INSTITUTE OF ADVANCED RESEARCH, GANDHINAGAR

PLANT LEAF DISEASE DETECTION SYSTEM USING MACHINE LEARNING

Prepared By: Sanskriti Sidola (UIAR/11316)

Department: Department of Engineering and Computing

Branch: Information Technology (FY)

Project Supervisor: Dr Ahsan Rizvi

Date: December 31, 2021

CERTIFICATE:

This is to certify that **Ms. Sanskriti Sidola B**, BTech (Information Technology) student from Institute of Advanced Research, Gandhinagar has done her project work titled **"PLANT LEAF DISEASE DETECTION SYSTEM USING MACHINE LEARNING"** in our Department of Engineering and Computing from 01-08-2021 to 10-12-2021 as a part of curriculum.

Date: 31-12-2021

Under Supervision of:

Dr Ahsan Rizvi

(Professor at Institute of Advanced Research, Gandhinagar.)

ABSTRACT:

Crop diseases are a noteworthy risk to sustenance security. Emergence of accurate techniques in the field of leaf-based image classification has shown impressive results. My proposed project includes various phases of implementation namely dataset selection, feature extraction, training the classifier and classification. The choosen dataset of diseased and healthy leaves are collectively trained to classify the diseased and healthy images. Overall, using machine learning to train the large data sets available publically, gives us a clear way to detect the disease present in plants on a colossal scale.

TABLE OF CONTENTS:

ACKNOWLEDGEMENT:

I am indeed grateful to many groups of people of who have helped me with various aspects of this study. I would also like to thank my supervisor Dr Ahsan Rizvi for guiding and motivating me to do this amazing project on “Plant leaf disease detection system using machine learning”.

INTRODUCTION:

One of the important sectors of Indian Economy is Agriculture. Employment to almost 55% of the countries workforce is provided by Indian agriculture sector. Farmer's economic growth mostly depends on the quality of the products that they produce, which in turn mostly relies on the plant's growth and the yield they get. Therefore, for the better yield in agriculture, detection of disease in plants plays an instrumental role. Plants are highly prone to diseases that affect the growth of the plant which in turn affects the profit of the farmer. In order to detect a plant disease at very initial stage and to avoid its spread, use of automatic disease detection technique is advantageous. The symptoms of plant diseases are commonly found on the leaf of the plants. Manual detection of plant disease using leaf images with the help of experts is a tedious job. Hence, it is required to develop computational methods which will make the process of disease detection and classification using leaf images automatically for better yield and profit for the farmers.

(i) PROBLEM STATEMENT:

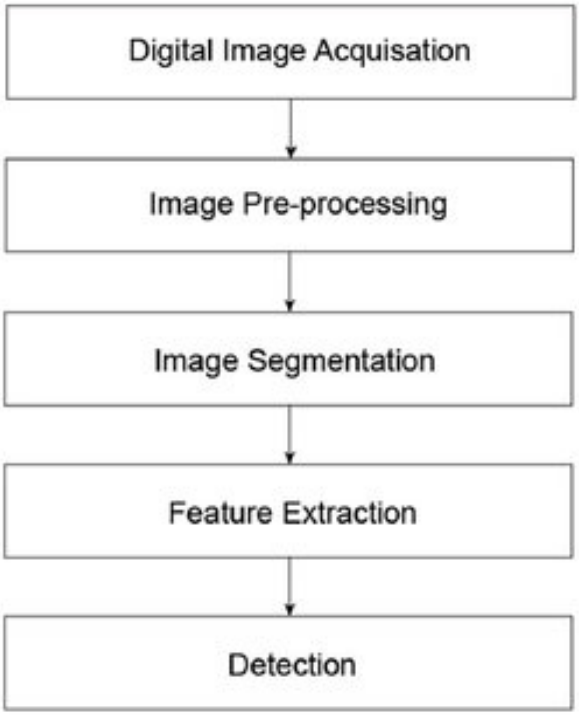
The most widely earlier-used method for plant disease detection was simply naked eye observation by experts through which identification and detection of plant diseases were done. For doing so, a large team of experts as well as continuous monitoring of experts were required, which costed very high when the size of the farms were large. At the same time, in some countries, farmers didn't have proper facilities or even idea that they can contact to experts. Due to which consulting experts even costed high as well as the whole process resulted to be time consuming too. In such a condition, the suggested technique is the automatic detection of the diseases by just seeing the symptoms present on the plant leaves. It makes it easier as well as cheaper.

