e-ISSN 2320 –7876 www.ijfans.org
Vol.11,S Iss.1, 2022
Research Paper UGC CARE Listed (Group -I) Journal

A NOVEL MACHINE LEARNING METHOD FOR IDENTIFYING PLANT DISEASE TO INCREASE YIELD

Vani V G[1], Mohd. Shaikhul Ashraf[2], Atul Kumar[3], Satyanarayan Padhy[4], Ganesh N Yallappa[5], Laxmi Biban[6], Pallavi Singh[7]

[1]Department of Computer Science and Engineering, Government Engineering College, Kushalnagar Kodagu Karnataka, India

[2]Department of Botany, HKM Govt. Degree College Bandipora, Bandipora, Kashmir, (UT of Jammu and Kashmir), India

[3]Dr. D. Y. Patil B-School, Pune, India

[4]Department of Electronics and Communication Engineering, Centurion University of Technology and Management, Odisha, India

[5]Department of chemistry, Jain Institute of Technology, Davangere, Karnataka, India

[6]Department of Botany, Government College, Karnal, Haryana, India

[7]Department of Biotechnology, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

Corresponding mail id: vanivg@gmail.com

ABSTRACT

Agricultural production is improving as a result of recent technological and scientific advances. In agriculture, identifying plant leaf diseases and improving plant leaf quality are critical. There are various laboratory procedures for identifying illnesses, but they are costly and time-consuming for farmers. Polymerase chain response, minimised food cultivation, pest control, and hyper spectral technologies are among them. Agriculture may produce more if machine learning (ML) technology is applied to generate unique, improved procedures and a wide range of systematic models. Researchers focused on how ML algorithms are presently used to diagnose leaf diseases to improve the accuracy of their findings. Each strategy has some potential and focuses on the direction that ML applications travel as well as agricultural challenges. The identification of leaf diseases is addressed in this paper utilising the Support Vector Machine and Random techniques. To supply farmers with more yield in less time and money, performance indicators such as the Root Mean Square Error (RMSE), the Maximum signal - to - interference Relation, the disease-affected part of the foliage by means of the Euclidian technique, and the efficiency of the outcomes are compared.

Keywords: Machine learning, Plant disease, Euclidian Distance technique, Agriculture

1. INTRODUCTION

The agriculture sector's productivity in India is a crucial driver of the country's economic development. Agricultural output accounts for around 70% to 80% of Indian GDP [1]. Agriculture is the primary source of food for all governments, whether in developing or affluent nations. Agriculture is influenced by a variety of components that depend on water, air, temperature, and other environmental elements, such as bacteria, viruses, fungus, and non-biotics [2]. Crop damage may consequently result in a significant loss of production and, as a result, have an impact on the economy. The leaves of diverse plants will be the most crucial portion for early symptom description. A variety of diseases harm plants, causing considerable output losses and jeopardising human lives. We are employing human visual inspection with our unassisted eyes to discover prevalent plant illnesses that have considerably impacted performance [3]. This is one of the most popular and often utilised challenges currently. Farmers use visual inspection to detect leaf-based diseases because there is a maximum occurrence of error in approximately cases, but there are numerous laboratory-based methods for disease detection, such as polymerase chain reaction, reduction in food cultivation, pest organization, and spectral techniques, but they are time-consuming and expensive for farmers [4]. The majority of agricultural fields in Iraq are located in village areas, causing farmers to travel long distances in quest of information, which is another issue. Image processing is more efficient than employing specialists since it is more accurate, faster, cheaper, and requires less time. Crops must be screened for diseases from the beginning of their daily life cycle until they are harvested. Plant problems will first be discovered with the bare eye, which is a laborious and high occurrence of error in the method. Several approaches have been developed to replace manual plant disease detection as a result of technological improvements; these methods must be automated in order to identify plant diseases more rapidly. Machine learning technology is used to develop an automated system that learns on its own deprived of being noticeably programmed. It employs both supervised and unsupervised learning methods [5]. The Support Vector Machine (SVM) identifier, a controlled knowledge model, is used to create a hyperplane (decision boundary) for classification. To create predictions, supervised learning, a kind of machine learning approach, employs known datasets. Figure 1 depicts some common illnesses that harm leaves:

