

Predicting the Tomato Plant Disease Using Deep Learning Techniques

Rishikesh Bhupendra Trivedi¹, Daksh Mittal², Anuj Sahani³, Clely Voyena Fernandes⁴, Somya Goyal^{5*}, Jyotir Moy Chatterjee⁶, Vanshika Mehta⁷

^{1,2,3,4,5,7} Manipal University Jaipur, Jaipur-303007, Rajasthan, INDIA

⁶ Lord Buddha Education Foundation, Kathmandu-44600, Nepal

bhupendtrivedi@gmail.com, dakshmesssi@gmail.com ,
clelyfernandes19@gmail.com, sahaniiianuj@gmail.com,
somyagoyal1988@gmail.com*, jyotirchatterjee@gmail.com,
vanshikamehta192001@gmail.com

Abstract. Plant Disease refers to the condition when the crop is not capable to yield fully. Plant Disease is the prevailing reason behind the damage of crops. Early prediction of plant diseases can save crops and can prevent the agricultural disasters. Deep Learning (DL) has made a huge contribution to classify the tomato crop with disease from the healthy ones based on the leave condition. This paper proposes VGG16 based model to predict nine types of disease in tomatoes. The publicly available dataset namely PlantVillage is used for experimental study. The proposed model achieves the accuracy of 99.7% and Area Under the Curve (AUC) of 93.3%. It can be concluded that the proposed model has potential to effectively predict the diseases in tomato plants.

Keywords: Tomato Plant Disease; Convolution Neural Network (CNN); Plant Village Dataset; Deep Learning; VGG16 and Area Under the Curve (AUC).

1 Introduction

Plant Disease is the one of the biggest reasons behind the destruction of the crops all around the world. It causes high economic loss too. In agricultural economy, the situation gets worse with the plant diseases. A survey suggests that 40% of total loss in crop-yield is due to the plant diseases [1]. The early prediction of plant disease can be a potential help in saving crop yield by issuing warning and precautionary measures. The crop production is needed to be increased to satisfy the food requirements of the increasing population [2]. Tomato is the second most popular crop in the

world after potato. The world needs approximately a hundred million tons of tomatoes on daily basis [3]. A wide range of Deep Learning (DL) techniques is available for predicting the tomato plant disease [4]. The leaves of tomato plant possess numerous indicators which are worthy for predicting the disease. Farmers detect diseases in their crops by inspecting the leaves of the plants through their eyes. Image processing using deep learning models revolutionized this detection mechanism [5]. The digital era has revolutionized day-to-day routine activities by deploying IoT, Sensors and ML [24-28]. Feature selection (FS) facilitates to enhance prediction accuracy by removing the non-contributing and irrelevant attributes [6-7]. Convolution Neural Network (CNN) has extensively been used to detect the diseases in tomato plant [8]. This work contributes a novel CNN based VGG16 model to predict the disease in tomato plant effectively.

The paper is organized as follows- the related literature work is discussed. Then, the experimental set-up is explained. Later, the results of experiments are analyzed and finally the conclusion is drawn.

2 Literature Work

The survey on the work carried out in the relevant literature applying the deep learning techniques has been summed up in this section. Although traditional ML techniques were also deployed for classification like SVM [9], KNN [10], ANN [11] and Ensembles [12]. Few works implemented feature selection too. [23]. But Deep Learning (DL) based classification models are prevailing in the domain. All the selected studies have deployed models over PlantVillage Dataset. The survey is summarized as Table 1.

Table 1. Related work in literature

S. No.	Study Reference	Year	Technique Used	Evaluation Criteria
1	[8]	2018	CNN + prepro- cessing	Accuracy
2	[13]	2020	CNN, AlexNet, Lenet	Computational time, Accuracy
3	[14]	2020	CNN	ROC
4	[15]	2021	CNN, SVM, En- semble	AUC, ROC, MCC
5	[16]	2018	CNN	Accuracy, Precision
6	[17]	2019	CNN	Accuracy, Precision
7	[18]	2020	Clustering, RCNN	Accuracy, Precision

8	[19]	2019	CNN	Accuracy, AUC
9	[20]	2020	CNN	Accuracy
10	[21]	2019	CNN	Accuracy

3 Research Methodology

The research methodology adopted for this work including the experimental methods, and set-up are briefed in this section.

The dataset used for the experiments is PlantVillage dataset [22]. The dataset has 9 varieties of diseases which are shown with the images of leaves for individual category as in Fig 1. The Convolutional Neural Network (CNN) based VGG16 model is proposed for detection of disease. Fig 2 depicts the proposed experimental model.

