LEAF DISEASE DETECTION OF CUCURBITS USING CNN

Isha Agrawal[1] *,Prada Hegde*[2] *Pooja Shetty*[3]*, and Priyanka Shingane*[4]

[1]Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
[2]Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
[3]Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
[4]Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India

Abstract. Identification of plant disease is tough in agribusiness arena. If it is inaccurate, there occurs a tremendous damage in the production and economical price. Leaf Disease detection requires huge amount of work, knowledge, processing time in plant disease. The most used and edible vegetable all over the world is from cucurbitaceous family. The crops under this family have great economic value in the food industry and its production is done in large scale. This family consists of 965 species. If any of these plants catch disease then there would be a tremendous loss in the production of this field yields. Thus, treating them at early stage is best way to prevent such losses. Hence, Deep Learning Algorithm like CNN can be used to detect the diseases of the plants. The leaves of the plants would be used as primary material for identification of the disease, as they are much more visible on the leaves.

Keywords -- Leaf Disease Detection, Cucurbitaceae family, CNN, Deep Learning.

I. INTRODUCTION

Agriculture is an essential factor in the financial development of the nation. Certain factors affect the condition of the crops. Due to varying weather constraints these plants are highly prone to infections. Some of the plant species are huge family , thus if the diseases are undetected then it may lead to huge loss. Thus detecting and classifying each and every plant is mandatory , even a single infected plant can expand the disease. In rural areas, farmers lack certain knowledge about the diseases and better treatments for them. Hence we have developed a Deep Learning based solution to solve all these problems.

The proposed system named “Leaf disease detection of cucurbits using CNN” is going to be very helpful for farmers and agriculture practitioners. Using this system one can easily detect diseases occurred to the plants from the Cucurbitaceae family. This system will be created using an algorithm which will benefit the farmhand to discover the disease in the initial stage and can give the required treatment. There are many different systems that detect diseases for tomatoes, cotton, rice, maize, soya bean but there is no such system that detects diseases for Cucurbit family.

Isha Agrawal: ishaagrawal2000@gmail.com
Prada Hegde: pahegde99@gmail.com
Pooja Shetty: p.shetty1406@gmail.com
Priyanka Shingane: priyankashingane@gmail.com

As agriculture is the most important aspect, using this system will help to reduce the loss in agricultural yields of cucurbit family due to diseased plants. This system can also give the required treatment.

Most of the plant leaves are infected due to bacteria, fungi and atmospheric pollutants. For using this system one must use an image of a leaf, of a diseased plant. After uploading the image the disease will be detected and appropriate treatments will be
suggested.

This project will help farmers to find out about the disease at its primary stage. This project will contain disease detection for some of the species of the family. The system detects the most common diseases The disease will classified based on texture , color and shape.

Fig 1. Pictures of some cucurbitaceae family fruits.

II. LITERATURE SURVEY

Different methods and classifiers are used by various researches to classify the diseases occurred. This paper is very precise for artificial intelligence and gives remedy for determining and categorizing discrete plant leaf diseases present, and uses convolutional neural network for classification grounds. It has four classifiers such as– logistic regression, KNN, SVM and CNN.[2]

K-means clustering is used for segmentation of images, GLCM is also used for texture feature extraction, comprising the mean for color feature extraction and KNN, Neural Network, and lastly SVM is for the final analysis of disease.[1]

Paper by Raghottam Kulkarni and Dr. A.V Sutagundar titled Plant Leaf Disease Management System (IEEE 2017). The main aim of the system was to analyze the plant species from its leaf and classify whether the leaf is in good physical condition or not and then if the leaf is detrimental, it warns about the disease the plant has and if it is in a good physical condition then it shows it is normal. Some of the main methods for detection of leaf diseases are: K-Means cluster, GLCM, and Neural Network Classifiers. Such methods successfully detect the healthy and unhealthy plant leaves. Hence, the project beneficial for identifying the plant disease.[3]

Disease detection was done on cotton plants using GLCM and SVM. To classify the diseases like Magnesium deficiency and Bacterial blight are differentiated by diverse class SVM classifier. Using SVM the features are pulled from the segmented images by considering the inputs given.[4]

The proposed methodology of image processing is achieved using MATLAB. Image processing is carried out on leaf images and features are pulled out from the diseased using k-means clustering algorithm. Variety of characteristics such as contrast, correlation, energy, homogeneity, mean, standard deviation and variance are taken out for tomato and cotton diseases.[5]

K-Medoid clustering and Random Forest classification algorithms are used to determine the diseases. The workflow of the project is: It stores an image dataset of unhealthy and healthy stalks. The whole database is separated into training and testing sets. Using this data the entire model is trained to detect the diseases.[6]

Based on RGB factors, the filters are put onto three channels. The LVQ are stored combining the resultant feature vector of convolution to train the network. The outputs confirm that the proposed methodology is very effective in recognizing the four varied types of tomato leaf diseases.[7]