The study discovered many manual methods for identifying infectious agents and diagnosing plant diseases. The older approach, which was discussed in literature, required more period and effort. The researcher also emphasised the position of automated agricultural technologies in increasing food output and improving accuracy [6]. Another study recommends a weed-control strategy that calls for applying less pesticide to the whole plant. The suggested approach thresholds or reduces the grey level in binary pictures by using Otsu's clustering algorithm. The suggested strategy increases food output while hastening weed clearance.

e-ISSN 2320 –7876 www.ijfans.org
Vol.11,S Iss.1, 2022
Research Paper UGC CARE Listed (Group -I) Journal

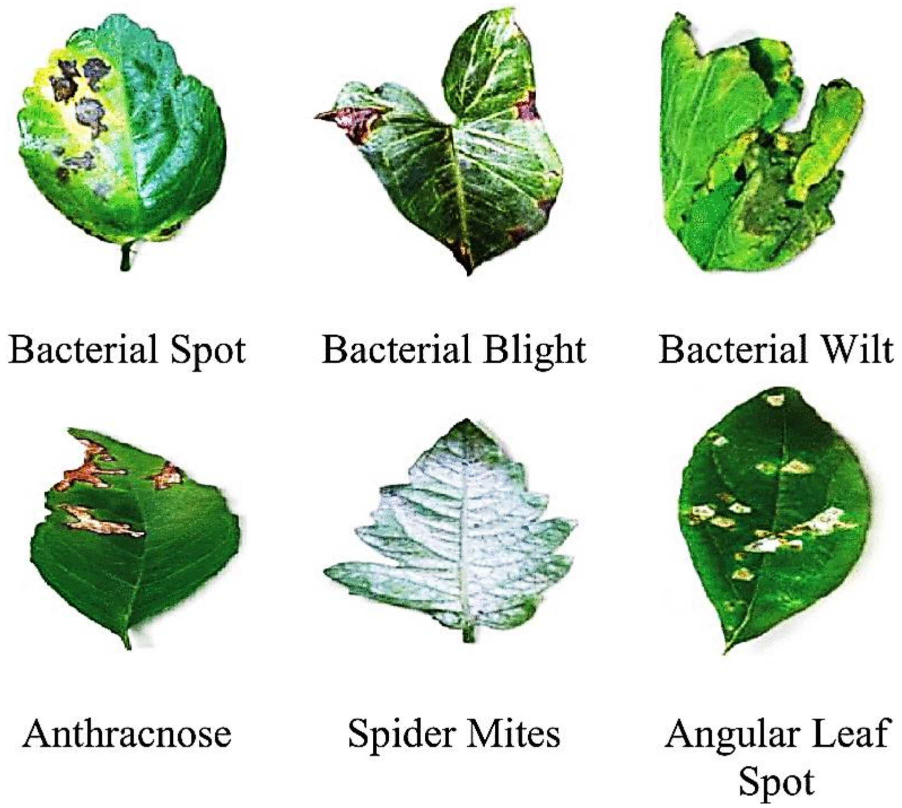

Fig. 1 Common disease in plant leaf

In order to avoid crop damage, they suggested an approach in this research that concentrates on early-stage interpretation of insect detection in a leaf picture. Several techniques, including DBSCAN and NN models, have stood reported for early pest identification. Founded on the pest dataset, the total correctness obtained, which is used in studies, is roughly 96%. In another study, a knowledge base and classification-based plant disease detection system were presented. Using 799 example pictures, an SVM classifier correctly recognised the plant disease in the proposed method, which creates knowledge.

Another research developed an automatic detection system and a segmentation approach to categorise leaf disorders. The effectiveness of several categorization approaches for diagnosing leaf diseases has been examined. In order to get the best results with the least amount of computational labour, the recommended technique categorised and recognised plant leaf diseases [7].