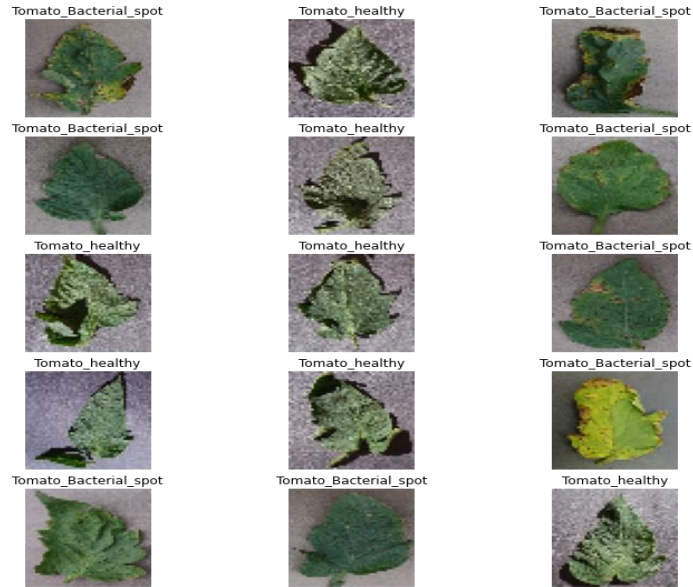


Fig. 1. Varieties of Leaves in PlantVillage Dataset

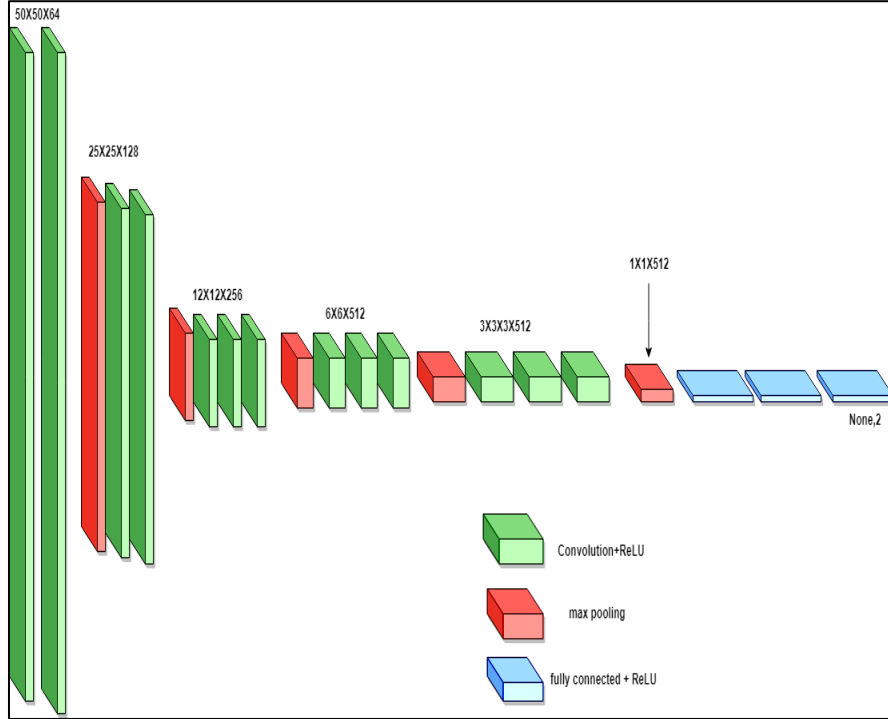


Fig. 2. Proposed VGG16 Model for Disease Detection

The disease detection is formulated as a binary classification problem. The dataset is split into nine sub-datasets corresponding to the nine diseases. For the performance evaluation, Confusion Matrix [10], ROC [11], AUC [12] and accuracy are considered [16-19].

4 Results

The experimental results are reported in this section along with the inferences drawn after analysis.

First up, the Confusion Matrix are recorded datasets for all the categories of plant disease and reported as in Fig 3. Next, the author records the accuracy measure (see Table 2).

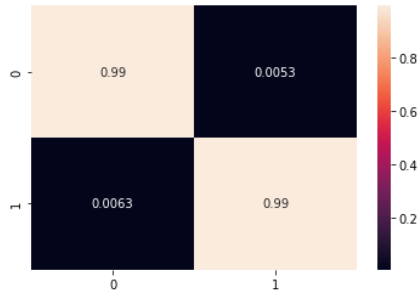


Fig 3a: Bacterial Spot.

