Diagnosis of Plant Diseases by Image Processing Model for Sustainable Solutions

Sasmita Pani[1], Jyotiranjan Rout[1], Zeenat Afroz[1], Madhusmita Dey[1], Mahesh Kumar Sahoo[2], and Amar Kumar Das[3](✉)

[1] Department of Computer Science and Engineering, Balasore College of Engineering and Technology, Balasore, Odisha, India

[2] Department of Computer Science and Engineering, Gandhi Institute for Technology (GIFT) Autonomous, Bhubaneswar, Odisha, India

[3] Department of Mechanical Engineering, Gandhi Institute for Technology (GIFT) Autonomous, Bhubaneswar, Odisha, India

amar.das120@gmail.com

Abstract. The first step in preventing losses in agricultural product output and quantity is to identify plant diseases. A significant loss in crop output and market economic value results due to incorrect identification. The farmers used their own eyesight or prior knowledge of plant illnesses to identify plant ailments. When farmers are doing this for a single plant, it is possible, but when it involves many distinct plants, it is much more challenging to detect and takes a lot of effort. Therefore, it is preferable to utilize image processing to detect plants diseases. Image acquisition, picture pre-processing, image segmentation, feature extraction, and classification are all processes in this approach to diagnose the plant disease. In this study, we would like to present the procedures for identifying plant diseases from their leaf photos. We have used VGG 19 model for efficient processing of trained data and test data. This paper aims to support and help the green house farmers in an efficient way.

Keywords: diseases · image processing · trained data

1 Introduction

India is an agricultural country in which agriculture employs the majority of the population. Agriculture aims to increase productivity and food quality while spending less and profiting more. A complex combination of soil, seed, and agrochemicals results in the agricultural production system. The most significant agricultural products are vegetables and fruits. A quality management control is essentially required in order to obtain more valuable products. Many studies show that plant diseases can reduce the quality of agricultural products. The normal state of the plant is impaired by diseases, which change or stop vital processes like photosynthesis, transpiration, pollination, fertilization, germination, etc. Pathogens, such as fungi, bacteria, and viruses, as well as poor environmental conditions, cause these diseases.

S. Nandan Mohanty et al. (Eds.): ICISML 2022, LNICST 471, pp. 181–192, 2023.
https://doi.org/10.1007/978-3-031-35081-8_15

As a result, diagnosing plant disease at an early stage is critical. This system is based on image processing technology and uses MATLAB as the primary processing tool. In addition, digital image processing, mathematical statistics, plant pathology, and other related fields are taken into account. In image segmentation and system construction, there are numerous innovations when compared to traditional image recognition. Users have a wide range of imaginative interactive options to meet their own needs in order to strengthen the division of the lesion. Meanwhile, linear regression models can be utilized.

Tomato problems are classified into two categories: bacteria or fungi causing 16 diseases and insects causing 5 other types of diseases. Bacterial wilt is caused by the bacteria Ralstonia solanacearum. This bacteria can survive in soil for a very long time and enter roots through wounds created naturally during the emergence of secondary roots, man-made during cultivation or transplanting, or even by insects. Disease development is favored by high humidity and temperature. By rapidly multiplying inside the plant's water conducting tissue, the bacteria fill it with slime. This has an effect on the plant's vascular system, while the leaves may remain green. A cross section of an infected plant stem appears brown with yellowish material coming from it. We have suggested a novel technique in the research study for identifying the disease in tomato crops after examining the leaf image data. The work will solve farmers' problems with plant disease identification without chasing down plant scientists. It will thus assist them in curing the plant's disease in a timely manner, increasing both the quality and quantity of food crops produced and thus contributing to an increase in farmer profit. We obtained the tomato leaves dataset from the village for the experiment. We developed a model to classify the images after downloading the dataset.

The model's performance was evaluated using a variety of parameters, including training accuracy, validation accuracy and loss, loss accuracy, and the number of trainable and number trainable parameters in reference to the pre-trained model.

2 Literature Survey

Even though many systems have been developed to date using various machine learning algorithms such as Random Forest, Naive Bayes, and Artificial Neural Network, the accuracy of those models is low, and the work using those classification techniques is done with the mindset of detecting disease for only one species of plant. Farmers continue to use their naked eyes to detect disease, which is a serious problem because the farmer has no idea what type of disease the plant is infected with. Farmers are still dealing with problems, and the methods they use to detect disease are tedious.

