A Deep Learning Approach for Detecting Multiple Plant Diseases

Md Delwar Shahadat Deepu, Indrani Sarker, Ariful Islam, Anamul Hasan, Farhan Shahriar Alvi and Raihan Rahman*
Department of Computer Science and Engineering, United International University
United City, Madani Avenue, Badda, Dhaka 1212, Bangladesh
Email: raiyan@cse.uiu.ac.bd.

***Abstract*—Bangladesh's economy is driven by the agriculture sector, which employs 42.7% of the workforce. But plant diseases often reduce a plant's ability to produce, so it is essential to diagnose them. Unfortunately, traditional techniques for detecting plant diseases are sometimes imperfect, complex, and time-consuming. In farming digitization, the ability to detect diseases using photographs of a specific plant's leaves has already been presented through recent advancements in computer vision. By examining photos of plants and their leaves, Machine Learning (ML) models may be trained to recognise specific disease signs and patterns, allowing them to make inferences and predictions about disease outbreaks. This study suggests a reliable method for identifying various plant diseases. To identify plant diseases for our research, we use image processing with several Artificial Neural Networks (ANN) that simultaneously identify the type of disease and the plant. From the "Plant Disease Merged Dataset," we gathered data on several plant diseases affecting four distinct types of plants: Potato, Tomato, Corn, and Grape. The suggested technique has a 99% accuracy rate for classifying the plants and, on average, 96% accuracy for detecting 21 distinct diseases of 4 popular plants.**

***Index Terms*—Machine Learning, Artificial Neural Network, and plant disease detection.**

I. Introduction

Plant diseases threaten the sustainability of the ecosystem and the world's food supply, leading to an overall reduction in agriculture and biodiversity [1]. Damages brought on by plant diseases may have detrimental economic effects on agriculture, compromising human health, national security, and food security. According to Food and Agriculture Organization(FAO) estimates, up to 40% of the world's yearly agricultural production is lost to insects that cause plant diseases, with a $220 billion global economic impact [2]. Plant diseases decrease crop productivity and quality, with losses corresponding to an average of 42% of the cultivation of the six essential food crops, including rice, wheat, corn, potatoes, cassava, and soybeans for human consumption [3]. To lessen the effects of plant diseases on agriculture and food security, it is essential to diagnose plant diseases and encourage sustainable agricultural practices.

Farmers or other experts constantly check the plants to find and recognize diseases, but this process is typically time-consuming, expensive, and inaccurate. They may adopt prompt treatments, secure their crops, and preserve agricultural productivity with the help of early and precise disease identification. Effective disease management and avoiding crop losses depend on promptly identifying plant diseases. A Deep Learning(DL) technique is available to identify plant diseases and define the kind of disease in early phases, enabling farmers to take necessary measures to stop the disease from spreading [4]. DL algorithms may speed up the disease detection process, saving time and money. They can correctly diagnose plant diseases by analyzing massive volumes of data, including photos of fatigued plants, environmental conditions, and other related data [5].

Numerous research methodologies have offered automated plant disease diagnostic methods, including pattern recognition [6], machine learning [7], and deep learning [8], since they have grown significantly in favour due to their dependability. Most of the approaches are tailored to certain plants, and in some cases, they are also disease-specific. In this field, multi-label categorization is utilized. Multiple classes may be produced simultaneously from a single input image via multi-label classifications [9]. Artificial Neural Network (ANN) is a form of machine learning model that can be shown to be useful in plant disease identification. It may be learned to recognize patterns and features that discriminate between healthy and unhealthy plants using massive databases of images of plants. ANNs may be employed to create a disease detection model that can classify the plant species and the kind of disease in plant leaves [10].

Based on the previously specified factors, We have presented a method that correctly identifies multiple plants and the associated various diseases of those plants. The proposed method employed multiple artificial Neural Networks (ANNs), including one for detecting multiple plants and further sub-ann models for detecting diseases in those plants.

