ALLIANCE COLLEGE OF ENGINEERING AND DESIGN

ALLIANCE UNIVERSITY, BENGALURU

"PLANT DISEASE DETECTION USING MACHINE LEARNING"

A DTM 003 PROJECT REPORT – FINAL REPORT

Submitted by

Group – C5

PAYAL (19030141CSE062)

ADITYA KESARI(19030141CSE038)

HAFFIS NAJEEB(19030141CSE009)

B. DHANUSH KRISHNA SAI REDDY(19030141CSE86)

VINAYAK M(19030141AE022)

In partial fulfillment of the Course- DTM in Semester VII of

BACHELOR OF TECHNOLOGY

Under the Guidance of

Dr. Shiela Mahapatra

Dr. Ashes Banerjee

Prof. Arpita Nath Boruah

Department of Computer Science and Engineering,

Alliance University,

Bangalore

January 2023

1

CERTIFICATE

This is to certify that the project work entitled "**PLANT DISEASE DETECTION USING MACHINE LEARNING**" is the bonafide work done by **Group – C5** submitted in partial fulfillment of the requirements for the award of the degree Bachelor of Technology in Computer Science and Engineering during the year 2022-2023.

Dr. Shiela Mahapatra   Dr. Ashes Banerjee,   Prof. Arpita Nath Boruah

Supervisor   **Supervisor**   **Supervisor**

HOD - EEE and,

Assistant Dean - Research   Assistant Professor   Assistant Professor

Dept. of EEE   Dept. of CE   Dept. of CSE

External Examiners:

1. Name: Signature

2. Name: Signature

ALLIANCE UNIVERSITY

Alliance College of Engineering and Design

DECLARATION

This is to declare that the Report titled "**Plant disease detection using Machine learning**" has been made for the partial fulfillment of the Course DTM in Semester VII. I confirm that this report truly represents our work undertaken as a part of our batch during the DTM 003 project session. This work is not a replication of work done previously by any other person. We also confirm that the contents of the report and the views contained therein have been discussed and deliberated with the academic supervisor.

NAME	REGISTRATION NO	SIGNATURE
PAYAL	19030141CSE062	
ADITYA KESARI	19030141CSE038	
B. DHANUSH KRISHNA SAI REDDY	19030141CSE086	
HAFFIS NAJEEB	19030141CSE009	
VINAYAK MB	19030141AE022	

<u>ACKNOWLEDGEMENT</u>

The satisfaction that accompanies the successful completion of the task would be put incomplete without the mention of the people that made it possible, whose constant guidance and encouragement crown all the efforts successfully. I am very thankful to my supervisors **Dr. Shiela Mahapatra, Dr. Ashes Banerjee, and Prof. Arpita Nath Boruah** for their sustained inspiring guidance and cooperation throughout the process of this project. Their wise counsel and valuable suggestions are invaluable. I avail myself of this chance to express my deep sense of gratitude and hearty thanks to the Management of Alliance University, for providing world-class infrastructure, a congenial atmosphere, and encouragement. I express my deep sense of gratitude because of the teaching and non-teaching staff in our department who stood with me during the project and helped me to make it a successful venture. I place my highest regards to my parents, my friends, and well-wishers who helped plenty in making the report of this project.

- **Group – C5**

ABSTRACT

Due to the increasing expansion of each nation's food supply as a result of population growth, automation of plant disease detection is a top priority. For many years, researchers have been investigating new strategies to combat the rising tide of diseases afflicting both plants and humans. A large number of parts are still needed to complete the detection and discovery device. This is the first in a series of sporting events meant to reduce disease transmission spread from one person or animal to another. This complicates the attack against some illnesses using learning and deep learning techniques. Plant diseases can be automatically identified.

TABLE OF CONTENTS

<u>INTRODUCTION</u>

Losses in modern agricultural productivity is caused due to plant diseases. The severity of a plant condition is a crucial component of a disorder diploma and can be used to predict yield and promote therapy. Reduced yield losses will result from the quick and precise assessment of disease severity[1]. In the past, professionals have used plant tissue samples to visually inspect them and grade the severity of plant disorders. The slow growth of modern agriculture is hindered by the high cost and poor effectiveness of human disorder assessment[2]. Precision agriculture, high-throughput plant phenotyping, smart greenhouses, and other digital camera and computer vision-based technologies have created a reasonably high demand for automated disorder evaluation models. The artwork suggests deep learning models for image-based totally automated evaluation of plant disease severity. Drawing inspiration from the use of the deep learning leap forward in image-based entirely plant disorder recognition, we also add severity descriptors to the apple healthy and black rot pictures found in the public PlantVillage dataset [3]. From scratch, we train shallow networks of various depths from scratch and fine-tune the pre-trained modern deep networks in order to determine the precise network form and schooling method. We compare the models' abilities to accurately identify the level of disorder severity. Our results represent the first step in the automatic assessment of the degree of plant disorder. Finding ways to improve the usage of modern technology is made possible by analyzing the data relevant to this area. The images of the leaves and unique plant components can be utilized to identify plant diseases. The technology can be completed by examining photographs of people who also show the presence of illnesses and calculating the degree of their destruction. Analyzing the pertinent facts in this field enables the discovery of methods to enhance the usage of contemporary technologies. Finding ways to improve the use of modern technology is made possible by analyzing the relevant data in this area.

