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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 8, August 2022

DOI: 10.17148/IJARCCE.2022.11808

Citrus Fruit and Leaves Disease Identification Using Deep Neural Network

Ms. Pashime V. U.[1], Prof. Bansode S. M.[2]

Student, Department of Computer Engineering, SGGSIE&T, Nanded, Maharashtra, India[1]

Professor, Department of Computer Engineering, SGGSIE&T, Nanded, Maharashtra, India[2]

Abstract: The identification of plant diseases is a quite difficult process in the field of agriculture. If the identification process is incorrect, then there will be a huge loss. The identification of Leaf diseases requires knowledge about plant diseases, a big amount of research study, research work, and more processing time. Vegetable and fruit plant supports the lives of approximately 7.5 billion people worldwide and plays a crucial role in the survival of the planet. The economic development of any nation depends on agricultural productivity. The livelihood of around 58 percent of India's population depends on agriculture which is the primary income source. A plant disease is an abnormal condition that alters the appearance and performance of the plant. It is a physiological process that affects some or all plant functions.

In this paper, a convolutional neural network (CNN) model is used to distinguish between healthy and diseased Citrus fruits and leaves. If the image is diseased then the proposed CNN-based model can identify the type of leaf or fruit disease. In the proposed method, we have tried to classify diseases from images of citrus fruit and leaves using the CNN model. Common citrus fruit and leaf diseases are black spot, canker, scab, greening, and melanose. The CNN Model performs better than the several traditional methods used for identifying citrus fruit and plant disease. This method is accurate and gives the results quickly. For farmers wishing to categorize citrus plant leaf or fruit diseases, the CNN Model is a helpful tool for decision-making with a test accuracy of 98.61 percent. This CNN model was checked on the Citrus dataset.

Keywords: deep learning, plant disease recognition, CNN, computer vision.

I. Introduction

Identification of plant leaf disease within time plays an important role in the effective growth of the crop. Plant leaf diseases like bacterial spots, black rot, black measles, chlorosis, etc. affect the quality of the crop, and the growth of plants hence it impacts the agriculture economy. To prevent the effect of these leaf diseases, the farmer uses expensive pesticides and other similar kinds of solutions. The heavy use of pesticides and other chemicals may damage the plant and the soil surrounding the plant [2]. Also, these kinds of solutions increase the production cost and farmers may face big economic losses. Thus, the early and accurate identification of plant disease plays an important role in the efficient management of plant disease. Usually, disease identification is done by human experts in the agriculture field. Now the technology has been improved, so the automatic identification of plant diseases is possible through artificial intelligence, deep learning, and computer vision.

Agriculture is monitored and managed in real-time using geographic information technology, remote sensing, global satellite positioning technology, and computer automatic control technology. These technologies benefit the field of agriculture in a variety of ways. Moreover, production can be enhanced by minimizing pollution and cost. Crop disease transmission can be substantially slowed by early detection and prevention by using less quantity of pesticides.

Plant leaf diseases are one type of natural disaster. It affects the regular growth of plants; also, can cause the death of plants during the entire growth process of plants. Deep learning involves advanced algorithms and techniques of ML (machine learning). This algorithm uses CNN which works like they are human brains.

The traditional techniques of ML use semantic features for the classification method. CNN is a model of deep learning that is broadly used in computer vision [1]. The proposed system can detect the following citrus plant diseases: Black spot, canker, scab, greening, and melanose.

The method used in this study contains three important stages involving the collection of image data, pre-processing of the images, and finally image classification. In this paper, four types of leaf diseases are identified along with healthy images of citrus plants. Pre-processing of the image dataset involves re-sizing and enhancement of images. This will be done before supplying images to the classification model.

II. Objective

The main objective of writing this paper is to identification of citrus plant (fruit and leaf) diseases from images as a

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 8, August 2022
DOI: 10.17148/IJARCCE.2022.11808

classification problem. In this paper we have proposed a classifier which can identify the disease of input image.

- To design an automated computational system for citrus leaf disease classification.
- To design an automated computational system for citrus fruit disease classification.

