COMPREHENSIVE SURVEY ON VARIOUS DISEASE DETECTION SYSTEM IN PLANTS USING ARTIFICIAL INTELLIGENCE BASED MODELS

M. Sangeetha
Department of Computer Science and Engineering, Cochin University of Science and Technology, India

Abstract

New diseases are discovered day after day in recent years, not only in people, but also in plants. The identification of plant diseases (through the use of digital image processing) is essential for the detection of diseases which assist agriculture, because diseases in plants are becoming quite prevalent. This article explains briefly how image processing is utilised by a number of processes to identify diseases in plants. The method for segmentation in this article is k-means clustering and GLCM extraction. Using SVM is done after extraction classification.

Keywords:

Disease Detection, Artificial Intelligence, Digital Image Processing, SVM

1. INTRODUCTION

The image consists essentially of a limited number of pixels called elements, each with a certain position and value. The processing of images is a method for extracting the necessary information from an image through a sequence of processes. Digital image processing is now one of the world's fastest-growing technologies. It also constitutes key areas of study in the fields of engineering and IT to support nuclear medicine, astronomical observation, the identification of signatures, the detection of plaque numbers, agriculture, etc. The agriculture sector uses digital image processing extensively because it is helpful for controlling harvests, fruit grading, identification of plant diseases, etc. [1].

Plant disease identification is one of the main applications of digital image treatment in agriculture. It contributes to the global incorporation of significant economic losses in agriculture. It is essential to monitor plant health and identify pathogens in order to prevent the transmission of diseases between plants. It is possible to identify diseases in a number of ways. Visually surveying symptoms is a classic way of doing this. However, this technique cannot identify all diseases, and only experienced farmers can detect naked-eye disease. Early stage diseases are often difficult to identify. In situations when diseases have no apparent signs, a microscopic examination should be carried out. An automated method has been created to identify plant diseases by utilising PCA and SVD algorithms from neural networks.

2. LITERATURE REVIEW

This section covers the work of scientists in many fields, such as fruit grading, weed identification, plant categorization, etc.

A disease detection software system based on images of several rice plants affected. Images from the rice plants infected with a digital camera are collected and analysed using image building and zooming methods for detecting infected plant sections. The diseased portion of a leaf has then been utilised by a neural network for categorization. The processing and soft-computing techniques used in this system are [6].

The suggested technique to detect and classify plant diseases accurately Identify green-coloured pixels in the first phase. Then pixels are masked using the Otsu technique and mostly green pixels are masked on the basis of particular threshold values achieved. The second extra step is to eliminate entirely the pixels with red, green and blue zeros and the pixels from the borders of infected clusters. The SGDM matrix for H and S is created and then GLCM is called to compute the characteristics. Results from experiments show that this approach is a powerful method for detecting diseases in leaves of plants.

Detection of remote areas of plant disease using image processing. A technique for diagnosis of illness in domestic malus utilising approaches like k-mean clustering, texture and colour analysis is presented in this article. The CCM technique is the algorithm used to segment the texture. As RGB images from the sheets are transformed into a representation of HSI colour space. This is then utilised to produce a matrix of co-occurrence (CCM). By comparing images of texture and colour, you may identify diseases of plants [8].

The image processing algorithm is used to segment the location of illness on the plant floor. Automatic identification and grading of plant diseases is carried out in the first phase [9]. A medium filter is used for image smoothing. To calculate the threshold, the Otsu technique is utilised. A noise-free algorithm was created and tests on various families of'monocot' and 'Dicot' plant leaves were performed, both with a noise-free (white) and a noisy backdrop. The impact of CIELAB, HSI and YCbCr on the process of spot detection of illness may be compared in this article.

The photo identification by utilising image processing technologies and the BP networks for two sorts: the grape disease of the grape (grape meldew and grape powdery meldew) and two sorts of wheat (white stripe rust and wheat leaf rust). The BP networks, built using different function combinations, were used as classificators for identifying grape diseases and wheat diseases in the use of combinations based on data from the extracted colour features, shape and texture features of disease images and their combined features. Image identification of pest conditions using BP networks was also performed on the basis of the dimensional data acquired from PCA for data processing [10].

Several methods of segmentation and classification are provided. It was determined from the systems mentioned that most researchers extensively utilised the K-means clustering technique for segmentation. GLCM and the SVM classifier were shown to perform better in classification and feature extraction than others [11].

The approach in [11] utilised histogram methods for the identification of edge detection and colour matching of plant diseases. In the first phase, two stages of the system are included, which contain all disease and healthy leaves.

Test samples will be presented as input to the system during the second phase. The RGB components are initially transformed in two stages into three layers: red, green and blue. The Canny Edge Detector is used for edge detection. Then, a histogram is compiled and saved in the systems for each component of the image of the healthy and sick leaf. A comparison of stored data is carried out with diseases that are infected or not affected plants [5].

