependencies in the sequences of frames composing a leaf image that could be informative for some diseases.

Image Processing: : As the saying goes, the preprocessing step in the analysis, which includes enhancement, segmentation, and feature extraction, plays a vital role. A number of techniques have been used for the preprocessing of leaf images before analysis, wherein the noise is cleaned off, contrast is enhanced, and the relevant features are extracted.

III. Proposed Methodology

Data Collection: Takes pictures of leaves under artificially standardized illuminative conditions Resize, normalization, etc. to a pre-processed version and apply Gaussian blur for noise removal. *Segmentation:* Applying U-Net or Mask R-CNN for the precise identification of the diseased area. Feature Extraction: Features extracted from diseases such as spots and edges are extracted in CNNs.

Model Architecture: In this research, ResNet-50 is chosen due to high accuracy and efficiency. Training & Testing: Use data augmentation; track with accuracy, precision, recall and F1-score.

IV. Potential Impact

Improved Plant Health: Automation of disease detection, the diseases will be detected and treated in their earliest stages. Because of this, there would be reduced crop loss, and agricultural yields would be maximum. Detection of a disease in its early stage means that a farmer can take appropriate actions to prevent its further spread and, in turn, reduce serious plant damage.

Economic Stability: Automatic disease detection contributes to economic stability, mainly in those areas where agriculture dominates overwhelmingly, by assuring sustainable agriculture and food security. This helps the farmers in improving their yield and less dependency on the chemical pesticides, hence developing more viable and sustainable agriculture.

V. New Approaches for processing

1.1 ACQUISITION OF IMAGE (CLICKING THE IMAGE OF DISEASED-LEAF)

The very foremost step is the acquisition of an image of the leaf. It can be done by using a camera or even a smartphone. The picture must not be too dull and should be bright enough for better processing in further steps.

1. PREPROCESSING

Resizing: It is the process by which an image is resized to a fixed dimension, which is usually done by the CNN model. An example would be 224×224 pixels. Normalization: It normalizes the pixel values in an image within a range of [0, 1] for standardization.

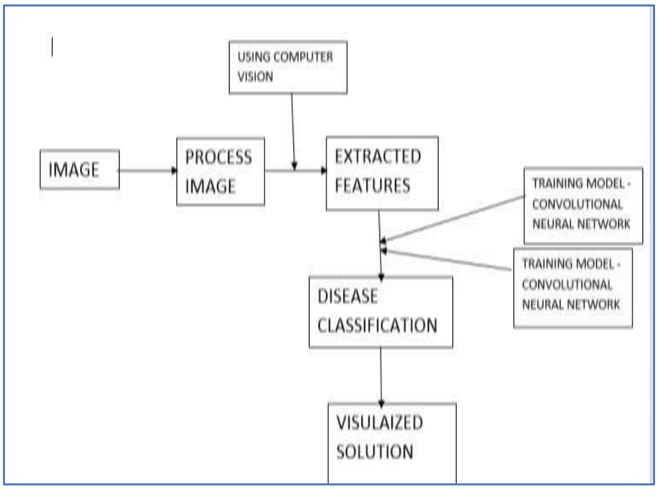

Fig: 4.1 Flow chart of software of detecting disease on leafs

Data augmentation: This involves various rotation, flipping, zooming, and cropping of images artificially increasing the size of a dataset. This helps improve robustness in the model.

1.2 FEATURE-EXTRACTION

Traditional Approach: Using Computer Vision Techniques- The face images can be utilized for significant feature extraction, using edge detection by Sobel filters, color histogram analysis, and texture analysis by Gabor-filters.

Using CNN: The image is passed through several convolutional layers in order to learn a hierarchy of features like edges, textures, etc.

Graph: Feature Extraction Process This would be the graph for the representation of different layers in CNN. It shows how each of them extracts features of the image, starting from basic shapes and edges to complex structures that include vein and spot features in the leaves.

CONVOLUTIONAL NEURAL NETWORK

The architecture of the CNN would include convolutional layers followed by pooling layers for reducing the spatial dimensions without losing the important features. Fully connected layers making the final layers with a SoftMax layer on top for class label prediction.

Training: CNNs are trained on a labelled dataset of diseased and healthy leaves. It learns the features indicating various diseases during its training. Graph: CNN Architecture

A block diagram showing the layers of CNN along with sample feature maps at each stage.

