Grenze International Journal of Engineering and Technology, Jan Issue

A Survey on Identification and Classification of Diseases in Crops using Image Processing Techniques

Prof.Gouri C.Khadabadi[1], Dr.Geeta R. Bharamagoudar[2], Dr.Shashikumar G. Totad[3] and Dr.V.B.Nargund[4]
[1]Gogte Institute of Technology, Computer science & Engg, Belagavi, India
Email: gourikhadabadi@gmail.com
[2]K.L.E. Institute of Technology, Computer science & Engg, Hubballi, India
Email: geetatotad@yahoo.co.in
[3]K.L.E. Technological University, School of Computer science & Engg, Hubballi, India
Email: totad@kletech.ac.in
[4]University of Agricultural Sciences, Dept. of Plant pathology, Dharwad, India
Email: nargund56@gmail.com

***Abstract*—Agriculture produce is of utmost importance which contributes to a country's economy. Identifying diseases in crop at an early stage of crop development can improve the quality and quantity of yield produced. The past method of disease identification and classification has been done manually by farmers or experts but the drawbacks have been inaccurate identification of disease or the process involves large amount of time since experts may be located at a distant location. Since a lot of advancement has taken place in the field of image processing and machine learning these techniques can be used in identification and classification of diseases in crops. This paper surveys the different methodologies available for disease identification and classification, the different techniques used for segmentation, feature extraction, and classification.**

***Index Terms*— Disease detection, Classification, image processing, machine learning, segmentation, feature extraction.**

I. INTRODUCTION

Agriculture plays a vital role in a country's economy. Around 70% of people in India are dependent on agriculture for their livelihood which contributes to over 60% employment to the population. Indian agriculture comprises of variety of crops where the staple food of people is wheat and rice. Apart from this the farmers also grow pulses, potatoes, sugarcane, oil seeds, cotton, tea, coffee, rubber, fruits etc. The yield produced has to be qualitative as well as quantitative. But many a times the crop is affected by diseases which results in huge economic loss and contributes to sub standardized crop yield. There are many factors contributing to the diseases which are broadly classified into two categories i.e. living agents-bacteria, viruses, fungi, insects and non-living agents like change in temperature, insufficient sunlight, poor nutrients in the soil, excess of moisture, amount of leaf wetness, air pollution.

The past method of disease identification and classification was done manually by observation through human eyes which was inaccurate and time consuming. Since experts in disease identification are not readily available nearby all the time. Decisions taken manually are not always correct because of cognitive limits in humans.

Grenze ID: 01.GIJET.7.1.506

Timely identification of disease is very important to farmers as it saves a lot of time and money.it can alert the farmers on what precautionary measures are to be taken in order to protect the field from pest attack. What pesticides are to be used and their quantity of usage. A lot of technical advancement has taken place in the field of image processing and machine learning, they can be applied to determine the disease affecting the crop and also classify the type of disease. In the past work has been done using image processing techniques in identification of diseases in plants, leaf, fruits and different types of crops. In this paper we present a survey on identification and classification of different types of diseases affecting the crop using image processing techniques [16]. The various stages of disease detection and identification are image acquisition, image pre-processing, image segmentation, feature extraction and classification.

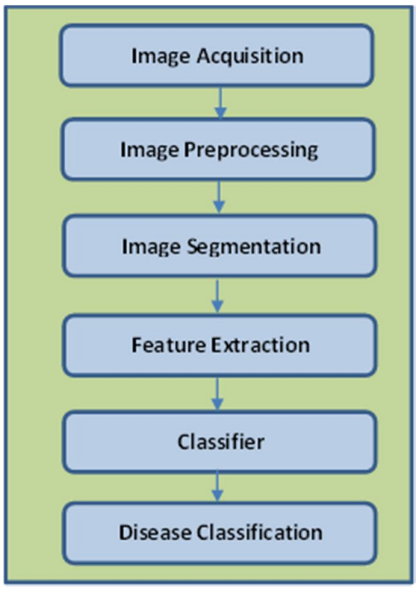

Figure 1. Steps to identify and classify disease from different crops

II. Related Work

The steps involved in image processing are as follows:

A. Image Acquisition

It involves basically capturing an image. Generally, the image acquisition stage involves pre-processing, such as scaling, etc. after the image has been obtained various methods of processing can be applied to the image to perform many different vision tasks required.

B. Image Pre Processing

Image Enhancement

It is the process of filtering image (removing noise, increasing contrast, etc.to improve the quality. The resulting image will be more suitable than the original image. Image enhancement is the procedure of improving the quality and the information content of original data before processing.

