Tomato Leaf Disease Detection Using Machine Learning

Rakesh Sharma [1], Ankita Panigrahi [2], Mamata Garanayak [3], Sujata Chakravarty [4], Bijay K. Paikaray [5] and Lalmohan Pattanaik [6]

[1234] *Dept. of CSE, Centurion University of Technology and Management, Odisha, India.*
[5] *School of Information & Communication Technology, Medhavi Skills University, Sikkim, India*
[6] *Faculty. of Management Studies, Sri Sri University, Odisha, India*

Abstract

India being the huge market of agriculture provides the appropriate environment for different varieties of crops. One of the highly produced staples of the market of India is the tomato crop with great commercial value. India is one of the largest countries in terms of production of tomato. However, it is a sad truth that the amount of production and the quality of production of tomato crop is decreasing day by day due to different diseases which affect the crop. This meets the farmer with heavy losses. To decrease this loss, it is very much necessary to have a complete supervision over the growth of the crop. There are various categories of diseases that harm these tomato leaf on a very large scale like Late Blight, Bacterial Spot, Early Blight, Septoria Spot, Mosaic Virus etc. So, it is necessary to get a solution to prevent these diseases beforehand. This could be done only if we can detect the disease when it is in its beginning stage. With the help of different ML algorithms, leaf disease detection can be done very easily and more precisely with great results. In this paper, Convolutional Neural Networks (CNN) and Res Net 50 are applied to the dataset. The CNN shows the best accuracy among both the applied algorithms.

Keywords

Tomato, Disease, Machine Learning, CNN, Res Net 50, Agriculture, Detection.

1. Introduction

India being an economical giant with more than 65% of population linked directly with agriculture or its products. Farmers face a huge loss due to plant diseases. Mostly tomatoes are produced on the soil which is well-drained. It is seen that out of every 10 farmers 9 grows tomatoes in their field. In India, the area of cultivation of tomato crop reaches around approximately 3, 50,000 hectares and the quantities of production is somewhat around 53,00,000 tons. For getting fresh tomatoes which would taste good many gardeners also grow it in their gardens. But many times, those farmers and gardeners don't get complete progress of the development of the crop. This increases the chances of getting diseases. Plant affected due to diseases makes 10 to 30% of the overall loss in crops. Detection of these diseases in plants is very much necessary for reducing the losses in production. It is a very difficult task to manually monitor the diseases because of its complex nature and also it consumes a lot of time. So, it is important to decrease the human effort applied in this task and increase the prediction accuracy and making the farmers live free of worries. The major objective of this paper is to detect diseases which harms the tomato leaves accurately. In this paper CNN and Res Net 50 have been implemented to work on the collected data.

ACI'22: Workshop on Advances in Computation Intelligence, its Concepts & Applications at ISIC 2022, May 17-19, Savannah, United States
EMAIL: 190301120079@cutm.ac.in (A. 1); 190301250002@cutm.ac.in (A.2); mamatagaranayak@gmail.com (A.3); chakravartys69@gmail.com (A. 4); bijaypaikaray87@gmail.com (A. 5); lalmohan.p@srisriuniversity.edu.in (A.6);
ORCID: 0000-0001-5843-0335 (A. 5)


