MANGO LEAF DISEASE DETECTION USING ANN

[1]Mr. Amit Kumar Pandey, [2]Saurabh Mehanti, [3]Yatharth Khosla , [4]Harshal Goyal, [5]Priyanka Garg
[1] Assistant Professor, Students [2,3,4,5]
Department of Information Technology
Akhilesh Das Gupta Institute of Technology and management, Delhi, India.

Abstract— Agricultural productivity is something on which economy highly depends. This is the one of the reasons that disease detection in plants plays an important role in agriculture field as having disease in plants are quite natural. Identification of plant disease is very difficult in agriculture. If identification is incorrect then there is huge loss on the production of crop and economic value of market. If proper care is not taken in this area then it causes serious effects on plants and due to which respective product quality, quantity or productivity is affected. The study of the plant disease means the studies of visually observable patterns seen on the plant. It is very difficult to monitor the plant disease manually. It require tremendous amount of work, expertise in the plant disease, and also require the expertise in the plant. For instance a disease named little leaf disease is a hazardous disease found in pine trees in United States. Detection of plant disease through some automatic technique is beneficial as it reduces a large work of monitoring in big farms of crops, and at very early stage itself it detects the symptoms of diseases i.e. when they appear on plant leaves. Disease detection involves the steps like image acquisition, image pre-processing, image segmentation, feature extraction and classification. This paper presents an algorithm for image segmentation technique which is used for automatic detection and classification of plant leaf diseases. It also covers survey on different diseases justification techniques that can be used for plant leaf disease detection. Image segmentation, which is an important aspect for disease detection in plant leaf disease, is done by using genetic algorithm.

Keywords— Tensorflow, Keras, Machine learning, Data Science;

I. INTRODUCTION

The agricultural land mass is more than just being a feeding sourcing in today's world. Indian economy is highly dependent of agricultural productivity. Therefore in field of agriculture, detection of disease in plants plays an important role. To detect a plant disease in very initial stage, use of automatic disease detection technique is beneficial. For instance a disease named little leaf disease is a hazardous disease found in pine trees in United States. The affected tree has a stunted growth and dies within 6 years. Its impact is found in Alabama, Georgia parts of Southern US. In such scenarios early detection could have been fruitful.

The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plant is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. This also supports machine vision to provide image based automatic process control, inspection, and robot guidance as well as printed text and to create user profiles based on their handwriting.

Machine Learning: Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. Machine Learning algorithms are classified as **Supervised** and **Unsupervised**. Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. Unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets. This project is based on supervised learning algorithm.

Image Processing: Digital Image Processing means processing digital image by means of a digital computer. We can also say that it is a use of computer algorithms, in order to get enhanced image either to extract some useful information. Different types of images are:

- **Binary image**: contains only two pixel elements 0 –black & 1-white.
- **Black and white image**: consists of only black and white colour
- **8-bit colour**: It has 256 different shades of colours in it and commonly known as Gray scale Image (0- Black, 255- white 127- gray).
- **16-bit colour:** It is a color image format. It has 65,536 different colours in it.

II. LITERATURE REVIEW

Our project presents survey on different classification techniques that can be used for plant leaf disease classification. For given test example, k-nearest-neighbor method is seems to be suitable as well as simplest of all algorithms for class prediction. If training data is not linearly separable then it is difficult to determine optimal parameters in SVM, which appears as one of its drawbacks. Authors in paper, describe that there are mainly four steps in developed processing scheme, out of which, first one is, for the input RGB image, a colour transformation structure is created, because this RGB is used for colour generation and transformed or converted image of RGB, that is, HSI is used for colour descriptor. In second step, by using threshold value, green pixels are masked and removed. In third, by using pre-computed threshold level, removing of green pixels and masking is done for the useful segments that are extracted first in this step, while image is segmented. And in last or fourth main step the segmentation is done. It presents the technique to classify and identify the different disease through which plants are affected. In Indian Economy a Machine learning based recognition system will prove to be very useful as it saves efforts, money and time too. The approach given in this for feature set extraction is the colour co-occurrence method. For automatic detection of diseases in leaves, neural networks are used. The approach proposed can significantly support an accurate detection of leaf, and seems to be important approach, in case of steam, and root diseases, putting fewer efforts in computation.

