IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ※ Peer-reviewed & Refereed journal ※ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13551

AN ENHANCED TOMATO PLANT DISEASE DETECTION AND CLASSIFICATION METHODOLOGY USING CNN

Adnan Pipawala[1], Dr. Naveen Choudhary[2]

Research Scholar, College of Technology and Engineering, Udaipur[1]

Professor, Department of Computer Science and Engineering, College of Technology and Engineering, Udaipur[2]

Abstract: Agriculture plays an important role in the growth of a country; also, economic growth relies on the quality of the crops produced which is proportional to the diseases occurring on it. The problem occurs when the leaves of the plants get affected by multiple diseases, which requires a solution by accurately detecting the disease. Here we will be taking tomato leaves with multiple diseases into consideration. The research is based on CNN based architecture VGG16 which helps to achieve accuracy above 92% when performed on tomato leaves dataset which consist of eleven classes. PlantVillage and Tomato Leaf Diseases are two dataset sources for the collection of images for the model.

Keywords: Convolution Neural Network, VGG16, Tomato leaf diseases detection, Tomato leaf diseases classification

INTRODUCTION

In the field of agriculture, identifying types of plant disease is extremely important and is considered a crucial issue. The diagnosis of leaf disease may pave the way for better decision making in managing agriculture production. Infected plants generally have obvious marks or spots on the stems, fruits, leaves, or flowers. Therefore, detection of tomato leaf diseases is important to identify the diseases before too much damage is done to the leaf.

Tomato is the most profit-making plant, and the farmers cultivate tomato and its leaves as they are widely used all over country. It would maximize the yield, minimize the economic losses, and will enhance the quality of tomato leaves.

Manual disease detection is very difficult and time consuming on crops like tomatoes and the accuracy depends upon the knowledge of an expert. The research in image understanding and machine learning in the last two decades suggests that automation can give solutions to the problems raised in manual systems. Effective tomato leaf disease detection using machine learning is crucial to predict the disease at a very early stage. Machine learning has found its application in various areas such as crop management, yield prediction, disease detection and weed detection crop quality.

The proposed methodology shall be able to detect multiple diseases on tomato leaves as the leaf part has its importance for healthcare and food industry.

So novel, and rapid methods for the timely detection of tomato leaf diseases will allow surveil and develop control measures with greater efficiency.

The key contribution of this paper is as follows:

1. To prepare a deep learning-based model for analyzing a variety of unique diseases like bacterial spot, early blight, late blight, leaf mold, Septoria leaf spot, spotted spider mites, target spot, yellow curl virus, mosaic virus and powdery mildew.
2. To design and develop a novel CNN machine learning based approach VGG16 for tomato disease detection with appropriate customization and classification for eleven classes.
3. To evaluate and validate the performance metrics with existing popular techniques of the domain with accuracy, sensitivity, and specificity.

LITERATURE REVIEW

In 2015, Mokhtar et al. [1] the study described in this paper consist of a method that applies Gabor wavelet transform techniques to extract relevant features related to image of tomato leaf in conjunction with using Support Vector Machines

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13551

(SVMs) with alternate kernel functions to detect and identify type of disease that infects tomato plant. In 2016, Sabrol and Satish [2] in the study, five types of tomato diseases i.e., tomato late blight, Septoria spot, bacterial spot, bacterial canker, tomato leaf curl and healthy tomato plant leaf and stem images are classified. In 2017, Shijie et al. [3] this paper explores the detection algorithm on leaf images and constructs the convolutional neural network model to detect tomato pests and diseases based on VGG16 and transfer learning.

In 2018, Luna et al. [4] the system used convolutional neural network to identify which of the tomato diseases is present on the monitored tomato plants using F-RCNN. Sardogan et al. [5] this paper presents a CNN model and learning Vector Quantization (LVQ) algorithm-based method for tomato leaf disease detection and classification. Suryawati et al. [6] in this paper, we evaluated the effect of different depth of CNN architectures on the detection accuracies of the plant disease detection. Tm et al. [7] this paper adopts a slight variation of the convolution neural network model called LeNet to detect and identify diseases in tomato leaves.