2. Literature Survey

In year 2019 Amrita S.Tulshan along with Nataasha Raul presented a paper here they worked on plant disease detection. They applied K Nearest Neighbor classification and got the satisfactory accuracy of 98.56% in predicting plant leaf diseases [1]. Arti N. Rathod, Bhavesh Tanawal, Vatsal Shah presented a research paper on detection of leaf disease in the year 2013 in which they described what are the various method to detect affected leafs using image processing techniques [2]. Prajwala TM, Alla Pranathi, Kandiraju Sai Ashritha, Nagaratna B. Chittaragi, Shashidhar G. Koolagudi in the year 2018 presented a paper in which they were studying tomato leaf for disease detection and they have applied a slight variation of CNN model known as Le Net to find and classify diseases in leaves of tomato and got an average accuracy with 94-95% [3]. In 2020 Surampalli Ashok, Gemini Kishore, Velpula Rajesh, S. Suchitra, S.G.Gino Sophia, B.Pavithra presented a paper working on tomato leaf disease detection, here they have applied Alex Net, ANN and CNN and got an accuracy of 95.75%, 92.94% and 98.12% accuracy respectively [4]. Mohit Agarwal, Abhishek Singh, Siddhartha Arjaria, Amit Sinha, Suneet Gupta presented a research paper in the year 2020 on tomato leaf disease detection using CNN and got an accuracy of 91.2% [5]. Halil Durmuú, Ece Olcay Güneú, Mürvet KÕrcÕ presented a paper on disease detection on the leaves of the tomato where they have applied Alex Net and Squeeze Net and received 95.65% and 94.3% accuracy respectively [6]. Konstantinos P. Ferentinos presented a paper in the year 2018 worked on plant leaf disease detection where they applied VGG model and achieved an accuracy of 99.48% [7]. In the year 2017 Alvaro Fuentes, Sook Yoon, Sang Cheol Kim and Dong Sun Park present a research paper on "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition", here they applied VGG 16 and got an accuracy of 83.06% as a result [8]. Geetharamani G., Arun Pandian J. presented a paper in the year 2019, here they were working on spoting of diseases in leaves of plants using a deep CNN with nine-layers and accied an accuracy of 96.46% [9]. In 2019 PENG JIANG, YUEHAN CHEN, BIN LIU, DONGJIAN HE, AND CHUNQUAN LIANG presented a paper on "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks" here they were using VGG-FCN-VD16 and VGG-FCN-S with average recognition accuracy 97.95% & 95.12%, respectively [10]. XIHAI ZHANG1, YUE QIAO, FANFENG MENG, CHENGGUO FAN, MINGMING ZHANG in the year 2017 presented a paper on "Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks", here they applied Google net model and got an accuracy of 98.9% [11]. Melike Sardogan, Adem Tuncer, Yunus Ozen presented a research paper in the year 2018 on "Plant Leaf Disease Detection and Classification based on CNN with LVQ Algorithm" and got an average accuracy of 86% [12]. Yang Lu, Shujuan Yi, Nianyin Zeng, Yurong Liu, Yong Zhang presented a paper in the year 2017, here they worked on "Identification of Rice Diseases using Deep Convolutional Neural Networks" and got an accuracy of 95.48% [13]. Jiang Lu, Jie Hu, Guannan Zhao, Fenghua Mei, Changshui Zhang in the year 2017 presented a paper on "An in-field automatic wheat disease diagnosis system" and applied VGG-FCN-VD16 and VGG-FCN-S and got the mean recognition accuracies of 97.95% & 95.12% respectively over 5-fold cross-validation [14]. Utkarsha N. Fulari, Rajveer K. Shastri, Anuj N. Fulari presented a paper on "Leaf Disease Detection Using Machine Learning" in the year 2020, here they applied CNN model on grape dataset and got an accuracy of 99.7% and when applied on strawberry dataset got an accuracy of 100% [15].

3. Data Collection

The dataset used in the paper is taken from the directory available on Kaggle [16]. In total there are 10 directories each containing 1000 images which make up the whole dataset for this paper which are mentioned underneath:

1. Leaf with Bacterial Spot: This directory has 1000 pictures of tomato leaf which have bacterial spot on it.
2. Leaf with Early Blight: This directory has 1000 pictures of tomato leaf which have Early Blight on it.
3. Leaf with Late Blight: This directory has 1000 pictures of tomato leaf which have Late Blight on it.
4. Leaf with Mosaic Virus: This directory has 1000 pictures of tomato leaf which have Mosaic Virus on it.

5. Leaf with Septoria Spot: This directory has 1000 pictures of tomato leaf which have Septoria Spot on it.
6. Target Spot: This directory has 1000 pictures of tomato leaf which have Target spot on it.
7. Leaf Mold: This directory has 1000 pictures of tomato leaf which have leaf Mold on it.
8. Yellow Leaf Curl Virus: This directory has 1000 pictures of tomato leaf which have Curl Virus on it.
9. Two Spotted Spider Mites: This directory has 1000 pictures of tomato leaf which have spider mites on it.
10. Healthy leaf: This directory has 1000 pictures of healthy tomato leaf.

