A Review on Various Approaches used in Leaf Disease Detection

Raushan Kumar
Computer Science & Engineering
SAM College of Engineering & Technology
Bhopal, Madhya Pradesh, India
raushan.sitaram@gmail.com

Pradeep Kumar Pandey
Computer Science & Engineering
SAM College of Engineering & Technology
Bhopal, Madhya Pradesh, India
pradeep65656@gmail.com

Abstract

Global sustenance depends heavily on agricultural output, which artificial intelligence greatly contributes to increasing. The automated identification of leaf diseases is one such use. It might be difficult to diagnose leaf diseases with normal eyesight since sick leaves can seem very normal. These illnesses have the potential to significantly lower crop output quality in the absence of prompt and appropriate care. Early identification is therefore crucial for increasing productivity and guaranteeing proper treatment. As a cornerstone of machine learning (ML), deep learning (DL) is increasingly essential for the early detection and categorization of plant diseases. This paper looks at the latest developments in crop disease detection using convolutional neural networks (CNNs), transfer learning, and deep learning. The investigation of deep learning architectures, data sources, and different image processing methods used to manage imaging data comes first. With the recent use of various DL architectures and visualization tools, plant disease classification and symptom identification have become critical tasks. The paper also discusses a few open issues that must be answered in order to create automated systems for the practical identification of plant diseases in the field.

Keywords— *Leaf Disease, Machine Learning, Convolutional Neural Network, Alternaria Alternata, Bacterial Blight, Cercospora, Leaf Spot .*

I. INTRODUCTION

These days, agricultural land serves as much more than a food source—it is essential to the economy, especially in India, where agricultural output has a big influence on financial stability. Early detection of plant diseases is therefore essential. It can be quite beneficial to use automated illness detection techniques for this. Plant disease identification is now mostly dependent on specialists physically examining the plants, which necessitates a big team and ongoing monitoring and is expensive for large farms. Furthermore, in many nations, farmers may not have access to specialists or are not aware of these resources, which results in consultations being costly and time-consuming [1]. The suggested approach of monitoring vast agricultural fields works well in these circumstances. Not only is automated disease diagnosis more economical, but it is also simpler to identify signs on plant leaves. Plant disease detection via visual means is a labor-intensive, less accurate method that is only practical in certain situations. On the other hand, automatic detection is more accurate and takes less time and effort. Brown and yellow spots, early and late blight, and several bacterial, viral, and fungal infections are examples of common plant diseases. Image processing is a useful technique for illness detection since it is used to assess the afflicted area and identify color changes in the sick parts [2].

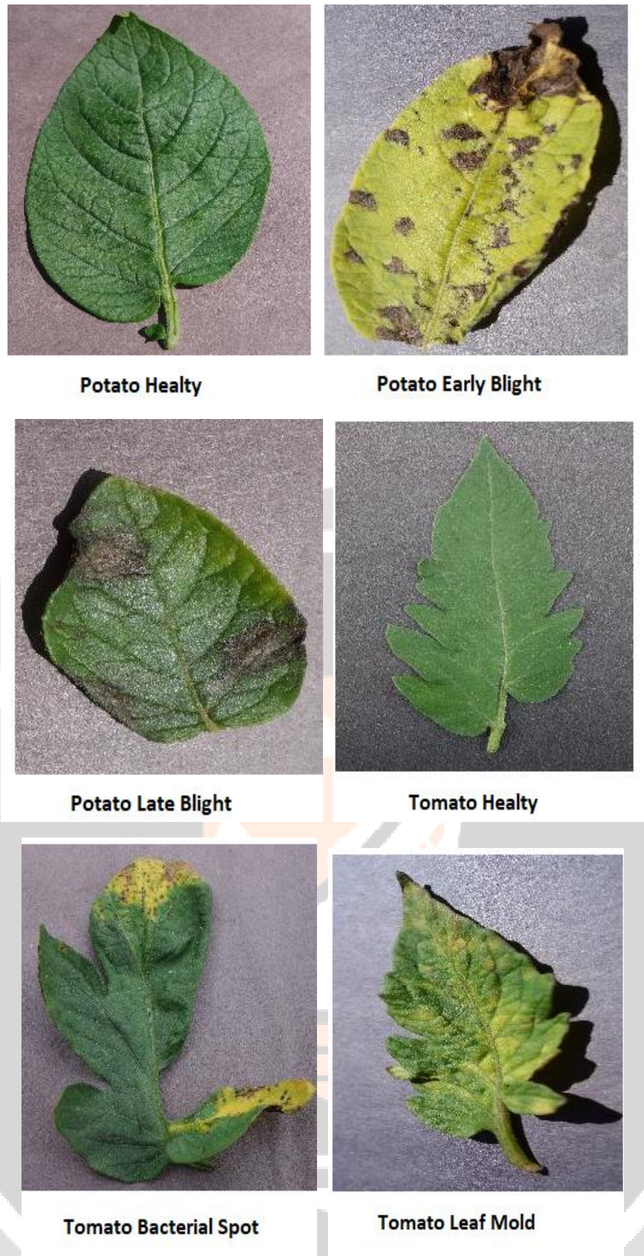


Fig. 1. Leaf Disease Example

Images of diseased tomato and potato leaves are displayed in Figure 1.