(ii) EXISTING SOLUTION:

In India, farmers still use the traditional method to detect the diseases in the leaf. Farmers or experts uses their naked eye observation to find the disease. This how the identification and detection of leaf diseases is done. So, in order to do this work, a team of knowledge person as well as a continuous monitoring of plant is necessary, which costs very high when we do with large farms and it can be done to the limited area. But in few countries, farmers do not have proper facilities or even idea that they could contact to experts. And also consulting experts even isn't a cost friendly as well as it is not time consuming. In such conditions, Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. Whereas if automatic detection technique is used it will take less efforts, less time and gives more accuracy. In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases. Image processing is used for measuring affected area of disease and to determine the affected area.

(iii) PROPOSED SOLUTION:

The proposed system makes use of one of the deep learning techniques called the Convolutional Neural Network (CNN). CNN are mainly used for classification problems. The proposed system involves transfer learning methodology i.e., using one of the pre-trained architectures to get better results. VGG19 is used as the transfer learning architecture which is composed of 19 deep layers. The system takes in an image of a leaf as an input through the device camera or its internal storage and goes through five level of classification described below:



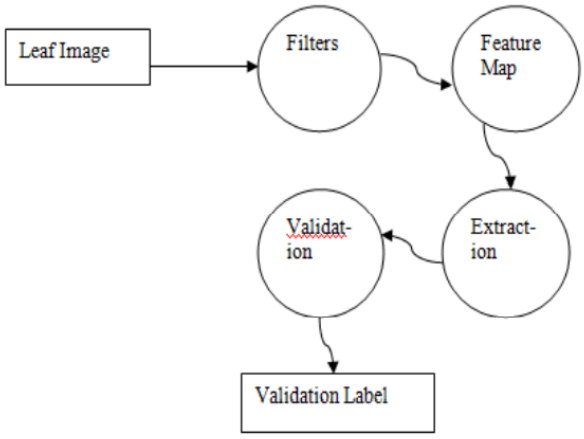
Once the disease is been finally classified, its related description about the disease is provided. The proposed system is capable of identifying the diseases such as Tomato Bacterial Spot, Tomato Early Blight, Tomato Healthy, Tomato Late Blight, Tomato Leaf Mold, Tomato Septoria Leaf Spot, Tomato Spider Mites Two-Spotted Spider Mite, Tomato Target Spot, Tomato Mosaic Virus, Tomato Yellow Leaf Curl Virus.

The proposed solution provides **93.1%** of accuracy.

METHODOLOGY AND REQUIREMENTS:

For the development of this whole project, I have used several platforms like Google Colab for the python coding, Kaggle for datasets, VS Code for coding and mapping, Anaconda Shell Prompt for running Streamlit and WordPress for making the website.

Now, talking about the methodology,



The system uses transfer learning technique i.e., VGG19 model has described below:

- VGG19 is a variant of VGG model which in short consists of 19 layers that includes 16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer. There are other variants of VGG like VGG11, VGG16 and others.

- A fixed size of (224 * 224) RGB image was given as input to this network which means that the matrix was of shape (224,224,3).

- They subtracted the mean RGB value from each pixel, which was the only pre-processing which was done and computed over the whole training set.
- Used kernels of (3 * 3) size with a stride size of 1 pixel, this enabled them to cover the whole notion of the image.
- Spatial padding was used to preserve the spatial resolution of the image.
- Max pooling was performed over a 2 * 2 pixel windows with stride 2.

I have used plant village dataset consisting of only one crop namely Tomato. The dataset is divided into 60% train and 40% test. 18345 images were used to train the model and 4585 images for testing. The fully connected layer is composed of 2 hidden layers consisting of 1280 and 1024 neurons respectively, The output layer is multi neuron output which gives out the final results. The system consists of 3 levels of classification. At the first stage the given leaf input is classified into the type of diseases mentioned in dataset. Secondly, image is passed to the next level of classification where the final disease has been outputted. Based on the type of disease identified the remedies for the same is given out.

