iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: VIII Month of publication: August 2020

DOI: https://doi.org/10.22214/ijraset.2020.31104

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VIII Aug 2020- Available at www.ijraset.com

Crop Disease Detection by Machine Learning

Mohammadi Iram[1], Maherabanu[2], Ayesha Siddiqua[3], Prof. Prakash Pathan[4]
[1, 2, 3]PDA College of Engineering, Aiwan-E-Shahi Area, Shambhognlli, Gulbarga, Karnataka 585102
[4]Asst. Professor, Computer Science and Engineering Department, PDA College of Engineering, Aiwan-E-Shahi Area, Shambhognlli, Gulbarga, Karnataka 585102

Abstract: Plants have different kinds of diseases which can be caused by various factors like insects, pests, pathogens etc. Many agriculture people who depends on agriculture for their needs are facing many problems, since it becomes very difficult to manage the large area of the agriculture. Hence a method to help the agriculturist must be provided. The proposed system of our project provides a way for the detection of plant diseases.
Keywords: Camera, Plant images, Anaconda Navigator, Tensor Flow

I. INTRODUCTION

Machine learning is an application of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

Anaconda navigator is a desktop graphical user interface included in Anaconda distribution that allows you to launch applications and easily manage packages, environments and channels without using command line commands.

India is a land of agriculture. India's agriculture is composed of many crops. Agriculture is the main source for the raw materials for the production management process, agriculture accounts to the growth of the nation, the nations progress and economy depends on our agricultural lands and the farmers who carry out agriculture.

India is vast country and many crops are grown here which has their own benefits. The crops are also grown based on the season which crops can be grown when according to their nature, process and type of weather.

The major crops grown in India are rice, wheat, millets, pulses, tea, coffee, sugarcane, oil seeds, cotton and jute etc. Different parts of India grow different crops depending on the type of their land, location, weather, methods, uses etc.

The model and the main crop of our project is cotton. Cotton is a soft, fluffy staple fiber that grows in a boll, or protective case, around the genus Gossypium in the mallow family Malvaceae.The fiber is almost pure cellulose. Under natural conditions, the cotton bolls will increase the dispersal of the seeds.

Cotton is an essential raw material of India and it is also use for the economy purpose. Cotton is essentially produced for its fibre, which is universally used as a textile raw material. Cotton is an important commodity in the world economy.

A. Problem Statement

There are many of the problems faced during the agriculture by our farmers. The problems depends upon the diseases of the crop, their yield, lack of agricultural products, water problems etc

The main problem is the diseases in the crops. Hence early detection in the disease would help us. Therefore, the development of the successful product or a system in the field of leaf based picture disease detection will have a better outcome.

B. Existing System

The various modern technologies have emerged to minimize postharvest processing, to the agricultural sustainability and also the minimize the productivity.

Various researches and 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.

The main disadvantage of the existing system is that its not cost effective.

C. Proposed System

The proposed system has two phases, first phase deals with the training data. In this both healthy and diseased leaf images are taken from the farm. Once the dataset is ready with both the healthy as well as infected image samples, the threshold is extracted for both aging and for diseased.

Periodically images are obtained by remote sensing. RGB values are taken and then monitored images are extracted and compared with threshold images. If the threshold is greater or less than given value, histogram analysis and the edge detection techniques are used to identify particular plant diseases.

Different types of crops given as the input to the training model. For each of the crops we will be having many of the healthy, Infected and diseased crop images. Hence, set the threshold value for each of the crop. Then, accordingly train you're model in such a way that it should give the proper result for all types of crops.

II. LITERATURE SURVEY

A. Some Supporting Research Papers

Authors have discussed leaf diseases on cotton plant must be identified early and accurately as it can prove detrimental to the yield. The proposed work presents a pattern recognition system for identification and classification of three cotton leaf diseases i.e. Bacterial Blight, Myrothecium and Alternaria. The images required for this work are captured from the fields at Central Institute of Cotton Research Nagpur, and the cotton fields in Buldana and Wardha district. Active contour model is used for image segmentation and Hu' s moments are extracted as features for the training of adaptive neuro-fuzzy inference system. The classification accuracy is found to be 85 percent. [1]