Imaging tests offer useful information that can assist anticipate diseases with a certain degree of accuracy, even if they cannot detect diseases on their own. Plant diseases seldom show symptoms in their early stages, thus prompt identification is essential for successful treatment. For the purpose of diagnosing leaf diseases, automatic diagnosis is quite helpful. CNN, RNN, ResNet, and other machine learning approaches have been the subject of much study. But for training and testing, these deep neural networks need large datasets and a lot of processing power [3]. In order to get effective outcomes with high accuracy, a network model should be lightweight and utilize modest intelligent filters in hidden layers. Using a classifier to differentiate between normal and aberrant cells and make judgments based on that distinction is an alternate method. Various classifiers can be used for classification to achieve better outcomes, including SVM, K-means clustering, and Naive Bayes. SVM is regarded as the best classifier among them because of its exceptional prediction ability to identify various cell patterns.

II. Related Works

Lesions have been successfully extracted from leaf photos by numerous researchers, with some false alarm rates and good accuracy. Jaskaran Singh et al. [4] proposed a study based on region-based segmentation and the KNN classifier.They came to the conclusion that diagnosing plant diseases is crucial to finding plant illnesses. They used the KNN classifier for illness prediction, the GLCM method for textural feature analysis, and k-means clustering for region-based segmentation in their work. The suggested model was put into practice using MATLAB, and the outcomes are shown in tables and figures. Eftekhar Hossain et al. [5] proposed a research study using the KNN classifier to detect diseases based on color and textures. Their method made use of KNN to recognize and categorize a number of illnesses that affect plant leaves, including blisters, bacterial blight, leaf spot, anthracnose, and Alternaria Alternata. The k-nearest neighbor classifier was utilized to identify the illness segment, and GLCM texture features were employed for classification. Plant disease identification with the KNN classifier-based segmentation was highly accurate, and the DSC, MSE, and SSIM metrics were used to assess the quantitative performance of their suggested algorithm.

Aamir Yousuf et al. [6] proposed a study based on an ensemble classifier, combining KNN and Random Forest classifiers to obtain results. But when it comes to classifiers, SVM is more advanced and better than the rest. SVM is the best classifier for identifying disorders associated with image processing since it can handle both linear and non-linear data with excellent prediction accuracy. Additionally, GLCM is utilized to extract the image's textural properties, which are then employed by classifiers to identify the condition. Ch. Usha Kumari et al. [7] proposed research based on K-means clustering and ANN. This study uses a neural network classifier to detect leaf illness, and K-means clustering is used for segmentation. For tomato and cotton illnesses, a number of characteristics are extracted, including Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, and Variance. Bacterial leaf spot, target spot, septoria leaf spot, and leaf form disease are among the infected leaves that were taken into consideration for the study. To identify and categorize the diseases, features from clusters 1 and 3 affected by the diseases are computed and given into the classifier. Nine out of the twenty cotton samples had the bacterial leaf spot diagnosis correctly, whereas one sample had the target spot classification incorrectly. Two samples were mistakenly identified as bacterial leaf spot, out of the eight samples that were classified as target spot. Abirami Devaraj et al. [8] proposed research based on an image processing approach. Feature extraction, classification, segmentation, preprocessing, and picture loading are all included in this method. The creation of an automated detection system that makes use of cutting-edge technology, such as image processing, helps farmers control diseases by helping them spot illnesses early on. They want to keep working to cover more ground in terms of illness detection.

Table I Models Comparison

Method	Finding
VGG-16	VGG16 has numerous weights and bi as parameters, making the model quit e large and heavy.
DenseNet	Excessive dense connections can com plicate the network and negatively im pact computational efficiency.
AlexNet	AlexNet is not a very deep model, whi ch limits its ability to scan for all feat ures, resulting in poor performance.
CNN	The conventional CNN model is inefficient in training and constru cting heavy networks, which directl y impacts execution time.
Ensemble Classifier	Ensemble classifiers are challenging t o interpret and involve high costs f or training, evaluation, and dep loyment of the model.