A coloured co-occurrence matrix is used to segment images, while another study offered a texture-based classification method that use statistics to identify leaf sickness. In order to get accurate findings, the texture-based parameters will be contrasted to values for typical leaf texture. Another study used image processing and feature-based extraction approaches to effectively gather colour information while using the k-means algorithm. Additionally, system segmentation for the detection of leaf illness was controlled using deformable models [8].

In another study, segmentation and extraction were suggested for the early diagnosis of illness. Only a few ailments can be treated using the described method. A separate study suggests that k-means clustering might be used to locate an infected region on a plant leaf. Additionally, color-based segmentation is used to identify ill zones, and the experimental findings are based on temporal complexity [9].

A different study recommends using computer vision technology to identify and measure early-stage plant diseases. An innovative technique for identifying and categorising hazardous rice in input photos was developed by another study. The author recommended using the centred feeding technique to get reliable results while using k-means clustering as a segmentation strategy. SVM is implemented using multiclass models on categories including shape, colour, and texture. A new picture identification system was developed utilising multiple linear regressions in a distinct study.

2. Methodology

It is feasible to determine if a plant leaf-based input picture is healthy or ill by using the procedures shown in Figure 2.

Figure 2's stages are described below: Inserting the photos is the first step. By submitting high-resolution sensor camera photographs in.jpeg format, users may find plant leaf datasets on the Kaggle website. After the first pre-processing step comes: The incoming picture datasets are reduced in size at this phase, making it simpler to extract features later. A Gaussian filter is used to decrease noise in the collected photographs, and scaling and other geometric adjustments are made to provide consistent picture dataset sizes.

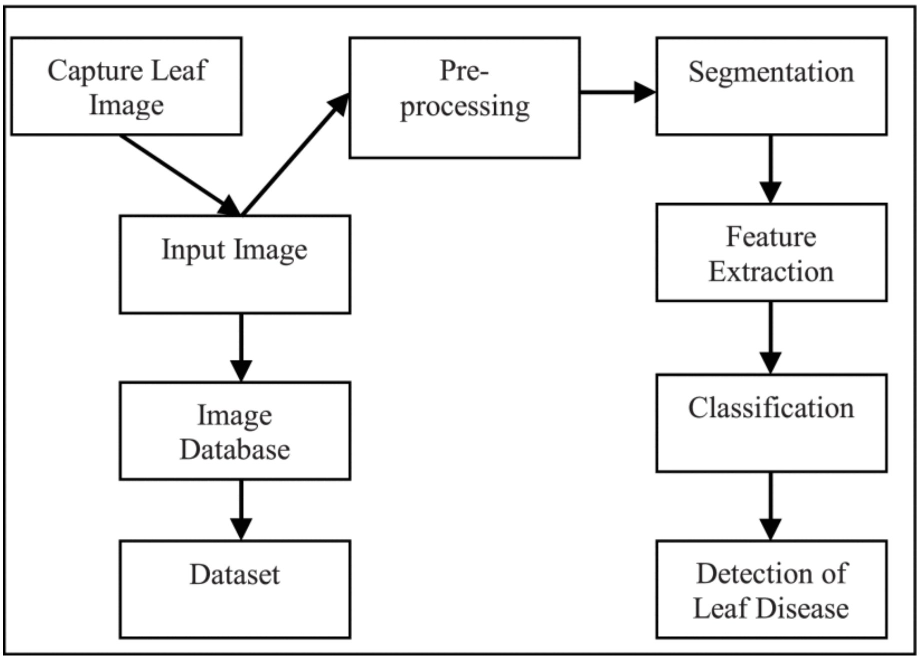

Fig. 2 Flow diagram

Segmentation comes as the third phase. The image segmentation stage translates the pre-processed picture from RGB Color Space to L* a* b* Color Space using the K-Means clustering algorithm. The L*, a, and b* chromaticity layers make up the majority of the L*, luminosity layer. The RGB colours in the a* b* colour interplanetary are then categorised using the K-means clustering method. We utilise the Euclidean Distance technique to split the picture into three groups based on distance measurements since it contains three colours.