III. LITERATURE SURVEY

In the past, many researchers have proposed a variety of methods for plant disease detection using leaf and fruit images. Fungi are the most common cause of plant illnesses, and they primarily target the leaves. Many others are caused by viral and bacterial infections. With the rising usage of machine learning and associated aspects in agriculture, precision has improved [2]. Agriculture's decreasing output quantity harms many people and animals, a problem that may be solved with contemporary technology. Because of its high accuracy and reduced problems and data duplication, the image-based detection method makes disease extraction and diagnosis easy. In some plants, such as tomatoes, using images to diagnose the diseases that afflict them, and the amount of damage is impossible unless the accuracy rate is high [3]. The results of the plant disease study demonstrate that a variety of factors influence how technology-based image detection is used. In other words, diseases that generate apparent dents and changes on plants may be recognized using this technique, as opposed to diseases that cause damages that are not evident in the photographs of the plants [5]. Plant diseases are frequently recognized when they begin to affect the outward look of the plants, according to the findings of this study. The biggest issue confronting agriculture is the decrease in productivity and the low quality of plant products. The problem arises as a result of the ineffective identification and control of plant diseases. The problem has also been expanded to include human beings in a variety of ways. Plant diseases limit plant cover, resulting in global warming, starvation, and poor air filtration. Hyperspectral imaging has proven to be a viable method of early detection of agricultural diseases . Unless diseases are discovered early, it is difficult to determine the variables that cause them. In other words, if a disease is diagnosed early enough, it is simple to link it to the causes that contributed to its occurrence. Scientists might, for example, assess if a change in weather or climate contributed to the disease's emergence.

According to [5], there is an insufficient database that may be utilized to offer background information for comparing the photos collected. Another issue is that the illnesses' symptoms and features vary, even if they may be comparable to some extent [4]. Many diseases, for instance, can cause leaves to wilt. The problem has yet to be solved since specialists continue to upload more and more images.

Another problem is the scarcity of appropriate devices for image detecting operations. Most field specialists lack the necessary technology to interpret the images they collect in the field, making it difficult for them to collect correct data and detect illnesses [4]. The other is that owing to restrictions put in place to maintain the authenticity and trustworthiness of the data from these studies, there is a low rate of implementation in some places. Following the 4th and 6th International Conferences on Machine Learning and Soft Computing, for example, there has been a slew of restrictions that might stymie the use of machine learning in specific areas [5].

The technology has been in use for some time. However, there are still numerous questions about its use that have yet to be answered. This information also poses another issue. Some crucial images that could aid in determining whether sickness occurs have not been recorded. The other issue is that the research's prospects are unclear, owing to the rising diversity of diseases that impact both humans and animals [2]. The rising diversity in the appearance of diseases has an impact on the use of image-based detection. Some of the diseases that afflicted plants only a few years ago have evolved into new forms, each with its own set of effects and results. It is tough to determine diseases and choose a solution just based on images. Some of the previous methods have also proven ineffective, decreasing the technology's effectiveness.

These difficulties demonstrate that image-based detection can be used in a variety of ways, but the difficulties limit its applicability. The first answer is to offer sufficient data that can be utilized to reliably diagnose diseases while avoiding confounding closely related diseases. Changes in the weather, global warming, and other factors have resulted in a slew of new diseases that have yet to be identified. The approach is to broaden the scope of scientists' work and promote more efficient data collection [4]. Another option is to provide training to scientists working in this sector so that they are prepared to collect data. Another option is to develop more efficient methods of capturing disease-related data. The problem of insufficient disease information can be remedied if the data-captioning procedure is modified to include detailed details of the photos taken and the differences that distinguish them [6]. The images should be thoroughly examined in order to identify those that are affected or infected.

Another option is to concentrate on using the most up-to-date technology that is both dependable and genuine. The confusion that arises with an insufficient database for use in disease detection stems from the existing systems' weak technology and storage capabilities. Most of the images are incorrectly saved, affecting the information's accessibility. It could be solved by utilizing modern data storage methods. The usage of cloud computing, for example, could help improve storage accuracy and accessibility. The other option is to train the personnel in charge of information research and analysis. A trained DL algorithm improves the technology's accuracy [10]. The other option is to learn about the phenotypes that are utilized to detect diseases [4]. Another option is to update the systems to verify that the data captured is current. The significant level of uncertainty in disease detection has an impact on how the technology is used. The

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋈ Impact Factor 7.39 ⋈ Vol. 11, Issue 8, August 2022
DOI: 10.17148/IJARCCE.2022.11808

employment of Bayesian DL, for example, relates to a number of uncertainties [11]. This indicates that if employed alone, this strategy is unreliable.