In 2019, Elhassouny and Smarandache [8] in their study proposed an efficient smart mobile application model based on deep CNN to recognize tomato leaf diseases. Kumar and Vani [9] performed experiments with the CNN architectures like- LeNet, VGGNet, ResNet50 and Xception for detecting diseases in tomato leaves for better accuracy. Militante et al. [10] proposed CNN on both vegetable and fruit plants. They take into consideration six crop plants such as apple, tomato, grapes, sugarcane, potato, and corn. They have managed to secure good accuracy using deep learning methods. Rahman et al. [11] study surveyed on tomato crop where four classes of tomato leaf are taken namely bacterial spot, late blight, septorial spot, and healthy leaf. They achieve better accuracy by using morphological methods.

In 2020, Ashok et al. [12] paper proposes to identify the tomato plant leaf disease using image processing techniques based on image segmentation, clustering, and open-source algorithms, thus all contributing to a reliable, safe, and accurate system of leaf disease of tomato plants. Kirti and Rajpal [13] they work on grape diseases on black rot over 400 images consist of 250 images. 150 images are available for testing purposes applied SVM classifier to achieve better accuracy. Marzougui et al. [14] they work on the plant image to classify it into either a healthy, disease-free crop plant or an unhealthy crop plant that is diseased contains 250 images each using augmentation with different models like AlexNet, ResNet, GoogleNet for better results.

In 2020, Rajasekaran et al. [15] proposed VGG16 architecture is considered which accounts for good accuracy of various image processing techniques are applied to computer vision on turmeric plant leaves. Zhang et al. [16] the paper focuses on improve the recognition model accuracy of crop disease leaves and locating diseased leaves, this paper proposes an improved Faster RCNN to detect healthy tomato leaves and four diseases: powdery mildew, blight, leaf mold fungus and ToMV.

In 2021, David et al. [17] paper introduces a comprehensive analysis of the disease classification and detection techniques implied for tomato leaf disease identification using computer vision-based technology. Guan [18] this study has developed a new plant diseases detection approach by combining four CNN models includes Inception, ResNet, Inception ResNet, and DenseNet were deployed, and the results of CNN models were processed by stacking method. Srinidhi et al. [19] applied their study to the apple fruit plant consisting of four classes named rust, scab, multiple diseases, and healthy. They achieved better accuracy using EfficientNet and DenseNet configuration.

In 2022, Omar and Jain [20] develop a detection model for tomato leaf diseases to identify whether the plant is well or not. This is performed over ten classes of tomato leaf using CNN.

In 2023, Vini and Rathika [21] This paper explains the application of a Kapur-based thresholding technique for the detection of mosaic, curl virus, and leaf mold infected regions on tomato leaves. And for the classification of these diseases, the features are extracted, and classification is performed using six different machine learning based classifiers: KNN, Ensemble Classifier, Discriminant Analysis, Decision Tree, SVM, and Naive Bayes for comparing results. The proposed method accurately detects the disease infected area and scores good accuracy through the KNN classifier, which outperforms other classifiers. So, while designing computerized diagnosis of mosaic, curl, and leaf mold viruses on tomato plant leaves, the proposed approach can be considered.

PROPOSED METHODOLOGY

To identify the disease leaves data needs to be collected, in our case we are using Kaggle data source comprises of 32,534 publicly available tomato leaf images, with ten diseases images class and one healthy images class dataset. The

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13551

distribution of images per class is as follows: 3,558 for bacterial spot, 3098 for early blight, 3905 for late blight, 3493 for leaf mold, 3628 for Septoria leaf spot, 2182 for spotted spider mites, 2284 for target spot, 2537 for yellow curl virus, 2737 for mosaic virus, 3856 for healthy and 1256 for powdery mildew.

