Biological Tomato Leaf Disease Classification using Deep Learning Framework

Ashwani Kumar Aggarwal
Department of Electrical and Instrumentation Engineering
Sant Longowal Institute of Engineering and Technology
SLIET, Longowal-148106, Punjab
India

Received: July 20, 2021. Revised: January 19, 2022. Accepted: February 17, 2022. Published: March 27, 2022.

Abstract: - Biological Tomato leaf classification is very important to decide the pesticide, insecticide, and other treatments needed for the plant to yield good crop. The images captured by handheld cameras or using drones are used by various machine learning algorithms to identify the diseases. Such methods need extraction of features from the images before the machine learning methods can be used for disease identification. In this paper, a deep learning framework is proposed that automatically extracts features in a hierarchical manner. The features are classified using neural networks to classify the leaves into three classes, viz. no disease, bacterial spot, and Septoria leaf spot. The performance of the model is tested using accuracy as the performance metric. The obtained performance metric validates the performance of the method. The method is useful for taking corrective measures to disease management of tomato plants.

Key-Words: - Biology, Precision agriculture, technology for biology and agriculture, leaf disease, classification, machine learning, feature detection, applied Biology.

I. INTRODUCTION

The biological tomato leave disease is very harmful for the growth of good quality of tomatoes. The diseases can be identified by shape, texture, color, and other features of tomato leaves. There are many categories of diseases that can be identified using leaf features. The common diseases are early blight, late blight, Septoria leaf spot, leaf mold, bacterial spot, among many others. Various computer vision techniques are used to identify the diseases by extracting features, describing area around the feature points by feature descriptors, and then matching the descriptors with those in the database [1]. The role of machine learning in classification is studied [2]. The use of handheld camera or cameras mounted on UAVs are used to take images of the tomato leaves. The UAVs can be used to mount not only cameras but also GPS sensors so that the location where the leaf image is taken can also be tagged [3]. The RGB images are converted into grayscale images. The image size is also made uniform for the images to 256x256. The various pre-processing tasks such as noise removal and histogram equalization can be done before features are extracted from the images. The feature extraction is done based on various geometric and illumination conditions. The commonly used feature point detectors and descriptors are Harris corner detector, Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Features from Accelerated Segment Test (FAST), Binary Robust Independent Elementary Features (BRIEF), and Oriented FAST and Rotated BRIEF (ORB) etc. There is some mismatching of feature descriptions which is taken care by RANSAC. The image matching can be performed using Manhattan distance, Euclidean distance, correlation, similarity index, and other template-based methods.

II. RELATED WORK

The biological tomato leaf disease detection has been attempted using convolutional neural networks [4]. The paper discusses application of slightly modified LeNet for tomato disease detection. The method has been shown to give an accuracy of 94-95% on a diverse dataset. The detection of tomato leave diseases with deep learning techniques is done [5]. The various image processing tasks, segmentation, clustering have been employed for tomato disease detection. The classification of tomato leaves in various disease classes has been attempted using LVQ algorithm and convolutional neural networks [6]. The dataset used in the paper consists of 500 images of tomato leaves with four different diseases. The image-based methods for detection of tomato diseases in done [7]. The dataset is taken from Plantvillage that contains 14903 images. The test accuracy obtained using the method is 99.25%. The ResNet is used for classification of tomato leaves. The method is validated using accuracy, precision, specificity, and F-score as performance metrics [8]. The leaf damage using infrared range of electromagnetic spectrum is used [9]. The ABCK-BWTR combined with B-ARNet is used for identifying the various diseases of tomato leaves [10]. The diseases such as late blight, bacterial spot leaves, and target classes have been classified using automatic methods [11]. The feature point extraction and describing the neighborhood around the feature points by feature descriptors is used for detection of tomato leaf diseases [12]. The support

vector machines for classification of tomato leaf diseases are done [13]. The segmentation of tomato leaf images using k-means algorithm is attempted [14]. A comparative study of machine vision and machine learning methods with deep learning methods for classification of tomato leaf diseases is done [15]. The classification of tomato leaf diseases using deep learning techniques with a mobile phone is attempted [16]. A transfer learning based deep learning approach to tarin and to classify the tomato leaf disease images are done [17].

III. METHODS AND MATERIALS

A. Dataset

The dataset of tomato leaves is taken from Plantvillage.