Image Restoration

It is the process of improving appearance (reducing blurring etc.) of an image by mathematical or probabilistic models. Image restoration is the operation of taking a corrupt/noisy image and estimating the clean, original image. Corruption may come in many forms such as motion blur, noise and camera miss-focus. More advanced image processing techniques must be applied to recover the object.

Applying filters on the image to remove noise

Image distortion is most pleasance problems in image processing. Image distorted due to various types of noise such as Gaussian noise, Poisson noise, Speckle noise, Salt and Pepper noise and many more are fundamental noise types in case of digital images.

It may produce at the time of capturing or image transmission. Noise means, the pixels in the image show different intensity values instead of true pixel values that are obtained from image. Noise removal algorithm is the process of removing or reducing the noise from the image. The four main types of filters are low-pass, high-pass, band-pass, and notch/band-reject (though there are also all-pass filters).

C. Image Segmentation

In digital image processing and computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image.
The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic.

D. Feature Extraction

Feature extraction involves reducing the number of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Many machine learning practitioners believe that properly optimized feature extraction is the key to effective model construction. Feature extraction is a dimensionality reduction process, where an initial set of raw variables is reduced to more manageable groups for processing, while still accurately and completely describing the original data set. When the input data to an algorithm is too large to be processed and it is suspected to be redundant then it can be transformed into a reduced set of features. Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

E. Classification of images

Classification allows to collect various attributes together into discernable categories, which can be used to draw further conclusions. Classification is used to classify each item in a set of data into one of a predefined set of classes or groups. Classification belongs to the category of supervised learning where the targets also provided with the input data. The training dataset is employed to obtain better boundary conditions which can be used to determine each target class; once such boundary conditions are determined, next task is to predict the target class. The purpose of classification is to break a subject into smaller, more manageable, more specific parts. Smaller subcategories help us make sense of the world, and the way in which these subcategories are created also helps us make sense of the world. A classification essay is organized by its subcategories. The advantages of classifying organisms are as follows: (i) Classification facilitates the identification of organisms. (ii) Helps to establish the relationship among various groups of organisms. (iii) Helps to study the phylogeny and evolutionary history of organisms. A disadvantage to classification is that many of the classifications themselves are based on subjective judgments, which may or may not be shared by everyone participating.

III. LITERATURE SURVEY

Mrunmayee Dhakate et al. [1] proposed an image processing and neural network methods to deal with the main issues of phytopathology i.e. disease detection and classification. The Pomegranate fruit as well as the leaves was affected by various diseases caused by fungus, bacteria and the climatic conditions .These diseases were Bacterial Blight, Fruit Spot, Fruit rot and Leaf spot. The system used some images for training, some for testing purpose and so on. The color images were pre-processed and underwent k-means clustering segmentation. The texture features were extracted using GLCM method, and given to the artificial neural network. The overall accuracy of this method was 90%.
Pranjali Padol et al. [4] used resizing, thresholding and Gaussian filtering for image preprocessing. The K-means clustering technique was used for segmentation of i mage then feature extraction was done using both texture as well as Muhammad Sharif et al. [10] proposed a system in which the citrus lesion spots were extracted by an

