Fruit Disease Prediction and Prevention Using Deep Learning Algorithms

Namrata D. Deshmukh and Sheetal Rathi

Research Journal of Agricultural Sciences
An International Journal

P- ISSN: 0976-1675
E- ISSN: 2249-4538

Volume: 12
Issue: 03

Res Jr of Agril Sci (2021) 12: 933–936

CARAS

Res. Jr. of Agril. Sci. (May-Jun 2021) 12(3): 933–936
www.rjas.org

Full Length Research Article

Fruit Disease Prediction and Prevention Using Deep Learning Algorithms

Namrata D. Deshmukh*[1] and Sheetal Rathi[2]

Received: 13 Apr 2021 | Revised accepted: 20 May 2021 | Published online: 28 May 2021

ABSTRACT

Agriculture plays an exceptional role in the economy of developing countries. Plant diseases and Plant protection are very compelling as they destructively effects quality as well as quantity of crops in agriculture production. The Advance Estimates released by Cooperation and Farmers Welfare, Department of Agriculture says that in year 2019-2020 production of fruits like Pomegranate, Banana, Grapes, Mango, Papaya to be lower by 2.27%. A survey showed that pests and diseases led to 26%-38% loss in yield production. Disease diagnosis is very crucial in early stage in order to control and cure them which will help farmers to increase the production. Contemporarily, most of the primary practices for fruit disease detection in India are performed by naked eye observation by an agriculture domain expert. Hence, there is need of an automatic fruit disease prediction system which can identify disease symptoms in early stages of disease. Proposed system aims to predict the two diseases of Pomegranate fruit (Anthracnose and Bacterial Blight) using fruit images and suggest preventive measures in order to control and cure diseases using Deep Learning Algorithms. The best achieved prediction accuracy using proposed CNN model is 78%.

***Key words*:** Agriculture, CNN, Hyperparameter tuning, Random search

Agriculture plays a vital role for the global economy. The sector employs around 50% workforce of India. Agriculture and its related activities have always meant a significant share in our national income. In recent years, the share of contribution of agriculture has declined gradually with the growth of other industrialized sectors in the country. Agriculture and its related activities had contributed over 59% of the total national income in year 1950-51. The number got decreased to 40% in year 1980-81 and then reported to 18% in year 2008-09. But as compared to many developed countries of the world, the agriculture share in India still remains very high [1].

India is one of the important fruits and vegetables producing countries in the world. Cultivation of fruit crops plays a significant role in the prosperity of any nation. It is generally considered that the standard of living of the people can be judged by per capital production and consumption of fruits. Fruits are rich source of vitamins and minerals. Fruits crops are capable of giving higher yield per unit area than other field crops. Fruit diseases are very compelling as they destructively effects quality as well as quantity of fruit crops production. Hence, fruit plant protection is very essential and challenging job in agriculture. Disease diagnosis is very crucial in early stage in order to control and cure them which will help farmers to increase the production.

Great improvements have been achieved in recent years with massive enthusiasm adding into the Deep Learning field. DL has drawn a lot of attention in agriculture domain. One of its applications in agriculture is image recognition, which has defeated a lot of barriers that limit fast development in robotic and mechanized agro-industry and agriculture. These improvements can be seen in many aspects of agriculture, such as plant disease detection, plant counting and weed control. An understanding of DL algorithms can relieve data analysis and thus enhance research in agriculture [2].

To date, researchers have focused on identification of diseases which are visible on leaves of many fruits and crops like soybean, apple, grapes and pomegranate, tomato etc. The area which is significantly less explored is prediction of diseases of pomegranate which are visible on fruit. Research done on pomegranate fruit is mostly on diseases on leaves using machine learning technique and diseases considered are usually one [3]. Pomegranate is the second largest produced fruits in India. Pomegranate is globally recognized as a Super-food for its nutritious characteristics. There is a decrease in the area of pomegranate cultivation in India. One of the reasons for this decrease is lacking in control of pests and diseases. This is where we need modern agricultural techniques and systems to detect and prevent the crops from being affected by different diseases. Hence, there is need of an automatic fruit disease prediction system which can identify disease

* Namrata D. Deshmukh
✉ ndeshmukh7777@gmail.com

[1-2] Department of Computer Engineering, Thakur College of Engineering and Technology, University of Mumbai, India

symptoms in early stages of disease, which can be done using image processing. Objective of this research is to focus on less explored area of agriculture which is early disease prediction of pomegranate fruit using image processing and CNN on real time image data. This research aims to build a system to predict the two diseases of Pomegranate (Anthracnose and Bacterial Blight) using fruit images and suggest preventive measures in order to control and cure diseases using Deep Learning Algorithms which can surely reduce the amount of loss in pomegranate fruit production.

