Detection of Tomato Plant Leaf Diseases on the basis of Deep Neural Network.

Pradeep S[1], Yogeesh A C[2], Madhu Chandra G[3], N Sathisha[4], Vidya J[5]

[1]Assistant Professor, Dept of Computer Science and Engineering Government Engineering College, Chamarajanagar; [2]Assistant Professor, Dept of Computer Science and Engineering Government Engineering College,Kushalnagar.[3]Associate Professor, Dept of Electronics and Communication Engineering, R L Jalappa Institute of Engineering, Doddaballapur; [4]Assistant Professor, Dept of Electronics and Communication Engineering Govt. SKSJTI (Sri Krishna Rajendra Silver Jubilee Technological Institute), Bengaluru;[5]Assistant Professor, Dept. of Computer Science and Engineering, GITAM School of Technology, Bengaluru.

Article History: **Received:** 20.10.2022 **Revised:**27.10.2022 **Accepted:** 12.11.2022

***Abstract*—** In order to provide the basic demands of food for the growing population worldwide, agriculture is a crucial industry. The growth of grains and vegetables, meanwhile, is essential to human nourishment and the global economy. As a result of their reliance on manual monitoring of grains and vegetables and their lack of correct knowledge and disease detection, many farmers cultivate in distant places of the world and incur significant losses. Digital farming techniques might be an intriguing way to swiftly and readily identify plant diseases. This research suggests a method for identifying plant leaf diseases and taking preventive action in the agricultural sector utilizing image processing and two well-known convolution neural network (CNN) models, such as ResNet-50. Numerous plant diseases have a substantial impact on the crop's quality and yield. Early identification of these disorders is therefore very beneficial. One of the significant crops that is produced in great amounts and has a high commercial value is the tomato. The rate at which various tomato diseases are affecting the crop is worrying. In this study, we implemented two ResNet-based convolution neural network models for the classification of tomato leaf diseases. With the help of deep learning, the proposed effort seeks to identify the most effective method to the problem of tomato leaf disease detection. To prevent production loss in tomato fields, the proposed technology can be employed for early disease identification.

***Keywords*—** **CNN and** Digital Farming Practices, hyperparameters; plant leaves dataset

DOI: 10.48047/ecb/2023.12.10.980

1. INTRODUCTION

The vast majority of the world's population relies on their robust economies. Growth of the economy also has a significant impact on a nation's GDP development. Agriculture is wholly dependent on the effects of this economy. However, various farming practices have an impact on the quantity and quality of grains and vegetables. Due to varying climatic and environmental conditions in various locations, these grains and vegetables are exposed to a variety of diseases. Therefore, cultivators in any nation suffer significant losses as a result of these diseases. The amount of agricultural yield affected by leaf disease is steadily declining. Finding the agricultural field's leaf disease and boosting output rates for both quality and quantity are the key challenges. To determine disease, it is first required to take into account the leaves of two crops. Two crucial crops that are used in our daily food and to restore nutrients in the human body are tomatoes and potatoes.

Any diseases that occurs naturally can have detrimental impacts on grains and vegetables as well as ultimately lower production, product quality, and output. In order to reduce agricultural erosion, appropriate classification and identification of leaf disease may be crucial. Different grains and vegetable leaves have various diseaseses, including bacterial, fungal, and viral ones. Alternaria Alternata, Anthracnose, Bacterial Blight, Cercospora Leaf Spot, Powdery Mildew, Black mold, Downy Mildew, and Rust are the most prevalent plant diseases. When an diseases affects a plant's leaf, the texture, color, shape, and size of the plant leaf reveal the infection's symptoms. Since the majority of symptoms are tiny, disease identification is impossible due to the limitations of human vision. However, it is essential to establish a highly effective method using scientific expertise and knowledge to identify disease symptoms. For this study, the tomato crop leaf photos that have been obtained are first gathered from Kaggle datasets. A standard digital camera or a high-resolution camera on a mobile phone might be used to take the pictures. The gathered tomato and potato leaves are then subjected to image processing. For the purpose of identifying plant diseases, a number of image processing techniques are used, including capture, pre-processing, restoration, segmentation, augmentation, feature extraction, and classification.