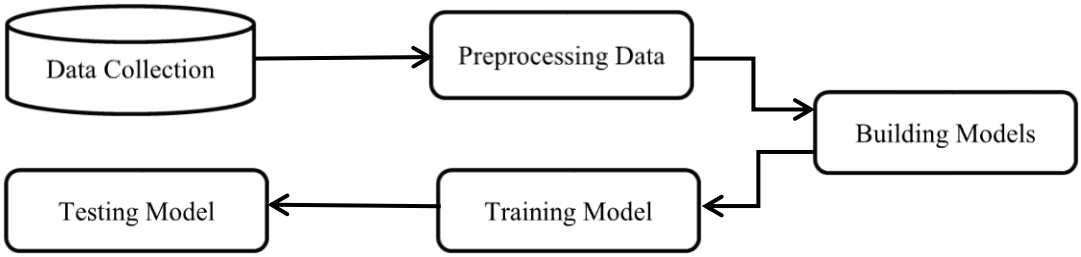

Figure 1. Dataset Collection Flowchart

4. Machine Learning Algorithms

4.1. Artificial Neural Network (ANN)

An artificial Neural Network is simply a Neural Network that resembles a biological Neural Network present in the brain of humans. It is designed in a way such that it would function the same way a human brain function. It is the collection of millions and millions of artificial neurons. These artificial neurons are the building blocks of the ANN model. Artificial Neuron consists of Inputs and their corresponding weights. An activation function is chosen which takes these inputs multiplies them to their corresponding weights and produces the output. Every Artificial Neural Network must have three layers: the input layer which takes the input, the hidden layer where all the computations take place, and the output layers which produce the output.

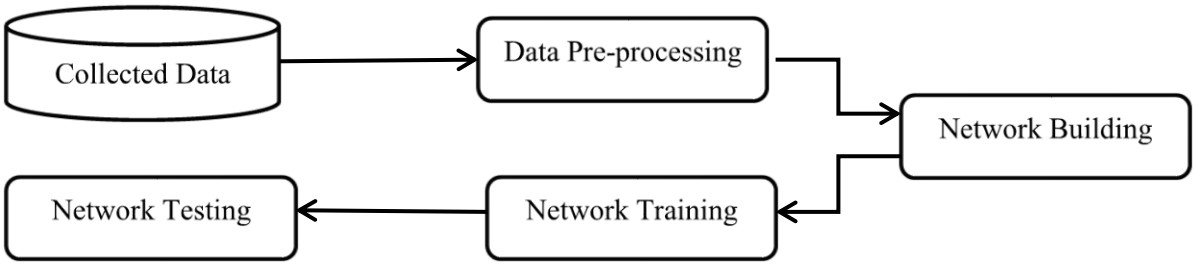

Figure 2. Flow Diagram of ANN Model

In the case of hidden layers, we have used Relu as Activation function with 20 and 10 for weights whereas Linear Activation Function with weight 1 is used in case of the final output. Here, adam optimizer is used [16].

$$Z_i = (\sum_{k=1}^{N_{j-1}} X_k^{j-1} W_{k,i} - b_k) \quad (3)$$

$$f(Z_i) = \frac{1}{1+e^{-Z_i}} \quad (4)$$

4.2. Resnet-50 Methodology

ResNet50 is the type of Resnet model of keras which contains 48 convolutional layers along with 1 MaxPooling and 1 Average Pooling layer. It has 3.8 x 10^9 Floating points operations. This is the most usable Resnet model. It can be also used for computer vision tasks like classification of images, localization of objects, and detection of objects [13]. This framework can also be applied to non-computational vision tasks to reduce the computational expenses and give them the benefit of depth.

5. Proposed Methodology

In the beginning, data is collected from a website named "Kaggle" in the raw form [17]. Then this raw information is pre-processed where first of all the data was resized to 150 for CNN and 224 for ResNet 50 then online data augmentation is done to prevent over fitting. After the pre-processing of the data, the required features are extracted as per the need, and then the data splits for training and testing purposes.