MATERIALS AND METHODS

Data collected and used in this research work is real time image data from farmer's field in period 2020 to 20201. Images of pomegranate fields were collected from farms located in Tuljapur and Osmanabad, District Osmanabad, Maharashtra, India. The image data is then validated from a domain expert who also helped and provided us preventive measures for two disease Anthracnose (Fig 1A) and Bacterial Blight (Fig 1B). Dataset contains total 3000 image; 1500 images of disease Anthracnose and 1500 images of disease Bacterial Blight are used to train the model. In this study, total 60 testing images are considered, 30 images for Anthracnose class and 30 images for Bacterial Blight class of diseased pomegranate fruit are used respectively.

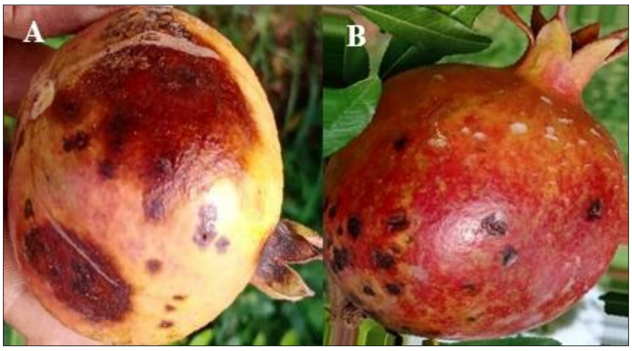

Fig 1 (A) Sample image of anthracnose disease (B) Sample image data of bacterial blight disease

Proposed methodology consists of phases included in the proposed research work (Fig 2). The first phase is data collection, in this real time image data of diseased pomegranate fruit is collected from farms and data is validated from domain expert. Then in second phase, image pre-processing is performed using image augmentation techniques like rotation, horizontal flip, zoom and vertical flip on collected diseased pomegranate images. Total images are then distributed in training and testing data for 2 classes of diseases.

CNN classification model building is third phase in the proposed methodology. Model Building is performed using Hyperparameter Tuning with Keras Tuner [4]. Random search technique is used while building proposed CNN model. In image classification, CNN models are build using various combinations of various layers like convolution layers, pooling layers, dropout layers and fully connected layers in last. We define different sizes of Kernals to extract feature maps and neurons also vary for different layers during the process of network building. So, we do not have any fixed rule to define number of layers, kernel size and neurons. Keras Tuner is a library which determines this problem and provides optimal parameters to effectuate high accuracy. Random Search technique used here is similar to randomized search cv which gives optimal parameters in machine learning.

The first step is to define tuner. Role of tuner is to select hyperparameter combinations which should be tested. The search function performs the iteration to evaluate certain number of hyperparameter combinations. Model evaluation is performed by computing model's accuracy on validation set (Fig 2). Finally, the best hyperparameter combination is chosen with respect to validation accuracy and test on test set.