The paper's overall contribution may be summarised as follows:

- We utilized a primary ANN model to identify four different plant types and other ANN sub-models for each plant's specific diseases.
- We used four distinct plants- potatoes, tomatoes, crops, and grapes—to construct our dataset, including 21 disease categories for those four plants.

- We assessed the model's effectiveness using precision, recall, f1-score, and accuracy metrics.
- We built a user interface that uses a picture of a leaf as input to predict the plant and disease of that plant.

II. Related Work

Several image segmentation approaches, feature extraction, and pattern recognition are suggested to detect and diagnose plant diseases. Deep learning algorithms have recently been used in various digital plant disease detection research.

P. Kulkarni et al. [11] suggested an innovative and effective method for identifying agricultural diseases capable of identifying 20 diseases in 5 popular plants. They chose 20 classes for experimentation using the public dataset PlantVillage, and for classification or detection tasks, a Random forest classifier has been utilized. Before training the model, data preprocessing and feature extraction are carried out; following model training, the accuracy is 93%. K Gowrishankar et al. [12] established the automatic recognition and classification of groundnut leaf diseases using image processing techniques, which involve gathering and preprocessing images of groundnut leaves, segmenting the preprocessed images by multi-threshold based color segmentation, and feeding the segmented images to feature extraction by Grey Level Co-occurrence Matrix (GLCM) and feature selection by rough set approach. After classifying the leaf diseases using ANN and SVM classifiers, performance measurements are created by contrasting the sensitivity and accuracy of the ANN and SVM classifiers to demonstrate the efficiency of ANN. AS Zamani et al. [13] provides a framework for identifying plant leaf disease using machine learning and image processing methods. The framework comprises removing noise from leaf photos before segmenting the picture with the K-Means method and extracting features with principal component analysis. After that, the photos are categorized utilizing methods like RBF-SVM, SVM, random forest, and ID3.

Kabir et al. [14] addressed the issue of detecting plant disease with multiple diseases from multiple plants. They've chosen 6 plants with 19 diseases in total. They've formed the dataset by collecting images of the selected plants using smartphone and specialized camera. To preprocess the dataset, to make the input leaf image more useful for subsequent processing, different image enhancing techniques and filters have been used in addition to color space conversion. Additionally, methods like histogram equalization are applied to keep the different lighting. Different Convolution Neural Network-CNN architectures have been implemented and benchmarked to perform multi label classification.

To identify cucumber leaf disease, Kawasaki et al. [15] presented CNN architectures and achieved 94.9% accuracy. For picture recognition in both small and large-scale datasets, CNN is the most practical classifier. It performed admirably in terms of images. classification and processing [16]. Deep learning model was trained by Mohanty et al. [17]. for accurately identifying 26 agricultural diseases and 14 crop species with 99.35% Architectures for GoogleNet and AlexNet. CNN is capable of feature extraction classifying images. Using a plant disease recognition method, Srdjan et al. [18] identify 13 various diseases and healthy leaves based on CNNs. The outcomes show that CNN is an appropriate choice for the job because of its reliable computing infrastructure diagnosis of plant disease.

CNN algorithms for the identification of plant diseases were examined by Abade et al. [19] . Between 2010 and 2019, more than 100 papers were published, which the authors examined. TensorFlow was found to be the most often used framework in this review, while PlantVillage was chosen as the most widely used dataset. The fundamental techniques of CNN models used to identify plant diseases using leaf pictures were described by Dhaka et al. [20]. Additionally, they contrasted frameworks, pre-processing methods, and CNN models. The datasets and performance metrics used to evaluate model efficacy are also examined in the study. Nagaraju et al.'s [21] analysis of the top datasets, pre-processing strategies, and DL algorithms for different plants was also included. They looked over and examined 84 studies on the use of DL for identifying plant diseases.

Murk et al. [22] applied a Convolution Neural Network for the detection of plant diseases. They've worked with 11 different plants with the dataset consisting of 15 classes in the first dataset and 38 classes in the second dataset. To expand the sample size, augmentation is first done to the dataset. Later, several convolution and pooling layers are employed with the Convolution Neural Network (CNN). Numerous diagnosis and identification techniques are proposed by the following image segmentation procedures, feature extraction, texture classification.