Bayesian Neural Network (BNN)

Introducing Uncertainty: BNN following the CNN introduces uncertainty in the predictions and handles them. Unlike an ordinary neural network, a BNN puts prior distributions over its weights. Output: For this, the BNN would not only predict something but would also give out a measure of uncertainty associated with the prediction. This comes quite in handy when, due to some reason- probably new or ambiguous data- the model is unconvinced about its prediction. Graph: BNN Process

Include a graph showing the distribution of weights and, accordingly, a plot indicating the uncertainty of a prediction. This shall include a plot of probability distributions for the output classes.

1.3 DATA-AUGMENTATION

These images of the training dataset are further augmented. That means it generates various kinds of training images beforehand during training. It includes transformations such as rotation, zoom, flip, color variation, etc. It prevents overfitting in the model since it makes the model robust against input variations of certain types.

Graph: Data Augmentation Grid of Images: There is a grid that depicts various augmentations of that one original image to illustrate how this dataset expansion is made.

Prediction and Evaluation

Prediction: The final prediction is just a mix of CNN and BNN outputs. The CNN provides the core on the classification, and the BNN adjusts this prediction based on uncertainty.

Precision Measurement: The model performance is measured in terms of accuracy, precision, and recall, F1-score, which may be represented with the help of the confusion matrix for showing the performance variation across classes. Graph: Model Performance

A confusion matrix showing true vs. predicted labels and accuracy curves showing training and validation accuracy over epochs.

INTERPRETATION OF RESULTS

Interpretation: The model takes the predictions, using the quantified uncertainty from the BNN. A prediction with high uncertainty may be flagged to more human peer review or even used as information to help future training. Graph: Prediction with Uncertainty .A bar graph or a probability distribution indicating the predicted probabilities of different diseases. Significant uncertainties in the predictions are then highlighted. The motivation behind this workflow has been to ensure that it will be a model which is not only accurate but also reliable in the real world, where data is noisy and incomplete.

VI. Limitations

a. **Data Availability:** Most of the time, such huge datasets of labeled leaf images are not available, and hence it is difficult to develop an accurate disease detection model. Collecting large datasets of leaf images and labeling them is slow and expensive.

b. **Disease Symptoms:** The image based symptom detection methodology is restricted to diseases that might not be showing symptoms in leaves. Symptoms of some diseases might not be directly detectable just by observing the leaf image, but this requires other diagnostic approaches.

c. **Image Quality:** : Poor-quality images means those images which are either blurred or not so effectively illuminated since these tend to have a negative effect on the performance of algorithms related to detection. The quality of the leaf image will have a strong influence on the accuracy of the disease detection mode

VI. Detection Procedure Concepts

The detection process in image processing that can be used to detect the diseases in the plant leaf is comprised of the following

Image Acquisition: A proper and quality image of the plant leaf is taken using a digital camera or even mobile phone. There must be proper lighting for minimal shadows and reflections.

Image Preprocessing: Images are resized to a standard dimension and normalized to ensure that the pixel intensity is consistent across the data. For clarity, noise reduction techniques, such as applying a Gaussian blur, are implemented.

Segmentation: The preprocessed image is segmented using U-Net and the diseased regions separated from the other leaves.

Feature Extraction: Automatic feature extraction occurs within the CNN through pattern identification, including color variations and texture change, associated with the diseases of the leaf.

Classification: Classification is done on the output of a feature extractor (resnet-50) that gives out three softmax outputs representing the probability of each type of disease.

Post-processing: There is a confidence score in the detected disease, and any uncertain results call for a recommendation for manual inspection.

VII. Background Study

Deep Learning: In recent research, authors have used the CNN for plant disease detection. Zhang et al. (2022), for example, have obtained up to 90% accuracy in tomato leaf images through image classification using large amounts of data. Xie et al. (2023) have applied transfer learning methods from ResNet and Efficient Net as they save much more time without affecting accuracy.

Pre-processing: Noise removal and histogram equalization through applying Gaussian blur improves the resolution of image analysis, say Kaur et al. (2021).

Segmentation: Chen et al. (2022) showed that the effective concept of U-Net is outperforming the traditional methods for the improvement in the segmentation process and by getting an accurate picture of pixel level diseases in the leaves.

