Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777608

LEAF DISEASE DETECTION USING MACHINE LEARNING TECHNIQUES

Dr. B. Chinna Rao[1], A. Prasanth[2], Bh. Lakshmi Gowri Prasanna[3], B. Chenna Kesava Reddy[4], D. Mahesh[5]

[1]Professor, Department of Electronics and Communication Engineering, Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India.

[2,3,4,5] UG Students, Department of Electronics and Communication Engineering, Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India.

Email address : profbcrao@gmail.com, 19981A0402@raghuenggcollege.in,19981A0423@raghuenggcollege.in, 19981A0422@raghuenggcollege.in, 19981A0459@raghuenggcollege.in

Abstract :

Agricultural productivity is a key component of the Indian economy. Therefore, the contribution of food crops and cash crops is highly important for both the environment and human beings. Every year crops succumb to several diseases. Due to inadequate diagnosis of such diseases and not knowing symptoms of the disease and its treatment, many plants die. This project aims to develop a system for the detection of leaf diseases using machine learning in MATLAB. The system involves collection a dataset of labeled images of leaves using machine learning and extracting relevant features from the images, selecting a suitable machine learning model, and training it on the extracted features. The performance of the model is evaluated on a test set of images, and the system can be deployed in a real-world application for detecting leaf diseases in new images. This project utilizes MATLAB's rich set of tools and functions for each step of the process, making it an efficient and effective solution for leaf disease detection.

The leaf disease detection system using machine learning in MATLAB aims to detect diseases in leaves using image processing and machine learning techniques. The system uses a dataset of labeled leaf images containing both healthy and diseased leaves. Relevant features are extracted from the images and used to train a machine learning model using algorithms such as SVM, KNN, or ANN. The model is tested and evaluated using metrics such as accuracy, precision, recall, and F1-score. Once the model is trained and evaluated, it can be deployed to detect leaf diseases in new images. The project has potential applications in agriculture and plant pathology, where early detection and identification of leaf diseases can lead to improved crop yields and reduced crop losses.

Keywords : Agriculture, Leaf Detection, Disease Identification, Accuracy, Disease Name, Support Vector Machine (SVM).

1.Introduction:

The agriculturist in provincial regions may think that it is hard to differentiate the malady which may be available in their harvests. It is not moderate for them to go to agribusiness office and

discover what the infection may be. Our principal objective is to distinguish the illness introduce in a plant by watching its morphology by picture handling and machine learning.
Pests and Diseases results in the destruction of crops or part of the plant resulting in decreased food production leading to food insecurity. Also, knowledge about the pest management or control and diseases are less in various less developed countries. Toxic pathogens, poor disease control, drastic climate changes are one of the key factors which arises in dwindled food production. Various modern technologies have emerged to minimize postharvest processing, to fortify agricultural sustainability and to maximize the productivity. Various Laboratory based approaches such as polymerase chain reaction, gas chromatography, mass spectrometry, thermography and hyper spectral techniques have been employed for disease identification. However, these techniques are not cost effective and are high time consuming.
In recent times, server based and mobile based approach for disease identification has been employed for disease identification. Several factors of these technologies being high resolution camera, high performance processing and extensive built in accessories are the added advantages resulting in automatic disease recognition. Modern approaches such as machine learning and deep learning algorithm has been employed to increase the recognition rate and the accuracy of the results. Various research have taken place under the field of machine learning for plant disease detection and diagnosis, such traditional machine learning approach being random forest, artificial neural network, K-means method, Convolutional neural networks etc…texture is used to get the texture of the leaves and color Histogram is used to represent the distribution of the colors in an image.