TABLE I. SUMMARY OF DIFFERENT IMAGE PROCESSING TECHNIQUES

Author	Title	Year	Segmentation Technique Applied	Feature Extraction Technique	Classifiers Used
Mrunmayee Dhakate et al.	Diagnosis of Pomegranate Plant Diseases using Neural Network	2015	k-Means clustering	Texture feature using GLCM method.	Artificial Neural Networks.
Gouri C.Khadabadi et al.	Identification and Classification of Diseases in Carrot Vegetable Using Discrete Wavelet Transform	2015	Canny edge detection technique	Discrete Wavelet Transform (DWT)	Probabilistic Neural Network
Bhavini J. Samajpati et al.	Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier	2016	k-Means clustering	Color and texture.	Random Forest classifier.
Pranjali B. Padol et al.	SVM Classifier Based Grape Leaf Disease Detection	2016	k-Means clustering	Color and texture.	SVM
Yang Lu et al.	Identification of rice diseases using deep convolutional neural networks	2017	k-Means clustering	Color and texture.	Deep convolutional neural networks (CNNs)
Alexander Johannes et al.	Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case	2017	Automated natural image leaf Segmentation and User input mask refinement	Local descriptors	Meta-classifier
Shanwen Zhanget et al.	Leaf image based cucumber disease recognition using sparse representation classification	2017	k-Means clustering	Shape and color	Sparse representation
Monzurul Islam et al.	Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine	2017	Entropy-based thresholding	Color and texture	support Vector machine.
Namrata R. Bhimte et al.	Diseases Detection of Cotton Leaf Spot using Image Processing and SVM Classifier	2015	Colors in a*b space using k-Means clustering	Color and texture using GLCM	upport Vector machine
Muhammad Sharif et al.	Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection	2018	Optimized weighted segmentation	Color, texture, and geometric features	Multi-Class Support Vector Machine (M-SVM)
Akshaya Aruraj et al.	Detection and Classification of Diseases of Banana Plant Using Local Binary Pattern and Support Vector Machine	2019	Not specified	Texture analysis using LBP	SVM and KNN
L. Sherly Puspha Annabel et al.	AI-Powered Image-Based Tomato Leaf Disease Detection	2019	Threshold based Segmentation was done.	Texture analysis	GLCM and random forest classifier
Gayatri Kuricheti et al.	Computer Vision Based Turmeric Leaf Disease Detection and Classification	2019	k-Means clustering	Texture analysis using GLCM	SVM
Abdul Hafiz Bin Abdul Wahab et al.	Detecting diseases in Chilli Plants Using K-Means Segmented Support Vector Machine	2019	k-Means clustering	Texture features and color features	Support Vector Machine (SVM).
Abirami Devaraj et al.	Identification of Plant Disease using Image Processing Technique	2019	k-Means clustering	Not specified	Random forest classifier

optimized weighted segmentation method, which was performed on an enhanced input image. Then, color, texture, and geometric color features. Then finally SVM classification technique was used to detect the type of leaf disease. In the experiment two classes of grape leaves were considered namely, Downy Mildew and Powdery Mildew. The given system gave 88.89% average accuracy for both Downey and Powderly grape leaf disease.

Yang Lu et al. [5] proposed novel rice disease identification method based on deep convolutional neural networks (CNNs) techniques. They used a dataset of 500 natural images of diseased and healthy rice leaves and stems captured from rice experimental field, CNNs were trained to identify 10 common rice diseases. Under the 10-fold cross-validation strategy, the proposed CNNs-based model achieved an accuracy of 95.48%. This accuracy was much higher than conventional machine learning model.

features were fused in codebook. Furthermore, the best features were selected by implementing a hybrid feature selection method, which consisted of PCA score, entropy, and skewness-based covariance vector. The selected features were fed to Multi-Class Support Vector Machine (M-SVM) for final citrus disease classification. The proposed technique was tested on Citrus Disease Image Gallery Dataset, Combined dataset (Plant Village and Citrus Images Database of Infested with Scale), and their own collected images database. They used these datasets for detection and classification of citrus diseases namely anthracnose, black spot, canker, scab, greening, and melanose. The proposed technique outperformed the existing methods and achieved 97% classification accuracy on citrus disease image gallery dataset, 89% on combined dataset and 90.4% on their local datasets.
Rahimi Zahari et al. [14] proposed an Artificial Intelligence based image processing algorithm to detect diseases on a Chilli plant using its leaves images. The proposed solution focused on using k means clustering algorithm for image segmentation and compared different Support Vector Machine (SVM) algorithm for classification. Computed images' features were extracted and used to classify these images into classes. Different parameters and different kernel functions were used to compute different SVM classification algorithms. The results were classified into background, healthy and unhealthy (Cucumber Mosaic) and could differentiate between health and unhealthy plant.

Research Gaps

1. There is a need for the development of cost effective Crop Disease Identification and Classification system.
2. Size of data sets was insufficient in most of the Research works, hence a large data set has to be considered for disease identification and classification for crops.
3. Need for combination of different features of image samples.
4. There is a great need to incorporate optimization techniques for Crop disease Identification and Classification.
5. Need for development of technological solutions to identify on-farm diseases for various crops.
6. Provide remedial suggestions and warning to farmers and agriculture experts for various crop diseases.

IV. CONCLUSION

Agriculture is an important sector which contributes to a nation's economy. With technical advancement in the field of image processing, machine learning, IOT etc. agriculture can be benefitted to a greater extent by use of these technologies. Many a times the crops are affected by diseases which lead in huge economic loss. A lot of work has been done to identify and predict diseases using image processing techniques. In this paper we have presented a survey on identification and classification of diseases for different types of crops using image processing and machine learning techniques .we have highlighted the summary of different segmentation, feature extraction and classification techniques used and also the research gaps have been mentioned. In future some of the techniques mentioned in the paper will be used for our research.

REFERENCES

[1] Mrunmayee Dhakate, Ingole A. B., " Diagnosis of Pomegranate Plant Diseases using Neural Network", 978-1-4673-8564-0/15/$31.00 ©2015 IEEE.

