International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-08, Issue-12, Mar 2023

A Novel Approach for Cotton Plant Leaves Disease Prediction System

Dr.N. Revathy[1], Dr.K.Padmavathi[2], M.Kavirajan[3],A.Pravin[4]

[1]Professor, [2]Assistant Professor,[3,4]Final MCA, [1,3,4] PG & Research Department of Computer Applications, [2]Department of UG Computer Science, PSGRKCW College for Women, [1,3,4] Hindusthan College of Arts & Science (Autonomous), Coimbatore, India

Abstract **Crop diseases are a noteworthy risk to sustenance security; however their quick distinguishing proof stays troublesome in numerous parts of the world because of the non attendance of the important foundation. Emergence of accurate techniques in the field of leaf- based image classification has shown impressive results. This paper makes use of Random Forest in identifying between healthy and diseased leaf from the data sets created. Our proposed paper includes various phases of implementation namely dataset creation, feature extraction, training the classifier and classification. The created datasets of diseased and healthy leaves are collectively trained under Random Forest to classify the diseased and healthy images. For extracting features of an image we use Histogram of an Oriented Gradient (HOG). Overall, usingmachine learning to train the large data sets available publicly gives us a clear way to detect the disease present in plants in a colossal scale.**

Keywords —Healthy and Diseased, HOG, CNN, Prediction, Machine learning.

I. INTRODUCTION:

Quality image of plant leaves. Data set need to be considered in large amount. Acquired images are affected by background data and noises. Segmenting the exact spot in a leaf into meaningful disease. Preparation of training and testing samples from input image. Classification plays a role in recognizing segmented spot into meaningful disease. Colour of plant leaf, size and texture are varying when climate changed. Regular observation is needed for particular plants. Identifying diseases for different plant leaves is challenging. Reviews suggest that image processing and machine learning techniques have more potential to find diseases so, there has to be improving in existing research. Many researchers had done research on various plants and their diseases also they had given some techniques to identify that disease. Automation of identifying disease entails the input date collected from different sources. In this review, considering all different research papers we are identifying and discussing key issues, challenges on disease and techniques are as follows.

II. MODULE DESCRIPTION:

Convolutional Neural Network (CNN):

Convolutional Neural Network knows as CNN consists of the Input layer, Middle layers, and the Output layer. Input layer is the one which accepts features as the input, in other words, images are given as input through this layer. The Middle layer consists of the desired number of nodes based on the application. Output layer produces an output. It performs a convolutional operation over the pixel values in it along with the kernel matrix. The kernel matrix is slid over the pixel matrix, and value is determined.

Augmentation:

In this stage, every image is augmented. This augmentation is done by using (Python), in which the different types of transformation are applied to the given image every time, every epoch. Since the transformation is done during the run time of the program, it does not require any additional memory space for the images to store, which is an added advantage. The Image Data Generator class is the PyCode which is responsible for the transformation of the data.

Dataset:

The data collection for our work is done in two ways. One is collecting data from the freely available open source dataset namely (Plant Village). This dataset will consist of different types of plants and their diseases. In order to exploit this model under a real-time scenario, we have added image manually other fields (Images belonging to classes listed in the primary data set – which exhibits different properties).

Evaluation metrics

Here the metric for ranking this model is Confusion Matrix. Based on the images supplied for testing, we have calculated the confusion matrix. There two different conditions for training and testing. One is under the lab

 DOI : 10.35291/2454-9150.2023.0059

conditions, which means that the model is tested with the images from the same dataset from which it is used for both training and testing.

III. SYSTEM STUDY

A. EXISTING SYSTEM:

During the early stages, the diseases of the plant are detected by early symptoms. Image processing technique was used to provide analytical result about the image of the plant from its color specifications. Along with the image processing technique K-means clustering algorithm has also been applied for the identification of the pixel of the disease and finally, Fuzzy Logic has been implemented for the classification of diseases.