The extraction of features is the fourth stage. By analysing the corner based, point-based, and line-based data of the leaf image, sickness leaf or healthy leaf attributes may be determined. At this point, a point-based approach is used to estimate the percentage of the leaf-based area affected by the disease. When classifying Stage 5: The SVM learning classification method is used at this level. This method is easy and efficient since it uses fewer picture sets and is more accurate. To correctly predict the outcomes, this model was tested on 30% labelled testing data and trained on 70% labelled training data.

According to the guidelines in the results chapter, the accuracy is assessed. Step 6: Efficiency metrics including RMSE, PSNR, and the disease-affected part of the foliage are utilised to forecast the detection of plant leaf illness by means of the Euclidian Distance method.

3. Results

Figures 3 (a) and 3 (b) following show the results of the MATLAB software tool:

Figure 3(a) and (b) employ yarn leaf dataset -1 and caffeine plant leaf dataset -2 as input imageries, respectively.

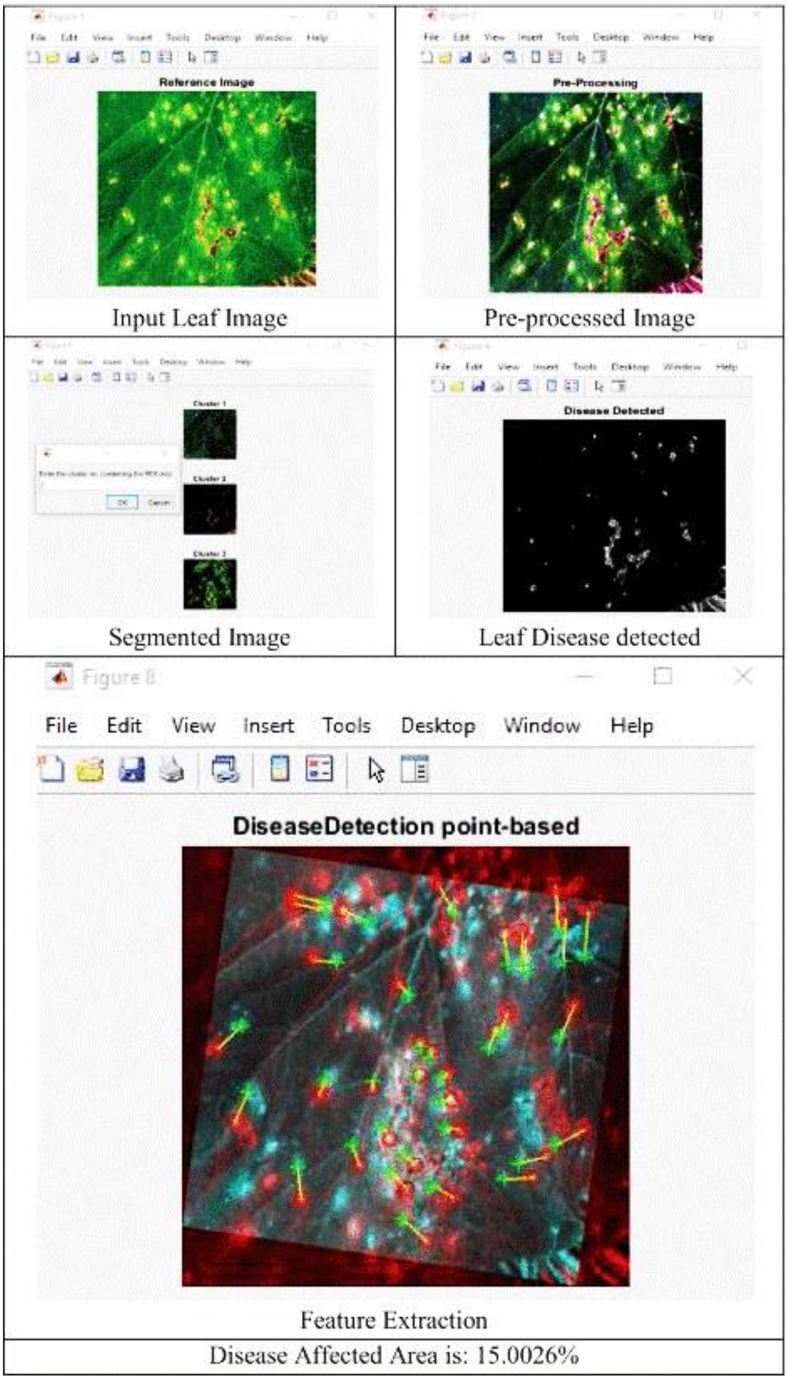

Fig. 3 Detected disease

The disease area is recognised with an accuracy of 87% in leaf dataset 1, but not with an accuracy of 98% in leaf dataset 2, suggesting that the leaf is healthy.