Meena et al. [23] addresses the vital issue of early plant disease detection using deep learning. The main highlight of the paper was to ensure global food security. The author of this paper clarifies the deep learning methodology. They also provide a clear overview of the convolution neural network(CNN) architecture for disease classifiers. The prediction of CNN is given in image-related tasks which is vital in declaration for various disease types. The paper would be more understandable and effective if it included explanation graphics and discussed any method limitation.

Vellaichamy et al. [24] offers a comprehensive approach to multi-class plant leaf disease classification. The use of DenseNet-121 architecture is a fitting choice, given its ability to capture complex features from images. The paper excels in explaining the architecture and methodology, along with robust experimentation. The reported results validate the effectiveness of the approach. However, incorporating visuals for architecture understanding and discussing limitations would enhance the paper's impact.

III. Methodology

A machine learning and image processing model for identifying leaf diseases is presented in the following subsection. Using an image of a leaf as input, the model can determine which plant the leaf belongs to and whether it is healthy or unhealthy. It will also predict which disease would be impacted if it is afflicted by one disease.

A. The overall approach for detecting plant diseases

The entire procedure is completed by gathering data for four plant species with multiple diseases, preprocessing the data, and training multiple ANN models for classifying the plants and detecting the diseases of those plants. Also designed a system that can identify plants with diseases by analyzing a picture of a plant leaf.

Figure1 represents the complete overview of the plant detection system. To classify the plants and their diseases, we trained a total of 5 ANN models. The system applied the optimal weights of a pre-trained model called VGG16 [25] to train new models. At first, the primary ANN model was trained with 8000 images of plant leaves to classify the plants, and the remaining Sub-ANN models were utilized to identify each plant's diseases. We need to perform image augmentation since some classes have insufficient data before training the sub-models to detect the disease of each plant.

Finally, we created an interface where users can upload images of the four plant leaf varieties. According to the provided photographs, the interface will determine which plant it belongs to and if it is healthy or unhealthy.

B. Dataset

We used a publicly available dataset named "Plant Disease Merged Dataset" to identify plant leaf disease. It includes photos of healthy and diseased plant leaves with 88 classes; however, we only chose 21 classes among them. We collected four distinct plant leaf types to build our dataset, each with 3 to 10 disease categories. The dataset contains approximately 8,000 images of four different plant leaves. Table 1 displays the distribution of plant species and diseases in the database, and the number of photos utilised for each class.

C. Dataset Preprocessing

We go through specific preprocessing methods with our data after collecting the dataset. Initially, we scaled the images into 240*240 pixels. We augmented the images of those classes whose sampling numbers were deficient. Some classes of plants require us to do augmentation, such as the Leaf_mold and mosaic_virus class of tomato, leaf_blight and healthy class of grapes, and so on. The augmentation technique involves rotating the photos by 20 degrees, shearing and enlarging the image by 20%, and then horizontally flipping the images. These techniques strengthen the model's ability to resist slight modifications and protect it from overfitting. The number of photos for each plant after augmentation is displayed in Table II.

TABLE I
SPECIFICATION OF THE DATASET

Plant Name	Plant Condition	Number of Samples
Potato	Healthy	152
	Early Blight	1096
	Late Blight	1093
Tomato	Mosaic Virus	382
	Leaf Mold	957
	Late Blight	1919
	Healthy	1598
	Early Blight	1009
	Bacterial Spot	2136
	Septoria Leaf Spot	1782
	Spider Mite	1676
	Target Spot	1404
	Yellow Leaf Curl Virus	3214
Corn	Northern Leaf Blight	1223
	Healthy	1162
	Gray Leaf Spot	1094
	Common Rust	1308
Grape	Leaf Blight	889
	Healthy	470
	Black Rot	11390
	Black Measles	1383

