IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

Plant Disease Detection Using GLCM and a Hybrid SVM-kNN

Er. Deepak Sidana
Research scholar
Punjabi University, Patiala

Er. Neelofer Sohi
Assistant Professor
Punjabi University, Patiala

Abstract: Plant disease detection is essential for reducing plant and agricultural product productivity and quantity losses. Controlling plant diseases by hand is a difficult task. It necessitates a tremendous deal of effort, knowledge of plant diseases, and, in many cases, an unnecessary amount of processing time. Image acquisition, image pre-processing, segmentation, feature extraction, and classification are the five primary phases in the proposed plant leaf disease detection system. Images of damaged leaves are being collected and image pre-processing is being used to improve them.

Keywords: ***Image Processing, Plant Disease Detection, SVM, kNN, Accuracy.***

I. INTRODUCTION

The process of performing specific operations on a picture in order to obtain or extract important information from a better image is known as image processing. It's a sort of signal processing where the input is a picture and the output is either an image or a function. Image processing is one of the most rapidly developing technologies. Enhancement, segmentation, extraction, classification, and other image processing techniques will be used to improve image processing efficiency for use in a variety of applications. The process of picture enhancement involves adjusting the brightness, changing the colour tone, removing noise, and sharpening the image. Picture segmentation [1] is a way of separating an image that is used to classify items in digital photographs. Image segmentation can be done in a variety of ways, including threshold, color-based, transform-based, and texture-based. Feature extraction is a type of dimensionality reduction in which key aspects of a picture are represented as a compact vector feature. When working with large photo files and limited flexibility, this strategy comes in handy for jobs like image matching and retrieval. Image classification is the classification of images into one of a number of specified categories. The classification is broken down into two categories: supervised and unsupervised. Picture data processing for storage, transmission, and representation for autonomous machine perception, as well as image information enhancement for human interpretation, are all significant parts of digital image processing [2].

The following are some of the most common digital image processing applications:

(a) to enhance pictorial information for visual comprehension by humans

(b) to process data depending on the perception of a machine

Remote sensing, business applications, medical image processing, acoustic image processing, forensic sciences, and industrial automation are just a few of the applications that have exploited these. Digital image processing is now used to solve a variety of problems in a variety of applications. They're also used in X-ray processing, ultrasonic scanning, electron micrographs, magnetic resonance imaging, nuclear magnetic resonance imaging, and other medical procedures. The photos can be used in biology to detect diseases in various plants and to identify plant categories. Implementations of digital image processing techniques [3] aid in the analysis and solution of problems using computer vision, resulting in positive outcomes.

Because illness in plants is quite natural, disease detection [4] in plants is critical in the agriculture economy. If this region is not properly cared for, it can have serious effects for plants, impacting product quality, quantity, and productivity. Plant disease detection is essential for lowering yield and quantity losses in agricultural products. Plant disease research comprises looking at patterns on plants that can be seen. Long-term agriculture necessitates plant health monitoring and disease detection. Manually monitoring plant diseases is quite tough. It necessitates a substantial amount of labour, plant disease knowledge, and a lengthy processing period. As a result, image processing is used to determine whether or not a plant is sick.

The following sections make up this paper: The first section provides a quick overview of how image processing can be used to detect plant diseases. The review of literature is depicted in Section II. The proposed scheme employed in the

IJFANS International Journal of Food And Nutritional Sciences
Official Publication of International Association of Food and Nutrition Scientists

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

dissertation work is described in Section III. The numerous parameters employed in the dissertation study, as well as the computed findings, are described in Section IV. Section 5 brings this dissertation's overall effort to a close.

II. REVIEW OF LITERATURE

A literature survey is a summary of what has been published on a given topic by academics and researchers. The literature review was divided into sections based on the detection of diseases on plant leaves. Feature extraction is the process of minimising the number of resources required to accurately describe a large amount of data. To uncover distinctive traits for identifying leaf disease, a combination of different features such as texture, shape, and colour are employed to evaluate the significant features.

Gavhale et al. [5] showed how to isolate the sick section of a leaf using a number of image processing techniques. For pre-processing, image augmentation is done in the DCT domain, and colour space conversion is done. After that, segmentation is carried out using the k-means clustering method. The GLCM Matrix is used to extract features. Citrus leaf canker and anthracnose disease are classified using SVM with radial basis kernel and polynomial kernel.