The implementation of deep convolutional generative adversarial networks, which aid in image identification and analysis [7], is another option. The contribution of adversarial networks improves the detection process' accuracy. The use of CNN methods could potentially be beneficial in dealing with disease identification inaccuracy and slowness. The approaches have been used to diagnose rice-related disorders and have several advantages [8]. Image-based detection necessitates many resources, which the authorities should make accessible to ensure that the activities run smoothly.

IV. METHODOLOGY

The flow chart of the proposed system is shown in Fig. 1 First, the user must provide a dataset to the CNN model. The image goes through several processing steps like pre-processing, feature extraction, selection of features, etc. In pre-processing, distortion in the input image is removed. CNN algorithm contains steps such as convolution, pooling, ReLU, and a fully connected layer. The model is properly trained using CNN and classification takes place.
The image dataset is divided into three parts. 80% of images from the dataset were provided to train the model, and 10% of images are provided to validate the model. After completing the training of the model 10 % of images are used to test the model. The comparison of the test image and trained model takes place followed by a display of the result. It detects the disease from a plant citrus leaf and fruit images.

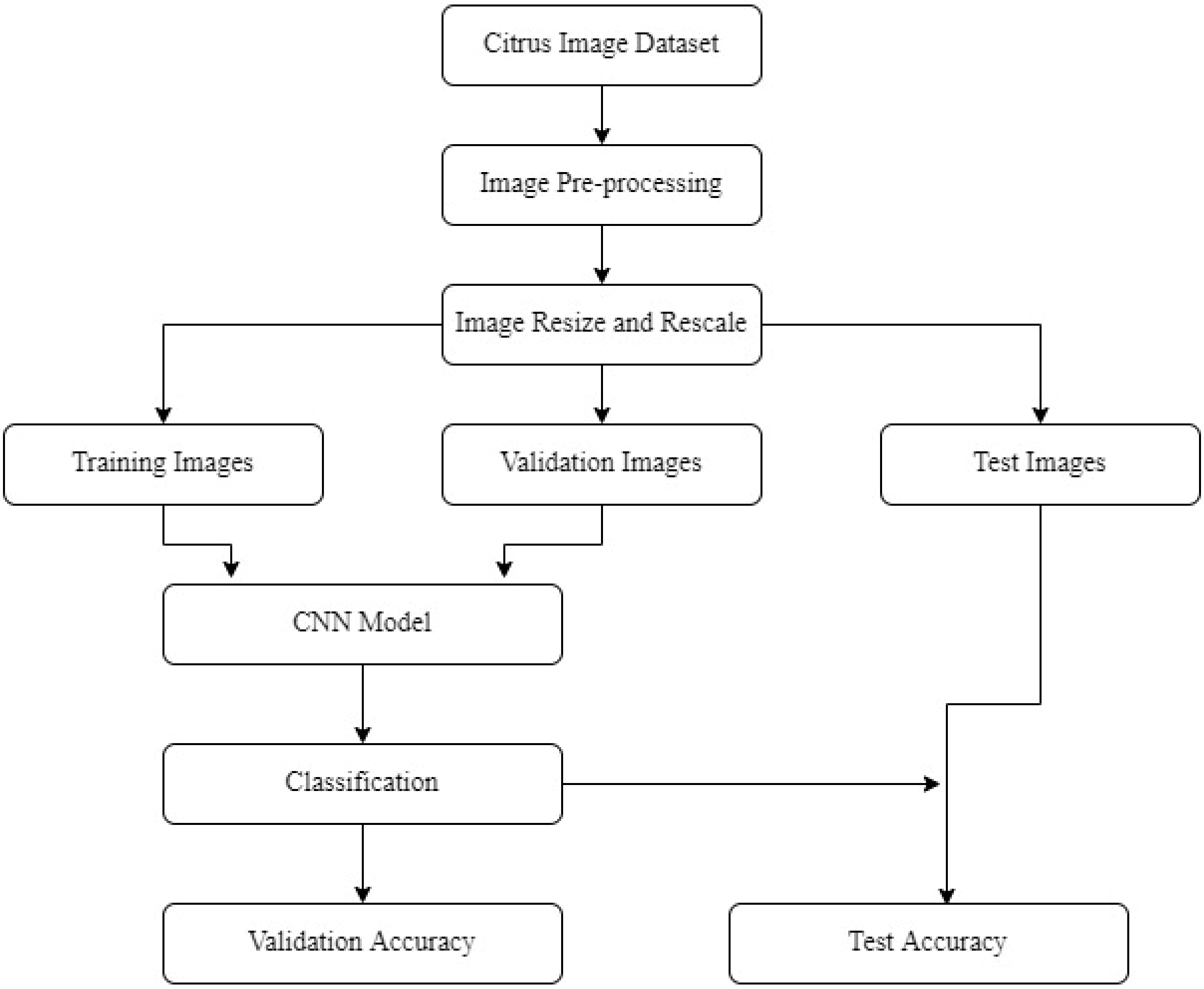

Fig. 1 Flow Chart of Proposed System

The multilayer convolutional neural network is suggested for the identification of healthy and diseased leaves. For this paper implementation, we have used Citrus Dataset, downloaded from https://data.mendeley.com/datasets/3f83gxmv57/2 [3]. This data set contains 759 images of citrus fruits and leaves. Table I contains the description of the

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⪫ Impact Factor 7.39 ⪫ Vol. 11, Issue 8, August 2022