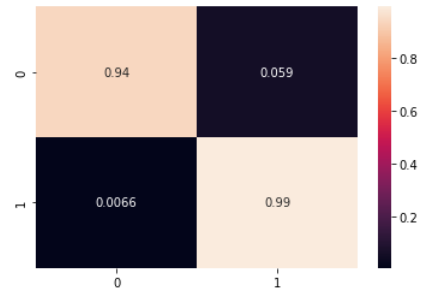


Fig 3b: Leaf Mold.

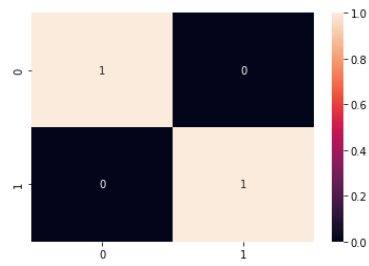


Fig 3c: Mosaic Virus

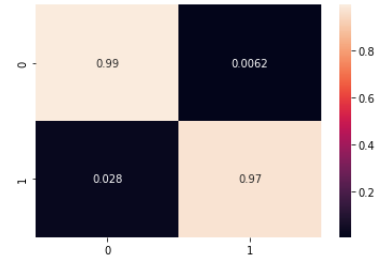


Fig 3d: Spider Mites

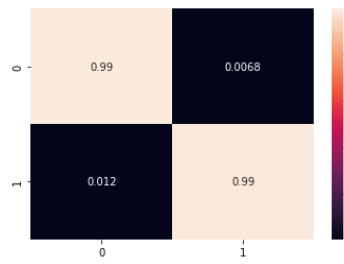


Fig 3e : Curl Virus

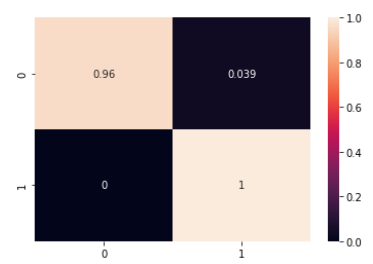


Fig. 3f: Early Blight

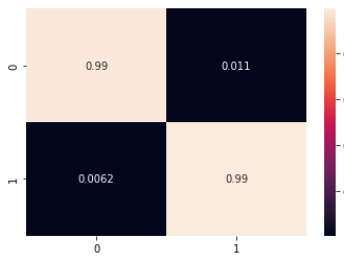


Fig 3g: Late Blight

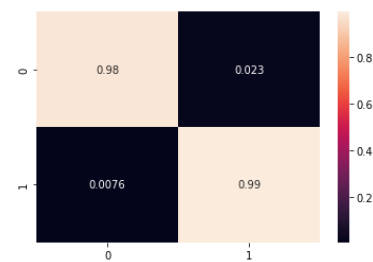


Fig 3h: Septoria Leaf Spot

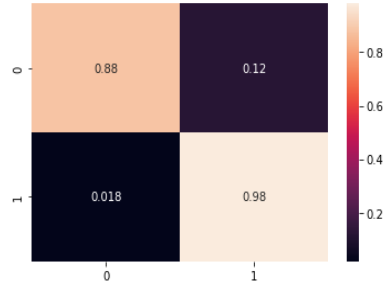
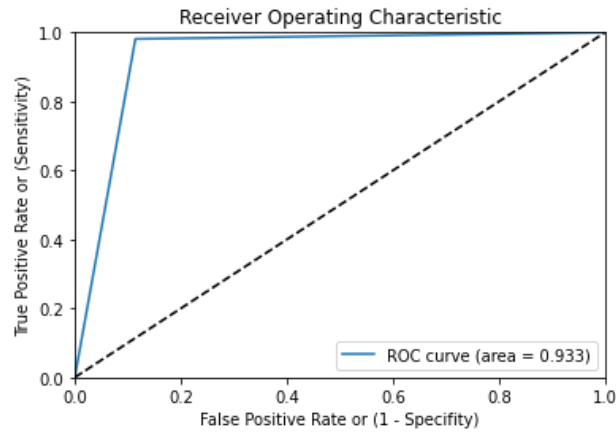


Fig 3i: Target Spot

Table 2. Results for Accuracy of Model over Varity of Diseases.

S. No.	Disease	Accuracy
1	Early Blight	98.53%
2	Bacterial Spot	99.76%
3	Leaf Mold	99.0%
4	Late Blight	98.57%
5	Mosaic Virus	99.45%
6	Spider Mites	98.05%
7	Curl Virus	99.45%
8	Target Spot	94.46%
9	Septoria Leaf Spot	98.35%

Next, ROC is considered for performance evaluation. The corresponding ROC plot are reported as Fig 4.