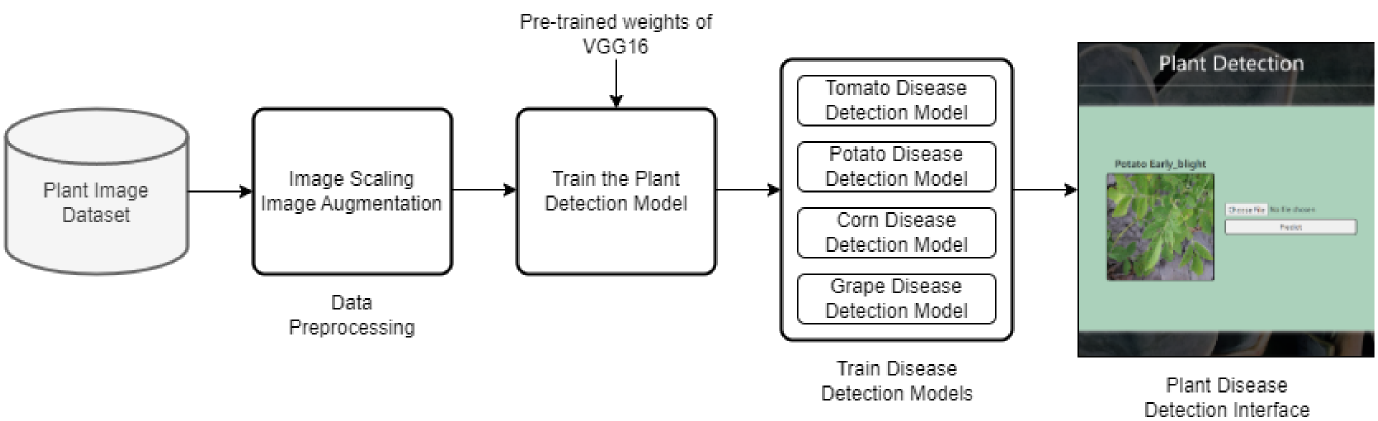


Fig. 1. Complete overview of plant disease detection system

TABLE II
NUMBER OF SAMPLES FOR EACH PLANT TAKEN FOLLOWING THE AUGMENTATION PROCESS

Plant Name	No of disease	Number of samples
Tomato	10	12000
Potato	3	3000
Corn	4	4000
Grape	4	4000

After all augmentation processes, the dataset is split into train and test sets with 80% and 20%, respectively.

IV. Result Analysis

The experiment used the Kaggle platform, which offers robust data science tools and resources. We utilised the 'Adam' optimizer and a batch size 32 for 50 epochs for training. The model's performance is assessed using evaluation metrics, including precision, recall, f1-score, and accuracy.

A. Evaluation Metrics

The performance metrics applied in the models include precision, recall, F1-score, and accuracy to evaluate how effectively the models work.

1) Precision: Out of all the actual positive values, it accurately counts the number of true positives. It is the ratio of true positives (TP) to the total of true positives (TP) and false positives (FP).

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

2) Recall: It evaluates the model's performance by accurately counting the number of true positives among all the actual positive values. It is the ratio of true positives (TP) over the sum of true positives (TP) and false negatives (FN).

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

3) F1-Score: It incorporates accuracy and recall scores into a single statistic to better assess model performance.

$$F1\_Score = \frac{2 * (Precision + Recall)}{Precision + Recall} \tag{3}$$

4) Accuracy: It calculates the ratio between predicted true positives and true negatives.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

B. Experimental Result

We achieved 99% accuracy after training the primary model to differentiate between the four varieties of plants. We get accurate results after configuring the submodels to identify various diseases affecting each plant. Table III displays those plants' precision, recall, f1-score, and accuracy.

TABLE III
PERFORMANCE METRICS FOR PLANT DISEASE DETECTION MODELS

Plant	Precision	Recall	F1-Score	Accuracy
Tomato	0.93	0.93	0.92	0.93
Potato	0.96	0.97	0.95	0.97
Corn	0.94	0.94	0.91	0.94
Grape	0.99	0.99	0.98	0.99

Figures 2 and 3 illustrate the graphical representations of the plant detection model and the disease detection models for each plant, respectively.