EXPERIMENTAL RESULTS:

The snapshots of the results are shown below. The system is capable of detecting the diseases with 93.1% accuracy.

The Figure (a) shows the accuracy of the detection.

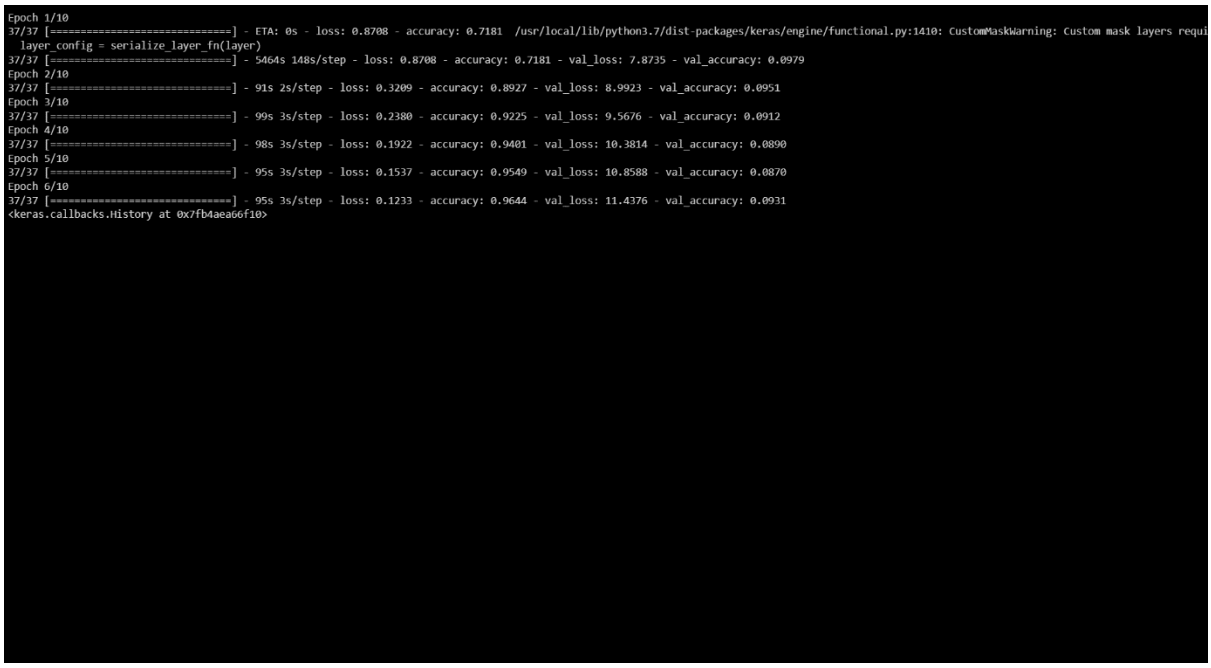


Figure (a)

The Figure (b) shows the classification for a given tomato leaf as input taken from file. It also shows the description and remedy associated with that disease.