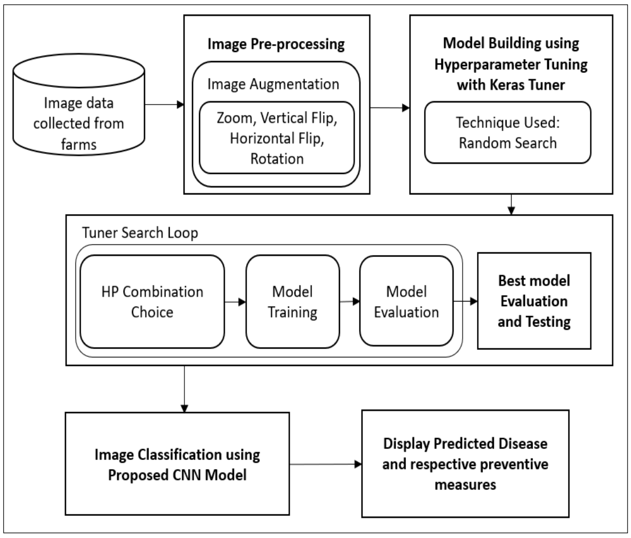

Fig 2 Proposed methodology

Hyperparameter combination choice specified for the proposed CNN model are; Input shape for images is set to 250 × 250 size, relu and sigmoid activation functions were used, range of filter size is set to min_value=16 and max value=128, Dropout set to 0.25. In the CNN model obtained using Hyperparameter Tuning with Random Search each input image will go through two convolution blocks (2 convolution layers followed by batch normalization and a pooling layer) after that a dropout layer for regularization. Eventually, every output is flattened and undergo a dense layer which classify the image into one of the 2 classes (Table 1).

The best model obtained is then evaluated and tested on test data, which is further used for image classification of two disease classes Anthracnose and Bacterial Blight. Once the disease is predicted, the results are displayed containing disease which is predicted and its preventive measures.

Table 1 Architecture of proposed CNN model

Layer	Function	Output shape	Filter size
conv 1	Convolution	248 x 248	16
pool 1	Max Pooling	124 x 124	16
dropout 1	Dropout	124 x 124	16
conv 2	Convolution	122 x 122	32
norm 1	Batch Normalization	122 x 122	32
pool 2	Max Pooling	61 x 61	32
dropout 2	Dropout	61 x 61	32
conv 3	Convolution	59 x 59	128
norm 2	Batch Normalization	59 x 59	128
pool 3	Max Pooling	29 x 29	128
dropout 3	Dropout	29 x 29	128
flatten	Flatten	-	-
dense1	Dense	-	128
dropout 4	Dropout	-	128
dense 2	Dense	-	2

RESULTS AND DISCUSSION

Plot of training process

Our aim was to improve accuracy performance of the model over the time. Plots of model accuracy scores obtained for 2 categories of classes containing training accuracy and validation accuracy (Fig 3). Here the training accuracy is fluctuating at some point but there is substantial fluctuation in the validation accuracy; reason being the size of dataset. The process can be improved using increased size of dataset.

Fig 3 Plot of accuracy scores

Pomegranate disease classification

In this work, total 3000 image data samples are considered to train the proposed model for 2 classes. Performance of the model is tested using 60 image data samples, considering 30 images for each class. The classification report consists of 4 evaluation parameters as precision, recall, f1-score and support (Table 2). Final classification accuracy achieved for 2 classes Anthracnose and Bacterial Blight is 78% with 0.21 error rate. The proposed CNN model predicts 21 images correctly out of 30 images for class Anthracnose and predicts 26 images correctly out of 30 for class Bacterial Blight (Table 3).

Table 2 Classification report of proposed CNN Model

Class	Precision	Recall	F1-Score	Support
Anthracnose	0.84	0.70	0.76	30
Bacterial blight	0.74	0.87	0.80	30

Table 3 Pomegranate disease classification result on test data

Disease classes	Classification result	
	Total testing data samples	Predicted samples using proposed model
Anthracnose	30	21
Bacterial blight	30	26

Confusion Matrix

In this work, Evaluation of behaviour and effectiveness of proposed CNN model is understood using confusion matrix. Confusion matrix is useful to visualize the summary of classification performance of a model with respect to some test data [5]. The confusion matrix obtained here consists of total 60 image data samples used for testing purpose for 2 classes of pomegranate diseases (Fig 4). Following two confusion matrices consisting of number of actual and predicted diseased fruit images (Fig 4A) and percentage of actual and predicted diseased fruit images (Fig 4B) can be referred to understand how model behaves while testing with new data samples.