DOI: 10.17148/IJARCCE.2022.11808

data set against each of its diseased and healthy leaf image classes. Table II contains the description of the data set against each of its diseased and healthy fruit image classes.

TABLE I Citrus Dataset leaf images disease classes

Name of Disease	No. of Images
Black Spot	171
Canker	163
Greening	204
Melanose	13
Healthy	58
Total Images	609

TABLE II Citrus Dataset Fruit images disease classes

Name of Disease	No. of Images
Black Spot	19
Canker	78
Greening	16
Scab	15
Healthy	22
Total Images	150

V. CONVOLUTIONAL NEURAL NETWORK

CNN has a complex interconnected structure and can accomplish complex operations. In the deep learning field, CNN is the popular model. Fig. 2 shows the model of a convolutional neural network is a combination of an input layer, convolution layer, pooling layer, full connection layer, and output layer [1]. It also shows the 2nd layer i.e. convolution layer and 3rd layer i.e. pooling layer can alternate each other several numbers of times. CNN is simply used to detect and categorize images. Also, it is well-organized in evaluating provided images and performs various operations on images for extracting the essential features. No full connection is required between the neurons of the convolution layer the and neurons of the pooling layer. The various structural features of CNN are used for image recognition, which makes CNN more popular than other techniques in the computer vision field.

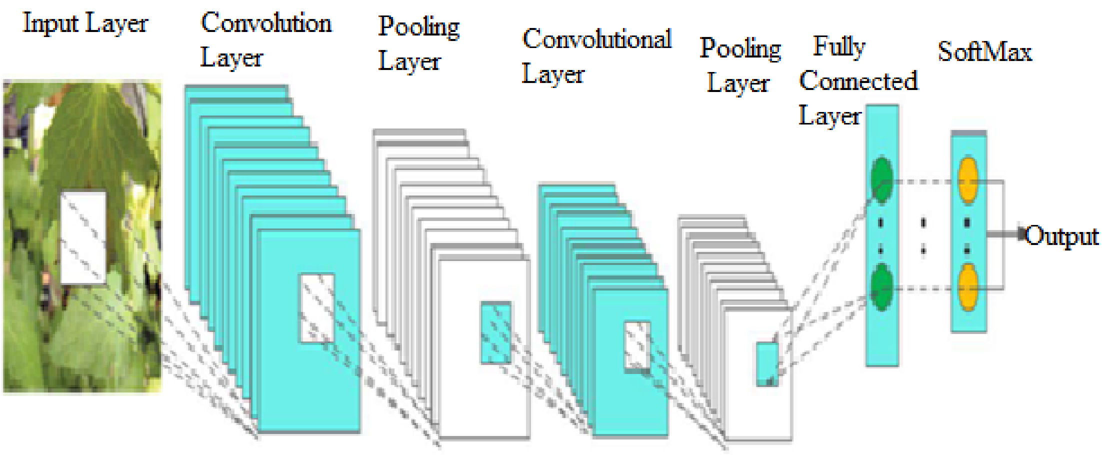

Fig. 2 Illustration of CNN Architecture

1. Convolutional Layer

The Convolutional layer is used to store the output provided by the previous layer to learn the weights and biases. The optimization function is applied to the provided output that represents the error-free data. In the convolution layer, a series of mathematical operations are performed to extract the various features of the provided input image [3]. The