Table I lists the limits of some models that have been found to be quite accurate in identifying leaf diseases. It contrasts several techniques used in autonomous leaf disease diagnosis, emphasizing the problems and shortcomings found in each model. The chart highlights the areas in which these models fall short and highlights the necessity of creating the perfect method for precisely identifying leaf diseases. A better strategy that uses fewer or lighter filters during training is needed for such a system in order to achieve high precision with a low number of false alarms.

M. Bhange et al. [] developed a web-based tool for identifying fruit diseases by uploading images to the system. From the uploaded fruit photos, the tool extracts characteristics including color, morphology, and CCV (color coherence vector). The k-means approach is used for clustering, while support vector machines (SVM) are used to categorize fruit as either infected or not. This method identified pomegranate illnesses with an 82% accuracy rate. J.D. Pujari et al. [] explored various crop types including fruit crops, vegetable crops, cereal crops, and commercial crops for detecting fungal diseases on plant leaves. They employed different methods tailored to each crop type. V. Singh et al. [] focused on automating the detection and classification of plant diseases by integrating genetic algorithms as the image segmentation technique. For training and testing, they used a limited collection of photos from four different plant species: rose, banana, beans, and lemon. The color co-occurrence approach was used to extract features, combining texture and color information. For the purpose of classifying diseases, the Minimum Distance Criterion via k-means clustering and SVM classifier was used, yielding accuracy rates of 86.54% and 95.71%, respectively. The accuracy was increased to 93.63% by combining the genetic algorithm and the Minimum Distance Criterion classifier. E. Kiani et al. [] aimed to identify disease-infected leaves in a strawberry field under outdoor conditions using a fuzzy decision maker. They achieved an overall accuracy of 97% for detecting and segmenting plant diseases, with a processing time of 1.2 seconds. H. Ali et al. [] aimed to apply the ΔE color difference algorithm to distinguish disease-affected areas and used color histograms and textural features for disease classification, achieving an overall accuracy of 99.9%. They used a variety of classifiers, including classifiers for boosted trees, bagged trees, fine KNN, and cubic SVM. For RGB, HSV, and LBP features, the Bagged tree classifier fared better than the others, yielding accuracy rates of 99.5%, 100%, and 100%, respectively. Additionally doing well were the fine KNN, Cubic SVM, and Boosted tree classifiers, which had accuracies of 88.9%, 90.1%, and 50.90%, respectively. G. Saradhambal et al. [] proposed an approach for developing an automated system for plant disease detection. Their study concentrated on utilizing Otsu's classifier and the k-means clustering approach to forecast the affected leaf area. The method entailed taking the leaves' shape and textural characteristics and extracting them. Area, color axis length, eccentricity, solidity, perimeter, and homogeneity were examples of shape-oriented features, whereas contrast, correlation, energy, homogeneity, and mean were examples of texture-oriented features. In this study, a classifier based on neural networks was used to conduct classification.

R Chouksey et al. [] proposed a system which is based on Polynomial Support Vector Machine and Euclidean Distance Metric. The suggested method categorizes and automatically diagnoses leaf diseases. Leaf photos usually contain a variety of noises, therefore in order to correctly diagnose lesion sites, background information needs to be either masked off or eroded. Image pre-processing is essential for improving photos and identifying certain areas. The system uses support vector machines (SVM) to categorize and sort cells or clusters according to patterns. The system specifically uses polynomial SVM, a non-linear technique that can handle non-linear data. It is better to use a non-linear classifier to get accurate precision in disease classification because of the intricate structure of leaf pictures.

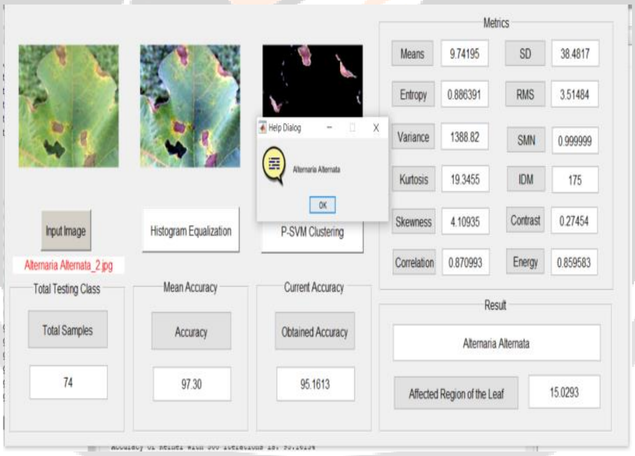


Fig. 2. Graphical User Interface

Figure 2 shows the graphical user interface of the proposed system, which includes steps such as histogram equalization and clustering. In the proposed system, the initial step involves retrieving the original image from the dataset and applying histogram equalization. Histogram equalization enhances the contrast and brightness of the image, thereby improving the visibility of lesions. This process aids in classifying the background, which may contain noise that can be reduced during edge detection.