IDENTIFYING THE PROBLEM

Typically, fungi are the source of diseases that affect vegetation and attack the leaves. Others are brought on by bacterial and viral infections. With the increased application of ML and its related aspects in agriculture, precision has advanced [13]. Many people and animals are harmed by the decreased agricultural production, which is a problem that needs to be solved by the current generation. Due to the image-based detection tool's high accuracy, reduced complexities, and reduced data duplication, extracting and identifying diseases is made simpler. When it comes to some tomatoes that resemble flora, using photos to determine the diseases that affect them and the extent of the damage can't be achieved until there is a high rate of accuracy [14]. The study of plant diseases reveals that many different factors influence how generation-based fully image recognition is used. In other words, illnesses that cause dents and changes in the vegetation are those that can be discovered using this generation rather than those that cause damages that cannot be detected using the photos of the plant [15]. According to the findings of these investigations, plant diseases are often recognised after they begin to affect the physical appearance of the plants.

The primary issue in agriculture is the deterioration in plant yield quality. The problem is brought on by inefficient plant disease identification and management. The challenge has been broadened to incorporate people in a number of ways. Plant diseases cause a decrease in plant cover, which results in famine, less air filtering, and global warming. Early crop disease identification has been successfully accomplished using hyperspectral imaging [16]. It is difficult to pinpoint the factors that cause illnesses unless they are discovered at an early stage. We may also state that if an illness is discovered early on, the causes can be rather easily determined.

Further investigation by [6] indicates an inadequate collection that would be utilized to offer background knowledge for assessing the captured photos. The unique difficulty is that the disorders' symptoms and characteristics might resemble good attributes to a large extent [17]. For instance, several infections should make leaves wilt. Since more and fresh photographs are being steadily submitted by professionals, the problem has not yet been solved.

Lack of suitable equipment for image detecting operations is another problem. It is challenging for most field specialists to gather accurate data and identify diseases because they lack the

technology needed to interpret the photos they take in the field [18]. Due to restrictions put in place to protect the accuracy and dependability of the data from these analyses, execution rates in some areas are low. There have been numerous rules on ML and soft computing that may prevent its use in a number of fields [19]. The regulations forbid some of the actions.

MACHINE LEARNING PROPOSALS

3.1. Deep Convolutional Neural Network

We compare architectures while also creating a shallow community from ground - up and altering the learning process by optimising the top layers of a deep community that has already been trained in order to find the best CNN structure for really well disease severity designation with minimal training data.

Shallow networks include a few layers, some filters included, and each layer is completely linked. A SoftMax algorithm is used for normalisation. In the shallow networks that we train in the system, there are convolutional layers of two, four, six, eight, and ten layers. Each layer has 32 filters of size 3 3. Each layer, with the exception of the last one, which consists of 64 units, is followed by a ReLU activation and a 2 2 "max-pooling." After sixty-four units of the first layer, which is fully linked, comes a layer with a 50% dropout ratio. In order to determine the opportunity result, the SoftMax layer receives four inputs from an absolutely connected layer that mirror the four instructions.

3.2. Transfer Learning

We will examine only a relatively small number of images. By fine-tuning parameters of a network that's already been trained on a large dataset, such as ImageNet, switch-gaining knowledge can build a powerful classification community with little data. Although the disorder severity classification is geared toward finer-grained photo category type issues than ImageNet, the lower layers can be adapted to most laptop vision tasks because they encode the simplest functions. For example, the first layer of the VGG16 version depicts the activations within the first layer, which displays direction and colour. Even though it has not been trained

on the plant disorder dataset, the version can be activated in opposition to diseased spots, leaves, and the historical past.

<div align="center">
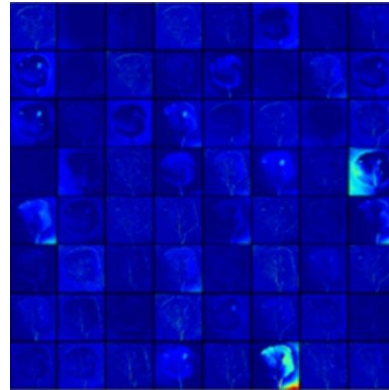
(a)   *(b)*

</div>

"Visualization of activations for a given input image in the 1st convolutional layer of the pre-trained VGG16 model:
(a)   original image; (b) the first convolutional layer output."

"Source: https://www.hindawi.com/journals/cin/2017/2917536/fig1/"

OBJECTIVE OF THE THESIS

Our objective of the thesis is to detect the various kinds of diseases in the leaves of different kinds of species of plants. In India, for the crops to have a better yield, everything depends on the quality of the crop. The present way of detecting leaf disease is through the naked eye which takes a lot of time and effort and manpower. This project's main objective is to make the process of detecting plant diseases easier and faster so that the quality of the plant is sustained before it becomes unhealthy for human consumption. So, we make use of CNN for that purpose. "Deep CNN-Supported Identification of Crop Diseases by Plant Disease Detection" is a new method for building a recognition model for crop disease detection using Deep CNN. A very efficient method is employed for the formation of the system required for this purpose using our proposed methodology. Using healthy leaves, the system will be able to identify 13 different types of plant illness. Agricultural experts have analysed the database and using that analysis, the system will identify the confidence of the disease in a leaf with the help of the image collection at the beginning. Our objective is to gain the maximum percentage of accuracy in this project. Below is the flowchart showing all the necessary steps in our methodology and the result that we expect to obtain.
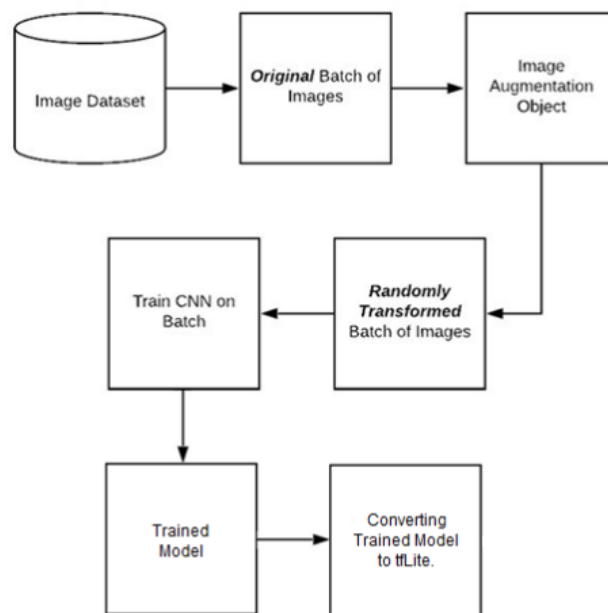
Fig. 1. Block Diagram Of Proposed System