Machine learning algorithms are used in a variety of fields, but feature engineering remains the key issue. With the advent of deep neural networks, promising results for plant pathology are now available without the need for laborious feature engineering. Deep neural networks design concerns classification accuracy significantly. This section describes various deep learning techniques used by plant disease researchers. Mohanty et al. [10, 11] used AlexNet [9] to train and classify previously unseen plant diseases. When testing under conditions that were different from those used for training, model accuracy was significantly decreased. Disease can appear on the upper or lower sides of the leaves. Rangarajan et al. [14] trained both AlexNet and VGG19net using the hyper-parameters minimum batch size, weight, and bias learning rate. In the instance of VGG19net, accuracy and minimum batch size are inversely associated. Convolution and pooling layers were combined and deployed as Inception V4 to the GoogleNet architecture for dimension reduction [16]. For fine-tuning, Too et al. added pre-trained weights from ImageNet to this architecture's average pooling layer of 8x8. Furthermore, DenseNets [6], which has 122 layers, is also optimized for the identification of plant diseases. This paper describes a colour image analysis-based method for identifying the visual signs of plant diseases. The RGB image of the sick plant or leaf is first transformed into the H, I3a, and I3b colour transformations as part of the processing procedure.

The I3a and I3b transformations were created as a result of altering the original I1I2I3 colour transformation to satisfy the needs of the plant disease data collection. The segmented image is then determined by looking at the intensity distribution in a histogram. This method is very helpful when the target in the picture data set has a wide range of brightness. The extracted region was post-processed in tests after the image had been segmented to get rid of pixel regions that weren't taken into consideration. In order to capture the disease-related component of the image and extract pertinent disease-related information, image processing techniques are used.

On the other hand, data mining techniques are used to extract pertinent hidden information useful for illness identification based on the derived features [1, 2]. Image processing offers more effective approaches to identify fungus, bacterium, or virus-related plant illnesses. The ability to diagnose diseases through simple eye observations is insufficient. Pesticide overuse, like improper washing, makes people more susceptible to dangerous chronic diseases. The quality of the nutrients that plants receive is also harmed by excessive use. For farmers, it means a significant loss in production. The use of image processing techniques to identify and categorize diseases in agricultural applications is therefore beneficial [3].

Our suggested article comprises a number of implementation phases, including dataset preparation, feature extraction, classifier training, and classification. To distinguish between healthy and unhealthy photos, the produced datasets of diseased and healthy leaves are combined and trained under Random Forest. The Histogram of an Oriented Gradient is used to extract features from an image (HOG). Overall, utilizing machine learning to train the sizable publically accessible data sets gives us a clear technique to detect the disease existing in plants on a massive scale [4]. The goal of this research is to create a computer programme that can mechanically locate and categorize diseases. A stage that involves disease detection is loading a picture and doing pre-processing operations including segmentation, extraction, and classification. For identifying plant diseases, leaf images are used. The majority of tomato plant diseases are discovered in the early stages because they first harm the leaves. Early disease detection on leaves will undoubtedly save impending loss. In this study, the Multi-class SVM algorithm and image segmentation are used to identify four major diseases. Image segmentation is employed to divide up damaged leaf tissue, and the multi-class SVM method is used to accurately classify diseases. Users are advised to seek treatment for disorders that have advanced to their last stages [5, 7]. Plant diseases can be found via image processing. Image acquisition, picture pre-processing, image segmentation, feature extraction, and classification are processes in the disease detection process. The techniques for identifying plant diseases using photographs of their leaves were covered in this essay. This article also included some segmentation and the feature extraction algorithm employed in the identification of plant diseases [8].

3 Proposed Model Architecture

3.1 System Architecture

Here the disease image is loaded into the model engine. Then the model engine will take all those data according to the disease image our model is used to evaluate the disease name. Some normal images are taken and those images are converted into a colored image using skimage. color so that the disease of that plant can be easily identified. VGG19 is used for creating our model (Fig. 1).

