JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Plant and Plant Disease Identification System Using Deep learning and Purchase Assistance System

[1]Srujanendhra Talla, [2]B.Deepika, [3]K.Sai Sumanth,[4] Dr.V.Saidulu
[1]Student, [2]Student, [3]Student, [4]Assistant Professer,[1]Electronics and Communication Engineering,
[1]Mahatma Gandhi Institute of Technology, Hyderabad, India

Abstract : Plant health is crucial for achieving higher agricultural production. Some plant species are more prone to diseases due to environmental changes, so it is important to take good care of agricultural plants and adopt effective farming practices. Most farmers are not aware of the plant diseases or the appropriate pesticides to use when a plant gets infected. This lack of knowledge often leads to reduced crop yield and financial loss.To address this problem, we have designed a system using deep learning CNN model that identifies and detects plant diseases from images. We have integrated CNN model into web application that facilitates disease detection from plant images and also provides recommendations for purchasing suitable plants and pesticides. Training of the models was performed with the use of an open database of about 87K rgb images of healthy and diseased crop leaves which is categorized into 38 different classes,[plant, disease] combinations, including healthy plants. Several model architectures were trained, with the best performance reaching a 96% success rate in identifying the corresponding [plant, disease] combination (or healthy plant). The model's notable success rate underscores its potential as an effective advisory and early warning tool. Moreover, this approach can be further extended to develop a comprehensive plant disease identification system capable of functioning under real-world cultivation conditions.

IndexTerms - ***CNN, Deep learning, leaf disease identification, Plant disease, Web application, Purchase assistance***.

I. Introduction

India is an agricultural country and depends on agriculture for around 70% of the total population. Farmers can select various crops for cultivation and also pesticides for them. Any disease to anything makes it weak. In the same way plants become weak when leaves have diseases and plant growth is also affected. Therefore, monitoring plants is an important role in cultivation of plants[1]. Farmers typically rely on visual inspection and past experience to identify plant issues, which can be inaccurate and inconsistent. Moreover, access to expert agricultural consultants is often limited in rural areas, making timely diagnosis difficult. For plant diseases to be treated specifically and their spread to be controlled, early detection is essential. Early detection helps farmers avoid extensive damage and wasteful resource use by implementing preventive measures like isolating infected plants or using the appropriate pesticide. Early diagnosis on large farms reduces labor and operating costs while also improving overall crop health. Fast, dependable, and scalable disease detection systems that can function well in a variety of real-world field conditions are becoming more and more necessary as agriculture grows in size and complexity.or using the appropriate pesticide. Early iagnosis on large farms reduces labor and operating costs while also improving overall crop health. Fast, dependable, and scalable disease detection systems that can function well in a variety of real-world field conditions are becoming more and more necessary as agriculture grows in size and complexity. Due to the visual similarity of symptoms across many illnesses, manual plant disease identification is frequently delayed and prone to errors, even for skilled professionals. With the increasing digitization of agriculture, automated solutions such as Convolutional Neural Networks (CNNs) provide precise, image-based plant species and disease identification, making diagnostics accessible through smartphones or embedded systems, even in remote locations. By recommending suitable pesticides, fertilizers, or instruments based on the identified disease, crop type, and local conditions, a purchase recommendation system improves the solution beyond detection. The system is a useful, scalable, and farmer-friendly instrument for contemporary agriculture since integration with online platforms or regional vendors guarantees prompt action.

II. LITURATUTE SURVAY

The purpose of this review is to show the capability of some commonly used machine learning approaches that can efficiently handle these different but closely related objectives. It also conducts a comparative study of various artificial intelligence techniques that are applied to the same task of plant leaf disease detection and diagnosis systems. Some technical aspects of the learning techniques that are used in the reviewed studies are discussed. Machine learning provides a flexible and powerful framework for decision making and the incorporation of expert knowledge into the system. These are some of the few advantages of machine learning algorithms that make them be extensively used in many fields, and they are greatly applicable in agriculture mechanization[2]. Intelligent computational learning algorithms of data mining, such as artificial neural networks, have recognized advantages over explicit modeling approaches for classification problems[3]. Another technique which incorporates the features that are extracted by the SVM with respect to the HLB or citrus greening disease of lemon trees. The technique improves the accuracy of the system with a final overall accuracy of 85% using multi-band imaging sensor inputs [4] and 92.8% using the fluorescence imaging technique [5]. Deep CNN approach for different plant identification tasks using plant leaf images and different amounts of data. Additionally, the identification of plant diseases and pests can be achieved using deep CNNs. This technique was implemented for tomato plant diseases and pest detection [6]. Most recently, a 13-layer convolutional neural network was developed in [7] for learning some high-level features in order to classify fruit images using data augmentation methods and stochastic gradient descent with momentum, and it reached a classification accuracy of 94.94% in the final experiment. Additionally, the classification of multi-temporal crops can be achieved using convolutional neural networks. This technique was implemented for economic crops and presented in [8]. The maximum classification accuracy using this approach was 85.54%. In our study, the K-NN Classifier, Decision Tree, SVM, and Deep CNN were trained and evaluated in order to design a plant disease identification model based on a plant leaf image dataset. The succeeding section presents the fundamentals of the implemented models and the training and testing datasets.