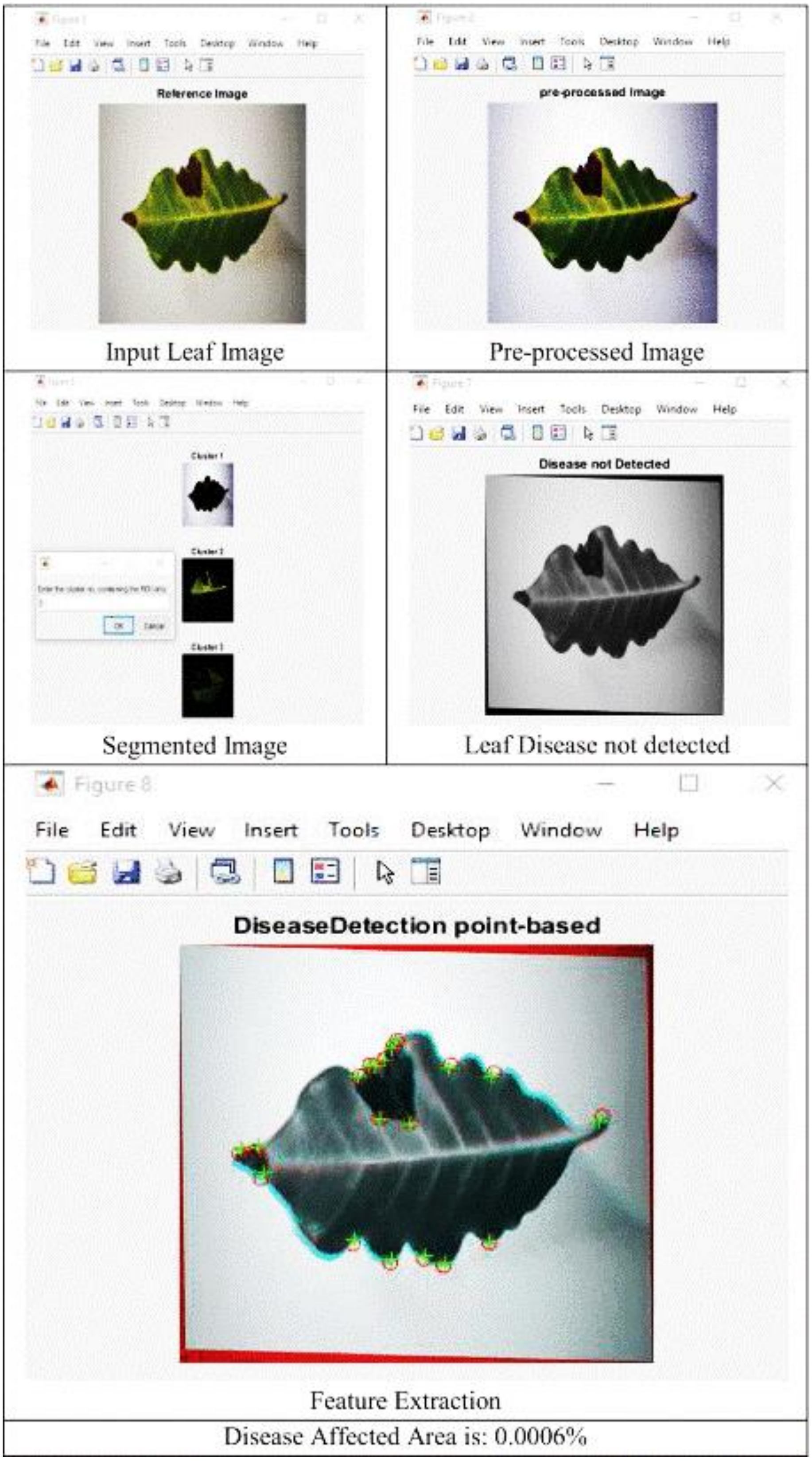

Fig. 3 Disease detected

CONCLUSION

The primary goal is to discover leaf disease in the photographs collected. Raising quality while generating for maximum crop yields is a key challenge in agricultural production. Farmers will benefit from increased agricultural yields made possible by technology and mechanisation. The majority of the study has been on automated detection and classification algorithms that support variable crop yields, as well as high resolve spectral, stereo, and spectral leaf images. We got better results when we employed the SVM classification ML algorithm model. However, in order to get more accurate findings, we will broaden our study into the automated diagnosis of leaf-based illnesses by utilising additional machine learning and deep learning algorithms over a wider range of datasets.

REFERENCES

[1] J. Azmi, M. Arif, M. T. Nafis, M. A. Alam, S. Tanweer, and G. Wang, "A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data," *Medical Engineering and Physics*, vol. 105, no. February, p. 103825, 2022, doi: 10.1016/j.medengphy.2022.103825.

[2] M. Ahammed, M. Al Mamun, and M. S. Uddin, "A machine learning approach for skin disease detection and classification using image segmentation," *Healthcare Analytics*, vol. 2, no. October, p. 100122, 2022, doi: 10.1016/j.health.2022.100122.

[3] N. Gobalakrishnan, K. Pradeep, C. J. Raman, L. J. Ali, and M. P. Gopinath, "A Systematic Review on Image Processing and Machine Learning Techniques for Detecting Plant Diseases," *Proceedings of the 2020 IEEE International Conference on Communication and Signal Processing, ICCSP 2020*, pp. 465–468, 2020, doi: 10.1109/ICCSP48568.2020.9182046.

