Plant Leaves Disease Detection System Using Machine Learning

Prof. Shailesh Kurzadkar; Achal Meshram; Aman Barve; Kajal Dhargave; Mruganayani Alone; Vijaya Bhongale

meshramachal86@gmail.com, amanbarve9@gmail.com, kajal.dhargave@gmail.com,
mrugnainaalone123@gmail.com, vijayabhongale@gmail.com

Dept. of Computer Technology, KDK College of Engineering, Nandanvan, Nagpur

DOI: 10.47760/ijcsmc.2022.v11i02.004

Abstract:

Plants and crops are one of the major human needs in day to day life. And detection of plant disease is also essential through some automatic disease detection technique. It will reduce a lot of work of monitoring crops in big farms and at early stage disease gets detected through the symptoms on leaves. The earlier disease detection may lead to longer survival of crops and plants. Data sets will be created for identifying diseased and healthy plants. The created datasets of diseased and healthy leaves will be trained to classify both types of images collectively. Overall using machine 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— Leaf Disease, Image Classification, Convolution Neural Networks (CNN)

I. INTRODUCTION

In agricultural field, detection of the plant diseases is important. Plant Leaf Disease Detection will be very helpful for the farmers and other users or gardeners. This project aims to create a system that identifies the plant disease and monitor a plant. Hence, image processing is used for detection of plant disease. This system will capture a picture of plant and will compare it with collected data in the hardware. With the help of convolutional neural networks (CNN) system it will classify the image and detect whether the plant is healthy or not, or diseased or not.

II. LITERATURE REVIEW

- [1] Anand H.Kulkarni et al. , wherever Gabor filter has been used for feature extraction and ANN primarily based classifier has been used for classification with recognition rate up to ninety one.
- [2] F. Argenti,et al. projected a fast algorithmic rule for scheming parameters of co-occurrence matrix by supervised learning and most probability technique for quick classification.
- [3] P. Revathi et al the extracted edge options area unit utilized in classification to identify the sickness spots. The projected unvaried peal investigating technique for cotton diseases detection (HPCCDD) algorithmic rule has been used for categorizing the diseases. They claim the accuracy of 98.1% over existing algorithmic rule.

III. ALGORITHM DESCRIPTION

1. Convolutional Neural Network (CNN)

NN (Neural network) has many types, one of which is the Convolutional Neural Network (CNN) . CNN is the best-known method used for processing images data compared to other types of NN. Therefore, in this study the identification of rice plants diseases was carried out using CNN method. From this study, better results were obtained compared to other methods, obtaining 100% accuracy for training data and 93.67% for testing data. The model obtained by the CNN method can be used for detecting 3 different types of Rice plant diseases, there are brown spots, leaf smuts, or bacterial Leaf blight disease based on the physical images of rice plant leaves.

2. Random Forest Algorithm

Here Random Forest Algorithm is used for learning method for Classification and other tasks. Random Forest could be a classifier that contains variety of decision trees on colorful subsets of the given dataset and takes the typical to enhance the prophetic delicacy of that dataset." rather than looking forward to one decision tree, the arbitrary timber takes the vaticinator from each tree and supported the bulk votes of prognostications, and it predicts the ultimate affair. Since the arbitrary timber combines multiple trees to prognosticate the order of the dataset, it's possible that some decision trees may prognosticate the right affair, while others might not. But together, all the trees prognosticate the right affair. Thus, below are two hypothetical's for an advanced Random timber classifier. Below are some points that designate why we must always use the Random Forest algorithm. It takes lower training time as compared to other algorithms. It predicts affair with high delicacy, indeed for the big dataset it runs efficiently. It may also maintain delicacy when an large proportion of knowledge is missing. There should be some factual values within the point variable of the dataset in order that the classifier can prognosticate accurate results rather of a guessed result.

3. Breadth First Algorithm (BFA)

Breadth initial Search (BFS) is associate algorithmic program for traversing or looking layer wise in tree or graph data structures Depth initial search (DFS) that explores the nodes as so much as doable (depth-wise) before being forced to get back and explore alternative nodes. The algorithmic program starts at the tree root (or any whimsical node of a graph referred to as 'source node'), and investigates all of the neighboring nodes (directly connected to supply node) at the present level before moving on to the nodes at the subsequent level. The strategy is perennial till the required results obtained.

4. Depth First Search (DFS)

Depth first searches (DFS) which explores the nodes as far as possible (depth-wise) before being forced to backtrack and explore other nodes. The algorithm starts at the tree root (or any arbitrary node of a graph called 'source node'), and investigates all of the neighboring nodes (directly connected to source node) at the present level before moving on to the nodes at the next level. The process is repeated until the desired result is obtained.