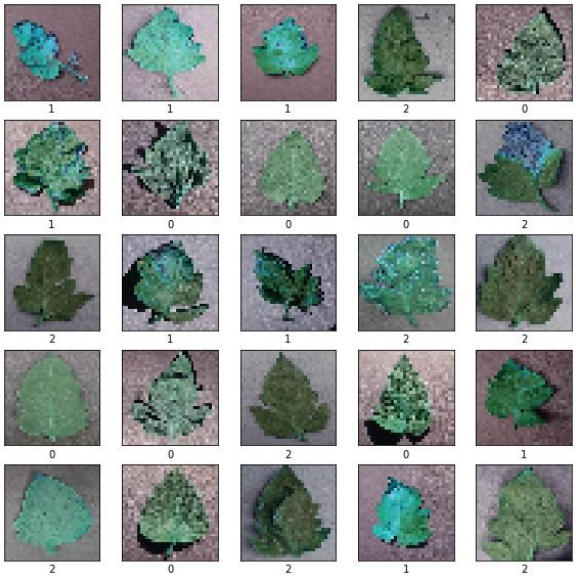

Figure 1. Dataset images of tomato leaves (0- healthy leave, 1- bacterial spot leaves, 2- Septoria spot leave)

It contains leaf images of healthy leaves, bacterial spot leaves, early blight, late blight, leaf mold, Septoria leave spot, spider mites, target spot, tomato yellow leaf curl virus, tomato mosaic virus. The total size of the dataset is 230.57 MB. The dataset contains healthy leave images and leaves of 9 different diseases. In this paper, the healthy leaf images, bacterial spot leaves, and Septoria spot leaves have been taken for classification.

B. Deep Learning Framework

The deep learning framework used consists of convolutional layers, max-pooling layers, flatten, dense layers, and dropouts. The Leaky ReLU is used as activation function. The dropout is used to avoid overfitting the model on the training data. The data is split into training and testing datasets in the ration of 0.8. Fig. 2. Shows the architecture of deep learning framework used for the classification of tomato leaves.

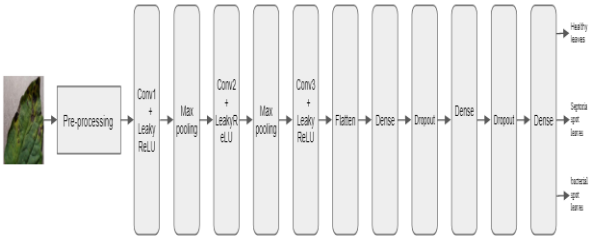

Figure 2. Architecture of deep learning framework

The activation function is a nonlinear function. The Sigmoid activation function follows an exponential relationship as is given in equation (1).

$$y = \frac{1}{1 + e^{-x}} \quad (1)$$

The ReLU activation function gives output equal to 0 when input is negative or zero and gives an output equal to input when it is positive. is given in equation (2).

$$y = max\ (0, x) \quad (2)$$

Where x is the input to the activation function and y is its output.

V. RESULTS AND DISCUSSION

The model is trained using the training images without applying the dropout and by applying the dropout. The layer details of the model are given in Fig. 3. The dropout is one of the options to overcome the problem of overfitting of model to the training set.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 118, 158, 32)	896
max_pooling2d_1	(MaxPooling2(None, 59,79,32)	0
conv2d_3 (Conv2D)	(None, 57, 77, 64)	18496

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 118, 158, 32)	896
max_pooling2d_1	(MaxPooling2 (None, 59, 79, 32)	0
conv2d_3 (Conv2D)	(None, 57, 77, 64)	18496
flatten_1 (Flatten)	(None, 280896)	0
dense_2 (Dense)	(None, 64)	17977408
dense_3 (Dense)	(None, 2)	130

Total params: 17,996,930
Trainable params: 17,996,930
Non-trainable params: 0

Figure 3. Layer details of the model

The size of the image is resized to 32x32 from 256x256 in order to reduce the training time. The model losss and accuracy in case of without dropout and with dropout are given in Table 1 and Table 2 respectively.