III. PROBLEM STATEMENT

Plant diseases significantly affect crop yield and quality, and manual detection methods are often inaccurate, time-consuming, and dependent on expert knowledge, which is limited in rural areas. Even when diseases are identified, farmers may lack the knowledge or access to appropriate treatments. This creates a gap between diagnosis and action, leading to delays and potential crop loss. There is a need for an automated, accurate, and accessible system that can detect plant diseases and recommend suitable treatments to support timely and effective intervention.

IV. METHODOLOGY

The complete process of training, validating and testing of plant and disease identification with using deep CNN is explained in detail. The whole process is explained from data collection to classification process.

4.1 Data acquisition and Prepossessing

The diseased and healthy plant leaf images were downloaded from the new plant disease dataset. The dataset containing 87.9K images of 14 different plant leaves was used for training and 33 test images is created for testing the proposed Deep CNN model. The database includes 38 different classes, and each class is defined as a healthy or infected plant using disease labels. All the 38 classes will undergo training and validation process. The sample images of the random classes are presented in Fig. 1.

(a) (b) (c) (d)

4.2 Environment Setup

To ensure a stable and efficient training environment, the system was set up using Python along with popular deep learning libraries such as TensorFlow and Keras. The model training was carried out on a system equipped with GPU acceleration to handle the large dataset and expedite training time. This allowed for faster matrix operations and reduced overall training time. The use of Jupyter Notebooks facilitated iterative testing, visualization, and debugging. Additional libraries such as Matplotlib and Scikit-learn were employed for visualization and performance analysis, while OpenCV was used for image preprocessing tasks.

4.3 Training and Evaluation

The CNN architecture was constructed using both custom layers and pre-trained models such as ResNet-50 and InceptionV3. These models were fine-tuned to adapt to the specific task of plant disease classification. The configuration involved resizing all input images to 128x128 pixels, applying batch normalization, using ReLU as the activation function. Dropout layers were used to reduce overfitting by randomly disabling neurons during training. The final output layer used Softmax activation to classify the input image into one of the 38 predefined categories, each representing a specific disease or healthy state. To improve model generalization and prevent overfitting, several strategies were employed. Data augmentation played a crucial role by artificially enlarging the training set through random transformations such as rotation, zooming, flipping, and brightness adjustments. This helped the model to become invariant to orientation and lighting conditions. Transfer learning was used to benefit from features learned by models trained on the ImageNet dataset. Initially, the base layers of these models were frozen, and only the top layers were trained. Later, selective fine-tuning of deeper layers was performed to optimize performance. The model was trained using early stopping and learning rate scheduling techniques to prevent overfitting and ensure optimal convergence.

4.4 Visualization and Model Testing

Training and validation accuracy and loss curves were plotted to assess the learning behavior of the model across epochs shown in Fig.2. These plots helped determine if the model was overfitting or underfitting. A consistent decrease in loss and increase in accuracy indicated a successful training process. The confusion matrix provided further insights by showing the distribution of true and predicted labels, allowing identification of specific classes that were prone to confusion and might require additional data or model tuning. A separate test set consisting of 33 real-world images was used to evaluate the model after training. The model maintained a high level of accuracy (~96%) on the test data, confirming its generalization ability. Errors occurred mostly in cases where leaf symptoms were ambiguous or visually similar across different diseases. Despite these minor inaccuracies, the model's overall performance was deemed sufficient for practical use, particularly when combined with a recommendation system that provides actionable outputs. The test phase served as a final verification step before integration into the web application interface. Web application is user -friendly and serves us with the whole process that has to be performed to know the status of the plant (disease or healthy) along with the plant and pesticide assistance system.