Example of all ten disease leaves images with one healthy leave image shown in Figure 1-

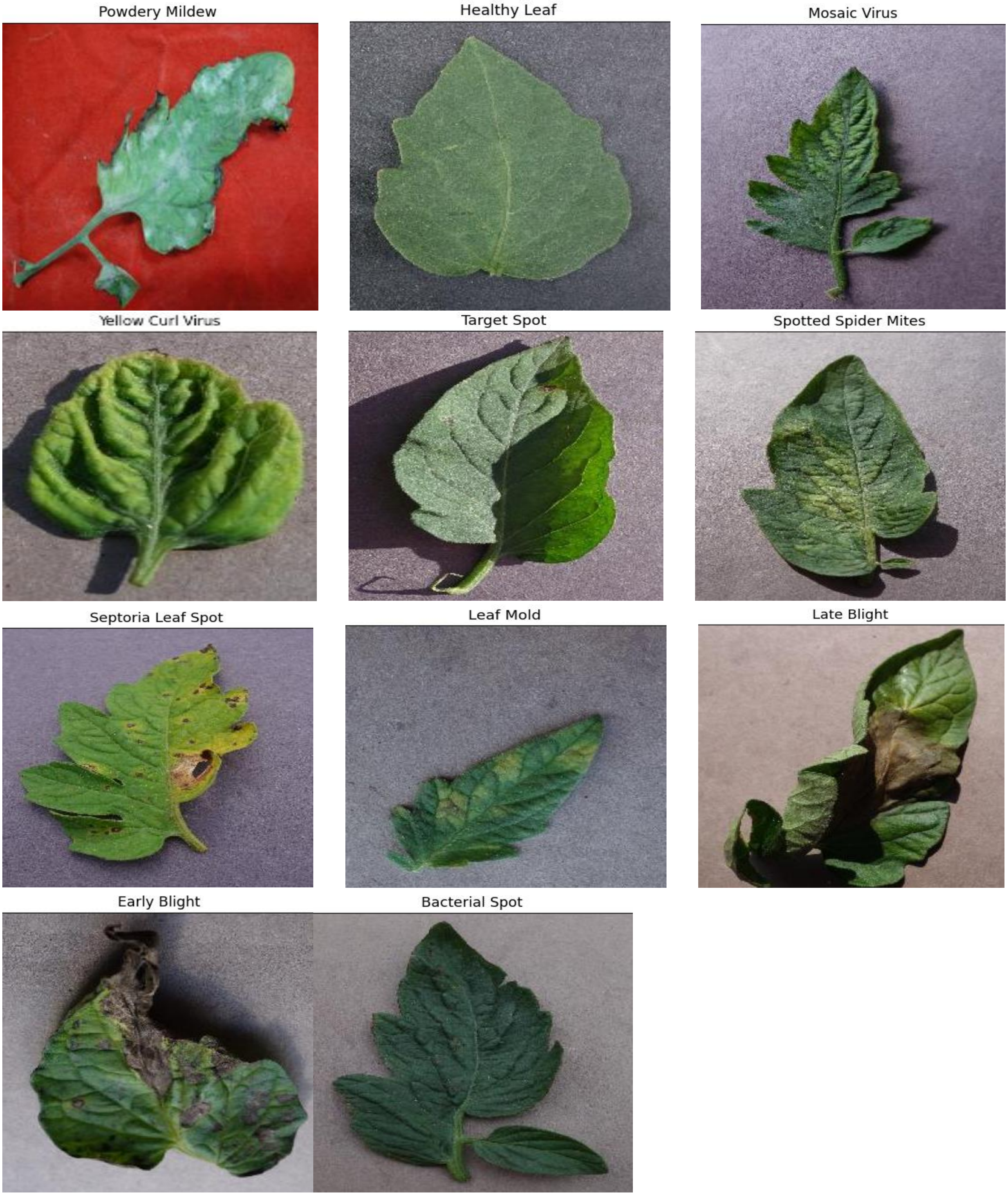

Figure 1. Sample Images

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13551

A. System Flow

A CNN model VGG-16 taken into consideration though it works perfectly well with image dataset. First most important part of CNN model is dataset here in this case we are using online available data source named PlantVillage and Tomato Leaves Dataset where there are only tomato leaves images will be used for our research work.