"Source: https://www.itm-conferences.org/articles/itmconf/pdf/2022/04/itmconf_icacc2022_03049.pdf"

METHODOLOGY

4.1. Data Material

Approximately fifty thousand photos of diseased and healthy plants with 38 folders are available online that is publicly accessible. Among all those jpg images, we will choose the pictures of apple leaves and those apple leaves which has black rot that has been formed because of the fungus botryosphaeria obtuse. Healthy, Early, Middle and End-stage are the types into which the images have been classified. At the healthy stage, there are spotless leaves. At early stage, a very little spots of round shape having a diameter of less than 5 mm are found. In the middle stage, there are a minimum of three irregular spots. The diseased leaves are so bad that the tree will lose them. Every image is studied and labeled carefully according to their situations. Some images that are inconsistent are not used by the professionals. Here, the abandoned images are 179 in number.
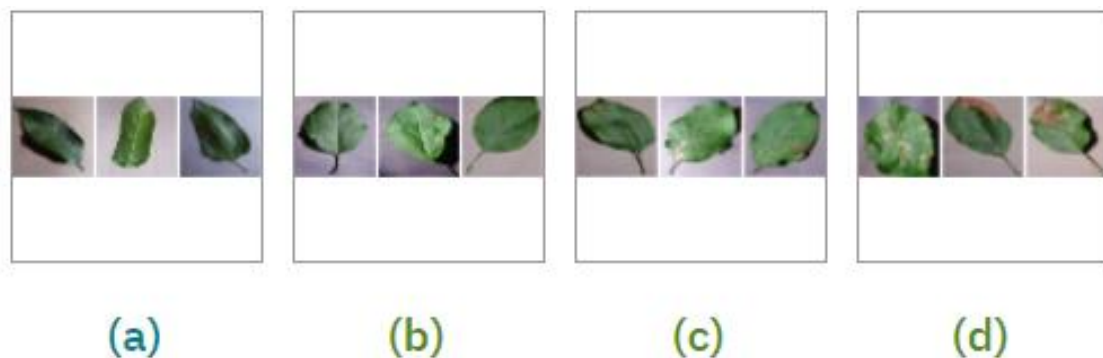

Figure 1

"Sample leaf images of the four stages of apple black rot: (a) healthy stage, (b) early stage, (c) middle stage, and (d) end-stage."

"Source: https://www.hindawi.com/journals/cin/2017/2917536/fig2/"

Since there are more healthy images than diseased images, there may be a bias towards the healthy staged images. So, we should have a balance towards all the stages of images so that the network doesn't produce inaccurate output. Apart from the healthy staged images, we are going to use eighty percent of the rest of the categories of images as the training dataset and twenty percent as the test dataset.

4.2. Neural Network Training Algorithm

The core of the "basic architecture of the DCNN" is composed of "many convolution layer," "pooling layers," and "completely linked layers." It determines the following:

$$X_{ic} = ReLU\ (W_i * x)$$

where $*$ = operation on convolution, W_i = the convolution kernels of the layer. $W_i = [W_i^1, W_i^2, \ldots, W_i^K]$, and K = the no. of convolution kernels. Every kernel W_i^K is an M X M X N weight matrix along with M being the window size and also the number of input channels.

ReLU is an abbreviation for "Rectified Linear Function" which is being used as an activate convolution operation in the model. ReLUs and deep CNN trains are much quicker than their contemporaries with inundating non-linearities.

The highest value of each convolution kernel output is being computed by max-pooling across non-overlapping rectangular regions. Pooling allows for position invariance across larger local regions while reducing output size.

4.3. Image Pre-processing

PlantVillage dataset is made up of RGB images of various sizes. Deep learning methods require only four basic image pre-processing phases. Due to their powerful learning method, to process images, the following stages are used: All images for shallow networks are rescaled to 256 x 256 pixels, 224 x 224, pixels for vgg19, resnet50, and vgg16, and 299 x 299 pixels for inception-v3 model enhancement and predictions are performed on these rescaled images. Second, the pixel values are divided by 255 to conform to the network starting condition. Finally, sample-wise normalization was carried out end-to-end. Training efficiency can be significantly improved by data normalization.