The color conversion technique is used at the pre-processing stage on RGB photos that are transformed into grayscale images. However, after removing various kinds of noise, a variety of contrast enhancement algorithms are utilized to boost the contrast of the images. These images can be transformed into aligned shapes using image augmentation techniques like flipping, cropping, and rotation, and other features like part, color information, or boundaries can be traced in the image. Additionally, the color image section can use the classification method to identify diseases. ResNet and other convolutional neural network (CNN) models, which are different categorization techniques, are used in this study. ResNet can identify several leaf illnesses and classify photos of healthy and unhealthy tomoto leaves. Additionally, many existing systems in the agricultural sector can identify some tomoto plant leaf diseases but do not offer a strategy for taking preventative steps. Due to this, the system suggested in this research uses a graphical user interface to both detect disorders and offer a preventive action. The major summary of the suggested framework is as follows: First, we use image processing techniques to detect disease in datasets of tomoto leaves. Second, we use ResNet-50 architecture to classify the tomoto leaf photos that have been processed. Thirdly, this study examines the accuracy of the overall classification of tomtoto leaf diseases. The graphical structure for illness identification with preventative measures is then evaluated and developed.

2. LITERATURE SURVEY

The color image part can be used to use the classification algorithm to identify diseases. Convolutional neural network (CNN) models like AlexNet and ResNet-50, which are various classification approaches, are used in this paper [1–7]. Images of healthy and ill leaves are categorized, and different leaf diseases are identified, using AlexNet and ResNet-50. In addition, many of the technologies already in use in agriculture can identify some plant leaf diseases but do not offer a strategy for taking preventative action. Due to this, the system suggested in this research uses a graphical user interface to both detect disorders and offer a preventive action.

To detect rice diseases and address issues like blurred image edges, noise, significant background interference, and low detection accuracy, Zhou et al. [8] proposed K-Means clustering algorithm and faster R-CNN Fusion algorithm using 3010 images taken by a camera. The speedier 2D-Otsu algorithm was also utilized to classify the photos of rice illness in order to get results.

In order to automatically detect and classify leaf diseases, Sharma et al. [9] introduced image pre-processing along with k-means clustering, segmentation, and four classifiers including logistic regression, SVM, KNN, and CNN. While logistic regression performs quite well due to classes, CNN provided the highest accuracy due to classification and detection of diseases using 20,000 images from GitHub and Kaggle.

With the use of IoT sensors, machine learning algorithms like SVM and NB, drone technology, and the Normalized Difference Vegetation Index, a strategy for generating heterogeneous data from NDVI was suggested in [10].

A CNN model and the Learning Vector Quantization (LVQ) approach were reported by Sardogan et al. [11] to identify and categorize leaf diseases in 500 tomato leaf pictures from the Plant Village dataset. Additionally, the various convolutional filters were employed to raise the level of recognition in the classification approach.

To detect leaf spot illnesses in sugar beet using imaging-based expert systems with 155 images, Ozguven et al. [12] created a modernised Faster R-CNN architecture with the altering parameters of a CNN model in order to get the greatest accuracy as well as to decrease the spending time and human error.

To identify and distinguish the rice diseases utilizing 500 real photographs from the rice empirical field, Lu et al. [13] suggested a new CNN-based algorithm with image preprocessing. The 10-fold cross-validation scheme, increased feasibility and economy, quicker convergence rate, and improved identification ability of the suggested model all contributed to its higher accuracy in comparison to the typical machine learning model. Kawasak et al. [14] introduced a novel disease detection model of viral plant based on CNN techniques with image pre-processing to automatically identify diseases using the 4-fold cross validation methodology with greater accuracy utilizing 800 cucumber leaf images taken by digital cameras.

In order to achieve higher accuracy and a workable solution, Jiang et al. [15] proposed a new model INAR-SSD that was applied in the Caffe structure of GoogLeNet Inception structure and Rainbow concatenation on the GPU platform. This model uses deep-CNNs to detect disease in real-time for apple leaves using an Apple leaf disease dataset that contains 26,377 images from the composition of lab and complex images of real field.

