Web Based GUI Detector to Recognise Tomato Plant Leaf Disease

Himanshu Pant[1], Janmejay Pant[2], Pawan Kumar Pant[3], Devendra Singh[4], Manoj Chandra Lohani[5]
[1]Graphic Era Hill University Bhimtal, Uttarakhand, India, himpant7@gmail.com
[2]Graphic Era Hill University Bhimtal, Uttarakhand, India, geujay2010@gmail.com
[3]Graphic Era Hill University Bhimtal, Uttarakhand, India, pantpawan@gmail.com
[4]Graphic Era Hill University Bhimtal, Uttarakhand, India, devendrasuno@gmail.com
[5]Graphic Era Hill University Bhimtal, Uttarakhand, India, getmlohani@gmail.com

ABSTRACT

Plants are the significant factors for the planet and all the liveliness. It creates habitats for numerous organisms. Plants are the main source of food for living things. All organisms use various parts of diverse plants as a food. Diverse symptoms and conditions of the plants avert to perform a maximum potential and healthy growth of the food. The main reason to decline the productivity of the crops is plant disease. We can visualize the symptoms of the disease of the numerous plant leaves which are used to recognize and categorized healthy plants on the basis of healthy leaves correctly. On the basis of accurate and early diagnosis of the plants leaves disease we can increase the productivity and quality of the crops and hence increase the farmers' incomes. Earlier, we can identify the plant conditions and quality of the leaves by only human visualization and observation by experts or testing procedures on the laboratories. These procedures required the plant or horticulture experts. Laboratory testing is a long enough and arduous approach hence outcome fails to come on time. To resolve these challenges, we have proposed and designed a graphical user interface based application which select, detect and predict plant disease classes using Convolutional neural networks (CNN). For this purpose tomato plants leaves are preferred, which is a group member of Solanum lycopersicum species. These tomato leave images are collected from the plant village dataset which are categorized in five classes: Bacterial Spot, Early Blight, Late Blight, Yellow Leaf curl virus and Healthy Leave. In this article Convolutional neural networks (CNN) offers a saner solution in numerous disease controls for tomato leaves with high precision of validation. We have made a web based GUI predictor to identify the tomato plant leaves class accurately so that farmers can easily select and predict for disease in the crop and take appropriate action for the remedy.

Key words : Accuracy, Convolutional Neural Networks, Graphical User Interface, Image classification, leaf disease, Tomato Leaves

1. INTRODUCTION

India is the thriving nation which has adequate amount of human and natural resources. Approximate 70% of Indian economy is based on farming and food production sector [1]. These sectors comprise numerous components. Plants condition and food security are closely linked together. The Indian Plants and food sector estimates that plant diseases can drop up to 20 ~40% of total food production. Adequate and healthy plants deliver healthy foods for living things. Infected leaves and diseases in plants are solemn threat for the maintenance of the natural resource. Plant Diseases are harmful to its health which may affect its growth and production. Economic losses increases and food production decrease due to plant syndromes in agriculture sector. Tomato (Solanum lycopersicum) is one of the widely cultivated crops throughout the world after potato [2]. It contains rich nutrition and healthy ingredients which are necessary for the living things. Its production plays a generous role in the agricultural and trade in whole world. The farming section of tomato crops in India holds around 3 lakh 50 thousands hectares and the manufacturing amounts around 53 lakhs tonnes. This manufacturing amount makes India third-biggest tomato producer in the world. Plant disease occurs due to various reasons like pathogens, living microorganisms, bacteria, fungi, nematodes, mycoplasmas, viruses and viroid. There are various tomato leaf diseases which may affects the tomato productions at an alarming rate. Tomato Bacterial Spot, Tomato Early Blight, Tomato Late Blight and Tomato Yellow Leaf curl virus are few of them [3]. These primary diseases of the tomato plant leave reduce the cultivation of the crop. Therefore, timely measurements are required for the assessment of early detection and identification of infected tomato leaves. Plant pathologies can be perceived by using numerous ways. Sometimes we cannot identify and recover the infected tomato leaves by visualize the symptoms because diseases on the plant leaves may envision too late. In old fashion we carried out detailed examination of the diseases in plants by the domain experts in laboratories by means of powerful optical microscope or electromagnetic spectrums which are invisible to naked eyes [4]. Therefore it is necessary to identify infected leaves from