RESULTS AND DISCUSSIONS

1. First we specify all the files in the dataset,i.e The different plant species diseases.

```
In [5]:  ▶  os.listdir('E:\sem7\plantDiseaseProject/archive/PlantVillage/')

   Out[5]:  ['Pepper__bell___Bacterial_spot',
             'Pepper__bell___healthy',
             'Potato___Early_blight',
             'Potato___healthy',
             'Potato___Late_blight',
             '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__Tomato_mosaic_virus',
             'Tomato__Tomato_YellowLeaf__Curl_Virus']

In [6]:  ▶  # total categires
             len(os.listdir('E:\sem7\plantDiseaseProject/archive/PlantVillage/'))

   Out[6]:  15
```

2. Then we count the number of images of particular plant species present in the dataset, and display it.

Out[11]:

|	disease	count_images
0	Pepper__bell___Bacterial_spot	997
1	Pepper__bell___healthy	1478
2	Potato___Early_blight	1000
3	Potato___healthy	152
4	Potato___Late_blight	1000
5	Tomato_Bacterial_spot	2127
6	Tomato_Early_blight	1000
7	Tomato_healthy	1591
8	Tomato_Late_blight	1909
9	Tomato_Leaf_Mold	952
10	Tomato_Septoria_leaf_spot	1771
11	Tomato_Spider_mites_Two_spotted_spider_mite	1676
12	Tomato__Target_Spot	1404
13	Tomato__Tomato_mosaic_virus	373
14	Tomato__Tomato_YellowLeaf__Curl_Virus	3209

3. The count is then displayed in a graphical format for the 3 plant species we are taking into consideration (pepper,tomato, and potato) and the frequency of their image count.

```
In [23]:  ▶ sns.barplot(x='Category', y = 'Count', data=ind_count)
             plt.show()
```
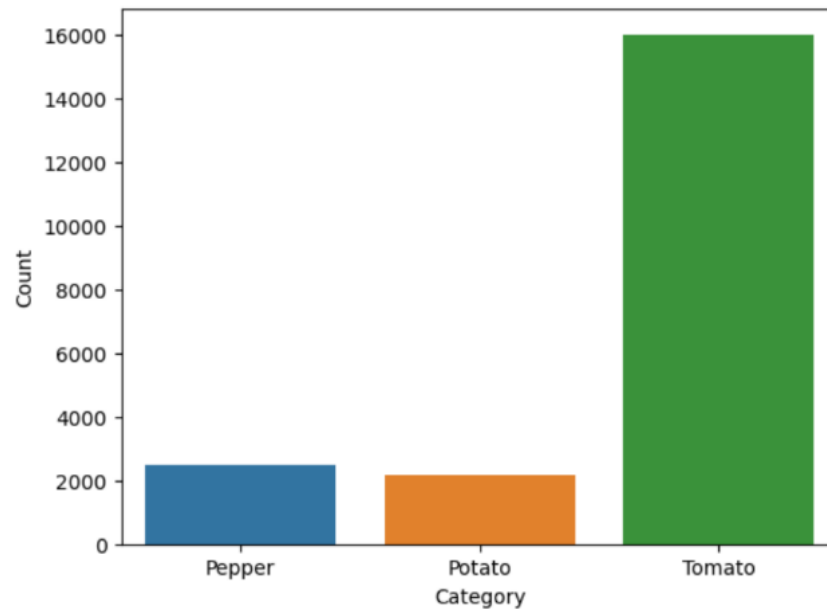

4. Then we are plotting the image dataset using the figure function with the mentioned height and   width   and   we   display   the   plot   without   grid.

```
In [28]:  ▶ rows = 3
             plt.figure(figsize=(16, 8))
             for i,name in enumerate(cat_list):
                 img = Image.open(images_list[i])
                 plt.subplot(rows,5,i+1)
                 plt.title(cat_list[i])
                 plt.axis('off')
                 plt.imshow(img)
```
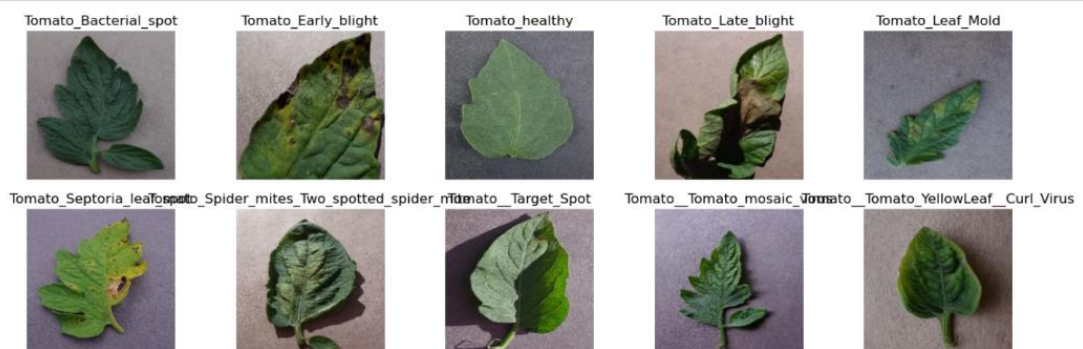
5. Here we are taking the image path of healthy leaf and appending with diseased one to show the dissimilarity.

```
In [29]:  ▶  p_cat_list = []
              p_images_list = []
              for t_p in pepper_diseases:
                  dir_path = os.path.join(fol_path, t_p)
                  img_path = os.listdir(dir_path)[0]
                  p_cat_list.append(t_p)
                  p_images_list.append(os.path.join(fol_path, t_p,img_path))

              rows = 3
              plt.figure(figsize=(16, 8))
              for i,name in enumerate(p_cat_list):
                  img = Image.open(p_images_list[i])
                  plt.subplot(rows,5,i+1)
                  plt.title(p_cat_list[i])
                  plt.axis('off')
                  plt.imshow(img)
```