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⪤ Impact Factor 7.39 ⪤ Vol. 11, Issue 8, August 2022

DOI: 10.17148/IJARCCE.2022.11808

operations of the convolution layer on a 5x5 input image and a result generated in a 3x3 filter that reduces the matrix size are shown in Fig. 3 Also, this figure shows the shifting of the filter beginning from the left upper corner of provided input image [1]. Then the multiplication of values at each step and the value of the filter is performed, and then these values are added to the result. A reduced size result matrix is formed from the provided input image.

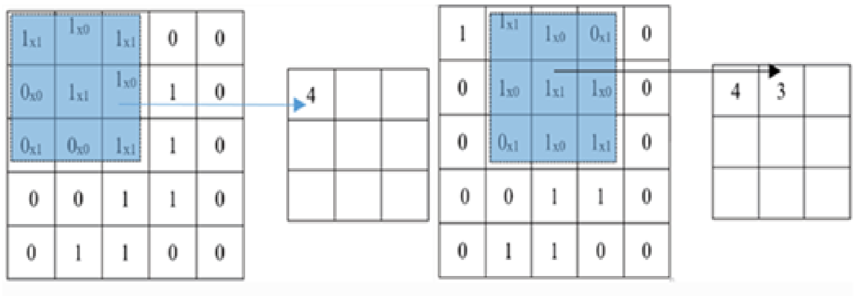

Fig. 3 Convolution layer filter operation

2. Pooling Layer

The pooling layer is used to reduce the overfitting and minimize the neuron size for the downsampling layer. The pooling layer minimizes the size of the feature map, the number of parameters, and time required for training, controls overfitting, and improves computational rate. To reduce the dimensions of the feature map ReLU and max pooling were utilized. The term overfitting means achieving a model using 100 % training dataset and 50 % from the test data.

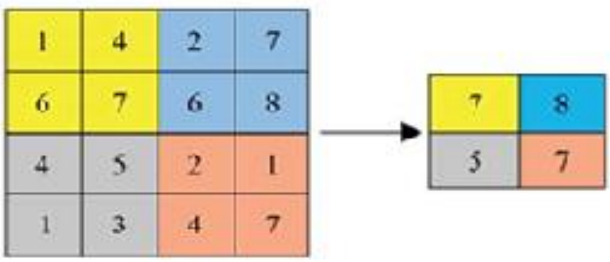

Fig. 4 Pooling operation

3. Activation Layer

Each convolution layer uses a non-linear Rectified Linear Unit (ReLU) activation layer. The dropout layer is used to prevent overfitting and is also applied in the activation layer [1].

4. Fully Connected Layer

A fully Connected Layer is used to analyze the probabilities of class and the classifiers technique is applied to the input to produce the output. The popular input classifier SoftMax is used for the classification and recognition of the Citrus plant leaf and fruit diseases.

VI. EXPERIMENTAL RESULT

The system was implemented using the Jupyter Notebook tool in Python programming language. We have mostly used methods that were available in Keras and TensorFlow libraries. The matplotlib library is used to plot the execution parameters and results in two-dimensional graphs. To train the model, we have executed this for 200 epochs, by providing the training dataset images in the batch size of 8. To test the trained model, the test data set images are provided to the trained model and it has given an accuracy of 98.61%.

The Fig. 5 and Fig. 6 are showing training and validation accuracy and loss graphs respectively. X axix shows number of epochs in both the figures. Whereas Y axis shows accuracy in fig 5, loss in fig 6.