the plant and diagnose their disease in an early stage. But due to lack of early detection of the plant disease, farmers tolerate a huge earning loss in every year. Therefore, the bacterial and fungal infection of the tomato leaves can be diagnose by every farmer in a fast, proficient, an automat, less expensive, precise and correct manner [5]**.** To restrain the plant loss and improvements the quality of the tomatoes, authors proposed an accurate, computerised and automate system that may increase economy of the country [6]. This article recommended a GUI based model to detect the disease on tomatoes leaves and classify the bacterial symptoms that predict infected and non-infected leaves. Horticulture experts professionally analyse the data and then distinguish the leaf disease present in the tomatoes plant [7]**.** Mostly there may be no domain professional available to provide accurate feedback to diagnose the disease in the particular cultivation area. Hence agriculturalists are required a rapid, economical, computerised, exact and automatic system to distinguish the tomatoes leaves disease efficiently. In this research, we have considered the five types of infected tomatoes leaves along with the healthy leaves images and proposed a web based graphical user interface which is based on deep convolutional neural network model (DCNN) to identify disease. We have also proposed a detection and prediction model that diminished the imprecise and inaccurate manual disease detection. This model improves the testing and validation accuracy, feasibility and efficiency of the algorithm. This paper influences by using numerous computer vision techniques to detect and classify five different set of tomato leaf disease including healthy as well as infected leaves from the tomato plants [8]. Baseline convolutional neural networks (CNN) are applied for the same. The primary objective of this exploration was the assortment of a sufficient tomato images data from existing open source datasets. This collected tomato plant leaves data is further classify in five different categories. The suggested GUI based system architecture of tomato plant leaves diseases identification and organisation is shown in figure-1.
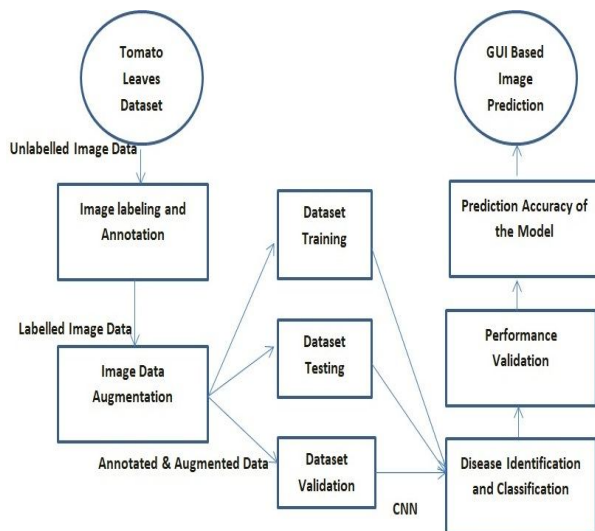
Now days, computer vision is the hot and trending topic among the researchers in which we train the computer to visualize the world. Image classification is one the important and demanding factor of computer vision. Convolutional neural network is one of the challenging algorithms due to its great performance in deep learning models. The CNN models build a relationship between the information present in the image and intermediate layers. These are convenient for the image classification and feature extraction [9]**.** Identification and classification is investigated for the of different rice crop diseases with the help of DCNN model. In the field of plant leaf disease detection, there are limited works on tomato plants using CNN. In this paper we have considered a total 9835 tomato leaves images with five different disease classes. We have used baseline CNN model in which three convolution layers as well as three stochastic layers for pooling and one softmax layer is applied. The 96.87% classification accuracy after augmentation of the dataset has been reported [10].