Pepper__bell___Bacterial_spot   Pepper__bell___healthy

6. Here we are taking the image path of healthy leaf and appending with diseased one to show the dissimilarity.

```
In [30]:  ▶  po_cat_list = []
              po_images_list = []
              for t_po in potato_diseases:
                  dir_path = os.path.join(fol_path, t_po)
                  img_path = os.listdir(dir_path)[0]
                  po_cat_list.append(t_po)
                  po_images_list.append(os.path.join(fol_path, t_po,img_path))

              rows = 3
              plt.figure(figsize=(16, 8))
              for i,name in enumerate(po_cat_list):
                  img = Image.open(po_images_list[i])
                  plt.subplot(rows,5,i+1)
                  plt.title(po_cat_list[i])
                  plt.axis('off')
                  plt.imshow(img)
```


Potato___Early_blight   Potato___healthy   Potato___Late_blight

7. Here we make use of sklearn library and split the data into testing and training dataset.

```
In [42]:  from sklearn.model_selection import train_test_split
          X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.2, random_state=0,shuffle=True)
```

```
In [43]:  X_train.shape
```

```
Out[43]:  (16510, 64, 64, 3)
```

```
In [44]:  import tensorflow as tf
          from tensorflow.keras.models import Sequential
          from tensorflow.keras.layers import Dense, Dropout, Conv2D, Flatten, MaxPooling2D
          from tensorflow.keras.optimizers import RMSprop, Adam, SGD
```

```
In [45]:  model = Sequential([
              Conv2D(16, (3,3), activation='relu',input_shape=(64,64,3)),
              MaxPooling2D((2,2),strides=2),
              Conv2D(16, (3,3), activation='relu'),
              MaxPooling2D((2,2),strides=2),
              Flatten(),
              Dense(64,activation='relu'),
              Dense(15, activation='softmax')

          ])
```

8. We are setting the batch size as 16 and epoch value as 20, so that the accuracy increases as the dataset is passed through the training algorithm.

```
In [50]:  history = model.fit(X_train, y_train, batch_size= 16, epochs=20, verbose=1, validation_data=(X_test, y_test),callbacks=[callb

Epoch 1/20
1032/1032 [==============================] - 33s 31ms/step - loss: 1.3461 - accuracy: 0.5689 - val_loss: 0.8170 - val_accura
cy: 0.7287
Epoch 2/20
1032/1032 [==============================] - 32s 31ms/step - loss: 0.6895 - accuracy: 0.7709 - val_loss: 0.6118 - val_accura
cy: 0.7895
Epoch 3/20
1032/1032 [==============================] - 33s 32ms/step - loss: 0.5313 - accuracy: 0.8233 - val_loss: 0.6081 - val_accura
cy: 0.7880
Epoch 4/20
1032/1032 [==============================] - 33s 32ms/step - loss: 0.4466 - accuracy: 0.8492 - val_loss: 0.5128 - val_accura
cy: 0.8299
Epoch 5/20
1032/1032 [==============================] - 31s 30ms/step - loss: 0.3790 - accuracy: 0.8725 - val_loss: 0.4062 - val_accura
cy: 0.8609
Epoch 6/20
1032/1032 [==============================] - 39s 38ms/step - loss: 0.3355 - accuracy: 0.8863 - val_loss: 0.4007 - val_accura
cy: 0.8660
Epoch 7/20
1032/1032 [==============================] - 35s 34ms/step - loss: 0.2971 - accuracy: 0.8994 - val_loss: 0.4219 - val_accura
cy: 0.8539
Epoch 8/20
1032/1032 [==============================] - 30s 29ms/step - loss: 0.2552 - accuracy: 0.9128 - val_loss: 0.3970 - val_accura
cy: 0.8687
Epoch 9/20
1032/1032 [------------------------------] - 34s 33ms/step - loss: 0.2358 - accuracy: 0.9191 - val_loss: 0.3732 - val_accura
```