Next comes the image pre-processing phase where data augmentation method is used to hold extra number of images in addition to images available from data sources where rescaling is done along with shear and zoom range for better accuracy.

Horizontal Flip is assigned to True value this all done for training dataset whereas for testing phase rescaling is done for images.

Then comes extraction of features is done for getting diseased and decayed part of the tomato leaves.

The next step is where CNN VGG-16 model is applied which has sixteen layers. Once the CNN model is applied to dataset then testing is done for the model to check the performance whether it is satisfactory or not. If satisfactory then move to the final stage of classification. If performance is of not satisfactory level, then it goes for hyper parameter tuning and this loop will continue till the fulfillment of aim of research as shown in Figure 2.

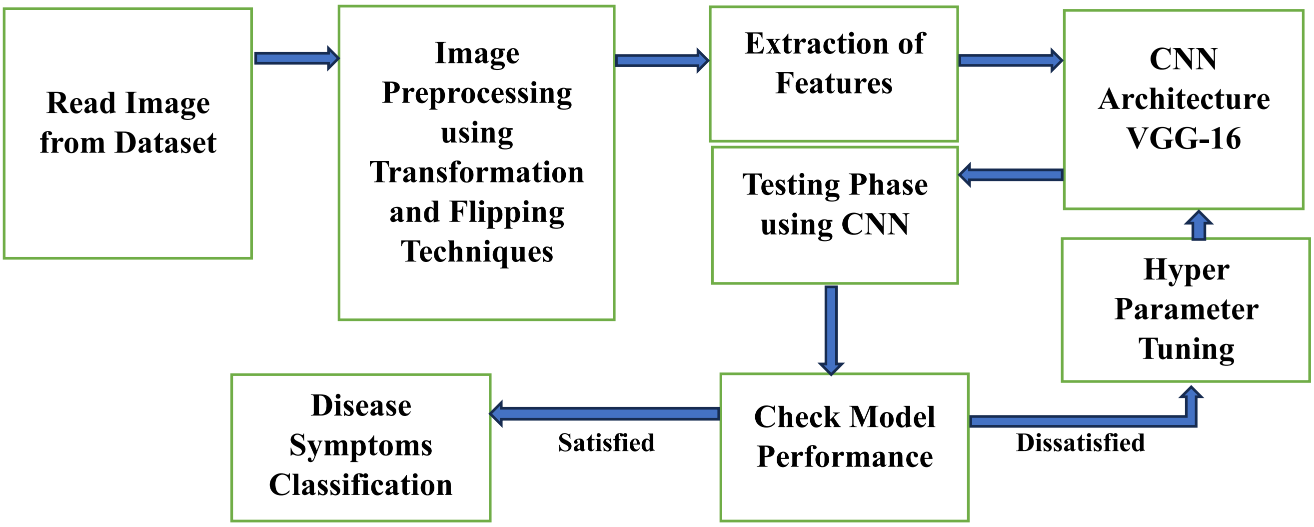

Figure 2. Flow of Proposed Model

B. Dataset

The database used for the algorithm is PlantVillage and Tomato Leaf Dataset. From dataset, training dataset have 25,851 images and 6,683 images is used for testing dataset. Ratio is of 80:20. Both training and testing dataset have a total of 11 classes where ten classes are of complete unhealthy leaf classes whereas one class is of healthy leaf class.

C. Convolutional Neural Network

As we are using VGG16 architecture, it expects the size of 224 by 224. We will set image size with the three-color stream is used red, blue, and green known as RBG. For applying the VGG-16, 16 layers huge deep neural network is used with batch size of 32. The database images are passed through CNN combinations then finally flatten and dense layer is applied this is applied on leaf part of tomato plant. Libraries of python involved with this algorithm are pandas, numpy, matplotlib, keras, seaborn, tensorflow et al. Adam optimizer is used with loss consisting cross-entropy with activation function softmax is used.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 5, May 2024

DOI: 10.17148/IJARCCE.2024.13551

D. Model Performance

In this model, there is usage of 30 epochs. While plotting the accuracy curve, an epoch is taken on the x-axis, and accuracy is taken on the y-axis. This shows increasing curve line represent the success for the model as shown in Figure 3. While plotting the loss curve, an epoch is taken on x-axis, and loss is taken on y-axis. This shows declining curve line represent the loss for the model as shown in Figure 4. For better understanding confusion matrix along with heatmap is shown in Figure 5 and 6 for the model. The performance of all 11 classes of tomato diseases leaves is calculated with the metrics such as precision, recall, f1-score, and support. Figure 7 shows the classification report for the model.