The given algorithm distinguishes disease in plants and classify them in a very accurate manner as compared to the existing techniques.[8]

The system provides an accurate result generated by using various cluster sizes, enhanced the experiment with image segmentation. The system gives viable evaluation and prediction of the disease causing agent with necessary safety measures.[9]

Summary of Leaf-based plant disease detection systems: It is a systematic study to find and classify the leaf disease. This paper was written by Ravindra Jogekar , Dr. Nandita Tiwari.[10]

III.PROPOSED METHODOLOGY

The system proposed has the following phases:

- Image Collection
- Image pre-processing
- Segmentation/Feature Extraction
- Selection of Classifier
- Analyzing Results
- Predicting the Disease
- Treatment Suggestion

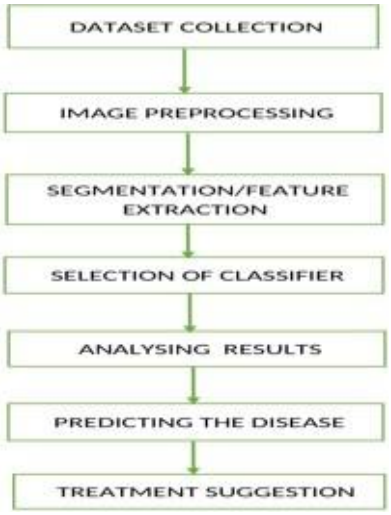

Fig 2. Workflow of the System

1. Image Collection:

In Image Collection, data-sets consist of both healthy and diseased leaves which has disorders like powdery mildew, downy mildew. Anthracnose, angular leaf spot, bacterial leaf spot. The images consist of cucumber, squash, pumpkin and watermelon. The pictures of the leaves were gathered from different online sites and were classified into healthy and diseased leaves.

2. Image Preprocessing:

In this, the images are resized into the resolution of 256x256 .The pre-processing is used to upgrade the image quality and to obtain the required features of images.

Formula for RGB to Grayscale:

Gray scale to RGB, have equivalent red, green and blue values.

$$B=G=R \quad (1)$$

For every image, pixels consisting red, green and blue of (R,G,B)

$$R' = G' = B' = (R+G+B) / 3 \quad (2)$$

3. Segmentation:

Segmentation of the image is carried out to separate the background from the leaves.

4. Selection of Classifier:

Here, the classification of the type of disease occurring on the leaves is done. There are many Deep Learning Algorithms which can be applied on the database. We have selected CNN as our classifier.

CNN (Convolutional Neural Network)

It is a complicated model that is used to categorize the diseases and is complex, thus it needs a good computational power. It is also a commonly used neural network which is referred to certain image classification problems. This system comprises of "Seven Layers" :

1. Input Layer (32 filters, Activation = 'relu')
2. Max pooling (strides = 2x2)
3. Convolutional Network Layer (64 filters, Activation ='relu')
4. Max pooling (strides = 2x2)
5. Flattening Layer
6. Fully connected Dense Layer (units = 120, Activation = 'relu')
7. Fully connected Dense Layer (units = 18, Activation = 'softmax')

The Database of the leaves from Cucurbitaceae family is created. It mainly consist of healthy and diseased leaves (i.e Angular leaf spot, Anthracnose, Bacterial leaf spot, Downey Mildew, Powdery Mildew) of Pumpkin , Cucumber, Watermelon and Squash. The images of these leaves are converted from RGB to Gray Scale for further process, as it helps in getting better accuracy. After the step of image pre-processing, using CNN a Deep Learning Algorithm training of the model is done.

The libraries used are open cv, keras, tensorflow, os, matplotlib.
The dataset was distributed in training set and validation set. Later the training data was generated by sending certain parameters like target_size, color_mode.

The the model was trained using Convolutional Neural network. There is a input layer with 32 filters and a convolutional network (2D) with 64 filters. Two max pooling layers of strides 2x2 are used, followed by a flattening layer. There are two fully connected dense layers with units 120 and 18.

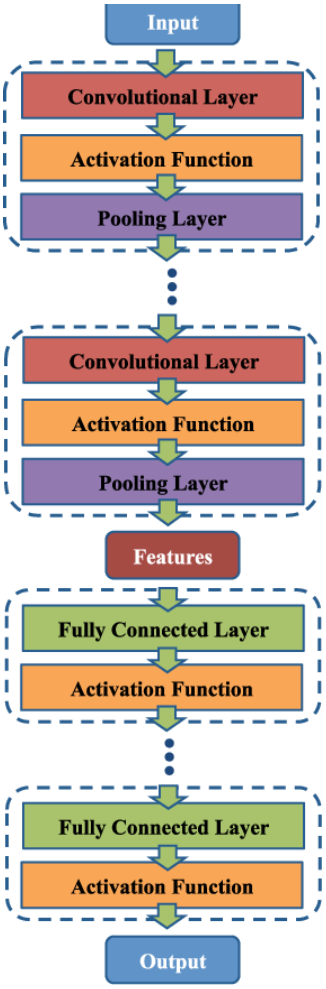

Fig 3. System Architecture

Max pooling in general, is an operation which calculates the maximum, or the largest value in each spot of feature map.