Following structure has been applied for this research study: In Section- 1 we describe the theoretical foundations and introduction of the proposed model. Dataset descriptions and data pre-processing are provided in section-2. Section-3 delivers the details of experimental performance on tomato leaves image dataset using convolutional neural network. In section-4 we proposed a graphical user interface based model to predict the correct plant leaves using their hidden feature which were already extracted in CNN model. Model conclusion and future scope for the project and research is described in section-5.

2. MATERIALS AND METHODS

This section performed the numerous stages and processes on the tomato leaves dataset. Figure 2 shows the various actions performed on the huge assorted image dataset and select healthy tomato leaves from the group of infected tomato leaves.
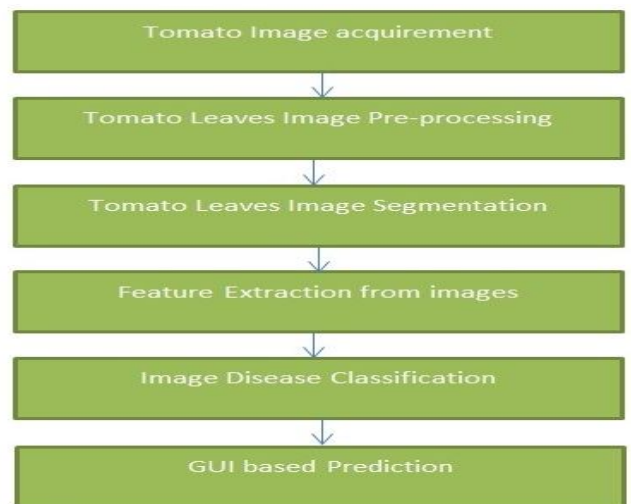

Figure 1: Tomato Plant Leaves Disease Identification and Classification Model

Figure 2: Tomato Leaves Disease Detection Architecture

2.1 DATASET DESCRIPTIONS

Correct data collection plays a challenging and an important role in convolutional neural network. An enormous quantity of the tomato plant leaves are required to perceive and categorize the syndromes connected with it. We have collected different images for the tomato leaf disease from standard open source datasets. We build our dataset by labelling those images. The concluding dataset consist of five different classes having 9835 images. Details of the five classes have been given in Table I. This table also shows the distribution of each class with training and testing dataset. The proposed work used four sets of infected images, namely Bacterial Spot, Early Blight, Late Blight, Yellow Leaf curl virus and one set of healthy images. Table-1 shows the distribution of tomato leaves dataset in training and testing phases.

Table 1: Distribution of Tomato Leaves in classes

| Tomato Plant Leaf class Name	Number of images in Training Set	Number of images in Testing Set	Total Leaves in each class
Tomato Bacterial Spot	1613	514	2127
Tomato Early Blight	742	258	1000
Tomato Late Blight	1446	463	1909
Tomato Yellow Leaf curl virus	2357	851	3208
Tomato Healthy Leaf	1168	423	1591
Total Images	**7326**	**2509**	**9835**

For the training and testing purpose the dataset is split into 75:25 proportions. Here all the images are in uncompressed JPG format and belong to the RGB color space.

2.2 DATA PRE-PROCESSING

The collected tomato leaves are originally un-annotated and unlabelled form. Five different directories for five types of tomato images are created and performed the annotation and labelling process of the images by their directory name with a word "TBS", "TEB", "TLB", "TYL" and "THL" respectively. Initially the collected images are in RGB colour form. All these images have different dimensions, figures and sizes. So it is necessary to reshaped all the images in the same size, shape and dimensions after annotation for image processing because features can be easily extract from the same sized images [11]. Keras and tensor flow image processing application programming interface are used to achieve the standard size of images in the dimension of 200×200 pixels photos. Some standard fix pixel sized sample images of all the five categories are shown in the table-2. To separate all five image class in to respected directories, an Image-data-generator class and flow_from_directory API tool

of keras are used. The API generates a separate test directories for train and test sets, and within each directory all five class directories are formed.