**Fig 4.** ROC Curve for Proposed Model over PlantVillage Dataset

Further, a comparison is made of the proposed model with the existing models in literature. The comparison is made over Accuracy criteria and reported as in Table 3.

Table 3. Comparison with the models from literature

S. No.	Study	Reported Accuracy (in %)
1	Tm et al. [8]	94.8
2	Ashok et al. [13]	98.0
3	Karthik et al. [14]	98.0
4	Kibriya et al. [15]	99.23
5	Sardogan et al. [16]	86.0
6	Proposed VGG16 Model	99.7

From the experimental results, it is seen that Proposed model shows the best accuracy for predicting the plant disease in tomato crop.

5 Conclusion

Plant disease is the biggest reason of loss of crop in the entire world. If it is predicted well in advance and the patient is fore alarmed, then the crops can be saved. ML classification algorithms are being used for predicting the tomato plant disease. In this paper, a novel CNN based VGG16 model has been proposed and implemented. The comparison is made among the performance of the proposed method and existing models. The authors conclude the work that the proposed VGG16 model performs best for tomato plant disease prediction.

In future, the authors propose to replicate the work with latest datasets to contribute more accurate plant disease prediction for multiple crops.

References

1. L. Han, M. S. Haleem, and M. Taylor, "A novel computer vision-based approach to automatic detection and severity assessment of crop diseases," in 2015 Science and Information Conference (SAI), 2015, pp. 638-644.
2. FAO, F., 2017. The future of food and agriculture—Trends and challenges. *Annual Report*, 296, pp.1-180.
3. Bosona, T., Gebresenbet, G.: Life cycle analysis of organic tomato production and supply in sweden. *J. Clean. Prod.* **196**, 635–643 (2018)
4. Shruthi, U., Nagaveni, V., Arvind, C.S. and Sunil, G.L., 2022. Tomato Plant Disease Classification Using Deep Learning Architectures: A Review. In *Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems* (pp. 153-169). Springer, Singapore.

5. X. E. Pantazi, D. Moshou, and A. A. Tamouridou, Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. *Computers and electronics in agriculture*, vol. 156, pp. 96-104, 2019.
6. Goyal S. (2022) FOFS: Firefly Optimization for Feature Selection to Predict Fault-Prone Software Modules. In: Nanda P., Verma V.K., Srivastava S., Gupta R.K., Mazumdar A.P. (eds) *Data Engineering for Smart Systems. Lecture Notes in Networks and Systems*, vol 238. Springer, Singapore. https://doi.org/10.1007/978-981-16-2641-8_46
7. Goyal, S., Software fault prediction using lion optimization algorithm. *Int. j. inf. tecnol.* (2021). <https://doi.org/10.1007/s41870-021-00804-w>
8. P. Tm, A. Pranathi, K. SaiAshritha, N. B. Chittaragi and S. G. Koolagudi. Tomato Leaf Disease Detection Using Convolutional Neural Networks. 2018 Eleventh International Conference on Contemporary Computing (IC3), 2018, pp. 1-5, doi: 10.1109/IC3.2018.8530532.
9. Goyal, S. Effective software defect prediction using support vector machines (SVMs). *Int J Syst Assur Eng Manag* (2021). <https://doi.org/10.1007/s13198-021-01326-1>
10. Goyal, S. Handling Class-Imbalance with KNN (Neighbourhood) Under-Sampling for Software Defect Prediction. *Artif Intell Rev* (2021). <https://doi.org/10.1007/s10462-021-10044-w>
11. Goyal, S. Predicting the Defects using Stacked Ensemble Learner with Filtered Dataset. *Autom Softw Eng* 28, 14 (2021). <https://doi.org/10.1007/s10515-021-00285-y>
12. Goyal, S. 2020. Heterogeneous Stacked Ensemble Classifier for Software Defect Prediction. 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, Solan, India, 2020, pp. 126-130, doi: 10.1109/PDGC50313.2020.9315
13. S. Ashok, G. Kishore, V. Rajesh, S. Suchitra, S. G. G. Sophia and B. Pavithra, "Tomato Leaf Disease Detection Using Deep Learning Techniques," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 979-983, doi: 10.1109/ICCES48766.2020.9137986
14. Karthik R., Hariharan M., Sundar Anand, Priyanka Mathikshara, Annie Johnson, Menaka R., Attention embedded residual CNN for disease detection in tomato leaves, *Applied Soft Computing*, Volume 86, 2020, 105933, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2019.105933H>.
15. Kibriya, R. Rafique, W. Ahmad and S. M. Adnan, "Tomato Leaf Disease Detection Using Convolution Neural Network," 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), 2021, pp. 346-351, doi: 10.1109/IBCAST51254.2021.9393311.
16. M. Sardogan, A. Tuncer and Y. Ozen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm," 2018 3rd International Conference on Computer Science and Engineering (UBMK), 2018, pp. 382-385, doi: 10.1109/UBMK.2018.8566635.
17. Ashqar, Belal & Abu-Naser, Samy. (2019). Image-Based Tomato Leaves Diseases Detection Using Deep Learning. *International Journal of Engineering Research*. 2. 10-16.
18. Y. Zhang, C. Song and D. Zhang, "Deep Learning-Based Object Detection Improvement for Tomato Disease," in *IEEE Access*, vol. 8, pp. 56607-56614, 2020, doi: 10.1109/ACCESS.2020.2982456.
19. Hasan, Mosin & Tanawala, Bhavesh & Patel, Krina. (2019). Deep Learning Precision Farming: Tomato Leaf Disease Detection by Transfer Learning. *SSRN Electronic Journal*. 10.2139/ssrn.3349597.