C. Prediction Displayed in Interface

Figure4 depicts the outcomes of the detection and identification of several plants randomly. The healthy tomato and grape leaves are displayed on the left, while the tomato and grape leaves affected by various diseases are shown on the right.

V. Conclusion

This work used deep learning to create a computerized approach for detecting plant diseases. This technique is based on a straightforward classification process examining numerous plant diseases. Data preprocessing techniques were used to strengthen our dataset so that our model would perform effectively. The plant identification model was trained using around 8,000 photos, and over 21,000 images were utilised to identify different diseases in each plant. To train the models, we made use of pre-trained weights of vgg16. We classified the plants with 99% accuracy and identified each plant diseases with an average accuracy of 95.5%. Lastly, we implemented a user interface that can accept a user-provided image of a leaf and predict the plant name and whether the leaf is healthy or unhealthy. The system can further anticipate the type of disease that will affect that plant if the leaf picture is deemed unhealthy. Our future research will focus on more plant diseases, and the system must be able to distinguish between images of plants and other objects.

References

[1] Singh, Brajesh K., Manuel Delgado-Baquerizo, Eleonora Egidi, Emilio Guirado, Jan E. Leach, Hongwei Liu, and Pankaj Trivedi. "Climate change impacts on plant pathogens, food security and paths forward." Nature Reviews Microbiology (2023): 1-17.

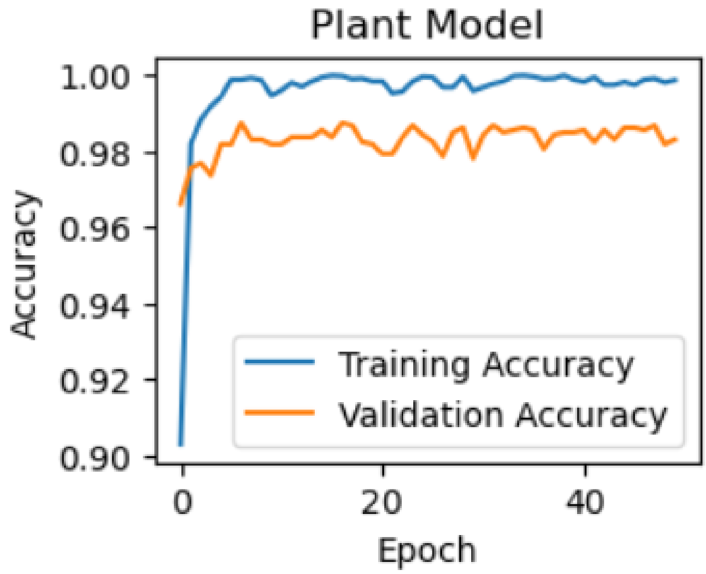
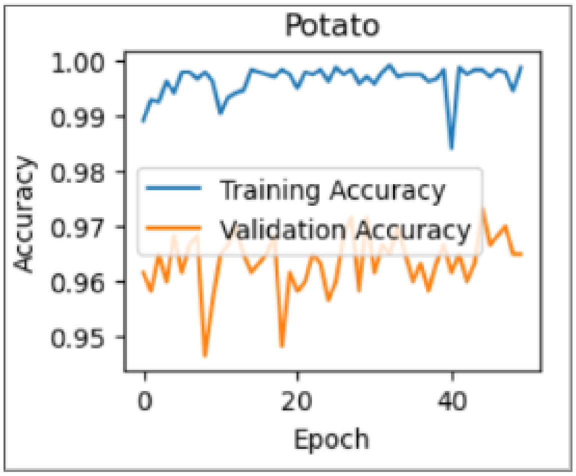
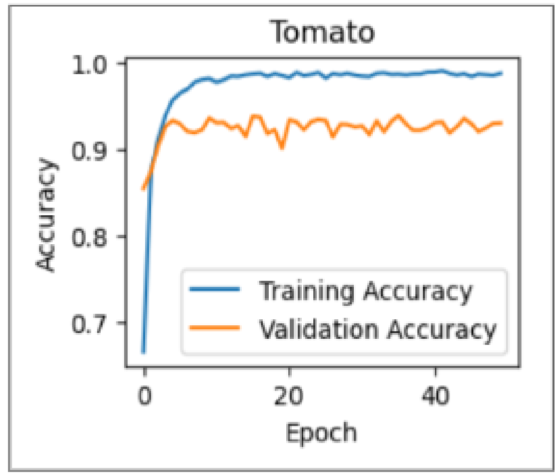