In order to properly detect crop illnesses in contemporary agriculture, Pallagani et al. [16] created the smartphone app Crop, which used a Deep Learning-based technique like AlexNet, ResNet-50, and ResNet-34 and utilized a public dataset of 54,306 photos of plant leaves. Additionally, a trained PyTorch model was transformed into a.pb file of tensor flow and loaded into the Crop app by exhibiting a workable solution in order to predict crop diseases in real time.

Using leaf photos from the PlantVillage database, CNN model with transfer learning method was used in [17] to diagnose diseases and identify pests of crops. This method reduced time and human labor while increasing accuracy.

3. SYSTEM DESIGN

System design is defined as the use of systems theory to the creation of a project. The convolution layer's architecture and parameters are determined by the system design.

3.1 System Architecture

The architecture of the proposed system is presented in fig 1.This process for building a model which can detect the disease associated with the tomoto leaf image The key points to be followed are:

1. Data Collection: The dataset taken was " Plant Village Dataset". It can be downloaded through the link "https://www.kaggle.com/vipoooool/new-plant-diseases-dataset". It is an Image dataset containing images of different healthy and

unhealthy crop leaves.

2. Model building: We have used pytorch for building the model. Our CNN model architecture consists of CNN Layer, Max Pooling, Flatten a Linear Layers. Also we used Transfer learning resnet34 Architecture.
3. Training: The model is trained by using variants of above layers mentioned in model building and by varying hyper parameters. The best model was able to achieve more percent of test accuracy.
4. Testing: The model was tested on total huge set of images

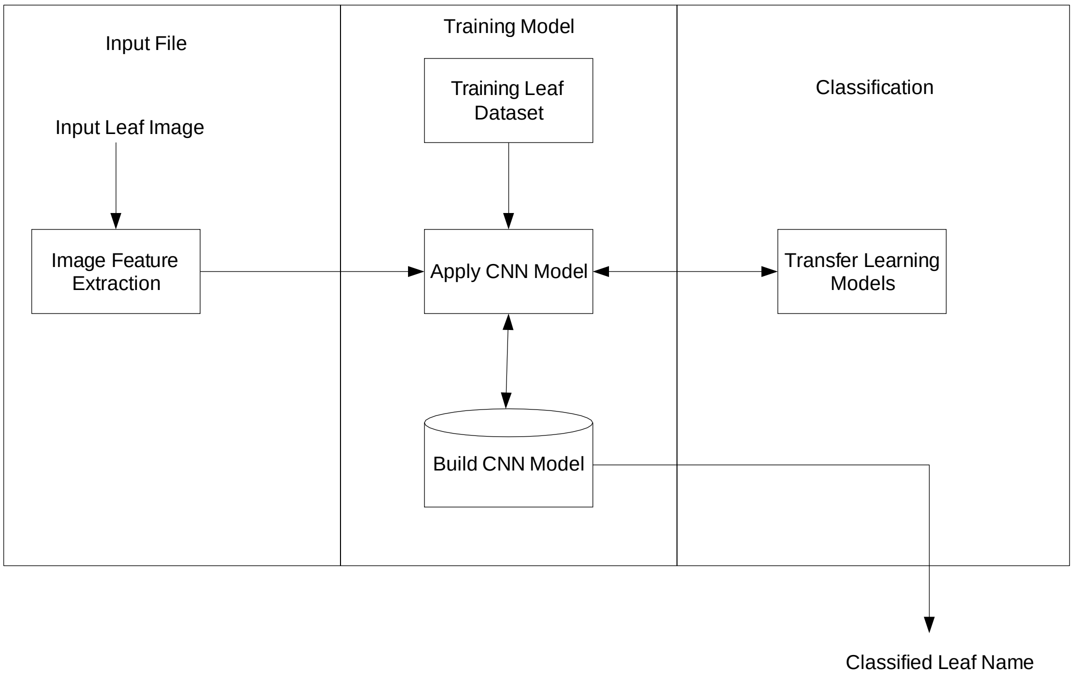

Fig. 1 Architectural Design of proposed system

The NVIDIA DGX v100 system has been used to run the suggested CNN model. The computer has 128 GB of RAM, 1000 TFLOPS speed, 5120 tensor cores, and 40600 CUDA cores. The class balancing data augmentation technique has been used since the images in the data set differ depending on the class. There are three convolution layers and a maximum pooling layer employed in the proposed CNN design. A different number of filters have been applied to each layer. The following Fig. 2 shows the proposed CNN model's design. The Panda technique has been used to obtain improved accuracy, and Table 1's description of the model's hyper parameters. Figure 3 shows the activation following each convolution layer.