The k-nearest neighbour test class is the simplest technique, but the time complexity of the prediction is one of the disadvantages of the k-NN method. Further, it is difficult to comprehend the algorithm structure [12]. Neural networks are resistant to noisy inputs. SVM has been discovered competing in high-dimensional data sets with the best known machine learning methods. The complexity of the SVM computation is reduced to a quadratic optimization problem and the complexity of the decision rules and frequency of mistakes may be controlled [13].

The algorithm starts with digital image acquisition of infected and uninfected plants; pre-processing the image, distinguishing the infected disease area from the uninfected area by means of a conversion of space, segmentation and extraction from the feature-based analysis, neural network, and support vector machine, fuzzy and rules-driven recognising image. For researchers in plant pathology and pattern recognition, this study was anticipated to be helpful [14].

Sugarcane leaf disease detection based on the extraction of the feature is suggested. This method is also possible using conventional approaches such as JAVA, but the efficient and effective results may be achieved using MATLAB. This program helps farmers and laboratories to quickly prevent illness in the early stages and boost output growth [15].

3. IMAGE ACQUISITION

This is the start of the detection process for plant diseases. The high-quality images are generated in this phase database. The image quality is very important because the method's effectiveness relies on the photos in the database. Images of the leaves are typically taken by cell phones or digital cameras like the Nikon D3200.

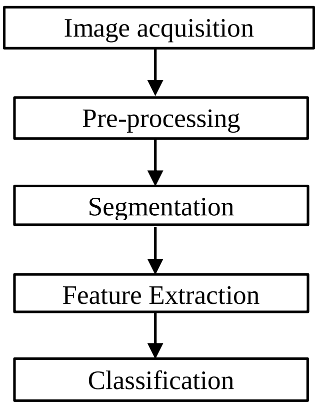

Fig.1. Block diagram of proposed approach

Image segmentation is described as the process of separating an image into different areas, each using the information to be extracted. Segmentation may be done on the basis of two characteristics of intensity values, one being interruption and another being similarity. These segmentation methods may be categorised according to the following properties:

3.1.1. Edge based Segmentation:

The method of edge detection, which is more commonly employed in the segmentation of images, results in a line drawing based on a rapid shift in intensity values. Edge detection methods are used in this segmentation. The initial step in segmentation is to determine the limit of the image to be segmented. Discontinuities are discovered when the edges are identified. The pixel value is determined by the edge tracing and compared to the adjacent pixels. They utilise the fixed and adaptive features of the vector support machine for this categorization (SVM). The primary goal of the technology consists of identifying regions where a significant variation in intensity values takes place, like canny and sobel, the most widely-used edge edge detection detectors, canny edge detectors for the detection of a broad range of edges [2].

3.1.2. Region based Segmentation

The concept of homogeneity is the basis of this method. It indicates that neighbouring pixels within an area have similar characteristics, but pixel characteristics in other regions vary. The objective of this method is to produce a larger region with comparable characteristics, i.e. a consistent area, and to reduce the number of areas. Segmentation based on region may be categorised as:

Region Growing: The process is to aggregate pixels or sub-areas via an examination of the surrounding pixels into bigger areas. For each limit pixel, it is repeated. The method will begin with the selection of a pixel called the seed point and area by taking the adjacent seed, the characteristics of which are similar to those of the chosen at the start [3].

Region Splitting and Merging: This method divides an area R into quadrants and sub-quadrants, e.g. smaller and smaller regions. The original assumption in this method is that the whole image is a single area, then the homogeneity test is calculated in order to verify certain characteristics. This procedure continues until additional divisions are required. Then these tiny square areas are combined if they are similarly irregular.

3.1.3. Partitioning Clustering:

An alternative technique called clustering may be used for segmentation. The total number of observations in this technique is split into K clusters, where K is the cluster number in the segmented region. K centroids, one for each cluster, must be determined via optimization of the sum of data distance from their respective centroids in an iterative way. The clustering of K-means is also known. This approach has the benefit of working on image information both locally and globally. In addition, it operates on image information both locally and globally.

3.1.4. Threshold Segmentation:

Segmentation of images is one of the most common techniques utilised, since it follows a simple segmentation strategy. First, the grey image is transformed into a binary image and the threshold values are calculated by using the histograms of the edges (in the original image). As objects and pixels below the threshold value are regarded as the background, the pixels above the threshold value. This method is helpful if there are significant differences in the intensity values between the item and the background, but it does not allow numerous objects to be recognised using this technique.

3.1 FEATURE EXTRACTION

Feature extraction is the way through which the data set to be used in the following phase is trained. The relevant information is extracted from the images processed. Various feature extraction methods are used [4]: extraction of colours, extraction of texture, and extraction of shape features.