Fig. 2. Training vs Vaalidation accuracy of the plant detection model

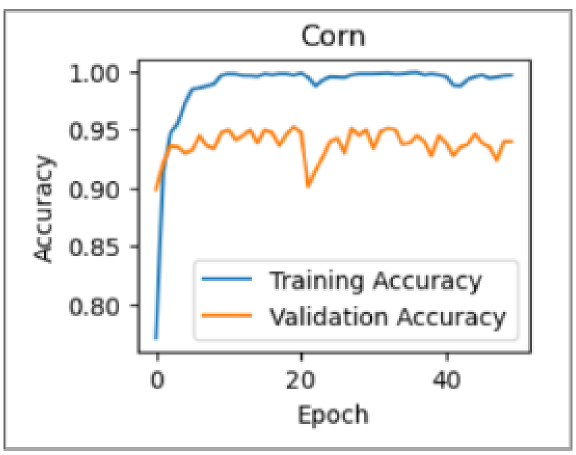
[2] FAO. ”Climate Change Fans Spread of Pests and Threatens Plants and Crops.” new FAO study (2021).
[3] Singh, Davinder, Grahame Jackson, Danny Hunter, Robert Fullerton, Vincent Lebot, Mary Taylor, Tolo Iosefa, Tom Okpul, and Joy Tyson. ”Taro leaf blight—a threat to food security.” Agriculture 2, no. 3 (2012): 182-203.
[4] Shoaib, Muhammad, Babar Shah, Shaker Ei-Sappagh, Akhtar Ali, Asad Ullah, Fayadh Alenezi, Tsanko Gechev, Tariq Hussain, and Farman Ali. ”An advanced deep learning models-based plant disease detection: A review of recent research.” Frontiers in Plant Science 14 (2023): 1158933.
[5] Harakannanavar, Sunil S., Jayashri M. Rudagi, Veena I. Puranikmath, Ayesha Siddiqua, and R. Pramodhini. ”Plant leaf disease detection using computer vision and machine learning algorithms.” Global Transitions Proceedings 3, no. 1 (2022): 305-310.
[6] Phadikar, Santanu, and Jaya Sil. ”Rice disease identification using pattern recognition techniques.” In 2008 11th International Conference on Computer and Information Technology, pp. 420-423. IEEE, 2008.
[7] Deepa, S., and R. Umarani. ”Steganalysis on Images using SVM with Selected Hybrid Features of Gini Index Feature Selection Algorithm.” International Journal of Advanced Research in Computer Science 8, no. 5 (2017).
[8] Kawasaki, Yusuke, Hiroyuki Uga, Satoshi Kagiwada, and Hitoshi Iyatomi. ”Basic study of automated diagnosis of viral plant diseases using convolutional neural networks.” In Advances in Visual Computing: 11th International Symposium, ISVC 2015, Las Vegas, NV, USA, December 14-16, 2015, Proceedings, Part II 11, pp. 638-645. Springer International Publishing, 2015.
[9] Kabir, Muhammad Mohsin, Abu Quwsar Ohi, and Muhammad F. Mridha. ”A multi-plant disease diagnosis method using convolutional neural network.” Computer Vision and Machine Learning in Agriculture (2021): 99-111.
[10] Jung, Minah, et al. ”Construction of deep learning-based disease detection model in plants.” Scientific Reports 13.1 (2023): 7331.
[11] Kulkarni, Pranesh, Atharva Karwande, Tejas Kolhe, Soham Kamble, Akshay Joshi, and Medha Wyawahare. ”Plant disease detection using image processing and machine learning.” arXiv preprint arXiv:2106.10698 (2021).
[12] Gowrishankar, K., and S. Lakshmi Prabha. ”An integrated image processing approach for diagnosis of groundnut plant leaf disease using ANN and GLCM.” (2020).