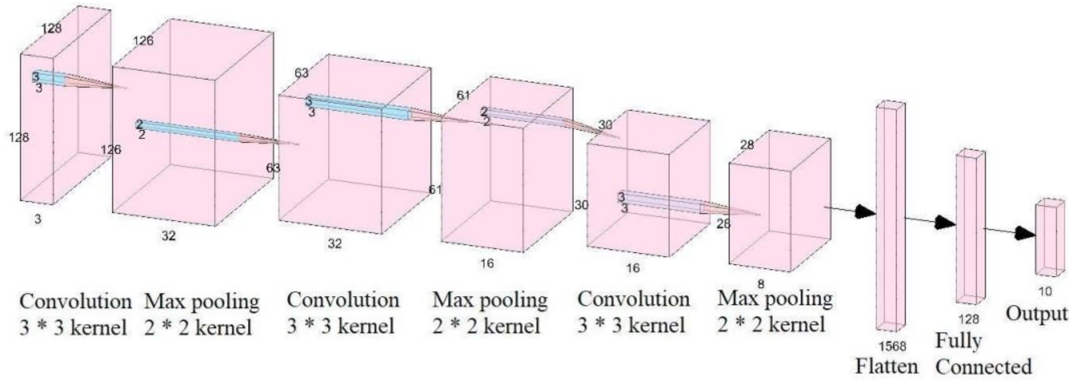

Fig. 2. Pictorial representation of the proposed convolution network.

Table 1. Hyper parameters for convolution neural network.

Hyperparameter	Description
convolution layers	3
max polling layers	3
Drop out rate	4
Network weight initialization	Glorot uniform
Activation function	Relu
Learning rate	0.001
Momentum	0.999
Number of epoch	1000
Batch size	32

3.2 Dataset

The Plant Village dataset has been used to extract images of tomato disease. Over 50,000 photos of 14 different crops, including tomatoes, potatoes, grapes, apples, corn, blueberries, raspberries, soybeans, squash, and strawberries, are included in the dataset. The crop we chose as our goal was tomato. The following are pictures of several tomato classes (see figure 3).

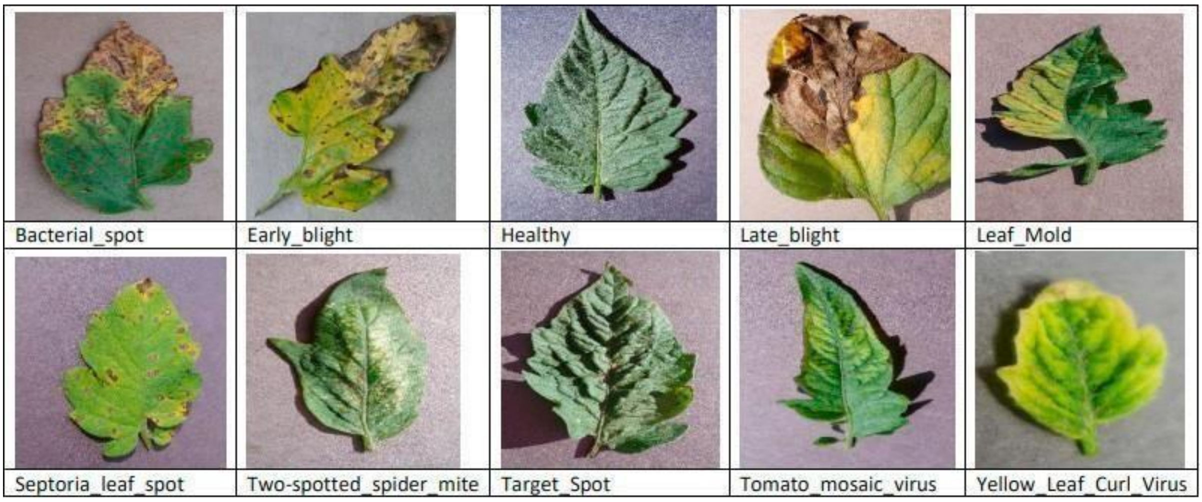

Fig.3. Class wise sample image of the dataset.