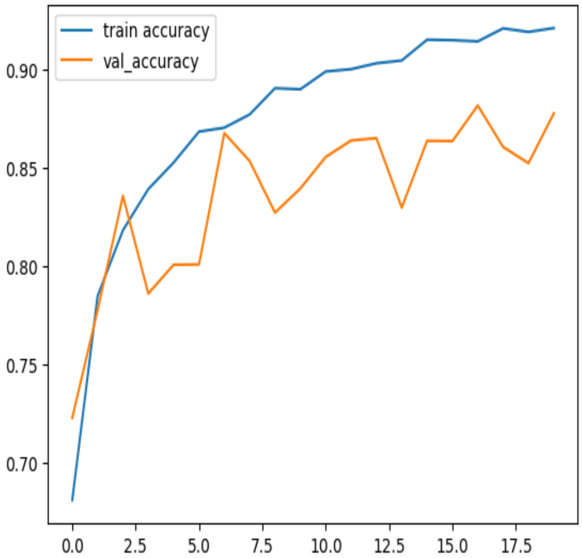

Figure 3. Training and Validation Accuracy Curve

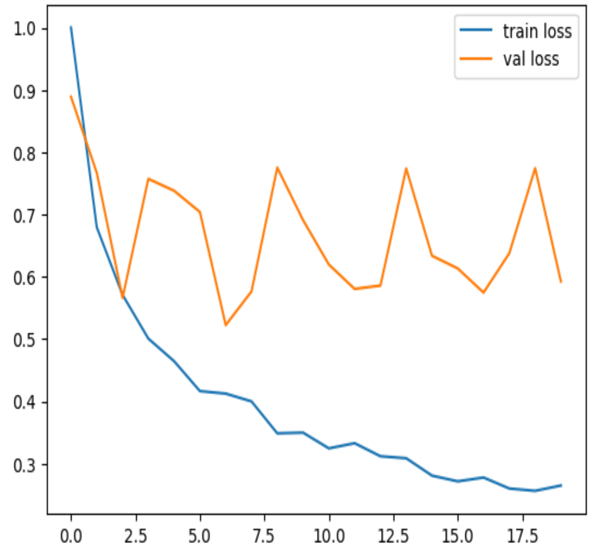

Figure 4. Training and Validation Loss Curve

```
Confusion Matrix
[[ 39  42  49  34  41  31  28  36  28  36  19]
 [121 123 171 142 138  92  80  98  95 119  42]
 [ 21  26  33  26  34  12  20  16  19  36  11]
 [162 124 153 153 159  86  90 103 119 168  60]
 [ 95  85  79  90  78  48  51  52  69 109  27]
 [ 71  52  68  67  74  42  43  43  58  86  23]
 [ 21  20  25  25  18  10  20  16  21  24   9]
 [ 39  31  48  55  50  29  36  27  42  53  13]
 [ 46  40  61  43  50  22  26  31  39  37  16]
 [ 72  68  62  69  79  39  44  44  68  92  23]
 [ 45  32  43  35  25  24  19  32  26  45   9]]
```