Table 2: Sample Tomato leaves Images of all Disease categories
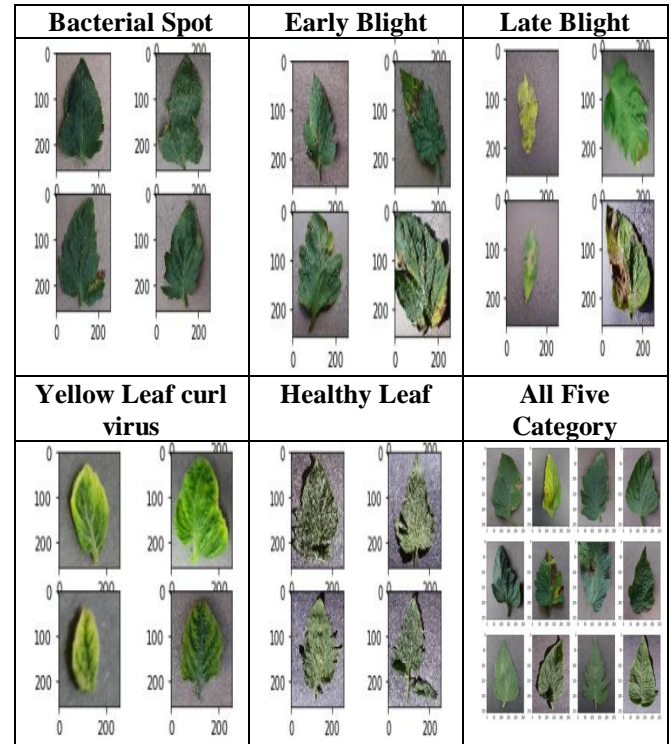
3. EXPERIMENT SETUP

The experimental performances for the tomato leaves dataset were evaluated by using three block baseline convolution neural network model. This model classifies and identifies healthy tomato leaves from the bunch of infected bacterial leaves. In the model we have used a deep learning technique for image classification, image data processing and recognition of the objects from an image. We have tracked the basic and standard VGG architectural principles for the simulation. For this we have proposed a baseline architecture which assembles convolutional layers having 3×3 small sized filters. The filters are trailed by a max pooling layer. So a convolutional block is built by convolutional layers and Max-pooling layers putting all together. When size of each convolutional block is increase by increasing the number of filters in the depth of the network. These blocks can be repeated in the entire neural network. We have used padding on the convolutional layers, to make sure the height and width shape of the output feature maps with respect to various inputs. VGG classification architecture is applied on the various healthy tomato leaves image and different infected tomato leaves image. This model performs a classification task and predicts healthy tomato images from the set of

numerous infected images. Each layer in this designed convolutional model is activated by a non-linear function known as rectified linear unit function (ReLU activation function).The considered baseline convolutional neural network is a sequential model having total 10,334,021 trainable parameters is used to fit as shown in **figure-3**. For Stochastic gradient descent optimizer, to train this sequential model, we have used the 0.001 as conservative learning rate and 0.9 as momentum.

```
Model: "sequential_4"

Layer (type)                    Output Shape              Param #
=================================================================
conv2d_10 (Conv2D)              (None, 200, 200, 32)      896

max_pooling2d_10 (MaxPooling    (None, 100, 100, 32)      0

conv2d_11 (Conv2D)              (None, 100, 100, 64)      18496

max_pooling2d_11 (MaxPooling    (None, 50, 50, 64)        0

conv2d_12 (Conv2D)              (None, 50, 50, 128)       73856

max_pooling2d_12 (MaxPooling    (None, 25, 25, 128)       0

flatten_4 (Flatten)             (None, 80000)             0

dense_7 (Dense)                 (None, 128)               10240128

dense_8 (Dense)                 (None, 5)                 645
=================================================================
Total params: 10,334,021
Trainable params: 10,334,021
Non-trainable params: 0
```