The following nine diseaseses affect tomatoes the most: Target Spot, Mosaic Virus, Bacterial Spot, Late Blight, Leaf Mold, Yellow Leaf Curl Virus, Two-spotted Spider Mite, Early Blight, and Septoria Leaf Spot are the diseases that affect plants. There are 10,000 photos in the training dataset, 7,000 images in the validation dataset, and 500 images in the testing dataset for the proposed study. 1000 photographs from the healthy group and 1000 images from each of the tomato disease categories listed above make up the 10,000 training images. Each class in the validation set has 700 images, while each class in the test set includes 50 images.

We removed 50 random photos from the training set from each class and used them in the testing set. We created our project's training dataset using the remaining training data by placing the identical number of photos (1000) in each class. When there were fewer than 1000 photographs in any class, we created some new images using the data augmentation technique. By rotating, flipping, cropping, and resizing the current photos, the Augmentor package of Python allows for the creation of similar new images. When there were more than 1000 photos in a class in the training dataset, the first 1000 were chosen. We used the same procedure for the validation dataset and gave each class 700 photos. This procedure is required to stop prejudice for any specific class during CNN training. All of the photographs are 256 *256 in size and are in the jpeg format.

4. Experimental Result

Additionally, we performed 1000 iterations of the proposed model, and Figure 4 shows the validation and training accuracy.The categorical cross entropy approach has been used to calculate loss. Equation 1 below represents the calculation's formula.

$$loss = -\sum_{y=1}^{X} log(p_{o,c}) \qquad (1)$$

where X - number of classes, y - binary indicator (0 or 1) if class label y is the correct classification for observation o and predicted probability observation o is of class c.

We conduct testing after evaluating the proposed model's performance. A total of 500 samples were used for testing, and the testing accuracy ranged from 74% to 100% for the different classes. The average accuracy of the suggested model is 92.3%.

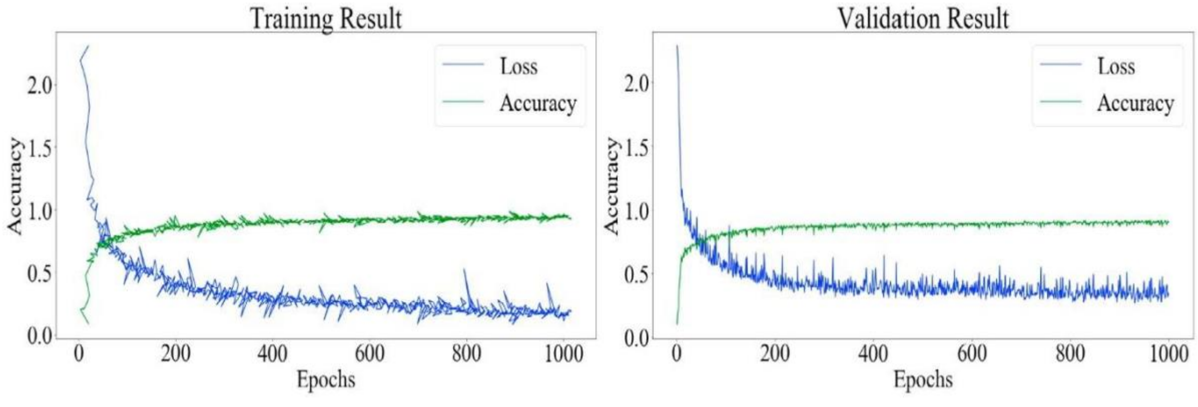

Fig. 4. (a) Training loss (b) Validation loss of proposed model.

The proposed technique is compared to pre-trained models, such as VGG16, MobileNet, and InceptionV3, in terms of the amount of trainable and non-trainable parameters. It is found that the new algorithm performs significantly better than the pre-trained models. The comparable bar charts are shown in figure 5 below. Table 2 also provides a description of the comparison. As can be shown, the suggested model requires substantially less storage space than pretrained models, making it better suitable for mobile devices with small storage capacities. In terms of accuracy, the proposed technique performs just as well as previously trained models (see figure 6). Figure 7displays the suggested model's performance in terms of Precision, Recall, and F1-score. Additionally, figure 8 shows the RoC-AUC curve for each of the 10 classes in relation to the suggested model.