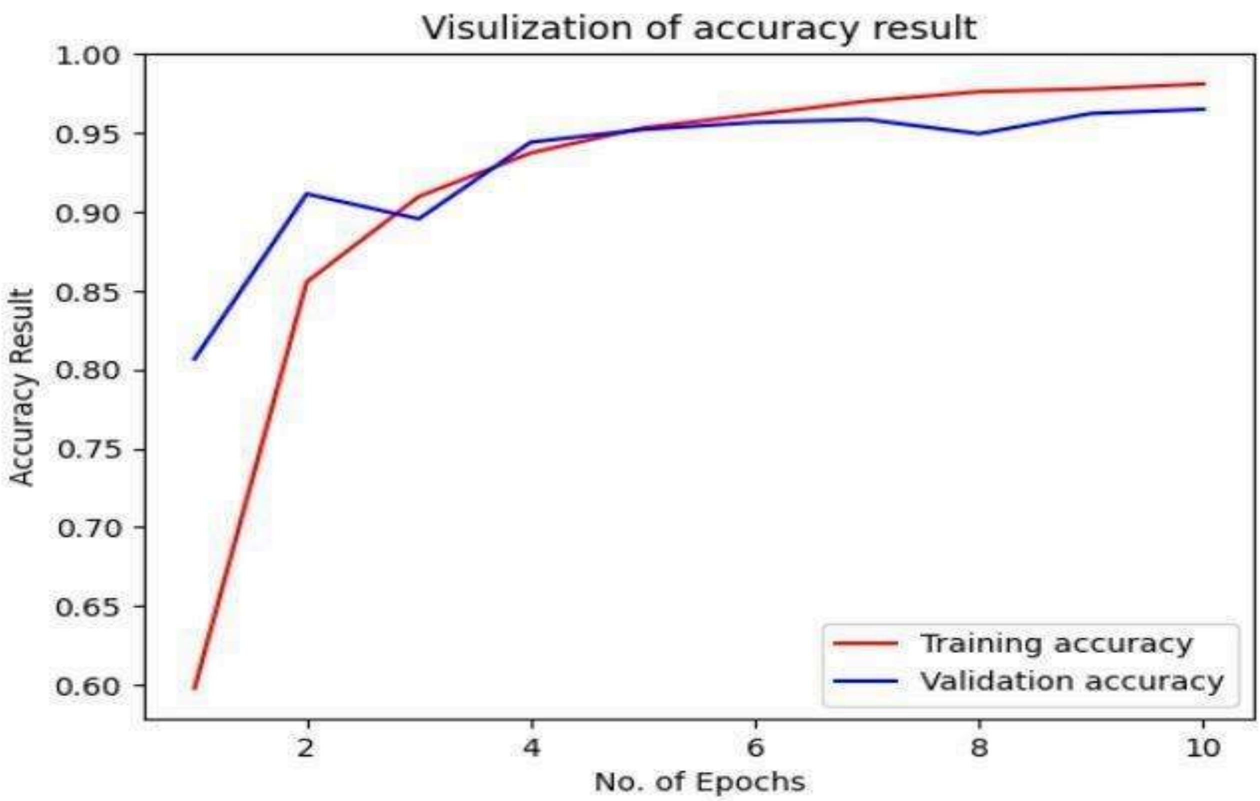

Fig.2. Training and validation accuracy curve

V. RESUTLS AND DISCUSSIONS

Fig 4.3 displays the precision, recall, F1-score, and support metrics for each plant disease and healthy category predicted by our CNN-based deep learning model. The dataset comprises 38 distinct classes, including various plant diseases and healthy conditions across crops like tomato, apple, corn, and grape. The model demonstrates high accuracy across most classes, with macro and weighted average precision, recall, and F1-scores all around 0.97, indicating a robust performance. Notably, classes such as Tomato Target Spot and Tomato Early blight had relatively lower scores, suggesting areas for further model tuning or data balancing.Fig.2. A graph titled "Visualization of Accuracy Result" illustrates the performance of the Convolutional Neural Network (CNN) model used in our project for plant identification and disease prediction. It presents the training and validation accuracy plotted over 10 epochs, providing insight into how effectively the model learns from the dataset. The red line represents training accuracy, while the blue line corresponds to validation accuracy. Initially, both metrics show a steep increase, indicating rapid learning of key features such as leaf shape, color, and texture. From epochs 4 to 6, the curves begin to converge and stabilize above 95%, reflecting strong model performance and minimal overfitting. By epoch 10, training accuracy reaches approximately 98%, while validation accuracy remains steady around 96%, showing that the model generalizes well to unseen data. This consistency between training and validation results demonstrates the robustness and reliability of the model, making it suitable for real-world deployment where it can assist users in accurately identifying plant species and diagnosing diseases, thereby guiding appropriate product recommendations in the purchase assistance module.