TABLE I. WITHOUT DROPOOUT

Epoch	Model Loss		Accuracy	
	Training	Testing	Training	Testing
Epoch #1	0.6755	0.3190	0.6969	0.8929
Epoch #2	0.2385	0.1930	0.9165	0.9362
Epoch #3	0.1410	0.1255	0.9520	0.9567
Epoch #4	0.1066	0.1732	0.9575	0.9317
Epoch #5	0.9693	0.0854	0.0830	0.9658
Epoch #6	0.0479	0.0708	0.9845	0.9681
Epoch #7	0.0410	0.0854	0.9824	0.9727
Epoch #8	0.0511	0.2402	0.9824	0.9226
Epoch #9	0.0718	0.0990	0.9748	0.9636
Epoch #10	0.0608	0.1013	0.9775	0.9704
Epoch #11	0.0447	0.0688	0.9839	0.9772
Epoch #12	0.0219	0.0526	0.9930	0.9863
Epoch #13	0.0115	0.0987	0.9961	0.9658
Epoch #14	0.0252	0.0564	0.9909	0.9727
Epoch #15	0.0143	0.0888	0.9951	0.9749
Epoch #16	0.0304	0.0581	0.9897	0.9863
Epoch #17	0.0139	0.0513	0.9967	0.9818
Epoch #18	0.0041	0.0418	0.9997	0.9863
Epoch #19	0.0027	0.0452	0.9994	0.9818
Epoch #20	0.0014	0.0388	1.0000	0.9863
Epoch #21	0.0006	0.0433	1.0000	0.9841
Epoch #22	0.0157	0.2586	0.9957	0.9294
Epoch #23	0.0701	0.0681	0.9763	0.9795
Epoch #24	0.0145	0.0547	0.9957	0.9863
Epoch #25	0.0134	0.0670	0.9961	0.9886
Epoch #26	0.0040	0.0550	0.9994	0.9863
Epoch #27	0.0125	0.0595	0.9961	0.9818
Epoch #28	0.1189	0.2135	0.9593	0.9362
Epoch #29	0.0311	0.0456	0.9882	0.9818
Epoch #30	0.0040	0.0471	0.9994	0.9818

TABLE II. WITH DROPOUT

Epoch	Model Loss		Accuracy	
	Training	Testing	Training	Testing
Epoch #1	0.8847	0.4757	0.5752	0.8155
Epoch #2	0.4374	0.2706	0.8187	0.8838
Epoch #3	0.3202	0.1811	0.8788	0.9408
Epoch #4	0.2440	0.2028	0.9095	0.9134
Epoch #5	0.2057	0.1938	0.9295	0.9248
Epoch #6	0.1755	0.1026	0.9399	0.9567
Epoch #7	0.1473	0.0581	0.9456	0.9841
Epoch #8	0.1155	0.0424	0.9660	0.9863
Epoch #9	0.1054	0.0943	0.9657	0.9590
Epoch #10	0.1512	0.1004	0.9544	0.9636
Epoch #11	0.0943	0.0463	0.9669	0.9841
Epoch #12	0.0790	0.1281	0.9742	0.9636
Epoch #13	0.0851	0.0321	0.9718	0.9863
Epoch #14	0.0687	0.9775	0.0923	0.9567
Epoch #15	: 0.0679	0.9784	0.0552	0.9795
Epoch #16	0.0571	0.9806	0.1174	0.9567
Epoch #17	0.0622	0.9821	0.0866	0.9681
Epoch #18	0.0473	0.9857	0.0261	0.9932
Epoch #19	0.0661	0.0255	0.9784	0.9886
Epoch #20	0.0715	0.0836	0.9760	0.9749
Epoch #21	0.0510	0.2089	0.9836	0.9385
Epoch #22	0.0458	0.0242	0.9851	0.9909
Epoch #23	0.0698	0.0529	0.9760	0.9886
Epoch #24	0.0281	0.1139	0.9930	0.9613
Epoch #25	0.0296	0.0238	0.9894	0.9886
Epoch #26	0.0394	0.0366	0.9882	0.9932
Epoch #27	0.0448	0.0537	0.9845	0.9818
Epoch #28	0.0309	0.0231	0.9921	0.9932
Epoch #29	0.0261	0.0360	0.9918	0.9909
Epoch #30	0.0366	0.0700	0.9891	0.9795