The key image processing algorithms used for diagnosing leaf diseases, according to Radha et al. [6], are k-means clustering and SVM. This procedure can substantially assist in determining the exact cause of leaf disease. The five procedures for recognising leaf illness include picture acquisition, image pre-processing, segmentation, feature extraction, and classification. By determining the amount of disease present in the leaf, an appropriate number of pesticides can be used to effectively control the pests, leading in increased crop output. Other segmentation and classification methods can be used to improve this methodology. Disease diagnosis for all sorts of leaves is done using this concept, and the user may also determine the afflicted part of the leaf in percent by accurately identifying the illness. By accurately diagnosing the illness, the user can easily and inexpensively repair the problem.

A machine vision-based agro medical expert system was proposed by Habib et al. [7]. A two-feature set is offered here to handle papaya disease recognition, with a total of ten characteristics. Image processing techniques were employed to extract the properties. The classification of papaya disorders was done using SVM, and the relative qualities of our study were determined by comparing the findings of similar studies. 90.15 percent accuracy is reached, which is both good and encouraging. Future study with a huge data collection of photos to cover a larger spectrum of papaya illnesses is still possible.

Hussein and Abbas [8] suggested a two-phase plant disease detection system. The knowledge base is built up in the first phase by introducing a collection of training samples into a number of processing steps that include cropping, resizing, fuzzy histogram equalisation, extracting a set of colour and texture properties, and usability testing. The classifier that was trained using the knowledge base is then used to detect and diagnose plant leaf diseases in the second step. The knowledge base is made up of 799 example photographs, with 80% of the time spent on training and 20% on assessment.

SVM multiclass picture segmentation is utilised by Islam et al. [9] to create an automated and easily accessible approach. Late blight and early blight, the two most deadly potato diseases, may be recognised with little computational effort. Farmers would be able to identify diseases in a practical, reliable, and time-saving manner using the proposed approach. The authors are working on expanding the approach to include other diseases from various plant species.

Ferentinos [10] created convolutional neural network models to recognise and diagnose plant ailments using simple-leaf photos of well and diseased plants using deep-learning methodologies. The models were constructed using an open database of 87,848 photos encompassing 25 distinct plants in 58 different groupings of [plant, illness] combinations, including healthy plants. Several model designs have been trained, with the acceptable (healthy plant) combination attaining the maximum performance, with a 99.53 percent success rate.

To segment agricultural ill findings, Shedthi et al. [11] use K-Means and Fuzzy C-Means. The proposed study compares the K-Means clustering method to the FCM clustering algorithm in terms of computing efficiency and clustering accuracy. The impact of Arecanut and Iris leaf spot disease is highlighted in these two ways, and the outcomes are noticeable. The result analysis is based on the consistency of the output image and the time necessary to obtain the output. Several samples are used for segmentation, and the outcomes are reviewed.

For the identification and classification of turmeric leaves, Kuricheti and Supriya [12] were used as models. For accurate findings, a variety of processing approaches and classification algorithms were used. A.mat file containing 200 leaf photos was used to create the data set. SVM was developed to diagnose illnesses of the turmeric leaf, and its accuracy was tested and determined to be satisfactory. Using an IOT platform with a huge

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

number of data sets, better automated algorithms can be deployed in the future.

Annabel and Muthulakshmi [13] suggest an autonomous tomato leaf disease detection system. The tomato leaf images were first transformed from RGB to grayscale. Masking and thresholding are also used to obtain ROI. Then, using GLCM and random forest, specific ailments in tomato leaves such as bacterial spot, late blight, and tomato mosaic are identified. The accuracy of these algorithms combined is higher when compared to other strategies.

III. METHODOLOGY AND PARAMETERS USED

The approach and parameters for the suggested algorithm are discussed in this section.

Recognizing plant diseases is essential for lowering agricultural product yield and quantity losses. Plant disease research comprises looking at patterns on plants that can be seen. Long-term agriculture necessitates plant health monitoring and disease detection. It's difficult to keep track of plant diseases by hand. It necessitates a substantial amount of labour, plant disease knowledge, and a lengthy processing period. As a result, image processing is used to diagnose plant ailments. The procedures of illness detection include image acquisition, pre-processing, segmentation, feature extraction, and classification.