Here the authors presents a survey on methods that use digital image processing techniques to detect, quantify and classify plant diseases from digital images in the visible spectrum. Although disease symptoms can manifest in any part of the plant, only methods that explore visible symptoms in leaves and stems were considered. This was done for two main reasons: to limit the length of the paper and because methods dealing with roots, seeds and fruits have some peculiarities that would warrant a specific survey. The selected proposals are divided into three classes according to their objective: detection, severity quantification, and classification. Each of those classes, in turn, are subdivided according to the main technical solution used in the algorithm. This paper is expected to be useful to researchers working both on vegetable pathology and pattern recognition, providing a comprehensive and accessible overview of this important field of research. [2]

Recognizing plants is a vital problem especially for biologists, chemists, and environmentalists. Plant recognition can be performed by human experts manually but it is a time consuming and low-efficiency process. Automation of plant recognition is an important process for the fields working with plants. This paper presents an approach for plant recognition using leaf images. Shape and colour features extracted from leaf images are used with k-Nearest Neighbor, Support Vector Machines, Naive Bayes, and Random Forest classification algorithms to recognize plant types. The presented approach is tested on 1897 leaf images and 32 kinds of leaves. The results demonstrated that success rate of plant recognition can be improved up to 96% with Random Forest method when both shape and colour features are used is discussed in [3]

the authors here studied , Accurate estimation of ventricular volumes plays an essential role in clinical diagnosis of cardiac diseases. Existing methods either rely on segmentation or are restricted to direct estimation of the left ventricle. In this paper, we propose a novel method for direct and joint volume estimation of bi-ventricles, i.e., the left and right ventricles, without segmentation and user inputs. Based on the cardiac image representation by multiple and complementary features, we adopt regression forests to jointly estimate the two volumes. Our method is validated on a dataset of 56 subjects with a total of 3360 MR images which shows that our method can achieve a high correlation coefficient of around 0.9 with manual segmentation obtained by human experts. With our proposed method, the most daily-used estimation of cardiac function, e.g., ejection fraction, can be conducted in a much more efficient, accurate and convenient way. [4]

Automatic methods for an early detection of plant diseases are vital for precision crop protection. The main contribution of this paper is a procedure for the early detection and differentiation of sugar beet diseases based on Support Vector Machines and spectral vegetation indices. The aim was (I) to discriminate diseased from non-diseased sugar beet leaves, (II) to differentiate between the diseases Cercospora leaf spot, leaf rust and powdery mildew, and (III) to identify diseases even before specific symptoms became visible. Hyperspectral data were recorded from healthy leaves and leaves inoculated with the pathogens Cercospora beticola, Uromyces betae or Erysiphe betae causing Cercospora leaf spot, sugar beet rust and powdery mildew, respectively for a period of 21 days after inoculation. Nine spectral vegetation indices, related to physiological parameters were used as features for an automatic classification. Early differentiation between healthy and inoculated plants as well as among specific diseases can be achieved by a Support Vector Machine with a radial basis function as kernel is discussed in. [5]

The authors studied that , this study describes an image-processing based method that identifies the visual symptoms of plant diseases, from an analysis of coloured images. The processing algorithm developed starts by converting the RGB image of the diseased plant or leaf, into the H, I3a and I3b colour transformations. The I3a and I3b transformations are developed from a modification of the original I1I2I3 colour transformation to meet the requirements of the plant disease data set. The transformed image is then segmented by analysing the distribution of intensities in a histogram. Rather than using the traditional approach of selecting the local minimum as the threshold cut-off, the set of local maximums are located and the threshold cut-off value is determined according to their position in the histogram. This technique is particularly useful when the target in the image data set is one with a large distribution of intensities. In tests, once the image was segmented, the extracted region was post-processed to remove pixel regions not considered part of the target region. This procedure was accomplished by analysing the neighbourhood of each pixel and the gradient of change between them. To test the accuracy of the algorithm, manually segmented images were compared with those segmented automatically. Results showed that the developed algorithm was able to identify a diseased region even when that region was represented by a wide range of intensities [6]

Here authors has gone through survey that, Tomato represents an important vegetable crop worldwide. During cropping cycle several diseases and abnormal conditions may affect tomato plants resulting on considerable losses of production. A precise identification of these pathologies in early phases is fundamental for the implementation of control strategies. Nevertheless, the right identification of symptoms of plants diseases require highly specialized knowledge and facilities, which are not available for small growers. Recently, computer vision tools have been proposed as an alternative for tomato diseases characterization. These works mainly focus on identification of affected regions and classification tasks. Nevertheless, non-specialists may lack of clarity about what they are looking for during the assessment. In these cases, Content Based Image Retrieval (CBIR) systems can be helpful as a complementary strategy to improve the quality of the search by allowing exploration of databases with supplementary information. This work presents a novel strategy for image retrieval of tomato leaves for greenhouse crops suitable to support disease diagnosis. The strategy is based on color structure descriptors and nearest neighbors. Experimental results show that the proposed approach can successfully characterize in several abnormal conditions, such as, chlorosis, sooty moulds and early blight. [7]