[13] AS Zamani et al.provide a framework for identifying plant leaf disease using machine learning and image processing methods. The framework comprises removing noise from leaf photos before segmenting the picture with the K-Means method and extracting features with principal component analysis. After that, the photos are categorised utilising methods like RBF-SVM, SVM, random forest, and ID3.
[14] Kabir, Muhammad Mohsin, Abu Quwsar Ohi, and Muhammad F. Mridha. ”A multi-plant disease diagnosis method using convolutional neural network.” Computer Vision and Machine Learning in Agriculture (2021): 99-111.
[15] Yusuke Kawasaki, Hiroyuki Uga, Satoshi Kagiwada, and Hitoshi Iyatomi. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks. In International Symposium on Visual Computing, pages 638–645. Springer, 2015.
[16] Andreas Kamilaris and Francesc X Prenafeta-Boldú. Deep learning in agriculture: A survey. Computers and electronics in agriculture, 147:70–90, 2018.
[17] Andreas Kamilaris and Francesc X Prenafeta-Boldú. Deep learning in agriculture: A survey. Computers and electronics in agriculture, 147:70–90, 2018.
[18] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. Deep neural networks based recognition of plant diseases by leaf image classification. Computational intelligence and neuroscience, 2016, 2016.
[19] Abade, A.; Ferreira, P.A.; de Barros Vidal, F. Plant diseases recognition on images using convolutional neural networks: A systematic review. Comput. Electron. Agric. 2021, 185, 106125. [Google Scholar] [CrossRef]
[20] Dhaka, V.S.; Meena, S.V.; Rani, G.; Sinwar, D.; Ijaz, M.F.; Woźniak, M. A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. Sensors 2021, 21, 4749. [Google Scholar] [CrossRef]
[21] Nagaraju, M.; Chawla, P. Systematic review of deep learning techniques in plant disease detection. Int. J. Syst. Assur. Eng. Manag. 2020, 11, 547–560. [Google Scholar] [CrossRef]
[22] Chohan, Murk, et al. ”Plant disease detection using deep learning.” International Journal of Recent Technology and Engineering 9.1 (2020): 909-914.
[23] Meena, M., and G. Monisha. ”Plant Diseases Detection Using Deep Learning.” 2022 1st International Conference on Computational Science and Technology (ICCST). IEEE, 2022.
[24] Swaminathan, Akshay, C. Varun, and S. Kalaivani. ”Multiple plant leaf disease classification using densenet-121 architecture.” Int. J. Electr. Eng. Technol 12 (2021): 38-57.
[25] Swasono, Dwiretno Istiyadi, Handayani Tjandrasa, and Chastine Fathicah. ”Classification of tobacco leaf pests using VGG16 transfer learning.” In 2019 12th international conference on information and communication technology and system (ICTS), pp. 176-181. IEEE, 2019.