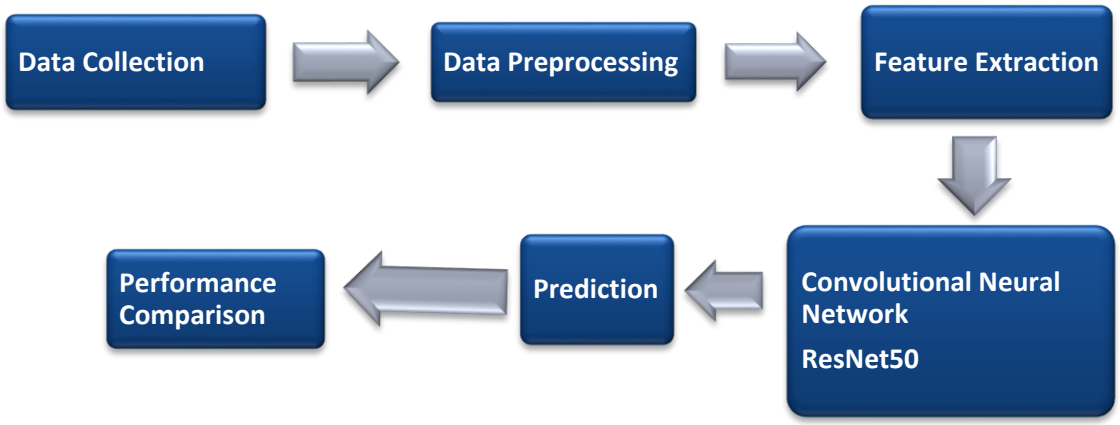

Figure 3. Proposed Methodology Architecture

6. Result Analysis

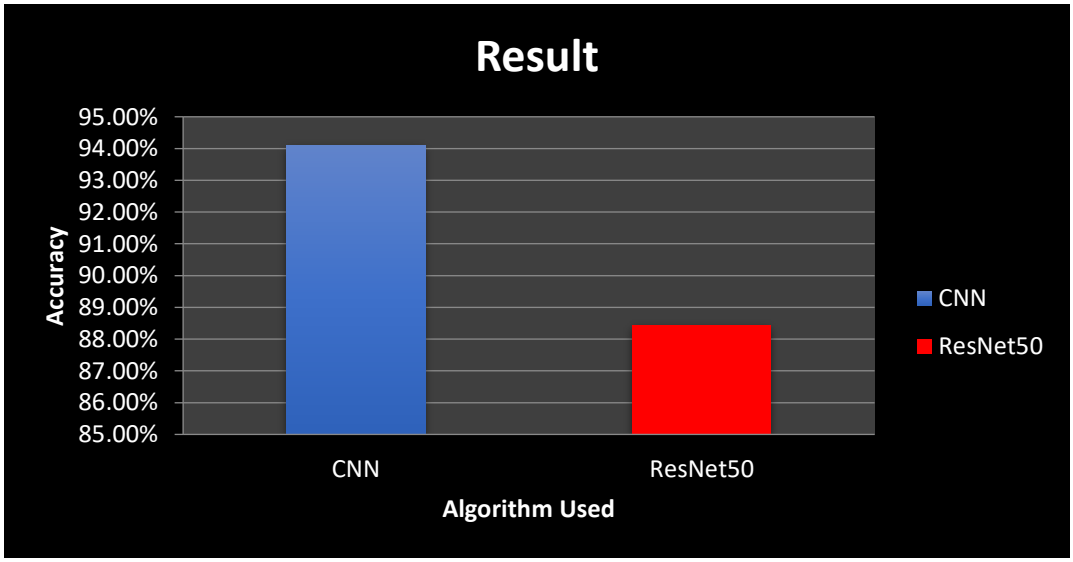

Figure 4. Accuracy of the Proposed Model

When we compare the output of the processed data on different proposed models, we get to know that the Convolutional Neural Network model gives 94.1% accuracy which is a good one followed by ResNet 50 with 88.44% accuracy which depict that our CNN model can predict more accurately.

The above graph showing result of the applied algorithms in the proposed model states that CNN is having the longest bar which means it is having the highest accuracy of 94.10% with compared to ResNet50 with the shortest bar and having the accuracy of 88.44%.