The authors discussed that ,this study describes an image-processing based method that identifies the visual symptoms of plant diseases, from an analysis of coloured images. The processing algorithm developed starts by converting the RGB image of the diseased plant or leaf, into the H, I3a and I3b colour transformations. The I3a and I3b transformations are developed from a modification of the original I1I2I3 colour transformation to meet the requirements of the plant disease data set. The transformed image is then segmented by analysing the distribution of intensities in a histogram. Rather than using the traditional approach of selecting the local minimum as the threshold cut-off, the set of local maximums are located and the threshold cut-off value is determined according to their position in the histogram. [8]

This paper reports research results on developing a machine vision system to assess the quality of row crop plants. Comparing to the prevalent machine vision system employed in agricultural industry for weed-crops classification as well as plant density evaluation, the proposed machine vision system is able to detect the location of plants (weed / crops) and calculate the leaves; area for plant quality assessment, even if the leaves are overlapped with each other. The developed machine vision system involves a camera system and an image processing system. The camera system uses a coaxial camera constructed by a RGB sensor and near infrared (NIR) sensor, which cooperate with a white front lighting and NIR front lighting respectively. Plants are firstly captured by the coaxial camera. The plants are segmented from background on RGB image; the overlapping edges of leaves are detected on NIR image. Afterwards the overlapping leaves are separated and assigned to the assessed stem position of plants. At last, based on the assigned leaves, the plants are separated, and the area of plant canopy is calculated. A set of experiments have been made to prove the feasibility of the proposed machine vision system is discussed in. [9]

Here the authors discussed, the latest generation of convolutional neural networks (CNNs) has achieved impressive results in the field of image classification. This paper is concerned with a new approach to the development of plant disease recognition model, based on leaf image classification, by the use of deep convolutional networks. Novel way of training and the methodology used facilitate a quick and easy system implementation in practice. The developed model is able to recognize 13 different types of plant diseases out of healthy leaves, with the ability to distinguish plant leaves from their surroundings. According to our knowledge, this method for plant disease recognition has been proposed for the first time. All essential steps required for implementing this disease recognition model are fully described throughout the paper, starting from gathering images in order to create a database, assessed by agricultural experts. Caffe, a deep learning framework developed by Berkley Vision and Learning Centre, was used to perform the deep CNN training. The experimental results on the developed model achieved precision between 91% and 98%, for separate class tests, on average 96.3%. [10]

Authors studied the prevention and control of plant disease have always been widely discussed because plants are exposed to outer environment and are highly prone to diseases. Normally, the accurate and rapid diagnosis of disease plays an important role in controlling plant disease, since useful protection measures are often implemented after correct diagnosis. This system is based on image processing technology and uses MATLAB as the main processing tool. Besides, digital image processing, mathematical statistics, plant pathology, and other relative fields are also considered. Comparing to the traditional image recognition, there are plenty of innovations in image segmentation and system construction. To strengthen the division of the lesion, users have different creative interactive options to meet their own needs. Meanwhile, linear regression model can be used in various types of plant disease.After the establishment of multiple linear regression model, the images from training libraries are placed into the multiple linear regression model; then the disease recognition system is constructed by using the least squares method. Use the images inside and outside the training libraries to test the accuracy of the system.[11]

Agriculture is the mainstay of the Indian economy. Immense commercialization of an agriculture has creates a very negative effect on our environment.

The use of chemical pesticides has led to enormous levels of chemical build-up inour environment, in soil, water, air, in animals and even in our own bodies. Artificial fertilizers gives on a short-term effect on productivity but a longer-term negative effect on the environment, where they remain for years after leaching and running off, contaminating ground water. Another negative effect of this trend has been on the fortunes of the farming communities worldwide. The central activity oforganic farming relies on fertilization, pest and disease control.