Crop diseases are a major source of food insecurity. Detecting them promptly, however, remains challenging in many parts of the world due to a lack of a necessary basis. In the realm of leaf-based picture categorization, the development of accurate techniques has generated exceptional results. The traditional approach proposed for distinguishing healthy and unwell leaves from the available data sets was Random Forest. The previous author proposed a number of implementation stages, including dataset preparation, feature extraction, classifier training, and classification. To categorise the infected and healthy photos, the created datasets of ill and healthy leaves are jointly trained using Random Forest. The Histogram of an Oriented Gradient (HOG) was used to extract attributes from an image. Overall, we have a clear path to identifying disease in plants on a broad scale by utilising machine learning to train massive publically available data sets. The traditional classifier has a number of flaws, including a long training period and poor performance on data with several dimensions. Second, there is no automatic method or methodology for detecting contaminated areas, and HOG feature extraction approaches also have significant disadvantages. The disadvantage is that it is extremely sensitive to picture rotation. As a result, HOG isn't a good fit for identifying textures or objects that are often mistaken for rotating images.

Plant-related diseases will be detected using a hybrid SVM-k-NN approach. Following the segmentation process, the characteristics from the diseased part must be calculated. The characteristics are utilised to identify the disease name. Using texture, shape, and colour to extract the features in this section. Texture features are calculated using the Gray Level Co-occurrence Matrix technique.

GLCM captures information about pixels of pairings because it is a second order statistic. GLCM shows how pixel brightness in a picture is created. At d=1 distance and angles in degrees, a matrix is created (0, 45, 90,135). Other measures such as entropy, energy, contrast, correlation, and so on are also available. These measurements are derived from a variety of angles. GLCM is a touch-related texture character profile that includes terms like smooth, silky, and rough.

On a Windows 10 operating system, the simulation is run in the MATLAB environment. The system runs on a laptop with an Intel Core i5 processor and 4GB of RAM.

The accuracy parameter is utilised to examine and measure the algorithm's performance in this study.

The proportion of correctly classified occurrences (TP + TN) to the total number of correctly identified cases is used to calculate accuracy.

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

where TP and TN stand for true positives and true negatives, respectively, and FN and FP stand for false negatives and false positives, respectively.

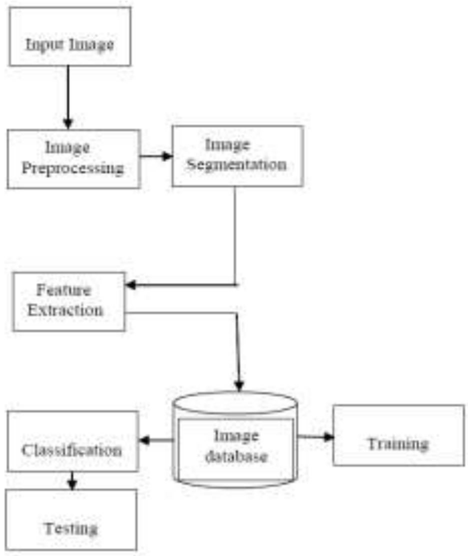

Figure 1. Flowchart of Proposed Work

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

IV. RESULTS AND DISCUSSION

The test results are presented in graphical form in this part, along with a commentary of the findings. The MATLAB GUIDE programme was used to produce the snapshots of the Application GUI for Leaf Disease System.

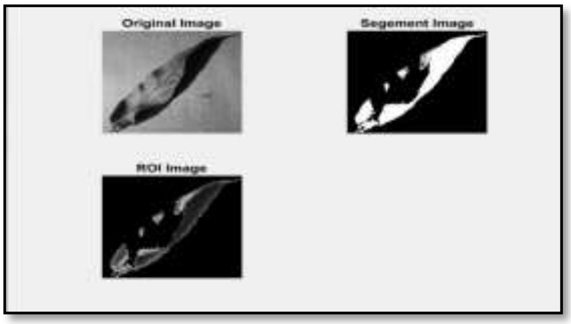

Figure 2. Original to Segment and ROI Image

A snapshot of a real leaf is shown in Figure 1. Before being transformed to a ROI image, the image is segmented using the kNN classifier. From the source image, the RoI (Region of Interest) is a suggested area. It's common to use a ROI to create a binary mask image.