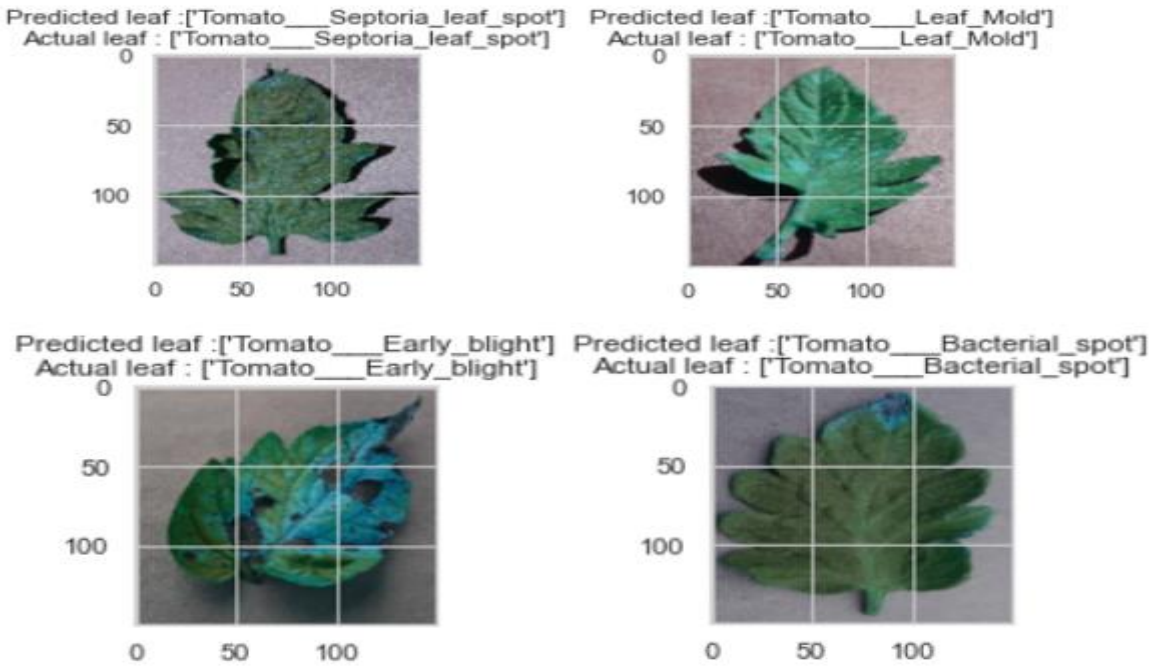

Figure 4. Actual Vs Prediction done by the model

The above images show that the predicted leaf disease is same as the actual one. Our model is successful to make the correct prediction.

7. Conclusion

Timely detection and identification of diseases which affect the leaves is very much necessary now days as it creates a lot of harm to the amount of production and quality of production of crops. This paper presents a model where a total of 10000 images have been pre-processed and then applied to machine learning algorithm i.e., CNN and pre-trained model i.e., ResNet 50 for image processing and prediction work. The acquired outcomes showed that the CNN algorithm got the highest accuracy. The accuracy of the ResNet 50 Model is satisfactory but, in the future, we can tune the parameters applied to increase the accuracy further.

8. References

[1] Tulshan, Amrita S., and Nataasha Raul. "Plant leaf disease detection using machine learning." 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2019.

[2] Rathod, Arti N., Bhavesh Tanawal, and Vatsal Shah. "Image processing techniques for detection of leaf disease." International Journal of Advanced Research in Computer Science and Software Engineering 3.11 (2013).

[3] Tm, Prajwala, et al. "Tomato leaf disease detection using convolutional neural networks." 2018 eleventh international conference on contemporary computing (IC3). IEEE, 2018.

[4] Ashok, Surampalli, et al. "Tomato Leaf Disease Detection Using Deep Learning Techniques." 2020 5th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2020.

[5] Agarwal, Mohit, et al. "ToLeD: Tomato leaf disease detection using convolution neural network." Procedia Computer Science 167 (2020): 293-301.
[6] Durmuş, Halil, Ece Olcay Güneş, and Mürvet Kırcı. "Disease detection on the leaves of the tomato plants by using deep learning." 2017 6th International Conference on Agro-Geoinformatics. IEEE, 2017.
[7] Ferentinos, Konstantinos P. "Deep learning models for plant disease detection and diagnosis." Computers and electronics in agriculture 145 (2018): 311-318.
[8] Fuentes, Alvaro, et al. "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition." Sensors 17.9 (2017): 2022.
[9] Geetharamani, G., and Arun Pandian. "Identification of plant leaf diseases using a nine-layer deep convolutional neural network." Computers & Electrical Engineering 76 (2019): 323-338.
[10] Jiang, Peng, et al. "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks." IEEE Access 7 (2019): 59069-59080.
[11] Zhang, Xihai, et al. "Identification of maize leaf diseases using improved deep convolutional neural networks." Ieee Access 6 (2018): 30370-30377.
[12] Sardogan, Melike, Adem Tuncer, and Yunus Ozen. "Plant leaf disease detection and classification based on CNN with LVQ algorithm." 2018 3rd International Conference on Computer Science and Engineering (UBMK). IEEE, 2018.
[13] Lu, Yang, et al. "Identification of rice diseases using deep convolutional neural networks." Neurocomputing 267 (2017): 378-384.
[14] Lu, Jiang, et al. "An in-field automatic wheat disease diagnosis system." Computers and electronics in agriculture 142 (2017): 369-379.
[15] D. Das, M. Singh, S. S. Mohanty, S. Chakravarty (2020), Leaf Disease Detection Using Support Vector Machine, International Conference on Communication and Signal Processing (ICCSP) (pp 1036-1040)
[16] S. Chakravarty, B. K. Paikaray, R. Mishra and S. Dash, "Hyperspectral Image Classification using Spectral Angle Mapper," 2021 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), 2021, pp. 87-90, doi: 10.1109/WIECON-ECE54711.2021.9829585.
[17] https://www.kaggle.com/kaustubhb999/tomatoleaf