Figure 3: Sequential Model Parameters and Neural Network Layers

Train iterator method is used to fit the proposed model and dataset can be validating by applying test iterator method. 115 steps are used with respect to each 50 epoch for the training and testing the convolutional sequential model. It Total number of training steps with respect to each epoch can be calculated by dividing the total number of training and testing tomato images (all five categories) in the train and test directories by the taken batch size (In this model authors have taken 64 batch size).

4. PERFORMANCE MEASUREMENT

4.1 THREE BLOCK BASELINE VGG MODEL

The performance of the sequential model can be evaluated by applying unseen testing dataset to the fitted trained model. The performance of the model can be measured in terms of validation accuracy. The validation accuracy of three block sequential VGG model can be evaluated by increasing the convolutional layer and pooling layer with respect to one block and two block VGG architecture. We have used three block VGG architecture for the same. When dataset is too less for the training then model required a technique to increase

the dataset without changing the base theme. This mage modification is done by image data augmentation technique. This approach allows us to increase the current training dataset devoid of collecting new training data. Image augmentation is a process that performs various operations like image cropping, image padding, image rotation, translation and scaling and horizontal flipping on existing dataset to create updated new dataset. This is done by randomly shifting the existing images in either horizontal direction or vertical direction.

Classification performance of this sequential convolutional neural model after augmentation can be measured mathematically by creating a matrix known as confusing matrix. Confusion matrix visualizes the performance of developed CNN model in the rows and column form [12]**.** The classification report is used to describe precision, recall, f1-score of a proposed classifier to show the high performance. In this article authors proposed multi class grouping problem to identify and classify the tomato plant leaves disease. The classification report for three block baseline VGG model after augmentation and the confusion matrix for the visualized classification with 96.87% accuracy are shown in the table-3.

Table 3: Multiclass Confusion Matrix and Classification Report

```
Confusion Matrix
[[100  53 113  90 158]
 [ 62  29  46  50  71]
 [ 82  55  71  77 143]
 [ 94  58 104  76 157]
 [177  94 153 150 282]]
```

```
Classification Report

              precision    recall  f1-score   support

       TBS       0.19      0.19      0.19       514
       TEB       0.10      0.11      0.11       258
       THL       0.15      0.17      0.16       428
       TLB       0.17      0.16      0.16       489
       TYC       0.35      0.33      0.34       856

  accuracy                           0.22      2545
 macro avg       0.19      0.19      0.19      2545
weighted avg     0.22      0.22      0.22      2545
```