20. M. A. Jasim and J. M. AL-Tuwaijari, "Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques," 2020 International Conference on Computer Science and Software Engineering (CSASE), 2020, pp. 259-265, doi: 10.1109/CSASE48920.2020.9142097.
21. C. U. Kumari, S. Jeevan Prasad and G. Mounika, "Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 1095-1098, doi: 10.1109/ICCMC.2019.8819750.
22. D. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," arXiv preprint arXiv:1511.08060, 2015.
23. Goyal, S. (2022). FOFS: Firefly Optimization for Feature Selection to Predict Fault-Prone Software Modules. In: Nanda, P., Verma, V.K., Srivastava, S., Gupta, R.K., Mazumdar, A.P. (eds) Data Engineering for Smart Systems. Lecture Notes in Networks and Systems, vol 238. Springer, Singapore. https://doi.org/10.1007/978-981-16-2641-8_46
24. Goyal, S. (2022). IoT-Based Smart Air Quality Control System: Prevention to COVID-19. In: Verma, J.K., Saxena, D., González-Prida, V. (eds) IoT and Cloud Computing for Societal Good. EAI/Springer Innovations in Communication and Computing. Springer, Cham. https://doi.org/10.1007/978-3-030-73885-3_2
25. Panwar, A., Bafna, S., Raghav, A., Goyal, S. (2022). Intelligent Traffic Management System Using Industry 4.0. In: Chakravarthy, V.V.S.S.S., Flores-Fuentes, W., Bhateja, V., Biswal, B. (eds) Advances in Micro-Electronics, Embedded Systems and IoT. Lecture Notes in Electrical Engineering, vol 838. Springer, Singapore. https://doi.org/10.1007/978-981-16-8550-7_34
26. Kumar, A., Gupta, R., Sharma, N., Goyal, S. (2022). Smart Quiz for Brain Stormers. In: Chakravarthy, V.V.S.S.S., Flores-Fuentes, W., Bhateja, V., Biswal, B. (eds) Advances in Micro-Electronics, Embedded Systems and IoT. Lecture Notes in Electrical Engineering, vol 838. Springer, Singapore. https://doi.org/10.1007/978-981-16-8550-7_38
27. Sobhani, S., Shirsale, S.B., Saxena, S., Paharia, V., Goyal, S. (2022). Emergency Bot in Healthcare Using Industry 4.0. In: Chakravarthy, V.V.S.S.S., Flores-Fuentes, W., Bhateja, V., Biswal, B. (eds) Advances in Micro-Electronics, Embedded Systems and IoT. Lecture Notes in Electrical Engineering, vol 838. Springer, Singapore. https://doi.org/10.1007/978-981-16-8550-7_33
28. Sinha, M., Chaurasiya, R., Pandey, A., Singh, Y., Goyal, S. (2022). Securing Smart Homes Using Face Recognition. In: Chakravarthy, V.V.S.S.S., Flores-Fuentes, W., Bhateja, V., Biswal, B. (eds) Advances in Micro-Electronics, Embedded Systems and IoT. Lecture Notes in Electrical Engineering, vol 838. Springer, Singapore. https://doi.org/10.1007/978-981-16-8550-7_37

□□□□□□□□□□□□□□□□