The proposed system are discussed to use recent detectors such as Faster Region-Based Convolutional Neural Network (Faster R-CNN), Region-based Fully Convolutional Networks (R-FCN) and Single Shot Multi box Detector (SSD) to detection and classification of plant leaf diseases that affect in various plants. The challenging part of our approach is not only deal with disease detection, and also known the infection status of the disease in leaves and triesto give solution (i.e., name of the suitable organic fertilizers)for those concern diseases. [12]

The author described as an in-field automatic wheat disease diagnosis system based on a weekly supervised deep learning framework, i.e. deep multiple instance learning, which achieves an integration of identification for wheat diseases and localization for disease areas with only image-level annotation for training images in wild conditions. Furthermore, a new in-field image dataset for wheat disease, Wheat Disease Database2017 (WDD2017), is collected to verify the effectiveness ofour system. Under two different architectures, i.e. VGG-FCN-VD16 and VGG-FCN-S, our system achieves the meanrecognition accuracies of 97.95% and 95.12% respectively over5-fold cross validation on WDD2017, exceeding the results of 93.27% and 73.00% by two conventional CNN frameworks, i.e. VGG-CNN-VD16 and VGG-CNN-S. Experimental results demonstrate that the proposed system outperforms conventional CNN architectures on recognition accuracy under the same amount of parameters, meanwhile maintaining accurate localization for corresponding disease areas.

Moreover, the proposed system has been packed into a real-time mobile app to provide support for agricultural diseased diagnosis is discussed in [13]

The authors discussed about convolution neural network models were developed to perform plant diseased detection and diagnosis using simple leaves images of healthy and diseased plants, through deep learning methodologies. Training of the models was performed with the use of an open database of 87,848 images, containing 25 different plants in a set of 58 distinct classes of [plant, disease] combinations, including healthy plants. Several model architectures were trained, with the best performance reaching a 99.53% success rate in identifying the corresponding [plant, disease] combination (or healthy plant). The significantly high success rate makes the model a very useful advisory or early warning tool, and an approach that could be further expanded to support an integrated plant disease identification system to operate in real cultivation conditions. [14]

The authors studied that , one of the major factors responsible for the crop destruction is plantdisease. Different plants suffer from different diseases. The main part of plant to examine the disease is leaf. The major categories of plant leaf diseases are based on viral, fungal and bacteria. The diseases on leaf can reduce both the quality and quantity of crops and their further growth. The easy method to detect the plant diseases is with the help of agricultural expert having knowledge of plant diseases. But this manual detection of plant diseases takes lot of time and is a laborious work. Hence, there is a need for machine learning method to detect the leaf diseases. Computer can play a major role to develop the automatic methods for the detection and classification of leaf diseases. There can be various pattern recognition and image processing techniques that can be used in the leaf disease detection. The leaf disease detection and classification of leaf diseases is the key to prevent the agricultural loss. Different plant leaves bear different diseases. There are different types of methods and classifiers to detect plant leaf diseases. It has made an attempt to study machine learning methods which are used by researchers for disease identification and classification of plants. These machine learning methods help agricultural experts in detection of disease in the plant in timely fashion, then the experts will suggest the medicines to the farmer. [15]

III. REQUIREMENTS

A. Hardware Requirements

1) *Processor:* Any processor above 500 MHz
2) *Ram:* 4 GB
3) *Hard Disk:*250 GB
4) *Input Device:* Standard Keyboard and Mouse
5) *Output Device:* High Resolution Monitor.

B. Software Requirements

1) *Operating System:* Windows 7 or higher
2) *Programming:* Python 3.6 and related libraries.

IV. PROPOSED ARCHITECTURE

To find out whether the leaf is diseased or healthy ,certain steps must be followed i.e. preprocessing,feature extraction, training of classifier and classification.

In pre-processing of image we will bring all the images size to reduce uniform size. Then HoG is a feature descriptor used for object detection. In this feature descriptor the appearance of the object and the outline of the image is described by its intensity gradients.

Here we made use of three feature descriptors.Hu moments: Image moments which have the important characteristics of the image pixels helps in describing the objects. Here Hu moments help in describing the outline of a particular leaf.

The first step involves converting RGB to Gray scale and then the Hu moments are calculated. Haralick Texture: usually the healthy leaves and diseased leaves have different textures. Here we use Haralick texture feature to distinguish between the textures of healthy and diseased leaf.

Colour Histogram: Colour histogram gives the representation of the colours in the image. RGB is first converted to HSV colour space and the histogram is calculated for the same. It is needed to convert the RGB image to HSV since HSV model is closely related with human eye discerns the colours in an image.