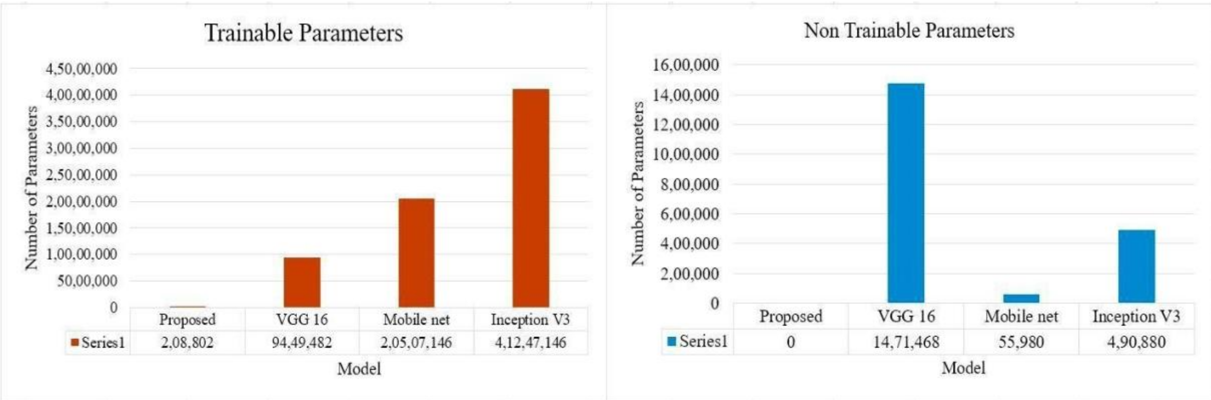

Fig. 5. Comparison of number of parameters in various Models.

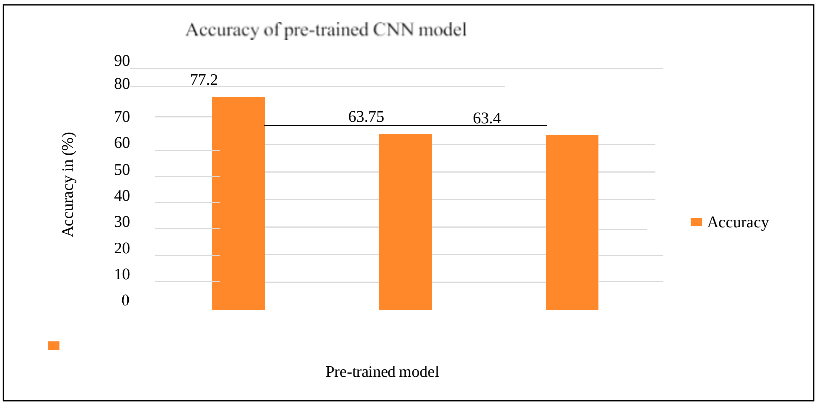

Fig. 6 Test accuracy using different pre-trained models i.e. VGG16, MobileNet and InceptionV3.

Table 2. Comparison of time and space complexity with Pretrained models.

S.No.	Model	Accuracy	Storage Space	Trainable Params	Non trainable Params
1	Mobilenet	62.75%	81,478 KB	20,307,346	559,808
2	VGG 16	78.2%	84,432 KB	9,559,482	13,784,678
3	InceptionV3	64.4%	1,53,734 KB	41,247,146	390,780
4	**Proposed Model**	**90.4%**	**1,586 KB**	**208,802**	**0**