$$P_n = \frac{number\ of\ Pixel\ Intensity\ n}{Total\ number\ of\ pixels}\ n = 0,1 \ldots .L - 1$$

Where P_n is the affected pixel value after histogram equalization. The histogram equalized image γ will be defined as

$$\gamma_{i,j} = floor\left((L-1)\sum_{n=0}^{f_{i,j}} P_n\right),$$

Where floor() can be considered as the nearest integer. It is equivalent to pixel intensity k;

$$k = floor\left((L-1)\sum_{n=0}^{f_{i,j}} P_n\right)$$

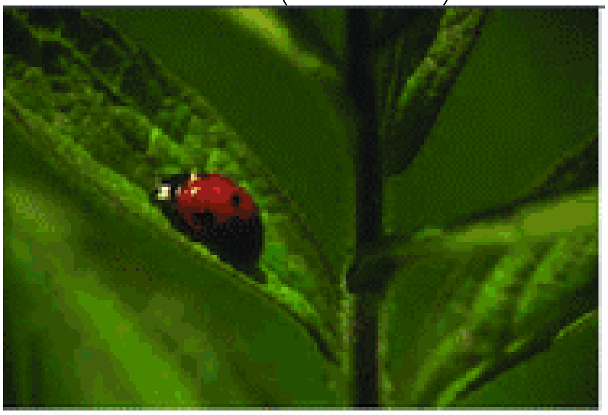


Fig. 3. Original Image

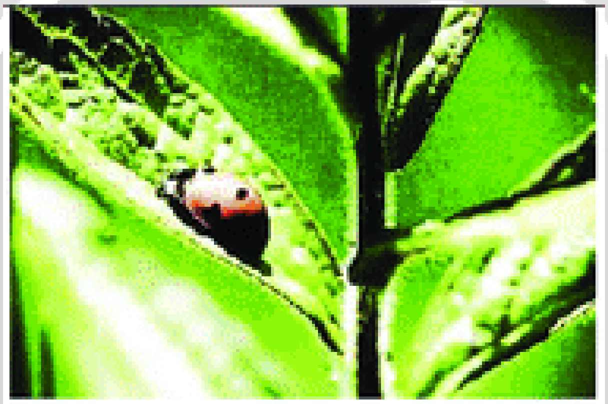


Fig. 4. Histogram Equalization affeccted Image

Figure 3 displays the original leaf image, while Figure 4 illustrates the image after histogram equalization, demonstrating enhanced visibility. This enhancement is crucial for improving the system's accuracy in detecting leaf diseases.

III. Experimental Result

True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) are the four metrics that are commonly used to assess experimental outcomes. When a picture falls into a disease category and the algorithm accurately classifies it as positive, it is indicated as a True Positive. When a picture does not fit into any disease category and the system accurately classifies it as negative (healthy), it is indicated as True Negative. When an image does not fit into any illness category but the algorithm mistakenly flags it as positive, this is known as a false positive. When a picture falls into a disease category but the system misclassifies it as negative (normal), this is known as a False Negative.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP}+\text{FN}} * 100\ \%$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP}+\text{TN}} * 100\ \%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} * 100\ \%$$

$$\text{Negative Prediction Rate} = \frac{\text{TN}}{\text{FN}+\text{TN}} * 100\ \%$$

$$\text{False Positive Rate} = \frac{\text{FP}}{\text{FP}+\text{TN}} * 100\ \%$$

$$\text{False Negative Rate} = \frac{\text{FN}}{\text{FN} + \text{TP}} * 100\ \%$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} * 100\ \%$$

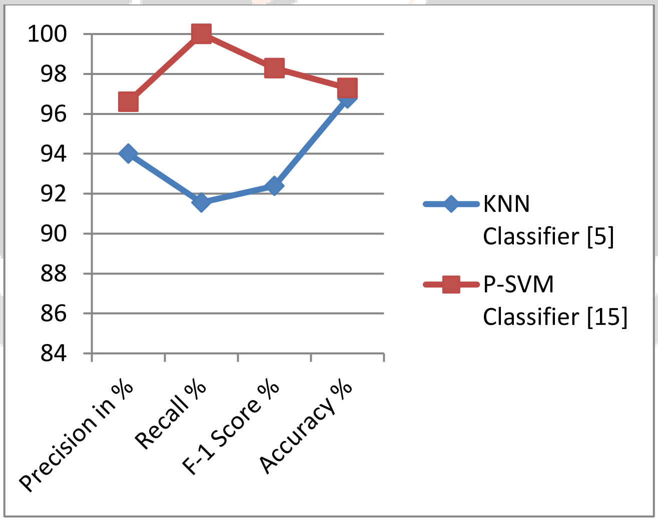
$$\text{F1} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} * 100\ \%$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} * 100\ \%$$