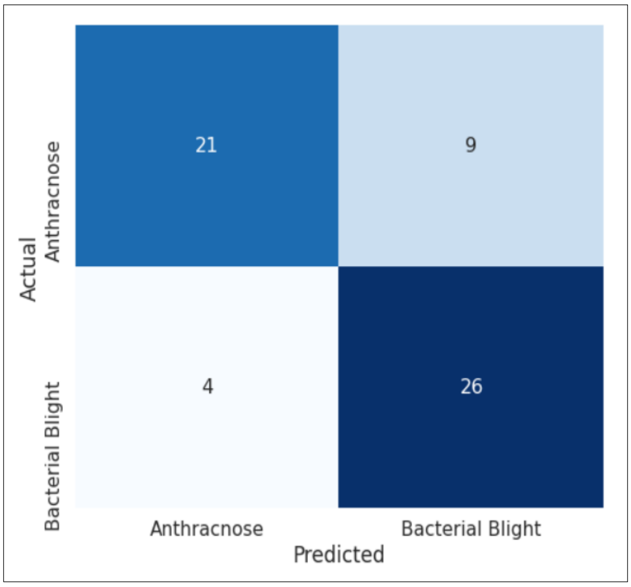

Fig 4 (A) Confusion matrix consisting number of samples predicted

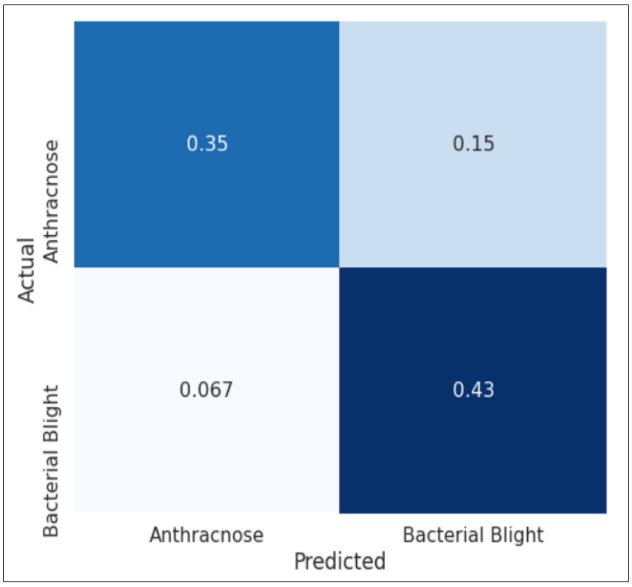

Fig 4 (B) Confusion matrix consisting percentage of samples predicted

Table 4 Comparison of classification report of proposed CNN Model

Class	Accuracy	Error Rate	Precision	Recall	F1-Score
VGG16	0.48	0.51	0.45	0.47	0.41
VGG19	0.53	0.47	0.59	0.52	0.44
ResNet	0.38	0.59	0.42	0.37	0.40
Proposed CNN Model	0.78	0.21	0.79	0.78	0.78

Performance comparison of proposed model with Some pretrained CNN models

Comparison of proposed model with other pretrained CNN models helps to understand how proposed model is performing better than pretrained models on same dataset. For this, 3 pretrained models are considered which are VGG16, VGG19 [6] and [7] ResNet. Following comparison is based on evaluation parameters accuracy, error rate, precision, recall and f1-score (Table 4).

Deployment

Proposed work is deployed using Flask web framework which can accessed and operated on computer system and mobile as well. The working flow of Pomegranate Disease Prediction System is first farmer will access the system. Second step is to choose image of diseased pomegranate fruit and at last system will display the output having predicted disease and its preventive measures. Following figures (Fig 5A, 5B) portrays pomegranate disease prediction system and the results displayed.

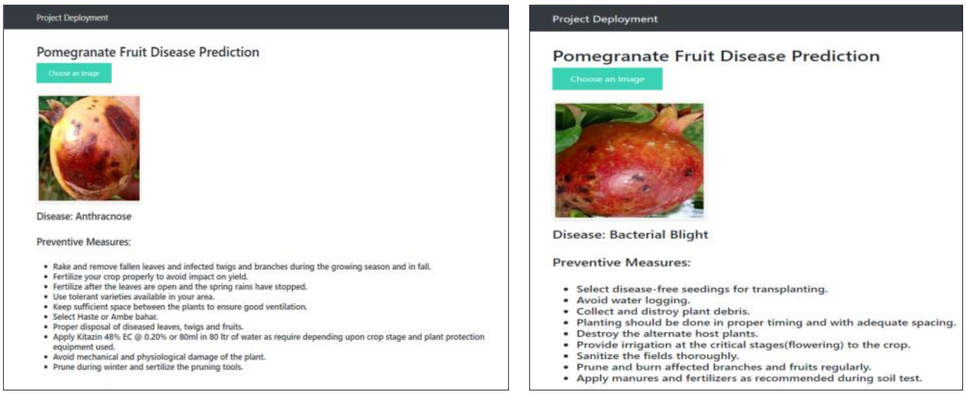

Fig 5 (A) Prediction result for anthracnose disease;
(B) Prediction result for bacterial blight disease