DISADVANTAGES:

Indian economy is dependent of agricultural productivity. Over 70% of rural homes depend on agriculture. Agriculture pays about 17% to the total GDP and provides employment to over 60% of the population. Therefore detection of plant diseases plays a vital key role in the arena of agriculture. Indian agriculture is composed of many crops like rice, wheat. Indian farmers also grow sugarcane, oilseeds, potatoes and non-food items like coffee, tea, cotton, rubber. All these crops grow based on strength of leaves and roots. There are things that lead to different disease for the plant leaves, which spoiled crops and finally it will effect on economy of the country. These big losses can be avoided by early identification of plant diseases. Accurate detection of plant disease is needed to strengthen the field of agriculture and economy of our country. Various types of Disease kill leaves in a plant. Farmers get more difficulties in identifying these diseases, they are unable to take precaution on those plants due to lack of knowledge on those diseases.

B. PROPOSED SYSTEM:

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. Preprocessing of image is bringing all the images size to a reduced uniform size. Then comes extracting features of a preprocessed image which is done with the help of HOG.

The algorithm here is implemented using random forests classifier. They are flexible in nature and can be used for both classification and regression techniques. Compared to other machine learning techniques like SVM, Gaussian Naïve bayes, logistic regression, linear discriminate analysis, Random forests gave more accuracy with less number of image data set.

ADVANTAGES:

The early detection of diseases is important in agriculture for an efficient crop yield. The bacterial spot, late blight, septoria leaf spot and yellow curved leaf diseases affect the crop quality of tomatoes. Automatic methods for classification of plant diseases also help taking action after detecting the symptoms of leaf diseases. This paper presents a Convolutional Neural Network (CNN) model and Learning Vector Quantization (LVQ) algorithm based method for tomato leaf disease detection and classification. The dataset contains 500 images of tomato leaves with four symptoms of diseases. We have modeled a CNN for automatic feature extraction and classification.

In this work, they have proposed a method for the classification of the plant diseases. All the classification was done based upon the images of the crop's leaf, which contains both the healthy and affected leaf. This model proposed by them was able to classify a count of 13 different types of crop diseases. The dataset used here consist of higher resolution, resized and cropped images. Augmentation process was also used for avoiding the overfitting issues in CNN. Their model has produced an accuracy of about 96.3%. In this Deeper Network architecture is implemented for the grading of plant species. In their work colored images were used, their model can be used only for colored images, it was able to identify weed plant species. The total number of images is over 10,000 with 22 different types of weeds and crops. Their result produced an accuracy of 86.2% which is considered to be less accurate.

IV. SOFTWARE DESCRIPITION

Requirements Specification

Use this Requirements Specification template to document the requirements for your product or service, including priority and approval. Tailor the specification to suit your project, organizing the applicable sections in a way that works best, and use the checklist to record the decisions about what is applicable and what isn't.

- The format of the requirements depends on what works best for your project.
- This document contains instructions and examples which are for the benefit of the personwriting the document and should be removed before the document is finalized.
- To regenerate the TOC, select all (CTL-A) and press F9.

Executive Summary

Describe this project or product and its intended audience, or provide a link or referenceto the project charter.

Purpose and Scope of this Specification

Describe the purpose of this specification and its intended audience. Include a description of what is within the

 DOI : 10.35291/2454-9150.2023.0059

ISSN : 2454-9150 Vol-08, Issue-12, Mar 2023

scope what is outside of the scope of these specifications. For example:

Out of Scope

The following items in phase 3 of Project A are out of scope: modification of Classification Processing to meet legislative mandate XYZ. modification of Labor Relations Processing to meet legislative mandate XYZ. (Phase 3 will be considered in the development of the requirements for Phase 2, but the Phase 3 requirements will be documented separately.)

Product/Service Description

In this section, describe the general factors that affect the product and its requirements. This section should contain background information, not state specific requirements (provide the reasons why certain specific requirements are later specified).

Product Context

How does this product relate to other products? Is it independent and self-contained? Does it interface with a variety of related systems? Describe these relationships or use a diagram to show the major components of the larger system, interconnections, and external interfaces.

User Characteristics

Create general customer profiles for each type of user who will be using the product.

Profiles should include:

- Student/faculty/staff/other Experience
- Technical expertise
- Other general characteristics that may influence the product

Assumptions

List any assumptions that affect the requirements, for example, equipment availability, user expertise, etc. For example, a specific operating system is assumed to be available; if the operating system is not available, the Requirements Specification would then have to change accordingly.