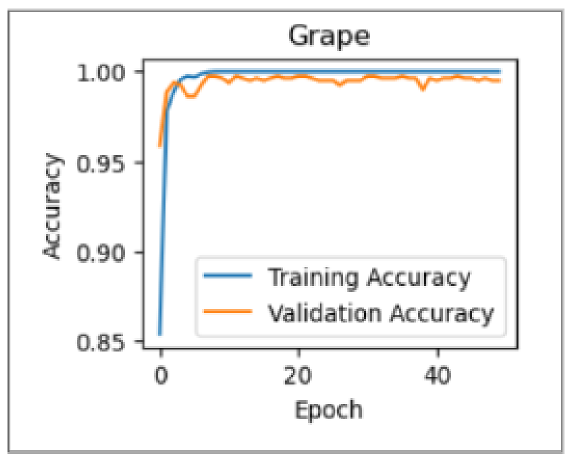
(a) Accuracy of Potato



(b) Accuracy of Tomato



(c) Accuracy of Corn



(d) Accuracy of Grape

Fig. 3. Training vs Vaalidation accuracy of the plant disease detection models

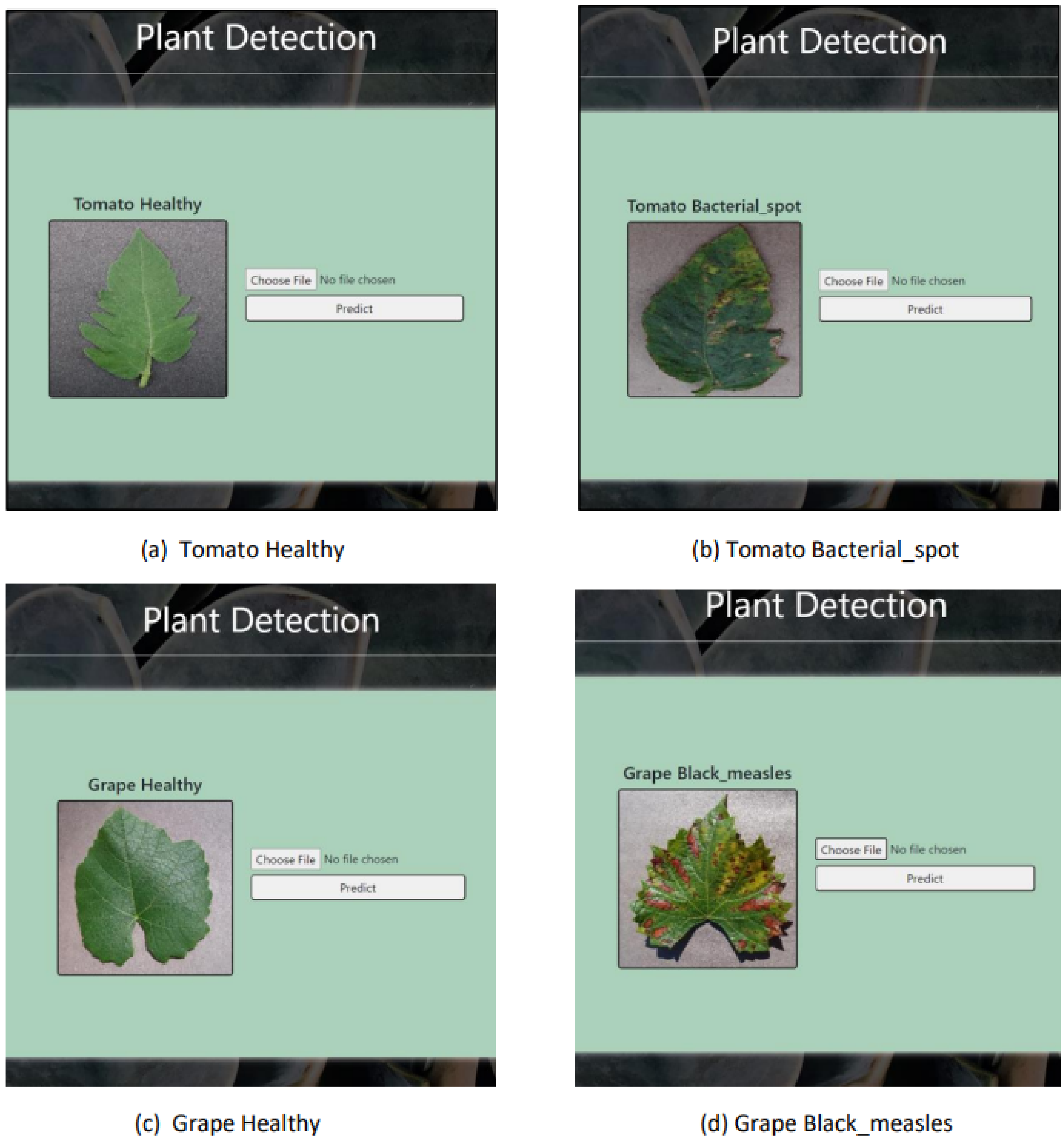


Fig. 4. Classification and Detection of Tomato and Grape plants