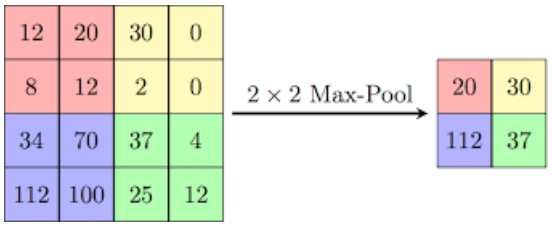

Fig 4. Example of Max Pooling

IV. EXPERIMENTAL RESULTS

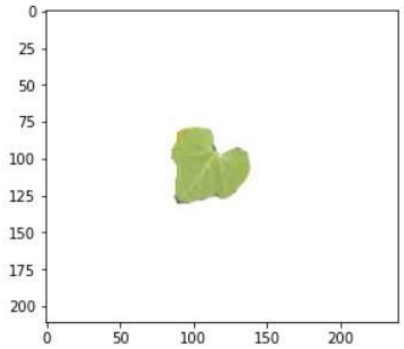

Fig 5. Image of leaf without background

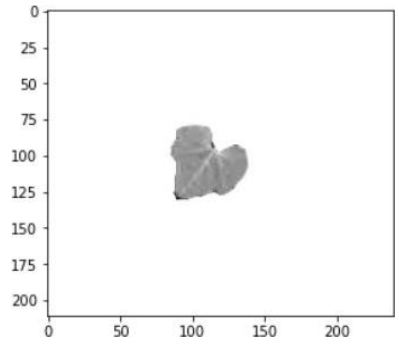

Fig 6. Conversion to grey scale

Fig 7. UI of Leaf Disease Detection of Cucurbits .

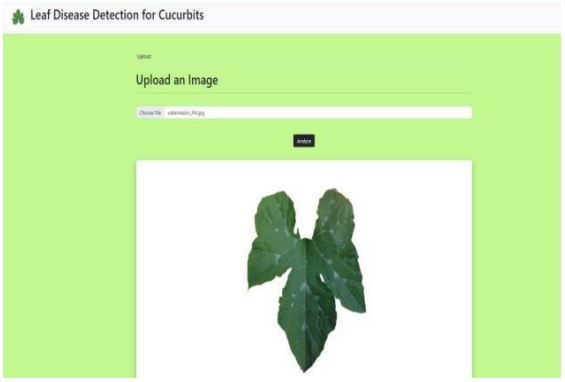

Fig8. Selecting an image for analysis (Diseased)

Fig 9. Displaying the result after analysis (Diseased)

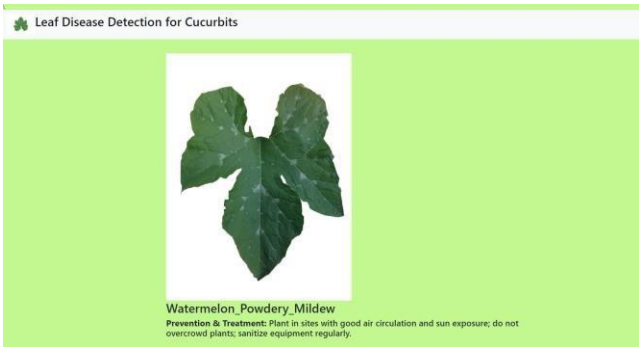

Fig 10. Treatment suggestion for the disease detected.

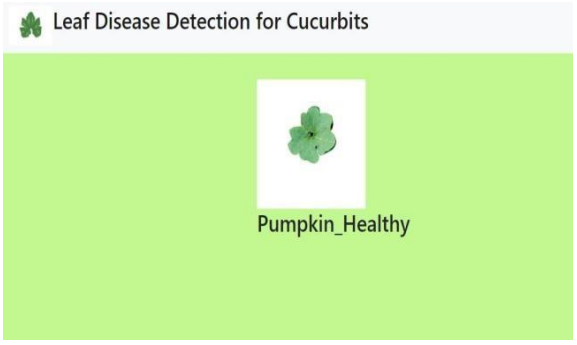

Fig12. Displaying the result after analysis (Healthy)