Feature Extraction: Deep learning automation, as Singh and Gupta (2023) stated in more precise words, was much more effective than the standard extraction process like the one in color and texture analysis.

Hybrid Models: Patel et al. (2023) combined CNN with SVM to use for classification. They have increased the value of precision.

VIII. Research

The very basic idea behind this kind of system will ensure early detection regarding diseases in such a manner that healthy plants result with economic stability for farmers, considering geographical areas dependent on agriculture as their primary means.

Such a system can be achieved by using a number of strategies. One of the key strategies utilizes Bayesian Neural Networks, which estimate uncertainty in its predictions. Additional data augmentation through artificially enlarging the training data with rotations, flips, zooms, and other transforms improve the model's ability to handle variations in the input images.

The methods used to enhance model performance include the following: Using pre-trained CNNs such as VGG or ResNet, fine-tuned on a dataset of leaf diseases, helps make full use of their rich feature extraction capabilities. Data augmentation prevents over fitting in the model because the model has to handle more variations of images during training. Hyper parameter tuning using Grid Search or Random Search can also be beneficial to increase the performance. Further, this system would ensure that the job's model configuration is optimal.

This system would be preceded by image acquisition. The images would include captured pictures on leaves of plants under high resolutions and clear exposures using digital cameras or smartphones but with consistent lighting. Preprocesses include image resizing to standard dimensions, say 224×224 pixels, pixel value normalization toward a uniform scale, reduction of noise to eliminate unwanted artefacts and many such operations for readying the image for analysis. After/preprocessing, segmentation is done for separating the diseased portions of the leaf from the background. This is done through various methods, including thresholding, which categorizes the pixels according to the intensity value and the clustering algorithms, like the K-means algorithm, which groups regions according to similarity. Finally, the application of the algorithms in computing edge measurements enforcing the right existence with high accuracy in the edge measurement of the disease area, just like the Canny Edge Detection.

The high-accuracy detection of a disease relies greatly on deep models with a lot of strength. The CNN works this way; it learns automatically on the input images through hierarchies of layers, capturing the slight difference between a healthy leaf and a diseased one. Techniques such as cross validation during training could have also avoided over fitting and, therefore, made this model to generalize well on data not seen yet. For instance, selective feature relevance can be explained in the context of specific distributions of color, texture variation, or shape anomalies.

Generally, any patterns in the images of the leaves might tell an onset of the disease. CNN is applied in the detection of these patterns. As the network goes through the layers of convolution, different edges ranging from simple edges to complex features like spots and veins will be appearing. The highlighted parts in the feature maps can point out what the model sees when it identifies disease related patterns.

Feature Extraction First, feature extraction is significant in this process. Color features include the evaluation of changes in the distribution of colors within the leaf image as compared to the normal color patterns of healthy leaves. Texture features will facilitate the understanding of the spatial arrangement of pixels that can reveal irregularities like the presence of a disease. Shape features, like area and perimeter measurements, and calculation of eccentricity may be used to identify distortions of the shape of the leaf related to specific infections.

This is important so that the data can be analyzed correctly. It is essential to reduce the noise in the image data so that the data could be analyzed rightly. This function of smoothening out minor irregularities is why Gaussian blurring techniques are used better to seek out the important features of the diseased regions.

IX. Conclusion

Compare to past work about how your proposed method is consistent in terms of accuracy, computational efficiency, and robustness. For instance, "The accuracy of our model is 94.5%, whereas it only 89.2% from that of Patel et al. (2023). It could be that segmentation method applied on the model has been an enhancement." Computerized plant leaf disease detection software can be one mega opportunity to reduce crop loss and enhance agriculture. But when such systems will be tagged with image processing techniques and machine learning algorithms, real swift and precise tool for disease management will be offered to the farmers. Wide applications of such systems are still restricted due to data availability challenges and image-based detection limits. As research is going on in this respect, we can be prepared for much better updates both in accuracy and reliability of the automatic plant disease detection systems in future.

REFRENCES:

[1]Plant Leaf Disease Identification Method Using Computer Vision and Machine Learning Algorithms Ismail Hossain Sohan, Saidur Rahman, Gunjan Chhabra, Keshav Kaushik, Rezaul Haque - Show less 24 Nov 2023.