Constraints

Describe any items that will constrain the design options, including parallel operation with an old system audit functions (audit trail, log files, etc.) access, management and security criticality of the application system resource constraints (e.g., limits on disk space or other hardware limitations) other design constraints (e.g., design or other standards, such as programming language or framework)

Dependencies

List dependencies that affect the requirements.

Examples:

This new product will require a daily download of data from X, Module X needs to be completed before this module can be built.

Requirements

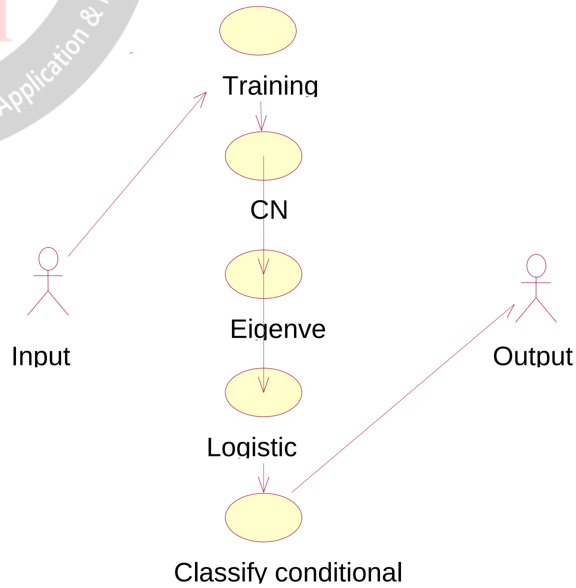
Describe all system requirements in enough detail for designers to design a system satisfying the requirements and testers to verify that the system satisfies requirements. Organize these requirements in a way that works best for your project. See 00, Organizing the Requirements for different ways to organize these requirements. Describe every input into the system, every output from the system, and every function performed by the system in response to an input or in support of an output. (Specify what functions are to be performed on what data to produce what results at what location for whom.)

Each requirement should be numbered (or uniquely identifiable) and prioritized.See the sample requirements in Functional Requirements, and System interface/**Integration**, as well as these example priority definitions.

User Interface Requirements

In addition to functions required, describe the characteristics of each interface between the product and its users (e.g., required screen formats/organization, report layouts, menu structures, error and other messages, or function keys).