9. We plot the graph between training accuracy and validation accuracy.
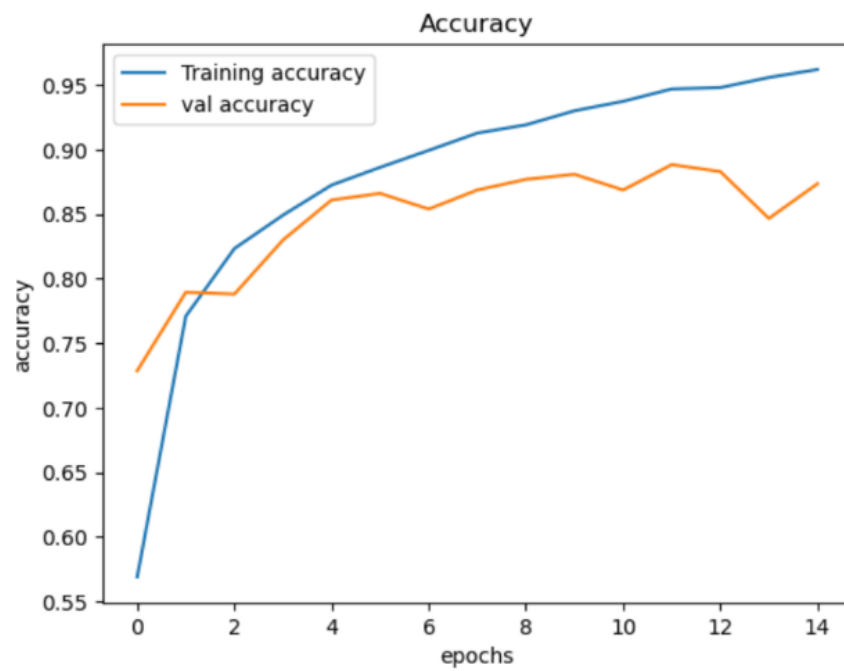
10. We plot the graph between training loss and validation loss.
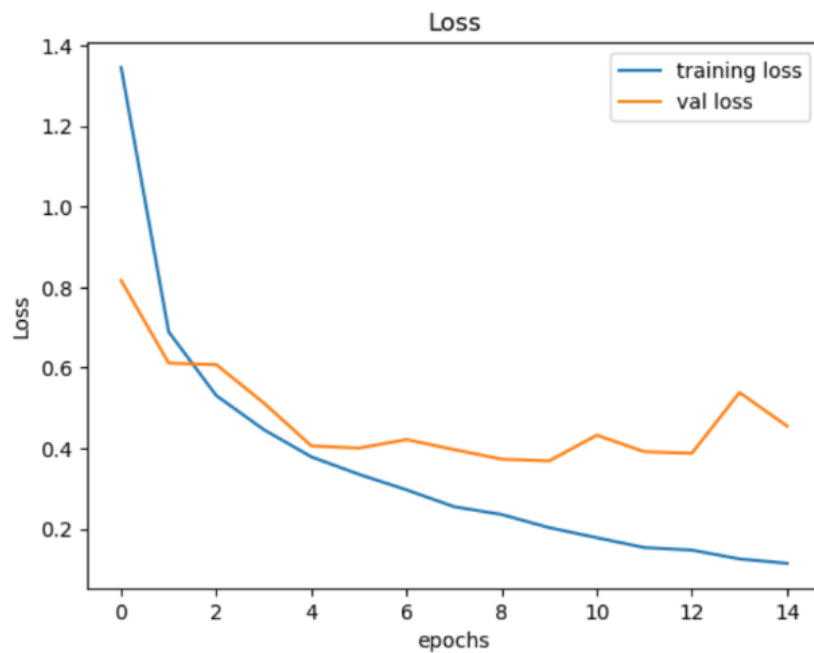
11. Then we evaluate the test score and maximum accuracy achieved.

```
In [54]:  ▶  score = model.evaluate(X_test, y_test, verbose = 1)
              print("Test Score: ", score[0])
              print("Test accuracy: ", score[1])

              129/129 [==============================] - 2s 16ms/step - loss: 0.4553 - accuracy: 0.8735
              Test Score:  0.4553268849849701
              Test accuracy:  0.8735465407371521
```

```
In [55]:  ▶  preds = np.round(model.predict(X_test),0)
              print('Rounded test labels',preds)

              129/129 [==============================] - 2s 16ms/step
              Rounded test labels [[0. 0. 0. ... 0. 0. 0.]
               [0. 0. 0. ... 0. 0. 0.]
               [0. 0. 0. ... 0. 0. 1.]
               ...
               [0. 1. 0. ... 0. 0. 0.]
               [0. 1. 0. ... 0. 0. 0.]
               [0. 0. 0. ... 0. 0. 0.]]
```

```
In [56]:  ▶  y_test

   Out[56]:  array([[0., 0., 0., ..., 0., 0., 0.],
                     [0., 0., 0., ..., 0., 0., 0.],
                     [0., 0., 0., ..., 0., 0., 1.],
                     ...,
                     [0., 1., 0., ..., 0., 0., 0.],
                     [0. 1. 0         0  0  0 ]
```

12. Then we display the whole measures in tables with all the species and their different accuracy measures (precision, recall, F1 score, support).

```
In [59]:  ▶  print(classification_report(y_target, pred, target_names = labels))
```

	precision	recall	f1-score	support
Pepper__bell___Bacterial_spot	0.75	0.83	0.79	203
Pepper__bell___healthy	0.87	0.98	0.92	315
Potato___Early_blight	0.80	0.96	0.87	174
Potato___healthy	0.90	0.68	0.78	28
Potato___Late_blight	0.78	0.80	0.79	191
Tomato_Bacterial_spot	0.88	0.97	0.93	432
Tomato_Early_blight	0.86	0.47	0.61	212
Tomato_healthy	0.96	0.96	0.96	335
Tomato_Late_blight	0.79	0.82	0.80	365
Tomato_Leaf_Mold	0.92	0.80	0.86	183
Tomato_Septoria_leaf_spot	0.86	0.79	0.82	373
Tomato_Spider_mites_Two_spotted_spider_mite	0.88	0.90	0.89	326
Tomato__Target_Spot	0.86	0.84	0.85	283
Tomato__Tomato_mosaic_virus	0.90	0.97	0.93	80
Tomato__Tomato_YellowLeaf__Curl_Virus	0.97	0.96	0.97	628
accuracy			0.87	4128
macro avg	0.87	0.85	0.85	4128
weighted avg	0.88	0.87	0.87	4128