2.Literature Survey:

Chaitali et al. segmentation of image is applied for background subtraction. The classification approach is carried out by ANN and SVM method. In KNN, it classifies samples using nearest distance between trained and testing subjects. Varun et al., has developed model for extraction thresholding technique and morphological operation. Then multiclass SVM is used as classifier. For segmentation, based on a set of marks generated by analysis of the color and luminosity components of different regions of image is L*A*B* color spaces. The GLCM is used for feature extraction. Vijai Singh et al., considered samples of plant leaves like rose/beans (bacterial disorder), lemon (sun burn disorder), banana (early scorch) and beans (fungal) that are captured using a digital camera. The green regions as background using thresholding algorithm. Finally, the genetic algorithm is used to get the segmented image. The color co-occurrence is adapted for useful extraction of features from the segmented images. The Minimum Distance Criterion and then SVM classifier is used for classification purpose. Sa'ed Abed et al., performed scaling and stretching (min-max linear) process for the input samples to improve the quality. The creation of HIS model is completed and the same is segmented later. The techniques of combined Euclidean distance and K-mean clustering is performed for segmentation of the samples. The GLCM and SVM are used for feature extraction and classification respectively. Arya et al., takes input RGB image and creates color transformation then conversion of the input samples to HIS format. Finally, segment the components using Otsu's method. Nema et al., images of 81 were included in the database and analysis was performed in L*a*b color space. Segmentation of the leaf disease was carried using k-means

clustering and the classification of the disease was performed using SVM. Statistical information such as mean, median, mode, standard deviation was used by authors to record their findings. Vidyashree Kanbur et al., developed the model for leaf detection disease using multiple descriptors. The model was tested on local leaf database and the performance of the model was superior., but it can be tested on publicly available dataset.

In addition to using different machine learning techniques, researchers have explored different digital image processing techniques to improve the performance of leaf disease detection systems. These techniques include segmentation, feature extraction, color normalization, and image enhancement.

3.Proposed Methodology:

The proposed method for leaf disease detection using machine learning is a promising approach that can significantly improve the efficiency and effectiveness of plant disease detection. It has the potential to revolutionize the agricultural industry by enabling early detection and intervention of plant diseases, ultimately leading to increased productivity and profitability for farmers. Further research and development in this area can help to refine and improve the performance of the system, as well as expand its applications to other areas of plant health and agriculture.

Furthermore, the same machine learning algorithms used in the proposed method can be applied to detect nutrient deficiencies in plants. By analyzing the color and texture of plant leaves, the machine learning model can identify nutrient deficiencies such as nitrogen, phosphorus, and potassium. This can enable farmers to address the nutrient deficiency and improve the overall health and productivity of their crops.

In summary, there are many areas of future research and development for the proposed method for leaf disease detection using machine learning. By continuing to refine and improve the system, we can help address a range of challenges faced by farmers in the agricultural industry and ultimately improve the sustainability and productivity of the industry.

A proposed method for leaf disease detection using digital signal processing and machine learning, involving capturing images of leaves and extracting features such as color, texture, and shape. These features are then processed using digital signal processing techniques and algorithms. Pretrained technique is used to detect the diseased and healthy plant images.

4.WORKING:

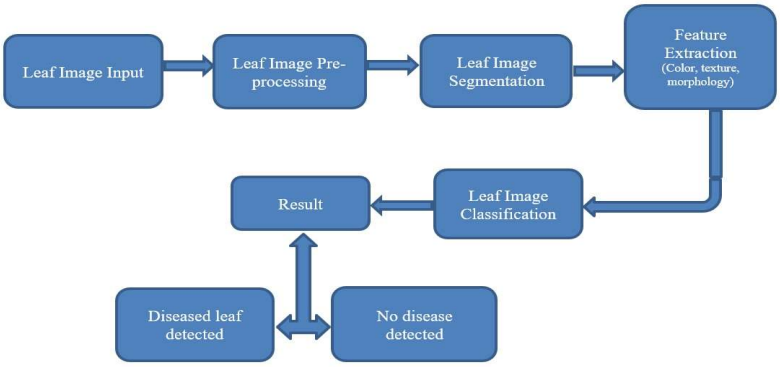

Fig1.Block diagram Representation

A leaf disease detection system using machine learning and digital image processing project works by acquiring images of leaves from plants in the field using a camera or from data base, and then preprocessing the images using digital image processing techniques to enhance the features of the leaves and improve image quality. Features such as color, texture, and shape are extracted from the preprocessed images using machine learning techniques. A machine learning algorithm, such as a convolutional neural network (CNN), is used to classify the images as healthy or diseased. The algorithm is trained using a large dataset of labeled images of healthy and diseased leaves, enabling it to accurately classify new images.

A real-time monitoring system uses a camera and a microcontroller to capture images of the leaves and analyze them using the machine learning algorithm. The system can provide alerts to farmers when a plant is showing signs of disease. The system also generates reports on the health of the plants, providing information on the types of diseases detected, their severity, and recommendations for treatment.

Fig 2. Enhanced Image

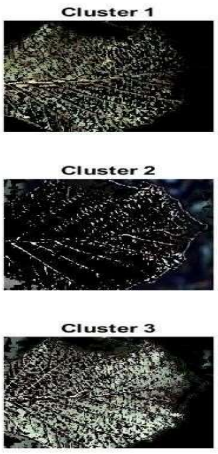

Fig 3:Image Clusters

5.Image Segmentation:

Leaf image segmentation is the process of separating the leaf object from the background in an image. Segmentation is a crucial step in developing a leaf disease detection system using machine learning, as it isolates the region of interest and enhances the feature extraction process.

K Means Clustering:

K-Means Clustering is an unsupervised machine learning algorithm which is used to solve the clustering problems in the machine learning. In real-world scenarios, the unlabeled data that might be exists to solve problems. In such cases, the K-means algorithm plays a vital role to solve the problem. Whereby taking unlabeled data dividing into subgroups. Where groups can be termed as a cluster. The grouping can be done by data having similar size, shape, measure,

characteristic. Finally, group the data in a separate cluster. Here come K values, how many clusters need to create. There is a separate technique to determine the k value.

K-Means Clustering is an Unsupervised Learning algorithm, which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on.

It is an iterative algorithm that divides the unlabeled dataset into k different clusters in such a way that each dataset belongs only one group that has similar properties.

It allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabeled dataset on its own without the need for any training. It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters.

Preprocessing: The image is preprocessed to remove noise, adjust brightness and contrast, and perform other image enhancement techniques.

Feature detection: Features are detected in the image using various algorithms, such as edge detection, blob detection, or corner detection.

Feature extraction: The relevant information is extracted from the detected features, such as color histograms, texture features, or geometric features.

	Feature	Expression
1	Mean =M	$\sum_{i=0}^{N-1} g(i) P(g(i))$
2	Standard Deviation =S	$\sqrt{\sum_{i=0}^{N-1} (g(i)-M)^2 P(g(i))}$
3	Entropy	$\sum_{i=0}^{N-1} P(g(i)) log_2 (P(g(i)))$
4	RMS	$\sqrt{\frac{1}{NxN} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (g(i,j)-I)^2}$
5	Variance	$\sum_{i=0}^{N-1} (i-\mu)^2 p(i)$
6	Smoothness	$\sum_i \sum_j \frac{1}{1+(i-j)^2} g_{i,j}$
7	Kurtosis	$\frac{1}{S^k} \sum_{i=0}^{N-1} (g(i)-M)^3 P(g(i))$
8	Skewness	$\frac{1}{S^3} \sum_{i=0}^{N-1} (g(i)-M)^3 P(g(i))$
9	Inverse Difference	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{p_{i,j}}{1+(i-j)^2}$
10	Contrast	$\sum_i \sum_j (i-j)^2 g_{i,j}$
11	Correlation	$\frac{\sum_i \sum_j (ij) g_{i,j} - \mu_x \mu_y}{\sigma_x \sigma_y}$
12	Energy	$\sum_i \sum_j g2_{i,j}$
13	Homogeneity	$\sum_i \sum_j \frac{1}{1+(i-j)^2} g_{i,j}$

Table 1: Feature Extraction and their expressions

Feature selection: The most relevant and informative features are selected from the extracted features using various techniques, such as principal component analysis (PCA), linear discriminant analysis (LDA), or genetic algorithms.