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 8, August 2022

DOI: 10.17148/IJARCCE.2022.11808

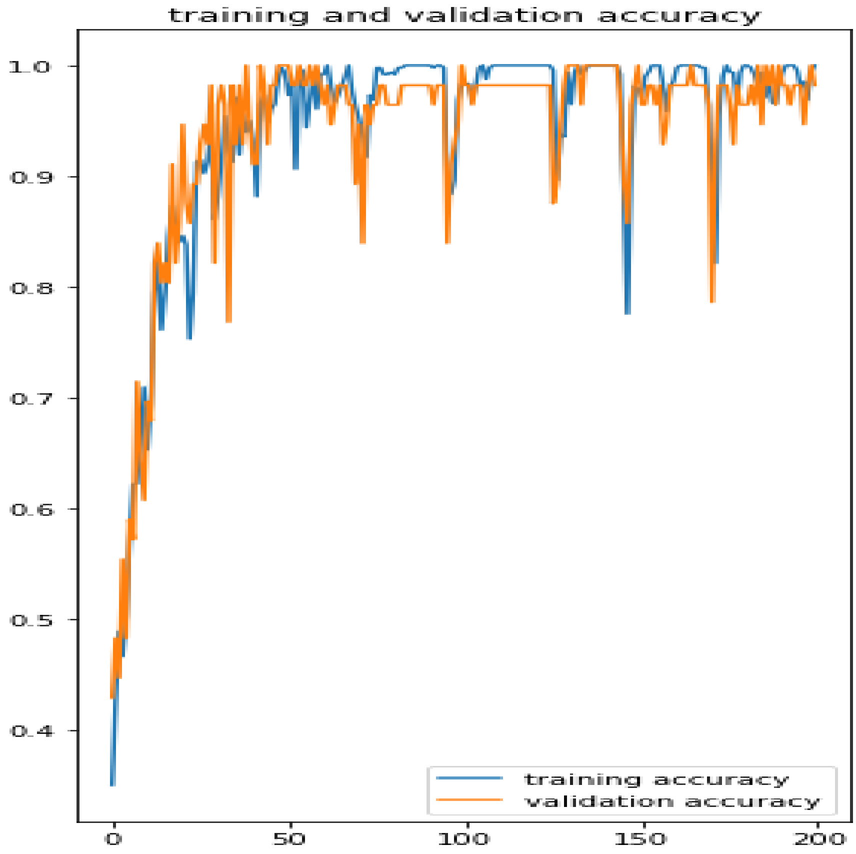

Fig. 5 Training accuracy and Validation accuracy

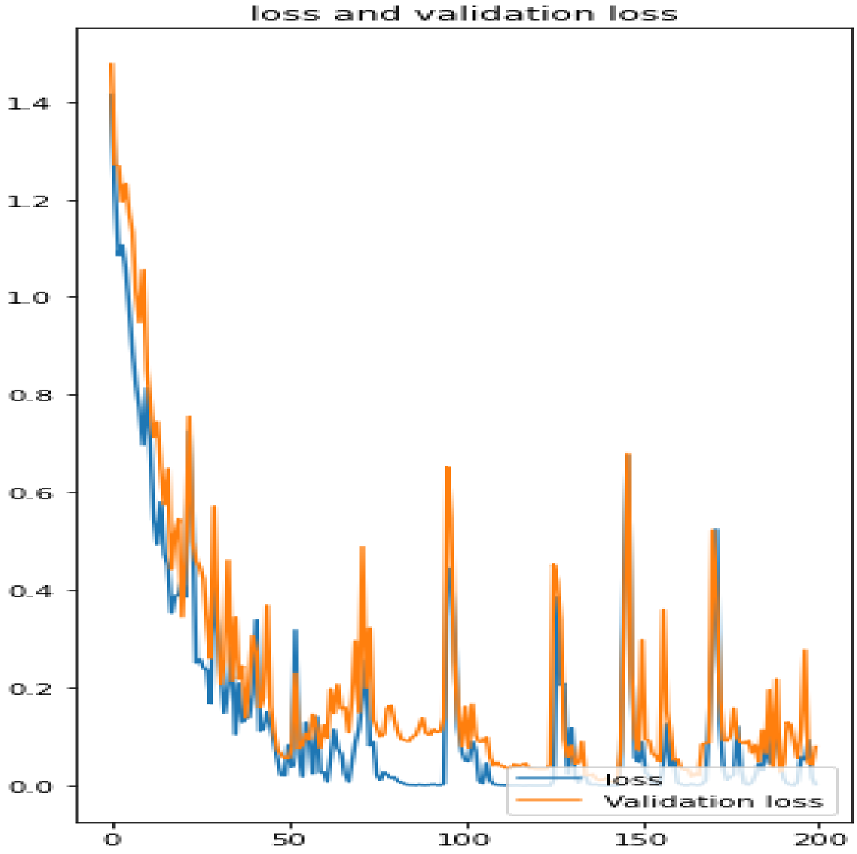

Fig. 6 Training loss and Validation loss

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 8, August 2022
DOI: 10.17148/IJARCCE.2022.11808

Effectively classified results from the test image of the leaves and fruit dataset are displayed in Fig. 7 and Fig. 8 together with their actual class or predicted class. The predicted class was returned along with its associated confidence percentage for a batch of provided images.

Fig. 7 Classification results on Test dataset images of leaves

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 8, August 2022
DOI: 10.17148/IJARCCE.2022.11808

Fig. 8 Classification results on Test dataset images of leaves

CONCLUSION AND FUTURE WORK

More than 60% of people in the world depend on the agriculture sector for survival where crops or plants are the basic need for food. Nearly 30% of the crop is wasted due to plant diseases, so early identification of plant leaf disease plays an important role in the effective growth of the crop. Healthy and diseased Citrus fruits and leaves can be distinguished using the proposed CNN-based model for leaf and fruit disease identification.

Hence, developing the automated system will help farmers to detect and identify the plat disease which will reduce the human errors in the manual classification and improve the accuracy of classification as well as reduce the time of classification. Early detection and prevention can effectively slow down the spread of crop diseases. At the same time, fewer drugs or pesticides can be used to prevent or control crop diseases ahead of the stage, which can reduce pollution to the environment and save the expenses of farmers. So, this automatic plant disease identification system will be useful

DOI: 10.17148/IJARCCE.2022.11808

for society and have a great impact on the agriculture sector. In classification tests for citrus fruit/leaf diseases, the proposed CNN performed better than existing classifiers in terms of accuracy, achieving 98.61%.