[4] Kirti and N. Rajpal, "Black rot disease detection in grape plant (vitis vinifera) using colour based segmentation machine learning," *Proceedings - IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2020*, pp. 976–979, 2020, doi: 10.1109/ICACCCN51052.2020.9362812.

[5] S. Dhanasekaran, D. R. Chigicherla, S. R. Tholeti, and G. R. Pole, "Intelligent System to Analyse Plant Diseases using Machine Learning Techniques," *Proceedings - International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2022*, no. Icaaic, pp. 150–154, 2022, doi: 10.1109/ICAAIC53929.2022.9793276.

[6] C. Jackulin and S. Murugavalli, "A comprehensive review on detection of plant disease using machine learning and deep learning approaches," *Measurement: Sensors*, vol. 24, no. August, p. 100441, 2022, doi: 10.1016/j.measen.2022.100441.

[7] D. P. Singh and B. Kaushik, "Machine learning concepts and its applications for

prediction of diseases based on drug behaviour: An extensive review," *Chemometrics and Intelligent Laboratory Systems*, vol. 229, no. August, p. 104637, 2022, doi: 10.1016/j.chemolab.2022.104637.

[8] J. Parraga-Alava, K. Cusme, A. Loor, and E. Santander, "RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition," *Data in Brief*, vol. 25, 2019, doi: 10.1016/j.dib.2019.104414.

[9] Z. Liu, R. N. Bashir, S. Iqbal, M. M. A. Shahid, M. Tausif, and Q. Umer, "Internet of Things (IoT) and Machine Learning Model of Plant Disease Prediction-Blister Blight for Tea Plant," *IEEE Access*, vol. 10, pp. 44934–44944, 2022, doi: 10.1109/ACCESS.2022.3169147.