Feature representation: The selected features are represented in a suitable format, such as a feature vector or a feature matrix, for further processing or analysis.

Feature extraction is a critical step in many image processing applications, such as object recognition, image classification, and image retrieval. It requires domain knowledge, experience, and careful selection of appropriate algorithms and techniques to ensure that the extracted features are relevant, informative, and robust to variations in the image data.

6.Conclusion:

In conclusion, leaf disease detection using machine learning in MATLAB is a promising approach for accurately and efficiently identifying plant diseases. However, it is important to note that the accuracy of the disease detection system depends on the quality of the image data, the feature extraction and selection methods, and the chosen classification algorithm. The purpose of this project is to implement a leaf disease detection system that can detect diseased leaves. As described system that can detect diseases of plants. With respect to the leaf disease detection system the following can be concluded:

Accuracy is far better than the previous methodology which have 95% and the proposed methodology gives 99% accuracy. The objective of the project is to recognize abnormalities that occur on plants in their greenhouses or natural environment. The use of Digital Image Processing is what made the disease detection easy in crops, though every farmer is unaware of every disease. The proposed system is capable of detecting the disease at the earlier stage as soon as it occurs on the crop, reducing the dependency on the expert to a certain extent.

7.References:

[1] Khirade SD, Patil AB (2015) Plant Disease Detection Using Image Processing. Int Conf on Comp Communication control and Automation IEEE 153: 768-771.

[2] Surya Prabha D, Satheesh Kumar J (2014) Study on Banana Leaf Disease Identification Using Image Processing methods. Int Jour of Res in Comp Sc and Inf Tech IJRCSIT 2(A): 89-94.

[3] Weizheng S, Yachun W, Zhanliang C, Hongda3 W (2008) Grading Method of Leaf Spot Disease Based on Image processing. Int Conf on Comp Sc and Soft Eng, IEEE 491-494.

[4] Phadikar S, Sil J (2008) Rice Disease Identification using Pattern Recognition Techniques, Process of 11th Int Conf on comp and Inf Tech, ICCIT, Khulna, Bangladesh, IEEE 420-423.

[5] Sabah B, Sharma N (2012) Remote area plant disease detection using image processing. IOSR J Electron Communication eng 2(6): 31-4.

[6] Wang Z, Chi Z, Feng D, Wang Q: Leaf Image Retrieval with Shape Features. Lecture Notes in Computer Science 1929. In: Laurini R ed. Advances in Visual Information Systems. Berlin: SpringVerlag, 477~487

[7] C. Yang, H. Wei, and Q. Yu, "Multiscale Triangular Centroid Distance for Shape-Based Plant Leaf Recognition," in European Conf. on Artificial Intelligence, 2016, pp. 269–276.

[8] B. Yanikoglu, E. Aptoula, and C. Tirkaz, "Automatic plant identification from photographs," Machine Vision and Applications, vol. 25, no. 6, pp. 1369–1383, 2014.

[9] G. Cerutti, L. Tougne, J. Mille, A. Vacavant, and D. Coquin, "Understanding leaves in natural images - A model-based approach for tree species identification," Computer Vision and Image Understanding, vol. 117, no. 10, pp. 1482–1501, 2013.

[10] O. Mzoughi, I. Yahiaoui, N. Boujemaa, and E. Zagrouba, "Semantic based automatic structuring of leaf images for advanced plant species identification," Multimedia Tools and Applications, vol. 75, no. 3, pp. 1615–1646, 2016.