After hyper-parameters and fine tuning this three block baseline sequential VGG model with augmentation, model achieves excellent performance with 96.87% training accuracy and 93.124% validation accuracy as shown in figure-4.

```
Epoch 40/50
115/115 [==============================] - 74s 645ms/step - loss: 0.1014 - accuracy: 0.9611 - val_loss: 0.2041 - val_accuracy: 0.9485
Epoch 41/50
115/115 [==============================] - 75s 651ms/step - loss: 0.1096 - accuracy: 0.9607 - val_loss: 0.0602 - val_accuracy: 0.9580
Epoch 42/50
115/115 [==============================] - 75s 649ms/step - loss: 0.0993 - accuracy: 0.9609 - val_loss: 0.1134 - val_accuracy: 0.9513
Epoch 43/50
115/115 [==============================] - 74s 646ms/step - loss: 0.0915 - accuracy: 0.9656 - val_loss: 0.1249 - val_accuracy: 0.9375
Epoch 44/50
115/115 [==============================] - 74s 645ms/step - loss: 0.1011 - accuracy: 0.9637 - val_loss: 0.1015 - val_accuracy: 0.9564
Epoch 45/50
115/115 [==============================] - 75s 648ms/step - loss: 0.0910 - accuracy: 0.9664 - val_loss: 0.0812 - val_accuracy: 0.9587
Epoch 46/50
115/115 [==============================] - 74s 646ms/step - loss: 0.0932 - accuracy: 0.9656 - val_loss: 0.0327 - val_accuracy: 0.9639
Epoch 47/50
115/115 [==============================] - 74s 645ms/step - loss: 0.0874 - accuracy: 0.9672 - val_loss: 0.0352 - val_accuracy: 0.9540
Epoch 48/50
115/115 [==============================] - 74s 647ms/step - loss: 0.0979 - accuracy: 0.9628 - val_loss: 0.5060 - val_accuracy: 0.9171
Epoch 49/50
115/115 [==============================] - 75s 653ms/step - loss: 0.0956 - accuracy: 0.9647 - val_loss: 0.1170 - val_accuracy: 0.9627
Epoch 50/50
115/115 [==============================] - 75s 651ms/step - loss: 0.0850 - accuracy: 0.9687 - val_loss: 0.2948 - val_accuracy: 0.9312

Accuracy-> 93.124
```

Figure 4: Training and Validation classification Accuracy

Over fitting and under fitting of this sequential model can be controlled by analysis of the classification accuracy and cross entropy loss with respect to each epoch is shown in figure-5.
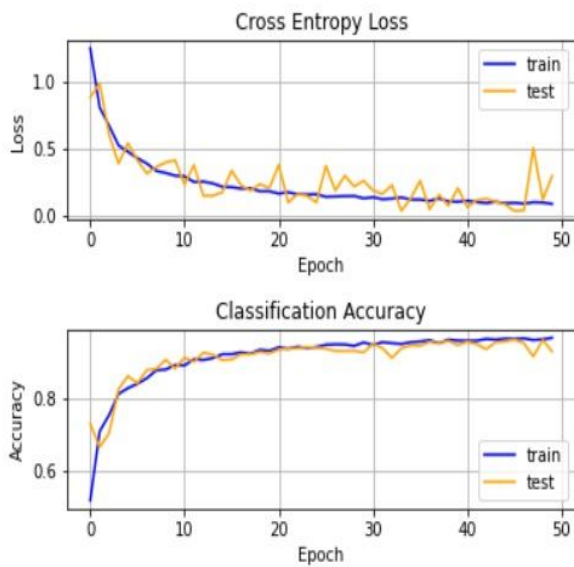

Figure 5: Classification model performance and Cross Entropy loss with respect to Number of epoch

4.2 WEB BASED GUI TOMATO LEAF DISEASE DETECTOR

Proposed deep neural network model using CNN to the end users or systems or farmers to predict accurate tomato leaf can be deployed using web application developed in keras and flask API. Convolutional neural network is used here as a classifying and predicting the model. Training part of the data can be feed by training the CNN model [13]. After training the CNN model the next part was to make a flask API which receives tomato plant leaves images through web GUI and computes the predicted and accurate leaf from our model [14]. This Graphical user interface is divided in to three sub modules: home page, select tomato image dialog box and predict desired leaf disease detection page.

4.2.1 Home Page

Once the baseline model is developed we need to make the HTML web pages using flask. We set the main page using the name index.html as shown in **figure-6.** Flask framework is a very minimal because it needs minimal configuration to execute and it has an inbuilt light-weighted web server. Flask server can be controlled by own model.
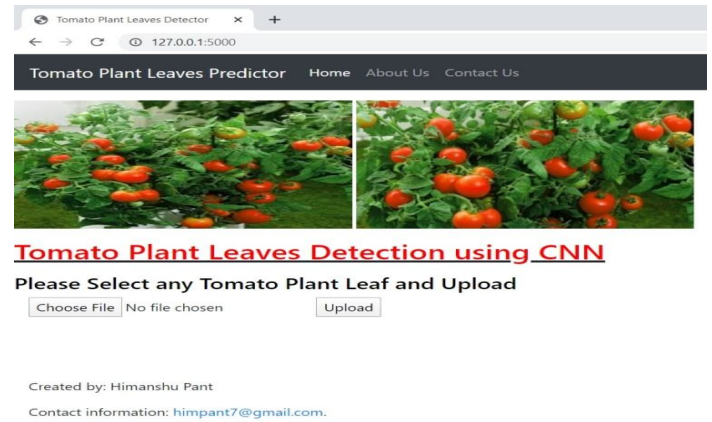