Figure 5. Confusion Matrix

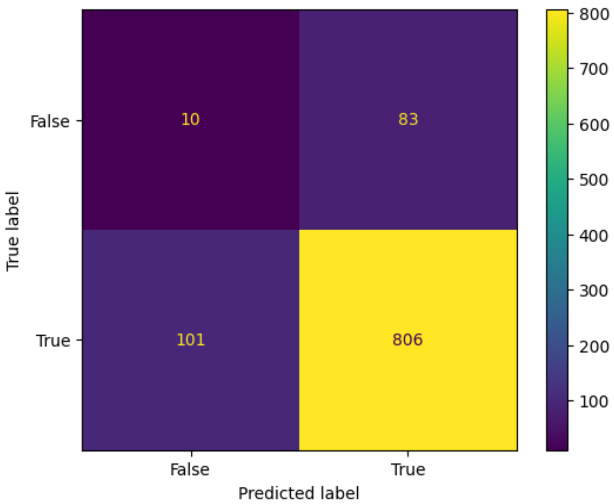

Figure 6. Heatmap

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 13, Issue 5, May 2024
DOI: 10.17148/IJARCCE.2024.13551

Classification Report

	precision	recall	f1-score	support
Tomato_Bacterial_spot	0.10	0.05	0.06	732
Tomato_Early_blight	0.10	0.19	0.13	643
Tomato_Late_blight	0.12	0.04	0.06	792
Tomato_Leaf_mold	0.11	0.20	0.14	739
Tomato Septoria leaf spot	0.09	0.10	0.09	746
Tomato_Spotted_spider_mites	0.06	0.09	0.06	435
Tomato_Target_spot	0.09	0.04	0.05	457
Tomato_Yellow_curl_virus	0.06	0.05	0.05	498
Tomato_Mosaic_virus	0.09	0.06	0.07	584
Tomato_Healthy	0.13	0.11	0.08	805
Tomato_Powdery_mildew	0.02	0.03	0.02	252
accuracy			0.83	6683
macro avg	0.08	0.08	0.07	6683
weighted avg	0.09	0.09	0.07	6683

Figure 7. Classification Report of the Model

RESULTS AND DISCUSSION

The dataset named PlantVillage and Tomato Leaf Dataset is used from online data source Kaggle. On this dataset CNN architecture VGG-16 is applied on a single type of crop named tomato. The VGG-16 algorithm is applied on the platform Google Colab and there is achievement of accuracy above 92% after applying the model on tomato plant.

FUTURE SCOPE

The work can be taken further by increasing the dataset and scope can be wider if more than one plant other than tomato taken like pepper, potato, corn etc. Various augmentation technique like threshold techniques, image enhancement models like HSV, grayscale etc. is used. Other than that, flipping techniques include vertical flipping and can also work on various parameters like sharpness, contrast, brightness, and activation functions etc. Instead of working on single leaves of the plant one can go with other parts like stem, flowers, fruits etc. Changes can be made further with hyperparameters, and the number of epochs can be increased for the better results. One can also use various types of advanced models instead of the above used model for better accuracy, sensitivity, and specificity.

CONCLUSION

Checking plant diseases manually is a very complex and difficult process. Instead of doing that farmers can go with automated robust models for spotting and identifying the diseases in plant. So, one of the automated VGG-16 model is suggested over here which is very user-friendly and makes tasks easier. The suggested model VGG-16 reports generated are very accurate which will save time and the effort of farmers. The yearly yield will increase after adopting this type of automated system which helps farmers to find the symptoms of diseases at a very early stage of plants growth and development.

REFERENCES

[1] Mokhtar, U., Ali., M. A. S., Hassenian, A. E. , Hefny, H. 2015. Tomato leaves diseases Detection approach based on support vector machines. 11th International Computer Engineering Conference (ICENCO). IEEE.
[2] Sabrol, H., Satish, K. 2016. The Tomato plant disease classification in digital images using classification tree. International Conference on Communication and Signal Processing (ICCSP). IEEE.
[3] Shijie, J., Peiyi, J., Siping, H., Haibo, S. 2017. Automatic detection of tomato diseases and pests based on leaf images.Chinese Automation Congress (CAC). IEEE.
[4] Luna, D. R. G., Dadios, E. P., Bandala, A.A. 2018. Automated image capturing system for deep learning based tomato plant leaf disease detection and recognition. TENCON Region 10 Conference, pp. 1414-1419. IEEE.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 5, May 2024
DOI: 10.17148/IJARCCE.2024.13551

[5] Sardogan, M., Tuncer, A., Ozen, Y. 2018. Plant leaf disease detection and classification based on CNN with LVQ algorithm. 3rd International Conference on Computer Science and Engineering (UBMK). IEEE.