To remove these research gaps a new methodology for automatic detection as well as classification of mango leaf diseases using image segmentation has been proposed. The advantages of proposed algorithm are as follows:

1. Use of estimators for automatic Initialization of cluster centers so there is no need of user input at the time of segmentation.
2. The detection accuracy is enhanced with proposed algorithm.
3. Proposed method is fully automatic while existing methods require user input to select the best segmentation of input image.
4. It also provides environment friendly recovery measures of the identified disease.

III. PROPOSED APPROACH

Digital camera or similar devices are use to take images of leafs of different types, and then those are used to identify the affected area in leafs. Then different types of image processing techniques are applied on them, to process those images, to get different and useful features needed for the purpose of analyzing later. Algorithm written below illustrated the step by step approach for the proposed image recognition and segmentation processes:

- Image acquisition is the very first step that requires capturing an image with the help of a digital camera
- Preprocessing of input image to improve the quality of image and to remove the undesired distortion from the image. Clipping of the leaf image is performed to get the interested image region and then image smoothing is done using the smoothing filter. To increase the contrast Image enhancement is also done.
- Mostly green coloured pixels, in this step, are masked. In this, we computed a threshold value that is used for these pixels. Then in the following way mostly green pixels are masked: if pixel intensity of the green component is less than the pre-computed threshold value, then zero value is assigned to the red, green and blue components of the this pixel.
- In the infected clusters, inside the boundaries, remove the masked cells.
- Obtain the useful segments to classify the leaf diseases. Segment the components using genetic algorithm.
- Computing the features using color co-occurrence methodology.

Image Acquisition:

- The images of the plant leaf are captured through the camera.
- This image is in RGB (Red, Green and Blue) form.
- Colour transformation structure for the RGB leaf images is created, and then, a device independent of colour space transformation for the colour transformation structure is applied.

Image Preprocessing:

- To remove noise in image or other object removal, different pre-processing techniques is considered.
- Image clipping i.e. cropping of the leaf image to get the interested image region. Image smoothening is done using the smoothening filter.
- Image enhancement is carried out for increasing the contrast and converting the RGB images into grey images using colour conversion.

Then, the histogram equalization which distributes the intensities of the image is applied on the image to enhance the plant disease images.

Before training the image is preprocessed and converted into Black and White and stored in Mat for more optimized and accurate results.

Mathematical representation of Adaptive Threshold is:

$$\mathtt{dst}(x,y) = \begin{cases} 0 & \text{if } \mathtt{src}(x,y) > T(x,y) \\ \mathtt{maxValue} & \text{otherwise} \end{cases}$$

Where T (x, y) = x (block size*block size) - c and block size is size of pixel neighborhood that is used to calculate a threshold value for the pixel and c is constant subtracted from the mean Mathematical representation of Dilation is:

$$\mathtt{dst}(x,y) = \max_{(x',y'):\,\mathtt{element}(x',y')\neq 0} \mathtt{src}(x+x', y+y')$$

Where src = input image; x and y are coordinates.

Image Segmentation:

Segmentation means partitioning of image into various part of same features. The segmentation can be done using various methods like otsu method, k-means clustering, converting RGB image into HIS model etc.

1. Segmentation using boundary and spot detection algorithm:-

The RGB image is converted into HIS model for segmenting. Boundary detection and spot detection helps to find the infected part of the leaf. For boundary detection the 8 connectivity of pixels is considered and boundary detection algorithm is applied.

2. K-Means Clustering:-

The K-means clustering is used for classification of object based on a set of features into K number of classes. The classification of object is done by minimizing the sum of the squares of the distance between the object and the corresponding cluster.

3. Otsu Threshold algorithm:-

Thresholding creates binary images from grey level images by setting all pixels below some threshold to zero and all pixel above that threshold to one. The algorithm exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

Feature Extraction:

Feature extraction plays an important role for identification of an object. In many application of image processing feature extraction is used. Color, texture, morphology, edges etc. are the features which can be used in plant disease detection. It considers color, texture and morphology as a feature for disease detection. They have found that morphological result gives better result than the other features.

Classification:

i) Using ANN

After feature extraction is done, the learning database images are classified by using neural network. These feature vectors are considered as neurons in ANN. The output of the neuron is the function of weighted sum of the inputs. The back propagation algorithm modified SOM; Multiclass Support vector machines can be used.

ii) Back Propagation

BPNN algorithm is used in a recurrent network. Once trained, the neural network weights are fixed and can be used to compute output values for new query images which are not present in the learning database.