Figure 6: Web Based Tomato plant leaves Disease Detector Home Page

4.2.2 Selection the random Images

In the home page of the graphical user interface, farmers can select any tomato plant leaf randomly from anywhere or they can select the particular unseen testing images from our dataset which shown like the **figure-7**.
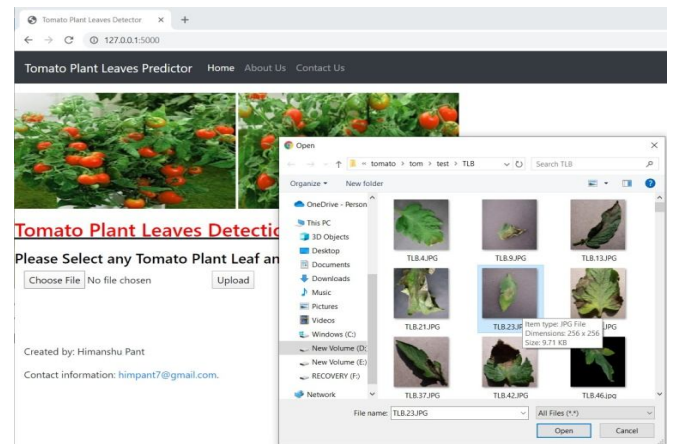

Figure 7: Web Based Tomato plant leaves Disease Detector Select test image Page

4.2.3 Prediction of accurate tomato leaf

After selecting a random image from the unknown source or our unseen testing dataset, model should predict the accurate result so that any farmers can easily predict and classify that the selected leaf is bacterial infected or healthy. In our training model the farmers can predict the approximate 97% accurate result [15]. The prediction of the image is look like as shown in figure-8.
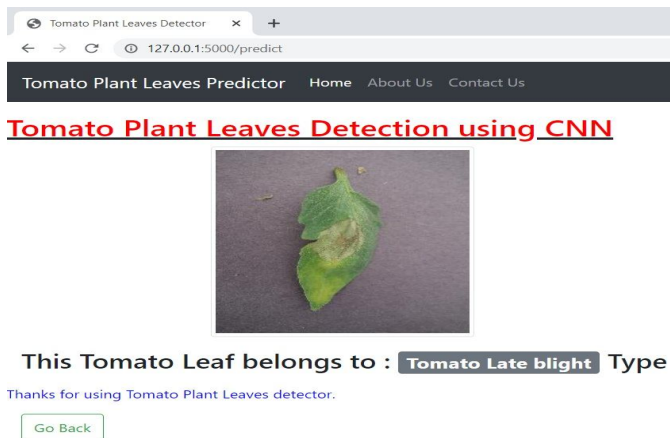

Figure 8: Web Based Tomato plant leaves Disease Detector predict Page

5. CONCLUSION

In this article baseline sequential convolutional neural network model is applied to select, discover, recognize and classify the accurate tomato plant disease in this research. A numerous researchers doing research for disease identification, classification and automation of tomato plant leaves. Sequential convolutional neural network with three convolutional block VGG architecture is applied to explore the grouping and identification of the tomato plant leaf disease. Moreover, on the basis of designed model we had also developed a graphical user interface based detector using flask and keras API on the training dataset to predict the model on real world environment. The projected convolutional model and GUI based detector are perfectly capable to classify bacterial infested and vigorous plants with training classification accuracy with 96.87% and validation accuracy with 93.124%. The proposed model classifies different types of tomato leaves in four bacterial infected classes and one healthy class. This CNN approach performed outstanding with compared to other classification algorithms by means of various statistical measures.