	precision	recall	f1-score	support
Apple___Apple_scab	0.98	0.95	0.96	504
Apple___Black_rot	0.95	0.99	0.97	497
Apple___Cedar_apple_rust	0.95	0.99	0.97	440
Apple___healthy	0.95	0.97	0.96	502
Blueberry___healthy	0.98	0.98	0.98	454
Cherry_(including_sour)___Powdery_mildew	0.98	1.00	0.99	421
Cherry_(including_sour)___healthy	0.99	0.98	0.99	456
Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot	0.93	0.94	0.93	410
Corn_(maize)___Common_rust_	1.00	0.99	0.99	477
Corn_(maize)___Northern_Leaf_Blight	0.95	0.95	0.95	477
Corn_(maize)___healthy	0.99	0.99	0.99	465
Grape___Black_rot	0.94	0.99	0.96	472
Grape___Esca_(Black_Measles)	0.99	0.97	0.98	480
Grape___Leaf_blight_(Isariopsis_Leaf_Spot)	1.00	0.99	0.99	430
Grape___healthy	0.95	1.00	0.97	423
Orange___Haunglongbing_(Citrus_greening)	0.97	0.99	0.98	503
Peach___Bacterial_spot	0.96	0.97	0.96	459
Peach___healthy	0.93	1.00	0.96	432
Pepper,_bell___Bacterial_spot	0.97	0.95	0.96	478
Pepper,_bell___healthy	0.96	0.96	0.96	497
Potato___Early_blight	0.98	0.99	0.99	485
Potato___Late_blight	0.97	0.93	0.95	485
Potato___healthy	0.99	0.93	0.96	456
Raspberry___healthy	0.97	0.99	0.98	445
Soybean___healthy	0.97	0.99	0.98	505
Squash___Powdery_mildew	0.99	0.97	0.98	434
Strawberry___Leaf_scorch	0.99	0.96	0.98	444
Strawberry___healthy	0.99	0.99	0.99	456
Tomato___Bacterial_spot	0.92	0.98	0.95	425
Tomato___Early_blight	0.86	0.92	0.89	480
Tomato___Late_blight	0.94	0.90	0.92	463
Tomato___Leaf_Mold	0.99	0.92	0.95	470
Tomato___Septoria_leaf_spot	0.98	0.81	0.89	436
Tomato___Spider_mites Two-spotted_spider_mite	0.97	0.95	0.96	435
Tomato___Target_Spot	0.89	0.95	0.92	457
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.99	0.97	0.98	490
Tomato___Tomato_mosaic_virus	0.99	0.99	0.99	448
Tomato___healthy	0.99	0.98	0.98	481
accuracy			0.97	17572
macro avg	0.97	0.97	0.96	17572
weighted avg	0.97	0.97	0.96	17572

Fig.3. support metrics for each plant disease

Once every process is completed and executed successfully we have integrated the whole system with a web application which is user friendly and easy to understand. As shown in Fig.4 the web application has a home page which describes, how to use the application toget the appropriate results based on data or image provided. And the following images Fig.5,6 depicts uploading an image and the prediction result on the user interface.