13. And here we compare the test image path with the trained algorithm and display the confidence.

```
In [60]:   model.predict(X_test)[1]

           129/129 [==============================] - 2s 16ms/step

Out[60]:   array([2.0816827e-03, 2.0601477e-03, 1.5524390e-06, 1.1552712e-04,
                  2.2907348e-01, 8.2969302e-05, 3.6951390e-03, 4.4757310e-12,
                  7.6048094e-01, 1.7497040e-06, 1.3004364e-03, 1.8163382e-07,
                  1.1056324e-03, 4.6458129e-11, 5.7471533e-07], dtype=float32)
```

```
In [62]:   from tensorflow import keras
           test_img_path = 'E:\sem7\plantDiseaseProject/archive/PlantVillage/Tomato_Spider_mites_Two_spotted_spider_mite/00fa99e8-2605-4
           img = keras.preprocessing.image.load_img(
               test_img_path, target_size=(64, 64)
           )
           img_array = keras.preprocessing.image.img_to_array(img)
           img_array = tf.expand_dims(img_array, 0) # Create a batch

           predictions = model.predict(img_array)
           score = tf.nn.softmax(predictions[0])

           print(
               "This image most likely belongs to {} with a {:.2f} percent confidence."
               .format(labels[np.argmax(score)], 100 * np.max(score))
           )

           1/1 [==============================] - 0s 71ms/step
           This image most likely belongs to Tomato_Septoria_leaf_spot with a 16.26 percent confidence.
```

<u>CONTRIBUTIONS</u>

Contributions by Payal:

My contribution to this project was to do a research study for the completion of this project and, to search for the already existing materials for this project and how we can improvise them. Basically, I did the literature survey and I made the list of all the challenges that may come while the implementation of this project. I searched thoroughly for the work that the researchers have already done on Plant disease detection, and I made a synopsis of it. I also contributed to the methodology part of our project. I have ensured that the used methodology will offer reliability and efficiency.

As this is a machine learning project and it requires technical knowledge and programming hands-on experience, being a Computer Science and Engineering student, I will be doing the partial coding along with my project partners for the implementation of this project and I will also be debugging the code wherever required. Also, I will be working on the implementation of the front-end design of our project. I will give my best effort so that my teammates and I are successful in making this project as useful as possible.

Contributions by Aditya Kesari:

The project on plant disease detection is highly based on the Machine Learning algorithms and the image dataset used. Thus, being a computer science undergraduate student and having some hands-on experience with machine learning algorithms, after reading about the different types of plant diseases that can lead to disruption in the production of these plants, I researched through and came up with different Machine Learning proposals that will be best suited with maximum accuracy. I came up with the algorithms like Deep Convolutional Network and Transfer Learning along with the methodology involved while performing these machine learning algorithms.

While implementing this project, I have contributed by researching on the different types of libraries that can be made use for the project. I wrote pieces of code parallelly with my project partners, did analysis, and debug the code while working on this project. Finding the open source datasets through various resources. The division of the project into the training

algorithm and validation algorithm with the help of these algorithms I provided the input images and validated with the already present dataset and matched the given images and displayed the different accuracy measures for the given image, and thereby performing it for more and more input images and taking the confidence of the accuracy into considerations. Thereby I and my team were able to achieve maximum confidence for our project and with increased accuracy.

Contributions by B. Dhanush Krishna Sai Reddy:

My contribution to this project was to collect the data set which was pre-existed and made it noise free, next part is finding out the algorithms which are used for the detection and next part is that I am working on training the data. I am using algorithms to train the dataset which contains very large number data so I did my research part for collecting the data as well as finding out the best algorithms and training the data using them.As this is a machine learning project and it requires technical knowledge and programming hands-on experience, being a Computer Science and Engineering student, I will be doing the partial coding along with my project partners for the implementation of this project and I will also be debugging the code wherever required. Also, I will be working on the implementation of the front-end design of our project. I will give my best effort so that my teammates and I are successful in making this project as useful as possible.

Contributions by Vinayak M:

My contribution to this project was to do a research study to ensure its success as well as to look for any pre-existing materials and consider how we may improvise with them. In essence, I created a list of all potential hurdles to this project's implementation together with the future scope. I looked everywhere for the study that has already been done on plant disease detection. I also contributed to the project's conclusion and future scope. I have made sure that the concepts adopted for the foreseeable future will be dependable and effective. Being an Aerospace Engineer, I will be doing some of the partial coding along with my project partners for the implementation of this project since it is a machine learning project that requires technical expertise and hands-on programming experience. I will also plan for an industry ready product with drones. Also, I will be working on the implementation of the front-end design of our project as I have some knowledge about HTML. I'll do all in my power to ensure that my team and I are successful in maximising the use of this project.

Contributions by Haffis Najeeb:

My contribution to this project was to collect the dataset and train both the training and validation dataset. After training it several times I found errors and issues with the dataset which I rectified, It took me several tries to get the accuracy up for both datasets. I also helped in the literature survey to find the challenges we may face during our execution.

As part of this project, I am going to work with my project partners to implement partial coding and debug the code as needed. Due to my technical expertise and hands-on programming experience, I am qualified to complete this machine learning project as a student of computer science and engineering. Moreover, I will work on implementing the front-end design of our project. In order to make this project as beneficial as possible, I will do everything in my power to ensure that my team and I succeed.

CONCLUSIONS AND SCOPE OF FUTURE WORK

Here are some of the different areas of work that can be explored with the ideation of crop disease detection using MLand deep learning concepts.

• *"Disease Identification"*. Out of every key topic in all research, plant disease detection is mainly the main thing to have been concentrated on by everyone. Every illness has several stages. Only one form of disease identification is focused on by the majority of researchers in their study yet none of these efforts specifically target any one disease identification.

• *"Estimations of a Disease"*. The recognition of a specific disease is an intriguing topic that needs to be researched. Despite extensive research in this field, few scientists have determined how much harm the illness has actually caused. They can be extremely beneficial because treatment options can be selected based on the severity of the disease. This type of quantification will determine the percentage of a particular culture that has the disease.