[2] Gouri C.Khadabadi, Prof. Arun Kumar and Dr.Vijay S. Rajpurohit, "Identification and Classification of Diseases in CarrotVegetable Using Discrete Wavelet Transform", International Conference on Emerging Research in Electronics, Computer Science and Technology – 2015, 978-1-4673-9563-2/15/$31.00 ©2015 IEEE.

[3] Bhavini J. Samajpati and Sheshang D. Degadwala, "Hybrid Approach for Apple Fruit DiseasesDetection and Classification Using Random Forest Classifier", International Conference on Communication and Signal Processing, April 6-8, 2016, India, 978-1-5090-0396-9/16/$31.00 ©2016 IEEE.

[4] Pranjali B. Padol and Prof. Anjali A. Yadav," SVM Classifier Based Grape Leaf Disease Detection", 2016 Conference on Advances in Signal Processing (CASP) Cummins College of Engineering for Women, Pune. Jun 9-11, 2016, 978-1-5090-0849-0/16/$31.00 ©2016 IEEE.

[5] Yang Lu , Shujuan Yi , Nianyin Zeng , Yurong Liu , Yong Zhang ," Identification of rice diseases using deep convolutional neural networks",http://dx.doi.org/10.1016/j.neucom.2017.06.023 0925-2312/© 2017 Elsevier B.V.

[6] Alexander Johannes , Artzai Picon , Aitor Alvarez-Gila , Jone Echazarra ,Sergio Rodriguez-Vaamonde, Ana Díez Navajas , Amaia Ortiz-Barredo,"Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case",http://dx.doi.org/10.1016/j.compag.2017.04.0130168-1699/_ 2017 Elsevier B.V.

[7] Shanwen Zhang a,b,⇑, Xiaowei Wuc, Zhuhong You a, Liqing Zhang," Leaf image based cucumber disease recognition using sparse representation classification", http://dx.doi.org/10.1016/j.compag.2017.01.014 0168-1699/_ 2017 Elsevier B.V.
[8] Monzurul Islam, Anh Dinh, Khan Wahid and Pankaj Bhowmik," Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine", 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 978-1-5090-5538-8/17/$31.00 ©2017 IEEE.
[9] Namrata R. Bhimte and V. R. Thool," Diseases Detection of Cotton Leaf Spot using Image Processing and SVM Classifier", Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018) IEEE Xplore Compliant Part Number: CFP18K74-ART; ISBN: 978-1-5386-2842-3,978-1-5386-2842-3/18/$31.00 ©2018 IEEE 340.
[10] Muhammad Sharifa, Muhammad Attique Khana, Zahid Iqbal, Muhammad Faisal Azam, M. Ikram Ullah Lali, Muhammad Younus Javed," Detection And classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection", 0168-1699/ © 2018 Elsevier.
[11] Akshaya Aruraj et al.," Detection and Classification of Diseases of Banana Plant Using Local Binary Pattern and Support Vector Machine",2019 International Conference on Signal Processing and Communication (ICSPC -2019), March. 29 – 30, 2019, Coimbatore, INDIA, 978-1-7281-1849-9 ©2019 IEEE.
[12] L. Sherly Puspha Annabel and V. Muthulakshmi," AI-Powered Image-Based Tomato Leaf Disease Detection", Proceedings of the Third International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2019) IEEE Xplore Part Number:CFP19OSV-ART; ISBN:978-1-7281-4365- 978-1-7281-4365-1/19/$31.00 ©2019 IEEE.
[13] Gayatri Kuricheti and Supriya P," Computer Vision Based Turmeric Leaf Disease Detection and Classification", Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019) IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439- 8
[14] Abdul Hafiz Bin Abdul Wahab and Rahimi Zahari," Detecting diseases in Chilli Plants Using K-Means Segmented Support Vector Machine", 2019 3rd International Conference on Imaging, Signal Processing and Communication, 978-1-7281-3663-9/19/$31.00 ©2019 IEEE.
[15] Abirami Devaraj, Karunya Rathan, Sarvepalli Jaahnavi and K Indira," Identification of Plant Disease using Image Processing Technique", International Conference on Communication and Signal Processing, April 4-6, 2019, India, 978-1-5386-7595-3/19/$31.00 ©2019 IEEE.
[16] Nilay Ganatra and Atul Patel," A Survey on Diseases Detection and Classification of Agriculture Products using Image Processing and Machine Learning", International Journal of Computer Applications (0975 – 8887) Volume 180 – No.13, January 2018.