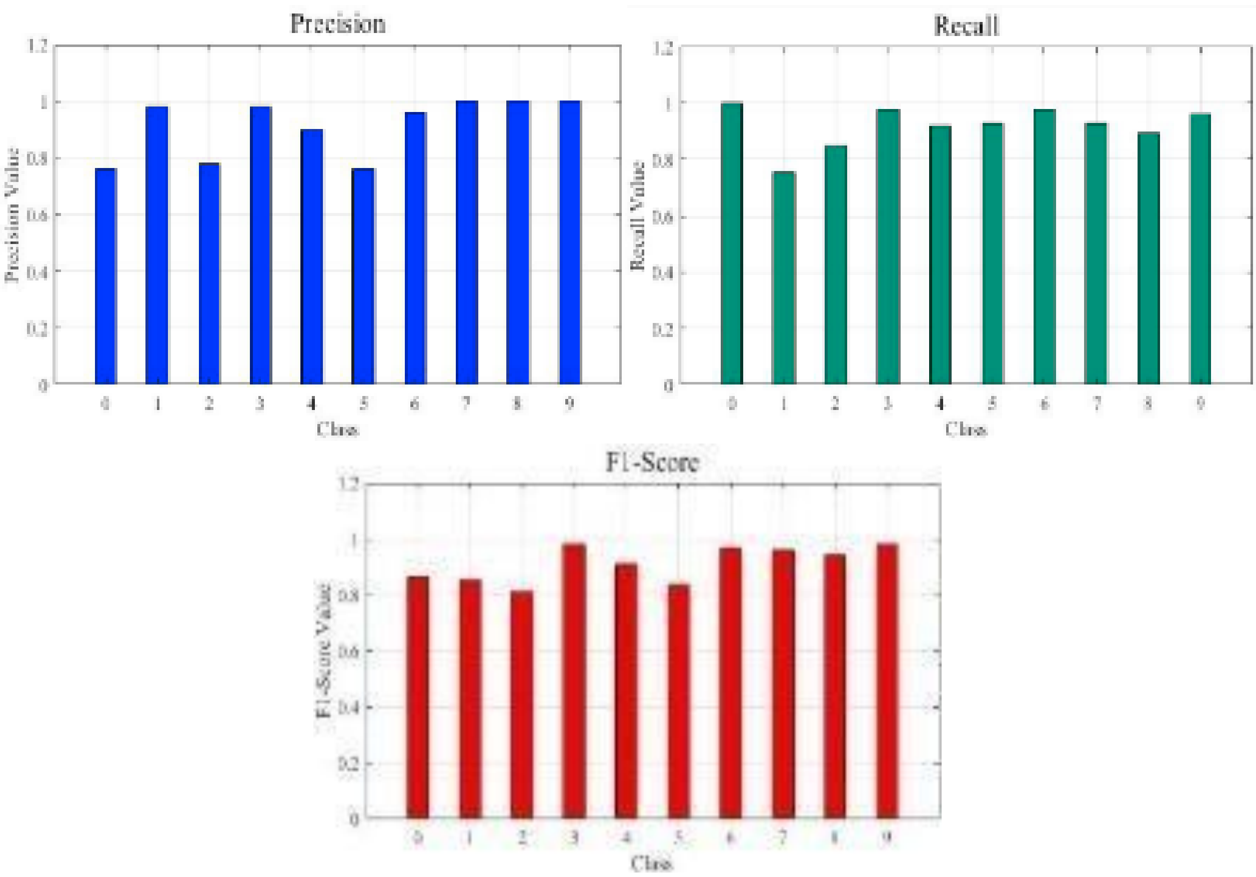

Fig. 7 Various performance evaluation metrics for proposed algorithm (a) Precision (b) Recall (c) F1-Score .

CONCLUSION

To identify the disease in tomato crops, a CNN-based model has been developed as part of the proposed endeavor. The suggested CNN-based design has three convolution and max pooling layers with various numbers of filters in each layer. For the experiment, we used the Plant Village dataset's data on tomato leaves. The photos inside the class have been balanced using augmentation techniques. Experimental results show that the model's class-specific testing accuracy ranges from 76% to 100%. Additionally, the model's testing accuracy is 91.2% on average. The suggested model requires storage space of about 1.5 MB, whereas pre-trained models require storage space of over 100 MB, demonstrating the proposed model's advantage over pre-trained models. We are trying to alter the model as part of ongoing research by using more images with a different crop. Additionally, because the testing accuracy is poor, we are striving to improve the same model on the same dataset.

REFERENCES

[1] H. Tao, M. Z. A. Bhuiyan, M. A. Rahman, G. Wang, T. Wang, M. M. Ahmed f, J. Li, “Economic perspective analysis of protecting big data security and privacy,” Future Generation Computer Systems, vol. 98, pp. 660–671, 2019.