• *"Mobile and Online Applications"*. Numerous approaches to the implementation of detection and diagnosis have been published in the literature. However, only a small number of the websites and smartphone applications are openly accessible online. Some of these programs, such as Leaf Doctor and Assess Software, are accessible to the general public. These programs, however, focus on leaving photos with a flat, all-black background.

• *"Investigating Transfer Learning to Expand Data Size"*. Similarly, it is not very pleasant to note that the plant disease detection problem is approached through all those complexities while the current scenario of innovations in cv is going very quickly towards dl. To investigate the given complexity of the dataset, the best is transfer learning, especially while training phase is going on. An adaptable heterogeneous domain strategy can be used to study knowledge transmission.

REFERENCES

1. C. H. Bock, G. H. Poole, P. E. Parker, and T. R. Gottwald, "Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging," *Critical Reviews in Plant Sciences*, vol. 29, no. 2, pp. 59–107, 2010.View at: Publisher Site | Google Scholar

2. A. M. Mutka and R. S. Bart, "Image-based phenotyping of plant disease symptoms," *Frontiers in Plant Science*, vol. 5, article no. 734, 2015.View at: Publisher Site | Google Scholar

3. "Plant Polyphenols, Prenatal Development and Health Outcomes," *Biological Systems: Open Access*, vol. 03, no. 01, 2014.View at: Publisher Site | Google Scholar

4. G. Wang, Y. Sun, and J. Wang, "Automatic image-based plant disease severity estimation using deep learning," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 2917536, 8 pages, 2017.

5. A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: a survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.

6. J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.View at: Publisher Site | Google Scholar

7. R. I. Hasan, S. M. Yusuf, and L. Alzubaidi, "Review of the state of the art of deep learning for plant diseases: a broad analysis and discussion," *Plants*, vol. 9, no. 10, article 1302, 2020.

8. K. Lin, L. Gong, Y. Huang, C. Liu, and J. Pan, "Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network," *Frontiers in Plant Science*, vol. 10, p. 155, 2019.

9. J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosystems Engineering*, vol. 172, pp. 84–91, 2018.

10. A. Ramcharan, P. McCloskey, K. Baranowski, et al., "A mobile-based deep learning model for cassava disease diagnosis," *Frontiers in Plant Science*, vol. 10, p. 272, 2019.

11. A. Dhakal and S. Shakya, "Image-based plant disease detection with deep learning," *International Journal of Computer Trends and Technology*, vol. 61, no. 1, pp. 26–29, 2018.

12. S. Wallelign, M. Polceanu, and C. Buche, "Soybean plant disease identification using convolutional neural network," in *Proc. Thirty-First International Florida Artificial Intelligence Research Society Conference (FLAIRS-31)*, pp. 146–151, Melbourne, FL, USA, 2018

13. M. Türkoğlu and D. Hanbay, "Plant disease and pest detection using deep learning-based features," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 27, no. 3, pp. 1636–1651, 2019.View at: Publisher Site | Google Scholar

14. A. Fuentes, S. Yoon, S. Kim, and D. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, article 2022, 2017.View at: Publisher Site | Google Scholar

15. M. Loey, A. ElSawy, and M. Afify, "Deep learning in plant diseases detection for agricultural crops," *International Journal of Service Science, Management, Engineering, and Technology*, vol. 11, no. 2, pp. 41–58, 2020.View at: Publisher Site | Google Scholar

16. K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathysubramanian, "Plant disease identification using explainable 3D deep learning on hyperspectral images," *Plant Methods*, vol. 15, no. 1, p. 98, 2019.View at: Publisher Site | Google Scholar

17. J. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Computers and Electronics in Agriculture*, vol. 153, pp. 46–53, 2018.View at: Publisher Site | Google Scholar

18. B. Ashqar and S. Abu-Naser, "Image-based tomato leaves diseases detection using deep learning," *International Journal of Engineering Research*, vol. 2, no. 12, pp. 10–16, 2019.View at: Google Scholar

19. H. Durmus, E. O. Gunes, and M. Kirci, "Disease detection on the leaves of the tomato plants by using deep learning," in *Proc. 6th International Conference on Agro-Geoinformatics*, pp. 1–5, Fairfax, VA, 2017.View at: Google Scholar

20. S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3289801, 11 pages, 2016.View at: Publisher Site | Google Scholar

21. A. Rangarajan, R. Purushothaman, and A. Ramesh, "Tomato crop disease classification using pre-trained deep learning algorithm," *Procedia Computer Science*, vol. 133, pp. 1040–1047, 2018.View at: Publisher Site | Google Scholar

paper

6% SIMILARITY INDEX **5**% INTERNET SOURCES **3**% PUBLICATIONS **2**% STUDENT PAPERS

PRIMARY SOURCES

1	www.hindawi.com Internet Source	**2**%
2	www.coursehero.com Internet Source	**1**%
3	scholar.archive.org Internet Source	**1**%
4	www.itm-conferences.org Internet Source	**1**%
5	Manish Kumar, Ahlad Kumar, Vinay S. Palaparthy. "Soil Sensors Based Prediction System for Plant Diseases using Exploratory Data Analysis and Machine Learning", IEEE Sensors Journal, 2020 Publication	**1**%
6	www.allfreeessays.com Internet Source	**1**%
7	Narendra Pal Singh Rathore, Lalji Prasad. "Hybrid Deep Learning Model to Detect Uncertain Diseases in Wheat Leaves", Journal of Uncertain Systems, 2022	**<1**%

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

paper

FINAL GRADE

/100

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20