6. FUTURE SCOPE

Prediction accuracy of the model can be improved by using large scale training dataset. The accuracy of tomato leaf disease prediction and classification can be increased by putting various pre trained transfer learning models like VGG 16, VGG 19, Alexnet etc. In future authors may apply other computer vision algorithmic techniques like plant object detection and leaf segmentation on the plant dataset. Disease prediction and detection from the infected plants can be performed in other categories of plants. Moreover in future a mobile phone android application for the plant disease detection may be developed so that all the farmers can easily predict the disease on the plant and take only healthy leaves for their purpose and improve their earning for selecting healthy crops for the farming.

REFERENCES

1. India economic survey: **"Farmers gain as agriculture mechanisation speeds up, but more R&D needed".** The Financial Express. January 2018.

2. Smith, A. F., "The Tomato in America: Early History, Culture, and Cookery". Columbia SC, US: University of South Carolina Press, p. 152- 1994.

3. Nowicki, Marcin, **"Potato and tomato late blight caused by Phytophthora infestans: An overview of pathology and resistance breeding"**, Plant Disease, Plant   Disease, ASP, 96:4–17, doi:10.1094/PDIS-05-11-0458, PMID 3 0731850, retrieved 2011-08-30

4. J. Garcia A. Barbedo **"Digital image processing techniques for detecting, quantifying and classifying plant diseases".** Springer Plus 2(1):660,-2013 DOI: 10.1186/2193-1801-2-660

5. K.A. Vakilian, J. Massah, **"An artificial neural network approach to identify fungal diseases of cucumber (Cucumissativus L.) Plants using digital image processing"**, Archives of Phytopathology and Plant Protection, Vol. 46, Issue 13, Taylor & Francis-2013, pp.1580-1588.)

6. Phadikar, S.Sil, Das A. K. **"Rice diseases classification using feature selection and rule generation techniques"**. Computers and Electronics in Agriculture, 90, pp. 76–85.-2013.

7. Camargoa, J.S. Smith, **"An image-processing based algorithm to automatically identify plant disease visual symptoms"**, Biosyst Eng., Vol 102:9–21, 2009).

8. Rothe, R. Kshirsagar, **"Cotton leaf disease identification using pattern recognition techniques,"** in International Conference on Pervasive Computing, Pune,- 2015.

9. S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk and D. Stefanovic, **"Deep neural networks based recognition of plant diseases by leaf image classification,"**- 2016.

10. F. Qin, Liu. D, Sun B., L. Ruan, Ma Z. and H. Wang, **"Identification of alfalfa leaf diseases using image"-2016.**

11. M. Dhakate and A. B Ingole, **"Diagnosis of Pomegranate Plant Diseases using Neural Network",** IEEE 5th –NCVPRIPG , Patna -2015.

12. Fuentes, A.Yoon, S.,SC Kim., and DS Park **"A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition."** Sensors 2017. doi: 10.3390/s17092022

13. Kundal Mondal, Dr Padmavati S, **Wild animal Detection and Recognition from Aerial Videos using Computer Vision Technique**,International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 7, Issue 5, may (2019).

14. Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez-Vaamonde, E., Diez, A. **"Automatic plant disease diagnosis using mobile capture devices applied on a wheat use case"**. Comput. Electr. Agri. 138,200–209.–2017doi:10.1016/j.compag.2017.04.01

15. Aditi P. Marathe, Avinash J. Agrawal, **Improving the Accuracy of Spam Message Filtering using Hybrid CNN Classification,**International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 8, Issue 5, may (2020).