[2] R. P. Narmadha and G. Arulvadivu, ”Detection and measurement of paddy leaf disease symptoms using image processing,” in ICCCI-2017, Coimbatore, INDIA, 2017.

[3] J. P. Shah, H. B. Prajapati and V. K. Dabhi, ”A survey on detection and classification of rice plant diseases,” in ICCTAC, Bangalore, INDIA, 2016, pp. 1-8.

[4] D. Kuznichov, A. Zvirin, Y. Honen and R. Kimmel, “Data Augmentation for Leaf Segmentation and Counting Tasks in Rosette Plants,” in IEEE Conference on CVPR Workshops, Long Beach, CA, USA, 2019.

[5] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos and P. de Geus, "Malicious software classification using transfer learning of ResNet-50 deep neural network," Proc. 16th IEEE Int. Conf. Mach. Learn. Appl., Cancun, MEXICO, 2017, pp. 1011-1014.

[6] S. Wang, X. Lv, D. Ye and B. Li, "Compressed Holistic Convolutional Neural Network-based Descriptors for Scene Recognition," 2019 4th Int. Conf. on Rob. and Auto. Eng., SINGAPORE, 2019.

[7] P. Khojasteh, L. A. P. Junior, T. Carvalho, E. Rezende, B. Aliahmad, J. ´ P. Papa, and D.K. Kumar, "Exudate detection in fundus images using deeply-learnable features," Comp. Biol. Med., vol. 104, pp. 62-69, 2019.

[8] G. Zhou, W. Zhang, A. Chen, M. He, and X.Ma, "Rapid detection of rice disease based on FCM-KM and faster R-CNN fusion," IEEE Access, vol. 7, pp. 143190-143206, Sep. 2019.

[9] P. Sharma ; P. Hans ; S. C. Gupta, "Classification Of Plant Leaf Diseases Using Machine Learning And Image Preprocessing Techniques," in Int. Conf. on Cloud Com., Data Sc. & Eng., Noida, INDIA, 2020

[10] U. Shafi, R. Mumtaz, N. Iqbal, S. M. H. Zaidi, S. A. R. Zaidi, I. Hussain, Z. Mahmood, "A multi-modal approach for crop health mapping using low altitude remote sensing, Internet of Things (IoT) and machine learning," IEEE Access, vol. 8, pp. 112708-112724, Jun. 2020

[11] M. Sardogan, A. Tuncer and Y. Ozen, "Plant Leaf Disease Detection and Classification based on CNN with LVQ Algorithm," in Int. Conf. on Comp. Sci. and Eng., Sarajevo, BOSNIA-HERZEGOVINA, 2018, pp. 382-385.

[12] M. M. Ozguven and K. Ademb, "Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms," Physica A: Statistical Mechanics and its Applications, vol.535, 2019

[13] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," Neurocomputing, vol. 267, pp. 378-384, Dec. 2017.

[14] Y. Kawasaki, H. Uga, S. Kagiwada, and H. Iyatomi "Basic Study of Automated Diagnosis of Viral Plant Diseases Using Convolutional Neural Networks," Lectures Notes in Computer Science, vol. 9475, pp. 638-645, 2015.

[15] P. Jiang, Y. Chen, B. Liu, D. He and C. Liang, "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks," IEEE Access, vol. 7, pp. 59069-59080, 2019.

[16] V. Pallagani, V. Khandelwal, B. Chandra, V. Udutalapally, D. Das and S. P. Mohanty," dCrop: A Deep-Learning based Framework for Accurate Prediction of Diseases of Crops in Smart Agriculture," in 2019 IEEE Int. Symposium on Smart Electronic Systems, 2019, pp.29-33.

[17] P. P. Patel and D. B. Vaghela, "Crop Diseases and Pests Detection Using Convolutional Neural Network," in 2019 IEEE Int. Conf on Electrical, Computer and Comm. Tech., Coimbatore, INDIA, 2019, pp. 1-4.

[18] M. A. Rahman, A. T. Asyharia, L. S. Leong, G. B. Satrya, M. H. Tao, M. F. Zolkipli, "Scalable Machine Learning-Based Intrusion Detection System for IoT-Enabled Smart Cities," Sustainable Cities and Society, vol. 61, pp. 1-43, 2020