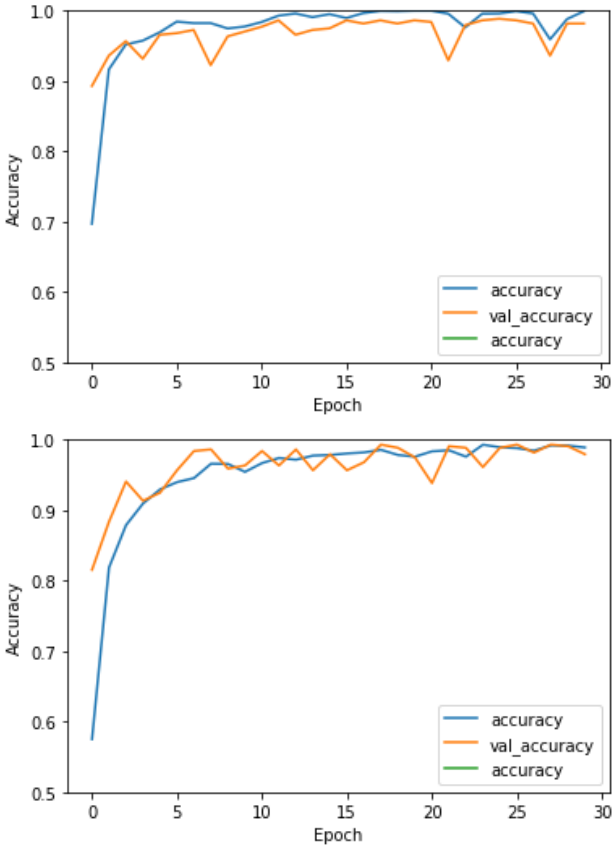

Figure 4. Accuracy plot (a) without dropout (b) with dropout

It is observed from the two plots that dropout improves the validation accuracy by overcoming the problem of overfitting of the model on the training dataset. The model validation can also be done using several other performance metrics such as precision, sensitivity, specificity, and AUC, etc. The method can be improved by performing some data augmentation techniques such as giving the images after histogram equilization, filtering the images with Gaussian filters, and by using GANs.

REFERENCES

[1] Arora, Kratika, and Ashwani Kumar Aggarwal. "Approaches for Image Database Retrieval Based on Color, Texture, and Shape Features." Handbook of Research on Advanced Concepts in Real-Time Image and Video Processing. IGI Global, 2018. 28-50.

[2] Chauhan, Sumika, Manmohan Singh, and Ashwani Kumar Aggarwal. "Data Science and Data Analytics: Artificial Intelligence and Machine Learning Integrated Based Approach." Data Science and Data Analytics: Opportunities and Challenges (2021): 1.

[3] Aggarwal, Ashwani Kumar. "Fusion and Enhancement Techniques for Processing of Multispectral Images." Unmanned Aerial Vehicle: Applications in Agriculture and Environment. Springer, Cham, 2020. 159-175.

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

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

[6] 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.

[7] Kumar, Akshay, and M. Vani. "Image based tomato leaf disease detection." 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2019.

[8] Kaur, Manpreet, and Rekha Bhatia. "Development of an improved tomato leaf disease detection and classification method." 2019 IEEE Conference on Information and Communication Technology. IEEE, 2019.

[9] Xu, H. R., et al. "Near-infrared spectroscopy in detecting leaf miner damage on tomato leaf." Biosystems Engineering 96.4 (2007): 447-454.

[10] Chen, Xiao, et al. "Identification of tomato leaf diseases based on combination of ABCK-BWTR and B-ARNet." Computers and Electronics in Agriculture 178 (2020): 105730.

[11] Lu, Jinzhu, et al. "Detection of multi-tomato leaf diseases (late blight, target and bacterial spots) in different stages by using a spectral-based sensor." Scientific reports 8.1 (2018): 1-11.

[12] Basavaiah, Jagadeesh, and Audre Arlene Anthony. "Tomato Leaf Disease Classification using Multiple Feature Extraction Techniques." Wireless Personal Communications 115.1 (2020): 633-651.

[13] Mokhtar, Usama, et al. "Identifying two of tomatoes leaf viruses using support vector machine." Information Systems Design and Intelligent Applications. Springer, New Delhi, 2015. 771-782.

[14] Tian, Kai, et al. "Segmentation of tomato leaf images based on adaptive clustering number of K-means algorithm." Computers and Electronics in Agriculture 165 (2019): 104962.

[15] Tan, Lijuan, Jinzhu Lu, and Huanyu Jiang. "Tomato Leaf Diseases Classification Based on Leaf Images: A Comparison between Classical Machine Learning and Deep Learning Methods." AgriEngineering 3.3 (2021): 542-558.

[16] Ngugi, Lawrence C., Moataz Abelwahab, and Mohammed Abo-Zahhad. "Tomato leaf segmentation algorithms for mobile phone applications using deep learning." Computers and Electronics in Agriculture 178 (2020): 105788.

[17] Thangaraj, Rajasekaran, S. Anandamurugan, and Vishnu Kumar Kaliappan. "Automated tomato leaf disease classification using transfer learning-based

[18] deep convolution neural network." Journal of Plant Diseases and Protection 128.1 (2021): 73-86.