Table 1. Accuracy Values for Various Algorithms

Technique	Accuracy (%)
LR	65.33
SVM	40.33
kNN	66.76
CART	64.66
Random Forest	70.14
Naïve Bayes	57.61
SVM-kNN (Proposed)	74.4361

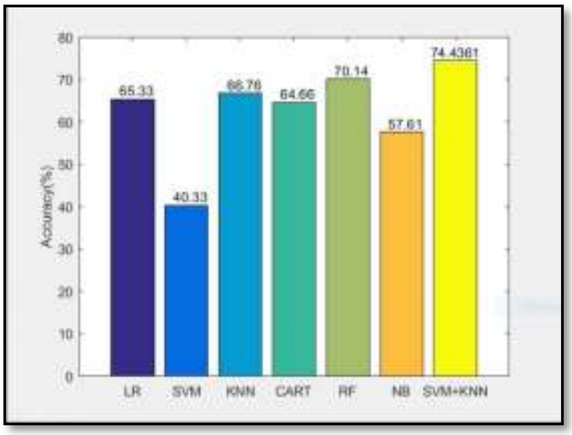

Figure 3. Comparison between Various Feature Extraction Techniques

Figure 3 shows a graph comparing different feature extraction techniques. SVM has the lowest percent accuracy of all the methodologies given, at 40.33, while the suggested hybrid SVM – kNN strategy has a percentage accuracy of 74.4361. The current hybrid SVM - kNN technique obviously outperforms the other feature extraction strategies, as seen in the above picture.

V. CONCLUSION AND FUTURE SCOPE

One of the most common methods for recognising and diagnosing plant leaf diseases is image processing. Using the SVM classification methodology, the plan intends to aid in the identification and diagnosis of leaf diseases. The color- and texture-characteristics of the diseased area are extracted once the knn classifier has detected the infected area. The categorization technique eventually identifies the leaf disease kind.

The k-NN clustering algorithm is used to segment collected leaf images in order to build clusters. SVM is used for disease classification and detection in plant leaves, whereas k-NN is used to extract GLCM. The k-NN method can be developed to detect a wide range of diseases on a huge scale. These algorithms were put to the test on a collection of photographs, some of which were diseased and some of which were not. As a result, the system's accuracy is improved.

Hybrid algorithms or other clustering methods may be developed in the future to improve the detection rate of the final classification procedure. It was also required to figure out how much disease was present on the leaf.

REFERENCES

[1] Zaitoun, N. M., &Aqel, M. J., "Survey on Image Segmentation Techniques," *Procedia Computer Science*, vol. 65, pp. 797–806, 2015.

[2] Moreno, C., Mancebo, I., Saa, A., & Moreno, M. M., "Image Analysis to Estimate Mulch Residue in Soil," *The Scientific World Journal*, vol. 2014, pp. 1–9, 2014.

[3] Joshi, A. A., &Jadhav, B. D., "Monitoring and controlling rice diseases using Image processing techniques," *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, pp. 71-476, 2016.

[4] Bhange, M., &Hingoliwala, H. A., "Smart Farming: Pomegranate Disease Detection Using Image Processing," *Procedia Computer Science*, vol. 58, pp. 280–288, 2015.

[5] Gavhale, K. R., Gawande, U., &Hajari, K. O., "Unhealthy region of citrus leaf detection using

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

image processing techniques," *International Conference for Convergence for Technology-2014*, pp. 1-6, 2014.

[6] Radha, S., Kumar, S., &Akhil, G. U., "Leaf Disease Detection using Image Processing," *Journal of Chemical and Pharmaceutical Sciences*, vol. 10, no. 1, pp. 670-672, 2017.

[7] Habib, M. T., Majumder, A., Jakaria, A. Z. M., Akter, M., Uddin, M. S., & Ahmed, F., "Machine vision based papaya disease recognition," *Journal of King Saud University - Computer and Information Sciences*, 2018.

[8] Hussein, M., & Abbas, A. H., "Plant Leaf Disease Detection Using Support Vector Machine," *Al-Mustansiriyah Journal of Science*, vol. 30, no. 1, pp. 105-109, 2019.

[9] Islam, M., Anh Dinh, Wahid, K., & Bhowmik, P., "Detection of potato diseases using image segmentation and multiclass support vector machine," *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, May 2017.

[10] Ferentinos, K. P., "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, Feb. 2018.

[11] Shedthi, B. S., Shetty, S., & Siddappa, M., "Implementation and comparison of K-means and fuzzy C-means algorithms for agricultural data," *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Mar. 2017.

[12] Kuricheti, G., & Supriya, P., "Computer Vision Based Turmeric Leaf Disease Detection and Classification: A Step to Smart Agriculture," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 545-549, Apr. 2019.

[13] Annabel, L. S. P., & Muthulakshmi, V., "AI-Powered Image-Based Tomato Leaf Disease Detection," *2019 Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 506-511, Dec. 2019.