In this study, we proposed a deep learning approach that predicts Anthracnose and Bacterial Blight diseases of pomegranate fruit. The benefits of Deep Learning are very encouraging to be used towards smarter, more secure food production and more feasible farming. The present study includes use of CNN deep learning algorithm build with Hyperparameter Tuning using Keras Tuner and Random Search Technique to predict pomegranate fruit diseases which are visible in fruit. On the other hand, in some recent research work based on identification of diseases which are visible on leaves of crops like cherry, soybean, apple, tomato etc., authors have predominantly opted for Transfer Learning using pretrained CNN models [8-12]. Dataset is improved over considered in research done on pomegranate fruit and ML techniques were used [3], as real time image data is collected, and size of dataset is increased as well during this research.

Our experimental results indicate that the proposed CNN model performs well and predicts 2 pomegranate diseases. The future work can endeavour to improve the performance accuracy of the model by increasing size of dataset and considering other diseases of pomegranate. Another future upgrade could be accomplished by considering fruits other than pomegranate for further exploration in agriculture domain.

Acknowledgement

This paper and the research behind it would not have been possible without the exceptional support of my guide, Dr. Sheetal Rathi. Her knowledge, enthusiasm and exacting attention to detail have been an inspiration and kept my work on track from the beginning of this research work to the final draft of this paper. We would like to express my gratitude to the domain expert for all his help and valuable time, Mr. Mahesh Chine Pathare from Nashik. For 10 years, he is advising farmers for pesticides and disease identification. We acknowledge the industry mentor, Mr. Gautam Sharma for his guidance and suggestions.

LITERATURE CITED

1. Della M. 2019. The Importance of the Agricultural Sector in Indian Economy. Retrieved from https://yourstory.com/mystory/the-importance-of-the-agricultural-sector-in-india
2. Magomadov V. 2019. Deep learning and its role in smart agriculture. *Journal of Physics: Conference Series*. 1399. 044109. 10.1088/1742-6596/1399/4/044109.
3. Bhange M, Hingoliwala HA. 2015. Smart farming: Pomegranate disease detection using image processing. *Procedia Computer Science* 58: 280-288. doi: 10.1016/j.procs.2015.08.022
4. Tom O’Malley. 2020. Hyperparameter tuning with Keras Tuner. Retrieved from https://blog.tensorflow.org/2020/01/hyperparameter-tuning-with-keras-tuner.html
5. Ting KM. 2017. Confusion matrix. Encyclopedia of machine learning and data mining, 260–260.
6. Zisserman A. 2014. Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556.
7. He K, Zhang X, Ren S, Sun J. 2016. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2016.90
8. Jadhav SB, Udupi VR, Patil SB. 2020. Identification of plant diseases using convolutional neural networks. *International Journal of Information Technology*. doi:10.1007/s41870-020-00437-5.
9. Sathya V, Rafidha H, Sumitha Rani G. 2019. Fruit and leaves disease prediction using deep learning algorithm. *International Research Journal of Multidisciplinary Technovation* 1(5): 8-16.
10. Ferentinos K. 2018. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture* 145: 311-318. 10.1016/j.compag.2018.01.009.
11. Rangarajan AK, Purushothaman R, Ramesh A. 2018. Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia Computer Science* 133: 1040-1047. doi:10.1016/j.procs.2018.07.070.
12. Mohammed B, Kamel B, Abdelouahab M. 2017. Deep learning for tomato diseases: Classification and Symptoms Visualization. *Applied Artificial Intelligence* 31(4): 299-315. DOI: 10.1080/08839514.2017.1315516.