In future, this work can be extended by developing an application where the farmer can click the photo of a fruit or leaf and identify the disease. Also, a similar type of model can be developed for other fruit trees.

REFERENCES

[1] A. Khattak et al., "Automatic Detection of Citrus Fruit and Leaves Diseases Using Deep Neural Network Model" in IEEE Access, vol. 9, pp. 112942-112954, 2021, doi: 10.1109/ACCESS.2021.3096895.

[2] S. V. Militante, B. D. Gerardo and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning", IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 579-582, 2019.

[3] Sabrol, H., and K. Satish. "Tomato plant disease classification in digital images using classification tree." International Conference on Communication and Signal Processing (ICCSP). IEEE, 2016.

[4] Barbedo, Jayme Garcia Arnal. "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification." Computers and electronics in agriculture 153 (2018): 46-53.

[5] Barbedo, Jayme GA. "Factors influencing the use of deep learning for plant disease recognition." Biosystems engineering 172 (2018): 84-91.

[6] Barbedo, Jayme Garcia Arnal. "Plant disease identification from individual lesions and spots using deep learning." Biosystems Engineering 180 (2019): 96-107.

[7] Kamal, K. C., et al. "Depthwise separable convolution architectures for plant disease classification." Computers and Electronics in Agriculture 165 (2019): 104948.

[8] Elangovan, K., and S. Nalini. "Plant disease classification using image segmentation and SVM techniques." International Journal of Computational Intelligence Research 13.7 (2017): 1821-1828.

[9] Shruthi, U., V. Nagaveni, and B. K. Raghavendra. "A review on machine learning classification techniques for plant disease detection." 5th International conference on advanced computing & communication systems (ICACCS). IEEE, 2019.

[10] Too, Edna Chebet, et al. "A comparative study of fine-tuning deep learning models for plant disease identification." Computers and Electronics in Agriculture 161 (2019): 272-279.

[11] Jadhav, Sachin B. "Convolutional neural networks for leaf image-based plant disease classification." IAES International Journal of Artificial Intelligence 8.4 (2019): 328.