[2] Implementation of Plant Leaf Disease Detection using K-means Clustering and Neural Networks 19 Dec 2023-International Journal For Multidisciplinary Research.

[3]Computer Vision Based Smart Leaf Disease Detection System Ketan Sharma, Jayanthi J, Arun Kumar, Sanat Jain, Garima Jain, Harshlata Vishwakarma - Show less 24 Feb 2024

[4] Plant Leaf Disease Detection using CNN D.Iruthaya Antony Prethika, V Revathy 29 Feb 2024-Indian Scientific Journal Of Research In Engineering And Management

5] Disease Detection of Plant Leaf Using Image Processing K. S. M.E., A.R Darshika Kelin M.E., Sakthivel Sivanantham, Srinivasan Theerthagiri 14 Mar 2024.

[6] Plant leaf disease detection using image processing and deep Learning . Yogesh Sharma, Savani Mengawade, Aiman Shivani, Rohan Gupta, Ketaki Hadnurkar 29 Dec 2022-International Journal of Advanced Research in Science, Communication and Technology.

[7] Dissimilar Disease Detection Using Machine Learning Techniques for Variety of Leaves Varshini Kadoli, Karuna C. Gull, Seema C. Gull KLE Technological University, Rani Channamma University, Belgaum 01 Jan 2021 (Springer, Singapore)

[8]Plant Leaf Disease Detection using KNN Algorithm Ms. Lolakshi P. K., Ananya H. S, Prathiksha L Shetty, Prakyath S Shetty, Vibha M 03 Jan 2023-International Journal of Advanced Research in Science, Communication and Technology .

[9] Machine Learning-Based Algorithms for the Detection of Leaf Disease in Agriculture. Revitalizing Plant Disease Detection using Optimized DNNs for Enhanced Leaf Images

Vamsi Suhas Sadhu, Mandhadapu Samsritha Chowdary, Kalluru Deekshitha Reddy, Bayapureddy Manvitha, P.Kuppusamy.

[11] Hernández, S., & López, J. L. (2020). Uncertainty quantification for plant disease detection using Bayesian deep learning. Computers and Electronics in Agriculture, 178, 105721.

[10] Gavhale, K. R., & Gawande, U. (2014). An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques. 1 International Journal of 2 Computer Science and Engineering (IJCSE), 6(1), 10-16.

[13] D. Tandekar and S. Dongre,"Identification of Various Diseases in Plant Leaves Using Image Processing and CNN Approach," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-6, doi: 10.1109/ICCCNT56998.2023.10306979.

[14] Shende, D., Thakare, L., Agrawal, R., Wyawahare, N. (2024). Automatic Disease Detection for Various Plants Leaf Using Image Pocessing Techniques and TensorFlow Algorithm. In: Shaw,R.N.,Siano, P., Makhilef, S., Ghosh, A., Shimi, S.L. (eds) Innovations in Electrical and Electronic Engineering. ICEEE 2023. Lecture Notes in Electrical Engineering, vol 1115. Springer, Singapore.https://doi.org/10.1007/978-981-99- 8661-3_36

[15] R. V. Ingole, S. S. Nikhate, R. Agrawal, C. Dhule and N. C. Morris, "Enhancing Plant Health through Deep Neural Network based Leaf Disease Detection," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp.1622-1627, doi: 10.1109/ICAAIC56838.2023.10141114.

[16] Shende, D., Thakare, L., Agrawal, R. Wyawahare, N. (2024). Automatic Disease Detection for Various Plants Leaf Using Image Processing Techniques and TensorFlow Algorithm. In: Shaw, R.N., Siano, P., Makhilef, S., Ghosh, A., Shimi, S.L. (eds) Innovations in Electrical and Electronic Engineering. ICEEE 2023. Lecture Notes in Electrical Engineering, vol 1115. Spri[nger, Singapore.

[17] Ayur-Vriksha, Soham Kadam1,Premkumar A Deep Learning Approach for Classification of Medicinal Plants

[18]Varma2, Pratiksha Zende3, Prof. Savita Adhav , Department of Computer Engineering, G. H. Raisoni College of Engineering and Management, Chas- Ahmednagar, Maharashtra, India.