- *Color Feature*: Another benefit is that the colour is not influenced by the image size, rotation and zoom.
- *Texture Feature*: Texture is defined as a pixel group with certain shared features. Two categories may be classified: statistical and syntactic. Our emphasis is on analysing statistical texture. Feature texture analysis characteristics are acquired in the statistical texture by monitoring the combinations of intensities in certain specific locations that relate. The cooccurrence matrix is a literary method. This is a way of obtaining statistical texture characteristics of second order.
- *Shape Features*: Shape is an essential image feature. The image that consists of: Identity identification, Translation, Rotation, and Invariance of the Scales, Affine invariance, and Noise Resistance, must have the following characteristics for efficient design. The form extraction methods may be divided into two categories: extraction from the area and extraction from the contours.

The contour-based technique is used to compute the boundary features except for the internal area, whereas the region-based extraction determines the feather of the region as a whole.

4. CLASSIFICATION

Following successful extraction of the function, features should be utilised for classification, identification and classification of plants based on leaf texture-related characteristics such as GLCM, via a decision-tablening classifier. It’s the moment where categorization is shown. Classification involves comparing the vectors with the different models and finding the closest match. The training data set to match the functional vectors should be utilised.

The SVM is a common classification technique and is also applied in this article. In one of the two distinct classes, SVM uses valuable information as an input. SVM’s fundamental concept is to identify the optimum hyperplane that divides the instances. The search for an optimum hyperplane is like the resolution of a quadratic problem. Most categorization models are developed on the basis of the concept of empirical risk reduction. This approach reduces the number of mistakes in training and the interval of superiority. The SVM is based on the idea of decision-making planes used to establish boundaries in decision-making.

5. CONCLUSION

In this study, it is concluded that the detection of plant disease is the method for the detection of diseased leaf portions. The detection of a plant disease consists of two stages. The first step is the division of the image, and the second step method is used for extracting features and classifying the diseases and normal parts of the image. This article reviews and discusses several methods of detection of plant diseases using different criteria.

REFERENCES

[1] Green Revolution in India, Available at: http://en.wikipedia.org/wiki/Green_Revolution_in_India, Accessed at 2020.

[2] R.C. Gonzalez, R. E. Woods and S. L. Eddins, “*Digital Image Processing using Matlab*”, Prentice-Hall, 2003.

[3] Yousra Ben Jemaa and Sana Khanfir, “Automatic Local Gabor Features Extraction for Face Recognition”, *International Journal of Computer Science and Information Security*, Vol. 3, No. 1, pp. 1-14, 2009.

[4] B.R. Kandukuri, V.R. Paturi and A. Rakshit, “Cloud Security Issues, in: Services Computing”, *Proceedings of IEEE International Conference on Services Computing*, pp. 517-520, 2009.

[5] X. Yang and T. Guo, “Machine Learning in Plant Disease Research”, *European Journal of Biomedical Research*, Vol. 3, No. 1, pp. 6-9, 2017.

[6] R. Ghadge, J. Kulkarni, M. Pooja, N. Sachee and R.L. Priya, “Prediction of Crop Yield using Machine Learning”, *International Research Journal of Engineering and Technology*, Vol. 5, No. 2, pp. 31-37, 2018.

[7] A.A. Sarangdhar and V.R. Pawar, “Machine Learning Regression Technique for Cotton Leaf Disease Detection and Controlling using IoT”, *Proceedings of International Conference on Electronics, Communication and Aerospace Technology*, pp. 449-454, 2017.

[8] Vaneesbeer Singh and Abid Sarwar, “Analysis of Soil and Prediction of Crop Yield (Rice) using Machine Learning Approach”, *International Journal of Advanced Research in Computer Science*, Vol. 8, No. 5, pp. 1254-1259, 2017.

[9] Sabri Arik, Tingwen Huang, Weng Kin Lai and Qingshan Liu, “Soil Property Prediction: An Extreme Learning Machine Approach”, *Proceedings of International Conference on Neural Information Processing*, pp. 666-680, 2015.

[10] Antonieta De Cal, Inmaculada Larena, Belen Guijarro and Paloma Melgarejo, “Use of Biofungicides for Controlling Plant Diseases to Improve Food Availability”, *Agriculture*, Vol. 2, No. 2, pp. 109-124, 2012.

[11] H. Benjamin Fredrick David and S. Antony Belcy, “Heart Disease Prediction using Data Mining Techniques”, *ICTACT Journal on Soft Computing*, Vol. 9, No. 1, pp. 1817-1823, 2018.

[12] H.B.F. David and A. Suruliandi, “Empirical Study of Ensemble Classifications on Benchmark Datasets”, *Journal of Analysis and Computing*, Vol. 12, No. 2, pp. 1-14, 2018.

[13] K.H. Miao, J.H., Miao and G.J. Miao, “Diagnosing Coronary Heart Disease using Ensemble Machine

Learning", *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 10, pp. 1-12, 2016.

[14] Eibe Frank, Mark Hall, Len Trigg, Geoffrey Holmes and Ian H. Witten, "Data Mining in Bioinformatics using Weka", *Bioinformatics*, Vol. 20, No. 15, pp. 2479-2481, 2004.

[15] Latha Parthiban and R. Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", *International Journal of Biological, Biomedical and Medical Sciences*, Vol. 3, No. 3, pp. 1-8, 2008.