Fig.4.Home Page of the Plant Disease Recognition System Interface

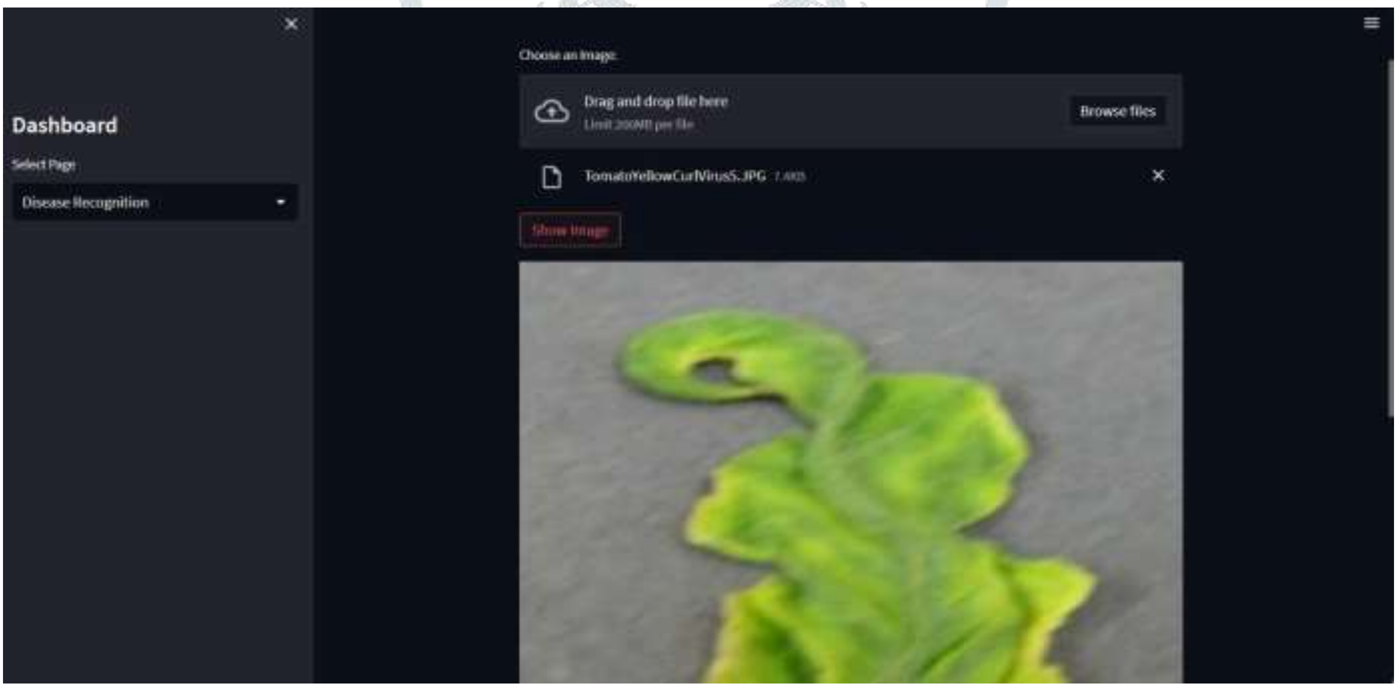

Fig.5.Disease Recognition Interface with Uploaded Leaf Image

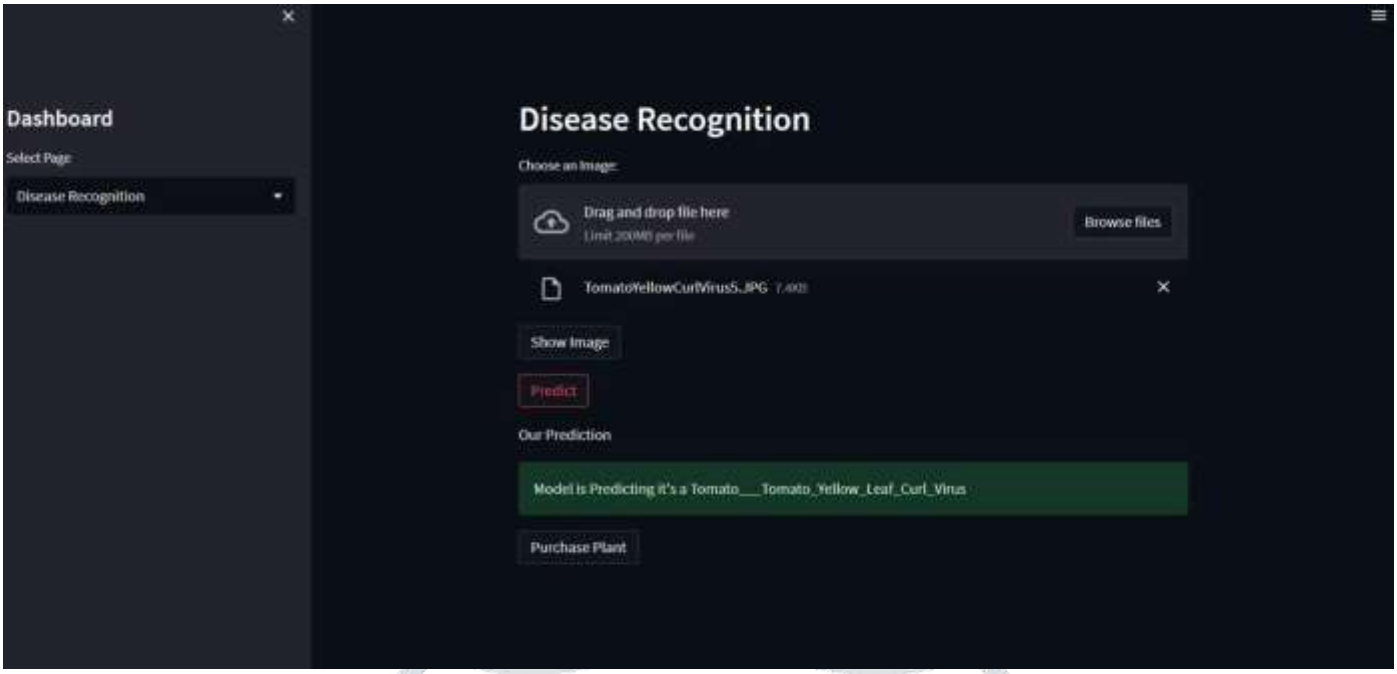

Fig.6. Disease Prediction Result Displayed on the User Interface

After prediction as shown in Fig.7 it will suggest to us some pesticides if the leaf is predicted as disease leaf along with the assistance of where to purchase the plant.

Fig.7. Final Disease Detection Result with Plant Purchase Assistance

VI. Conclusion

This paper presents a comprehensive deep learning-based system that streamlines the process of plant disease identification and offers practical solutions through purchase assistance of plants and pesticides. The system is designed with an intuitive web-based interface where users can upload images of affected plants. Using a trained convolutional neural network model, the system efficiently identifies the plant species and accurately detects the presence of specific diseases such as Tomato Yellow Leaf Curl Virus. Upon prediction, the interface provides a clear, easy-to-understand output and even offers a direct link for users to purchase healthy plants of the same type online. This combination of intelligent disease recognition and actionable recommendations addresses key challenges in agriculture, such as late disease diagnosis and unavailability of reliable recovery options. The project ultimately empowers users, especially farmers and home gardeners, to take timely action, enhance crop health, and minimize losses, contributing toward smarter and more sustainable agricultural practices.