[6] Suryawati, E., Sustika, R.,Yuwana, R.S., Subekti, A., Pardede, H. F. 2018. Deep structured convolutional neural network for tomato diseases detection. International Conference on Advanced Computer Science and Information Systems (ICACSIS). IEEE.

[7] Tm, P., Pranathi, A., Saiashritha, K., Chittaragi, N. B., Koolagudi, S. G. 2018. Tomato leaf disease detection using convolutional neural networks. Eleventh International Conference on Contemporary Computing (IC3). IEEE.

[8] Elhassouny, A., Smarandache, F. 2019. The smart mobile application to recognize tomato leaf diseases using Convolutional Neural Networks. International Conference of Computer Science and Renewable Energies (ICCSRE).

[9] Kumar, A., Vani., M. 2019. Image Based Tomato Leaf Disease Detection. 10 th International Conference on Computing, Communication & Networking Technologies (ICCCNT). IEEE.

[10] Militante, Sammy, V., Gerado, Bobby, D., Dionisio, Nanette, V. 2019. Plant Leaf Detection and Disease Recognition using in Deep Learning. IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 579–582.

[11] Rahman, A. M., Islam, M. M., Mahdee, S. G. M., Kabir, U. W. M. 2019. Improved Segmentation Approach for the Plant Disease Detection. 1 st International Conference on Advances in Science, Engineering and Robotics of Technology (ICASERT) , pp. 1–5.

[12] Ashok, S., Kishore, G., Rajesh, V., Suchita, S., Sophia, S. G. G., Pavithra, B. 2020. Tomato Leaf Disease Detection Using Deep Learning Techniques. 5 th International Conference on Communication and Electronic Systems (ICCES). IEEE.

[13] Kirti, Rajpal, N. 2020. Black Rot Disease Detection in Grape Plant (Vitis vinifera) Using Colour Based Segmentation and Machine Learning. 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN).

[14] Marzougui, F., Elleuch M., Kherallah, M. 2020. A Deep CNN Approach for Plant Disease Detection. 21st International Arab Conference on Information Technology (ACIT).

[15] Rajasekaran, C., Arul, S., Devi, S., Gowtham, G., Jeyaram, S. 2020. Turmeric Plant Diseases Detection and Classification using Artificial Intelligence. International Conference on Communication and Signal Processing (ICCSP), pp. 1335–1339. IEEE.

[16] Zhang, Y., Song, C., Zhang D. 2020. Deep learning based object detection improvement for tomato disease. IEEE Access 8: 56607-56614.

[17] David, H.E., Ramalakshmi, K., Gunasekaran, H., Venkatesan, R. 2021. Literature Review of Disease Detection in Tomato Leaf Plant using Deep Learning Techniques. 7 th International Conference on Advanced Computing and Communication Systems (ICACCS).

[18] Guan, X. 2021. A Novel Method of Plant Leaf Disease Detection Based on Deep Learning and Convolutional Neural Network. 6 th International Conference for the Intelligent Computing and Signal Processing (ICSP).

[19] Srinidhi, V. V., Sahay, A., Deeba, K. 2021. Plant Pathology Disease Detection in Apple Leaves Using Deep Convolutional Neural Networks : Apple Leaves Disease Detection using EfficientNet and DenseNet. 5 th International Conference Computing Methodologies and Communication (ICCMC).

[20] Omar, S., Jain, R. 2022. Classification of Disease Symptoms in Tomato Leaf with the help of Convolutional Neural Network. International Conference on Computation Intelligence and Sustainable Engineering Solutions (CISES). IEEE.

[21] Vini, S.L., Rathika, P. 2023. Thresholding based Tomato Leaf Disease Classification. 4 th International Conference on Signal Processing and Communication (ICSPC). IEEE.