IV. PROPOSED METHODOLOGY

Our proposed system is a web-based software. We have a web-based software which simply takes the image of the plant and uploads it to the mobile device. Then this image is shipped through a Convolutional Neural Network which encodes this image into a numerical array and classifies it with the opposite numerical arrays within the model. The model may be a tensor flow model which is formed into a tensor flow life model due to the massive size of the conventional tensor flow model. This model helps classify the uploaded image numerical values into the dataset values. When a numeric array matches it calculates the arrogance and display the worth which has the best confidence. In this way we are able to make sure that we always have the very best confidence value showing within the results.

The process of detection of plant disease involves basically four phases:

1. Image Acquisition
2. Segmentation
3. Feature Extraction
4. Classification

1. Image Acquisition:

In this section, pictures of plant leaves are gathered with the help of digital media, cameras, with desired resolution and size. The images will even be taken from net. The formation of information of pictures is completely dependent on the appliance system developer. The image information is responsible for higher potency of the classifier at intervals the last phase of the detection system.

2. Image segmentation:

This section aims at simplifying the illustration associate Image such it becomes a lot of meaning and easier to analyze. As a result of the premise of feature extraction, this phase is also the fundamental approach of image process. There is varied strategies victimization that pictures are going to be segmental such as Random Forest Algorithm, Otsu's algorithmic rule and thresh holding etc. Here the Random Forest Algorithm is used for image segmentation and classification. Random Forest is a learning method which classifies data better compared to other methods.

3. Feature Extraction:

After segmentation, the aftereffect so far achieved is that the realm of interest. Hence, throughout this step the options from this space of interest have to be compelled to be extracted. These options are demanded to determine that means of a sample image. Options are usually based on color, shape, and texture. Lately, utmost of the experimenters are intending to use texture options for discovery of plant disorders. There are varied strategies of feature extraction which is able to be used for developing the system similar as Argentine- position co-occurrence

matrix (GLCM), color co-occurrence system, special grey- level dependence matrix, and. bar chart primarily based feature extraction. The GLCM methodology may be statistical method for texture.

4. Classification:

The classification section implies to work out if the input image is healthy or pathologic. If the image is found to be diseased, some existing works have further classified it into a number of diseases. For classification, a code routine is required to be written in MATLAB, additionally noted as classifier. Variety of classifiers are utilized in the past few years by researchers like k-nearest neighbor (KNN), support vector machines (SVM), artificial neural network (ANN), back propagation neural network (BPNN), Naïve Bayes and call tree classifiers. The foremost usually used classifier is found to be SVM. Though each classifier has its advantages and drawbacks, SVM is straightforward to use and robust technique.

V. CONCLUSION

The detection and classification of leaf diseases accurately is that the key to stop the agriculture loss. There are an inventory of methods and classifiers to detect plant diseases. But during this project work, we are using CNN algorithm technique for the detection of plant leaf diseases. A Flask and D-jango framework for the design of plant disease detection is presented her with its work flow. Dataset for the proposed concept is ready with the plant leaf suffered with diseases Leaf Spot, Bacterial Blight. We've also analyzed the disease affected portion of leaf. We are using random forest algorithm for search all the info without delay. As per the designed framework, we will say that proposed concept of CNN (convolution neural network) is efficient enough to detect the plant leaf diseases.

VI. FUTURE SCOPE

This system considers solely the leaf of the plant to observe the malady of that crop. It'll be a lot of convenient if the opposite elements of the crop like roots, stem, branches etc. that will increase the detection accuracy over current one. Additionally, image categorization also will be done to envision whether or not the given leaf is of most efficient class or not. If a model given input aside from leaf image, then additionally it shows some name of malady for it.

REFERENCES

- [1]. Khirade, Sachin D., and A. B. Patil. "Plant Disease Detection Using Image Processing." In Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on, pp. 768-771. IEEE, 2015.
- [2]. Anand H. Kulkarni, Ashwin Patil R. K., Applying image processing technique to detect plant diseases, International Journal of Modern Engineering Research, vol.2, Issue.5, pp: 3661-3664, 2012.
- [3]. F. Argenti, L. Alparone, G. Benelli, "Fast algorithms for texture analysis using co-occurrence matrices" Radar and Signal Processing, IEE Proceedings, vol. 137, Issue 6, pp:443-448, No. 6, December 1990.
- [4]. P. Revathi, M. Hemalatha, Classification of Cotton Leaf Spot Diseases Using Image Processing Edge Detection Techniques, IEEE International Conference on Emerging Trends in Science, Engineering and Technology, pp-169-173, Tiruchirappalli, Tamilnadu, India, 2012.
- [5]. Rothe, P. R., and R. V. Kshirsagar. "Cotton leaf disease identification using pattern recognition techniques." In Pervasive Computing (ICPC), 2015 International Conference on, pp. 1-6. IEEE, 2015. Plant disease.
- [6]. Kaur, Rajleen, and Sandeep Singh Kang. "An enhancement in classifier support vector machine to improve detection." In MOOCs, Innovation and Technology in Education (MITE), 2015 IEEE 3rd International Conference on, pp. 135-140. IEEE, 2015.
- [7]. Dandawate, Yogesh, and Radha Kokare. "An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective." In Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on, pp. 794-799, IEEE, 2015.