REFERENCES

[1] Pranali K. Kosamkar, V.Y.Kulkarni, Krushna Mantri, Shubham Rudrawar, Shubhan Salmpuria, Nishant Gadekar "Leaf disease detection and recommendation of pesticides", 2018 Fourth ICCUBEA 2018.

[2] Identification of plant leaf diseases using a nine-layer deep convolutional neural network

[3] Yang CC, Prasher SO, Enright P, Madramootoo C, Burgess M, Goel PK, et al. Application of decision tree technology for image classification using remote sensing data. Agric Syst 2003;76(3):1101–17. doi:10.1016/S0308-521X(02)00051-3.

[4] Garcia-Ruiz F, Sankaran S, Maja JM, Lee WS,Rasmussen J, Ehsani R. Comparison of two aerial imaging platforms for identification of Huanglongbing Infected citrus trees. Comput Electron Agric 2013;91:106–15.doi:10.1016/j.compag.2012.12.002.

[5]] Wetterich CB, de Oliveira Neves RF, Belasque J, Ehsani R, Marcassa LG. Detection of Huanglongbing in Florida using fluorescence imaging spectroscopy and machine-learning methods. Appl Opt 2017;56(1):15–23. doi:10.1364/AO.56.000015

[6] Fuentes A, Yoon S, Kim SC, Park DS. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors 2017;17(9):2022. doi:10.3390/s17092022.

[7] Zhang Y-D, Dong Z, Chen X, Jia W, Du S, Muhammad K, et al. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimed Tools Appl 2019;78(3):3613–32.doi:10.1007/s11042-017-5243-3.

[8] Zhong L, Hu L, Zhou H. Deep learning based multi-temporal crop classification. Remote Sens Environ2019;221:430–43.doi:10.1016/j.rse.2018.11.0 32

[9] New Plant Diseases Dataset from Kaggle