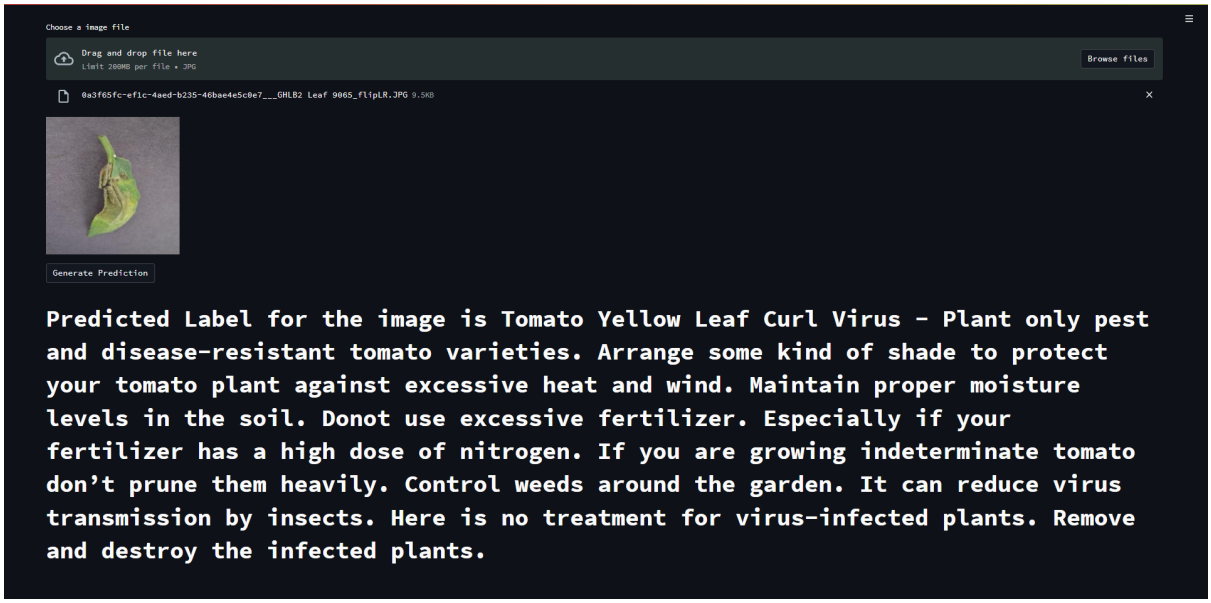


Figure (b)

The Figure (c) shows that besides choosing the image from the file, an image can also be taken as input from camera as well. It will also show the description and remedy associated with that disease.

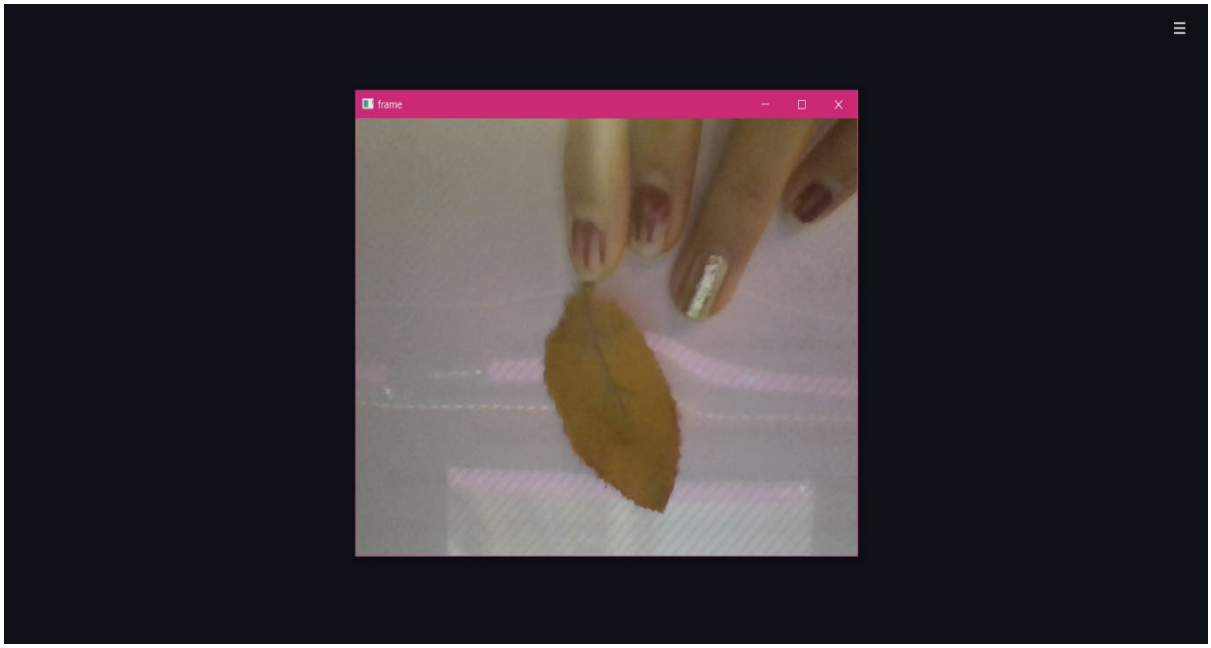


Figure (c)

The Figure (d) shows the website made for the purpose of taking this project to a web-application form. Here is the screenshot of the website's home page that contains two options- either to upload the leaf image from the file or by clicking a new image with the help of camera.