Fig 13. Suggestion for a healthy leaf

CODE SNIPPETS :

```
import tensorflow as tf
from tensorflow.keras import datasets, layers, models
model = tf.keras.models.Sequential()

model.add(layers.Conv2D(32,(3,3),input_shape = (256,256,1),activation= 'relu'))
model.add(layers.MaxPooling2D(pool_size=(2,2)))
model.add(layers.Dropout(0.05))

model.add(layers.Conv2D(64,(3,3),activation= 'relu'))
model.add(layers.MaxPooling2D(pool_size=(2,2)))
model.add(layers.Dropout(0.2))

model.add(layers.Flatten())

model.add(layers.Dense(120,activation= 'relu'))
model.add(layers.Dropout(0.3))

model.add(layers.Dense(18,activation= 'softmax'))
model.compile(loss = tf.keras.losses.CategoricalCrossentropy(),
              optimizer = 'adam',
              metrics = ['accuracy'])

history = model.fit(train_generator, epochs= 20,batch_size = 10, validation_data = test_generator)
```

Fig14. Layers applied in the model.

```
from tensorflow.keras.preprocessing import image
import numpy as np
import matplotlib.pyplot as plt
import cv2
def prediction(test):
    #plt.imshow(test)
    test7 = cv2.resize(test,(256,256))
    test8 = cv2.cvtColor(test7, cv2.COLOR_BGR2GRAY)
    test_img1=image.img_to_array(test8)
    test_img1=np.expand_dims(test_img1,axis=0)
    result=model.predict(test_img1)
    a=result.argmax()
    #print('a:',a)
    classes=train_generator.class_indices
    #print(classes)
    #print(len(classes))
    category=[]
    for i in classes:
            category.append(i)
    for i in range(len(classes)):
            if(i==a):
                output=category[i]
    return(output)
```

Fig15. Prediction Funtion

Name	Leaf Disease Detection using CNN	Other Systems
Remedies	In our system , treatment will be suggested according to the disease of the leaves at early stage.	In some systems, remedies are not suggested.
Detection	Disease detection is done on some of the species of cucurbitaceae family.	There is no disease detection system for Cucurbits
Accuracy	Moderate percent of accuracy is obtained	Less or moderate percent of accuracy is obtained.

Table 1 Comparison of our system with other existing system.

V. CONCLUSION

In the agriculture industry, there is great risk of the diseases that attacks the plants. Cucurbitaceae being the largest family and the most used vegetable , it is important to treat the plants in this family. This sums up the major image pre-processing and identification of leaf diseases using CNN Algorithm in the backend and UI is designed using Flask. The Training accuracy achieved is 72% and testing accuracy achieved is 12% .This approach can significantly helps in accurate detection of leaf disease.
By analyzing the amount of disease present in the leaf, one can use sufficient amount of fertilizers to control and treat the disease. Also the user can understand the affected area of leaf and solve the problem with ease.

VI. FUTURE SCOPE

The accuracy can be improved by trying different algorithms or applying some unsupervised learning technique such as k-means clustering. The quantity of dataset needs to be increased in order to achieve accurate predictions. This system can be improved by converting the proposed web app to android app in order to increase its usability . Additional species can also be added to cover the entire family.

REFERENCES

[1] Neha Kurale and Madhav Vaidya, "*Classification of Leaf Disease using Texture using feature and Neural Network* Classifiers", ICIRCA, 2018.

[2] Pushkar Sharma, Pankaj Hans , Shubash Chand Gupta, "*Classification of Leaf Disease using Machine Learning and Image Processin",* IEEE 2020

[3] Raghottam Kulkarni and Dr. A.V Sutagundar, "*Plant Leaf Disease Management System",* ICCMC, 2017.

[4] Namrata Bhimte and V.R Thool "*Disease Detection of cotton leaf spot using Image Processing and SVM Classifier"*, ICICCS, 2018.

[5] Ch. Usha Kumari, S. Jeevan Prasad, G. Mounika, "*Leaf Disease Detection : Feature Extraction with K-means Clustering and Classification using ANN"*, ICCMC, 2019.

[6] Indumati R, Saagari N, Thejuswini V, Swarnareka R *"Leaf Disease Detection and Fertilizer Suggestion"*,ICSCAN, 2019.

[7] Melike Sardogan, Adem Tuncer , Yunus Ozen, *"Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm"*, ICCSE (UBMK), 2018.

[8] Sandeep Kumar, KMVV Prasad , A. Srilekha, T.Suman , B. Pranav Rao , J Naga Vamshi Krishna, "*Leaf Disease Detection and Classification based on Machine Learning*", ICSTCEE, 2020

[9]Nikita Goel, Dhruv Jain , Adwitiya Sinha, *"Prediction model for Automated Leaf Disease Detection & Analysis"*, IACC, 2018.

[10] Ravindra Jogekar, Dr. Nandita Tiwari, *"Summary of Leaf-based plant disease detection systems: A compilation of systematic study findings to classify the leaf disease classification schemes"*, WorldS4, 2020.