Table V Result Comparison

Methods	KNN Classifier [5]	P-SVM Classifier [15]
Precision in %	94.00	96.61
Recall in %	91.56	100
*F*1-Score in %	92.39	98.28
Accuracy in %	96.76	97.30

Graph I Result Comparison



IV. CONCLUSION & FUTURE SCOPE

Proposed system is based on Polynomial Support Vector Machine that classifies the abnormal and normal cells and on the basis of that decision can be make effectively with high accuracy. System is bit efficient in all parameters and achieved higher result as compare to the KNN classifier. Proposed classification technique is bit powerful to classify the normal and abnormal cells and take decision whether the image or disease belongs to respective category. The classification result is better than the previous model. System has been tested with kaggle benchmark and achieved better result. In future system can be tested with different benchmarks that contain several images or data. The accuracy can be enhanced by using certain modern pre-processing approaches.

REFERENCES

[1] Vijai Singh, A.K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques, Information Processing in Agriculture, Volume 4, Issue 1, 2017, Pages 41-49.

[2] Savita N. Ghaiwat, Parul Arora, Detection and classification of plant leaf diseases using image processing techniques: a review Int J Recent Adv Eng Technol, 2 (3) (2014), pp. 2347-2812.

[3] Sanjay B. Dhaygude, Nitin P. Kumbhar, Agricultural plant leaf disease detection using image processing, Int J Adv Res Electr Electron Instrum Eng, 2 (1) (2013)

[4] Singh, Jaskaran & Kaur, Harpreet. (2019). Plant Disease Detection Based on Region-Based Segmentation and KNN Classifier. 10.1007/978-3-030-00665-5_154.

[5] E. Hossain, M. F. Hossain and M. A. Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1-6, doi: 10.1109/ECACE.2019.8679247.

[6] Yousuf, Aamir & Khan, Ufaq. (2021). Ensemble Classifier for Plant Disease Detection. International Journal of Computer Science and Mobile Computing. 10. 14-22. 10.47760/ijcsmc.2021.v10i01.003.

[7] C. U. Kumari, S. Jeevan Prasad and G. Mounika, "Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 1095-1098, doi: 10.1109/ICCMC.2019.8819750.

[8] A. Devaraj, K. Rathan, S. Jaahnavi and K. Indira, "Identification of Plant Disease using Image Processing Technique," 2019 International Conference on Communication and Signal Processing (ICCSP), 2019, pp. 0749-0753, doi: 10.1109/ICCSP.2019.8698056.

[9] Bhange, M., Hingoliwala, H.A., 'Smart Farming: Pomegranate Disease Detection Using Image Processing', Second International Symposium on Computer Vision and the Internet, Volume 58, pp. 280-288, 2015

[10] Pujari, J.D., Yakkundimath, R., Byadgi, A.S., 'Image Processing Based Detection of Fungal Diseases In Plants', International Conference on Information and Communication Technologies, Volume 46, pp. 1802- 1808, 2015

[11] Singh, V., Misra, A.K., 'Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques', Information Processing in Agriculture, Volume 8, pp. 252-277, 2016

[12] Kiani, E., Mamedov, T., 'Identification of plant disease infection using soft-computing: Application to modern botany', 9th International Conference on Theory an d Application of Soft Computing, Computing with Words and Perception, Volume 120, pp. 893-900, 2017

[13] Ali, H., Lali, M.I., Nawaz, M.Z., Sharif, M., Saleem, B.A., 'Symptom based automated detection of citrus diseases using color histogram and textural descriptors', Computers and Electronics in Agriculture, Volume 138, pp. 92-104, 2017

[14] Saradhambal, G., Dhivya, R., Latha, S., Rajesh, R., 'Plant Disease Detection and its Solution using Image Classification', International Journal of Pure and Applied Mathematics, Volume 119, Issue 14, pp. 879-884, 2018

[15] Ritika Chouksey, Preety D. Swami."Leaf Disease Detection using Polynomial SVM and Euclidean Distance Metric ", Volume 10, Issue VII, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 1926-1935, ISSN : 2321-9653