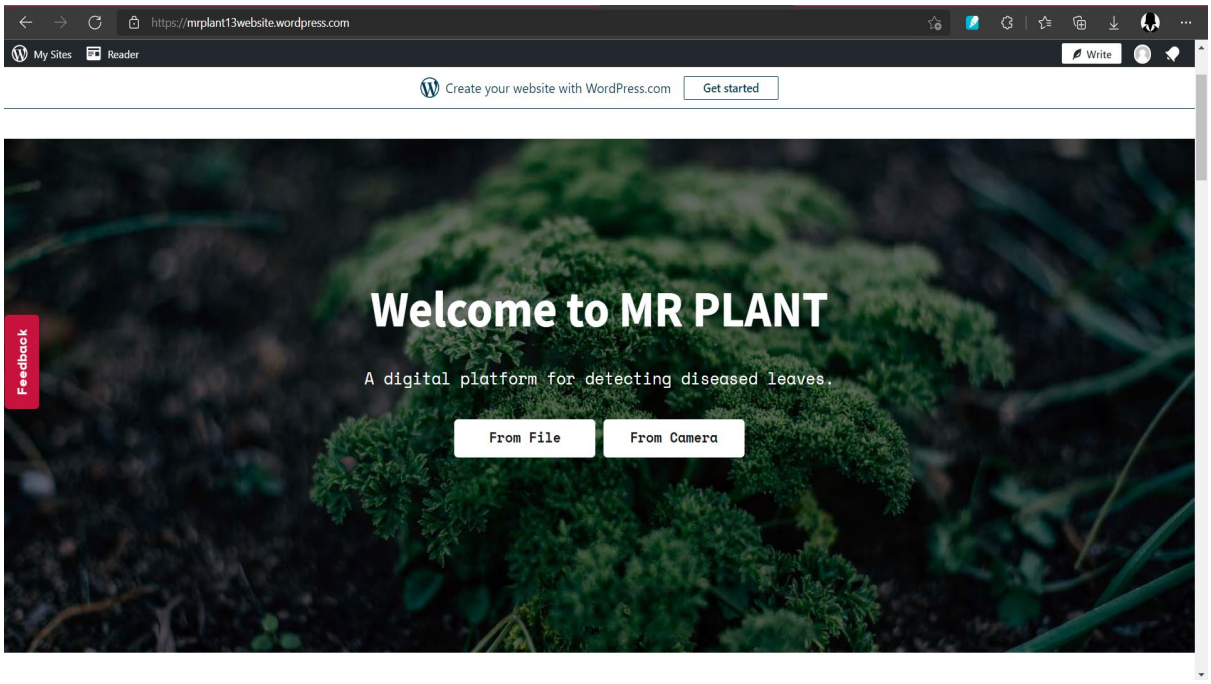


Figure (d)

(i) CONCLUSION:

Detection of the plant disease at its early stages proves to be one of the best methods to avoid the spread of the disease further hence, we have come with the system that automatically detects the disease affected to a plant with the help of the leaf image. The symptoms of the disease appear on the leaf region and hence taken as an input for further detection and classification. These images are classified through transfer learning methods by using VGG19 architecture. The system is capable of finding disease with 93.1% accuracy.

FUTURE SCOPE:

In future the same technique can be extended to multiple crops to detect more number of diseases. Also, many other deep learning techniques can be used to enhance the accuracy. More number of datasets can be used to get better results.

REFERENCE:

https://ijtre.com/wp-content/uploads/2021/10/2020071107.pdf

homyden.com/bacterial-spot-control-identify-get-rid-bacterial-spots/

https://growingthehomegarden.com/treating-early-blight-on-tomatoes

https://www.gardeners.com/how-to/prevent-late-blight/7262.html

https://www.gardeningchannel.com/tomato-diseases-how-to-fight-leaf-mold/

www.thespruce.com/identifying-and-controlling-septoria-leaf-spot-of-tomato-1402974

www.gardeningknowhow.com/edible/vegetables/tomato/target-spot-on-tomato

www.wikihow.com/Prevent-Spider-Mites-on-Tomato-Plants

www.almanac.com/pest/mosaic-viruses-gardening-abc.com/tomato-leaf-c