Text Recognizer Model

```
20/20 [==============================] - 60s 3s/step - loss: 0.1416 - accuracy: 0.9486
Epoch 6/25
20/20 [==============================] - 56s 3s/step - loss: 0.1164 - accuracy: 0.9597
Epoch 7/25
20/20 [==============================] - 59s 3s/step - loss: 0.1285 - accuracy: 0.9486
Epoch 8/25
20/20 [==============================] - 58s 3s/step - loss: 0.1196 - accuracy: 0.9645
Epoch 9/25
20/20 [==============================] - 57s 3s/step - loss: 0.1342 - accuracy: 0.9450
Epoch 10/25
20/20 [==============================] - 58s 3s/step - loss: 0.0999 - accuracy: 0.9699
Epoch 11/25
20/20 [==============================] - 56s 3s/step - loss: 0.0904 - accuracy: 0.9670
Epoch 12/25
20/20 [==============================] - 58s 3s/step - loss: 0.1403 - accuracy: 0.9521
Epoch 13/25
20/20 [==============================] - 57s 3s/step - loss: 0.1232 - accuracy: 0.9592
Epoch 14/25
20/20 [==============================] - 56s 3s/step - loss: 0.0846 - accuracy: 0.9663
Epoch 15/25
20/20 [==============================] - 56s 3s/step - loss: 0.0886 - accuracy: 0.9652
Epoch 16/25
20/20 [==============================] - 59s 3s/step - loss: 0.0792 - accuracy: 0.9787
Epoch 17/25
20/20 [==============================] - 59s 3s/step - loss: 0.0883 - accuracy: 0.9699
Epoch 18/25
20/20 [==============================] - 56s 3s/step - loss: 0.0799 - accuracy: 0.9734
Epoch 19/25
20/20 [==============================] - 59s 3s/step - loss: 0.0986 - accuracy: 0.9699
Epoch 20/25
20/20 [==============================] - 58s 3s/step - loss: 0.0934 - accuracy: 0.9707
Epoch 21/25
20/20 [==============================] - 55s 3s/step - loss: 0.1145 - accuracy: 0.9628
Epoch 22/25
20/20 [==============================] - 59s 3s/step - loss: 0.0837 - accuracy: 0.9645
Epoch 23/25
20/20 [==============================] - 58s 3s/step - loss: 0.0868 - accuracy: 0.9663
Epoch 24/25
20/20 [==============================] - 55s 3s/step - loss: 0.0969 - accuracy: 0.9652
Epoch 25/25
20/20 [==============================] - 58s 3s/step - loss: 0.0649 - accuracy: 0.9752
```

IV. RESULTS AND ANALYSIS

An accuracy of 97.5 percent was achieved after training the model.

V. CONCLUSION

The project has been successfully completed according to the Initial requirements and specifications. The dataset has been modified to reduce character redundancy for similar looking letters such as lowercase and uppercase 0, m, w etc. and thus reducing the number of classes and conflicts within prediction set.

The project has been successfully completed according to the initial requirements and specifications. With room for improvement in the application, the following are the functionalities and modules to be implemented during the major phase of the project:

1. A neural network trained which will will give better accuracy thus helping in giving more accurate results.

2. To deploy the machine learning model to a RESTful API for smoother updates.

3. Hosting the database to a cloud-based service such as mongodb.

4. Developing an android/Flutter app to upload the image to the software as part of the image acquisition part.

REFERENCES

1. Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, Dongyan Zhang, "New Optimized Spectral Indices for Identifying and Monitoring Winter Wheat Diseases", *IEEE journal of selected topics in applied earth observation and remote sensing*, vol. 7, no. 6, June 2014.

2. K. Thangadurai, K. Padmavathi, "Computer Visionimage Enhancement For Plant Leaves Disease Detection", *World Congress on Computing and Communication Technologies*, 2014.

3. Monica Jhuria, Ashwani Kumar, Rushikesh Borse, "Image Processing For Smart Farming: Detection Of Disease And Fruit Grading", *Proceedings of the 2013 IEEE Second International Conference on Image Information Processing*. Tensorflow Tutorials – Official guids.

4. Saitoh, T.; Kaneko, T., “Automatic recognition of wild Flowers”, Pattern Recognition, Proceedings. 15th International Conference on, vol.2, no., pp.507-510 vol.2, 2000.