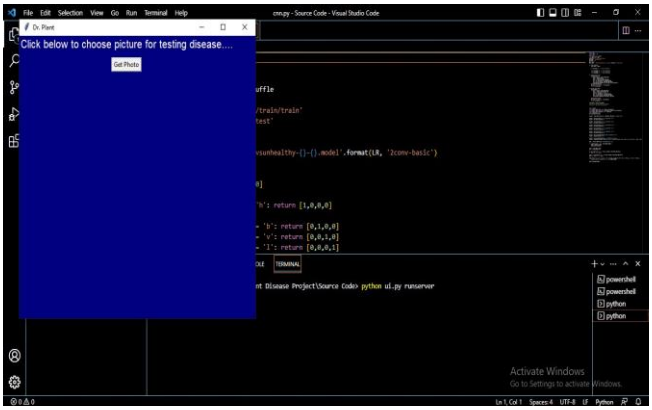
V. BLOCK DIAGRAM

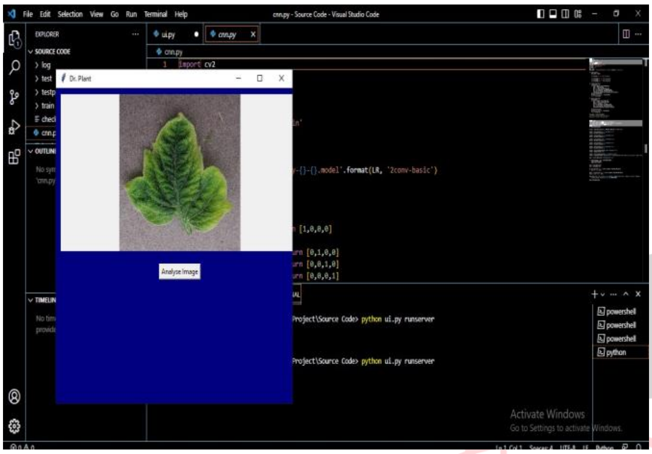


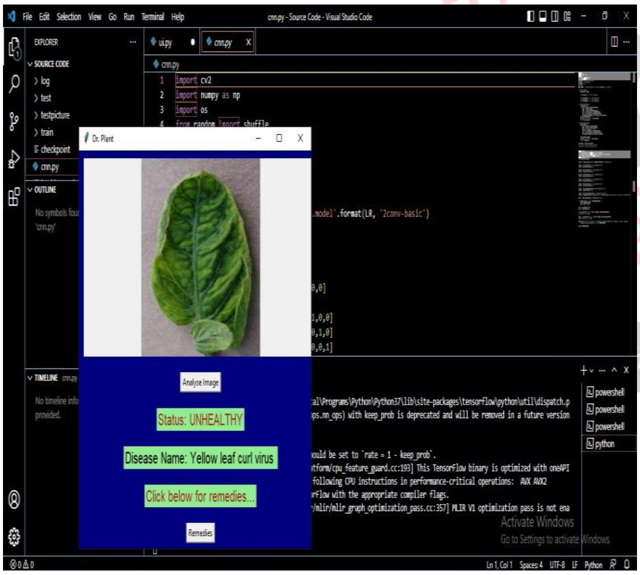
DOI : 10.35291/2454-9150.2023.0059

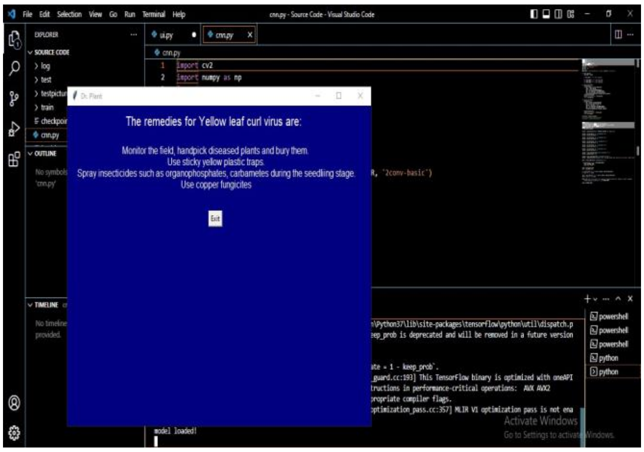
ISSN : 2454-9150 Vol-08, Issue-12, Mar 2023

VI. RESULT









VII. CONCLUSION

Thus a model built for the identification of disease affected plants and healthy plants is done and this proposed work is focused on the accuracy values during the real field conditions, and this work is implemented by having Ten number of plant disease classes. Overall this work is implemented from scratch and produces an accuracy of 86%. The future work is to increasethe number of classes present in the open database (Plant Village) and to modify the architecture in accordance with the dataset for achieving better accuracy. The other condition is that field condition; this means that our model has tested with the images taken from the real world conditions (land). Since the lighting conditions and background properties of the images are totally different when we take samples from the real field, there is a chance that our model to produce a very low accuracy, when comparing to the accuracy values acquired during the lab conditions. So to overcome this impact, we had an idea of having a mixed variety of images during the training phase (heterogeneity).

REFERENCES

1. User Interfaces in C#: Windows Forms and Custom Controls by Matthew MacDonald.

2. Applied Microsoft® .NET Framework Programming (Pro-Developer) by Jeffrey Richter.

3. Practical .Net2 and C#2: Harness the Platform, the Language, and the Framework by Patrick Smacchia.

4. Data Communications and Networking, by Behrouz A Forouzan.

5. Computer Networking: A Top-Down Approach, by James F. Kurose.

6. Operating System Concepts, by Abraham Silberschatz.

7. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D.

8. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley, Tech. Rep. USB-EECS-2009- 28, Feb 2009.

9. "The apache cassandra project," http://cassandra.apache.org/

10. L. Lamport, "The part-time parliament," ACM Transactions on Computer Systems, vol. 16, pp. 133–169, 1998.

 DOI : 10.35291/2454-9150.2023.0059