V. EXPERIMENTAL RESULTS

We prepared the software system and programmed the plant images to first get the images of both healthy and diseased images. Then we use different techniques to perform the experimentation. When the image is captured, it is loaded and adjusted according to the platform based on the type of the crop and its disease. We then get the result whether the crop is healthy, infected or diseased. Therefore, from that we can use different remedies to protect our crops from those diseases.

VI. CONCLUSION AND FUTURE SCOPE

In our project we have discussed the method how we can detect the crop disease using machine learning.

We first classify the crop based on its type, then depending upon it, we classify and identify the cotton leaf.

This paper presents the design of an ML based system which provides easy accessible real-time local environmental data in rural crop fields. The data are pushed in real-time to easily accessible cloud storage, providing researches and crop field managers with accurate environmental data without the need for visits to the crop field to retrieve local data

REFERENCES

[1] P. R. Rothe , R. V. Kshirsagar , “Cotton Leaf Disease Identification using Pattern Recognition Techniques”, International Conference on Pervasive Computing (ICPC), 2015, pp.9-14.
[2] J. G. A. Barbedo, “Digital image processing techniques for detecting quantifying and classifying plant diseases”, Springer Plus, vol. 2, no. 660, 2013, pp. 1-12.
[3] A. Caglayan , O. Guclu , A. B. Can, “A plant recognition approach using shape and color features in leaf images”, International Conference on Image Analysis and Processing, pp.161-170.
[4] X. Zhen, Z. Wang, A. Islam, I. Chan, S. Li, “Direct estimation of cardiac bi-ventricular volumes with regression forests”, Accepted by Medical Image Computing and Computer Assisted Intervention MICCAI 2014, pp.10-20.
[5] T. Rampf, A. K. Mahlein, U. Steiner, E. C. Oerke, H.W. Dehne, L. Plumer, “Early detection and classification of plant diseases with Support Vector Machine based on Hyperspectral reflectance”, Computers and Electronics in Agriculture, vol. 74, no. 1, October 2010, ISSN 0168-1699 , pp.91-99.
[6] A. Camargo, J. S. Smith, “An image processing based algorithm to automatically identify plant disease visual symptoms”, Biosystems Engineering, vol. 102, no. 1,2008, pp. 9-21.
[7] D. Baquero, J. Molina, R. Gily, C. Bojacáy, H. Franco, A. F. Gómez, “An image retrieval system for tomato disease assessment”, Symposium on Image Signal Processing and Artificial Vision (STSIVA) XIX,2014, pp. 1-5.

[8] A. Camargo, J. S. Smith, "An image processing based algorithm to automatically identify plant disease visual symptoms", Biosystems Engineering, 2008 ,vol. 102, no. 1, pp. 9-21.
[9] M. Weyrich, Yongheng Wang, M. Scharf, “Quality assessment of row crop plants by using a machine vision system”, IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society,2013, pp. 2466-2471.
[10]]Srdjan Sladojevic et al., “Deep neural networks based recognition of plant diseases by leaf image classification”, Computational intelligence and neuroscience, vol. 2016, pp.12-20.
[11] Guiling Sun , Xinglong Jia , and Tianyu Geng “Plant Diseases Recognition Based on Image Processing Technology” 21 May 2018 , pp.15-22.
[12] “Detection and Classification of Plant Leaf Diseasesby using Deep Learning Algorithm” International Journal of Engineering Research & Technology (IJERT)ISSN: 2278-0181,2018. pp.9-16.
[13] Jiang Lu, Jie Hu, Guannan Zhao, Fenghua Mei, Changshui Zhang, “Anin-field automatic wheat disease diagnosis system, Computers andElectronics in Agriculture” 2017, pp. 369–379.
[14] Konstantinos P. Ferentinos “Deep learning models for plant diseasedetection and diagnosis Computers and Electronics in Agriculture”(2018) pp.311–318.
[15] Goutum Kambale, Dr.Nitin Bilgi “A Survey Paper on Crop Disease Identification and Classification Using Pattern Recognition and Digital Image Processing Techniques” 2017,IOSR Journal of Computer Engineering (IOSR-JCE)e-ISSN: 2278-0661,p-ISSN: 2278-8727,pp. 14-17.
[16] Anuradha Badge, "crop disease detection using machine learning: Indian Agriculture",vol:05 Issue:09 sept 2018.
[17] Shima Ramesh, Nivedita M, Pooja R, Prasad Bhat N, Shashank N, Mr.Ramachandra Hebbar,Mr,P V Vinod, "plant disease detection using machine learning"