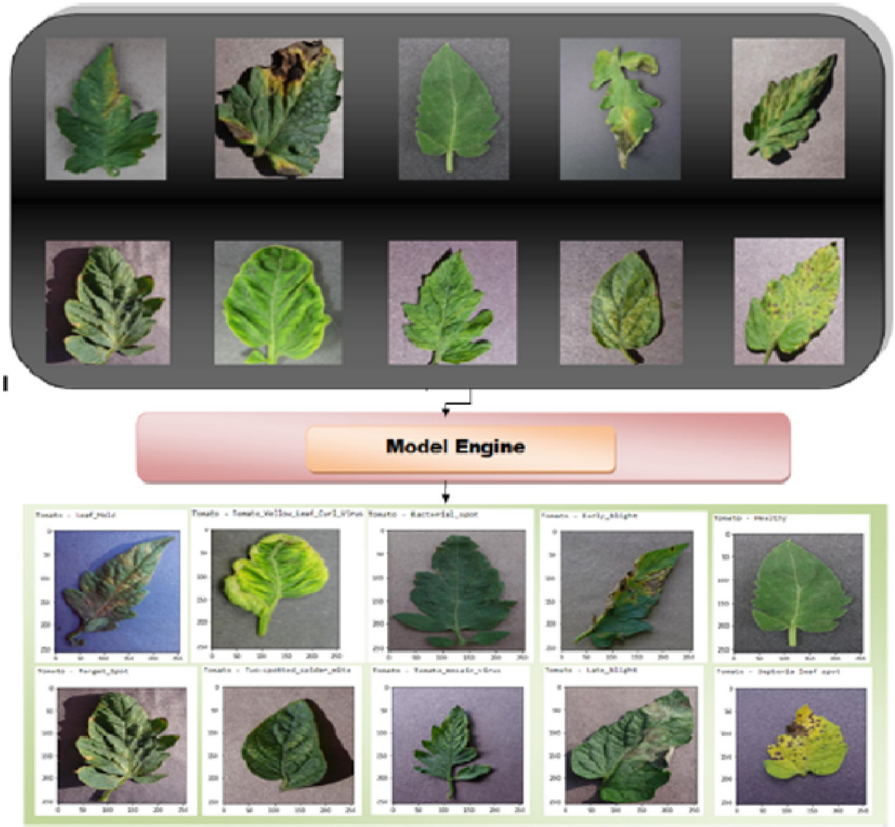

Fig. 1. System Architecture

3.2 Model Engine Architecture

Here all those disease images are taken and datasets are created. First one is the training data set and the second one is the validating data set. Then training data set as well as the all the dataset is loaded using machine learning and deep learning tools. One model is created that is called Tomato base model.h5. H5 is the extension of this model and tomato base model is the name of our model. Then all our data sets are loaded to train our model for prediction of disease.

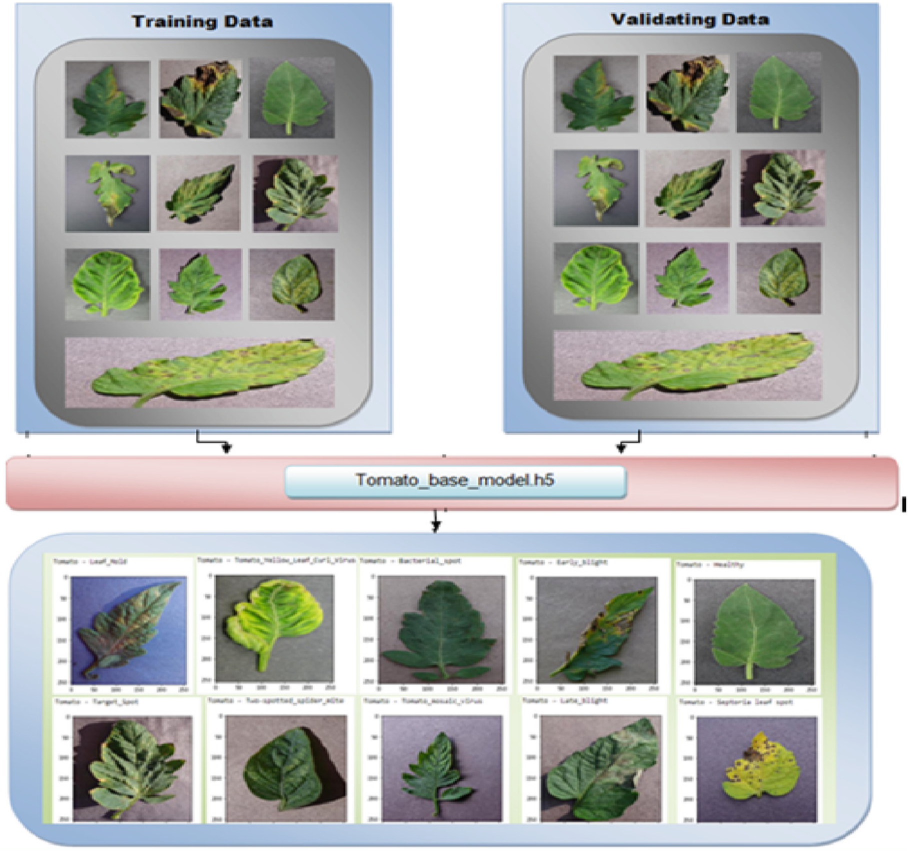

Fig. 2. Model-Engine Architecture

Figure 2 first, all diseases photos and data sets have been compiled here. The training data set is the first, and the validation data set is the second. Then, using machine learning, both our training data set and the val data set were loaded. The Tomato base model.h5 model is the only one that has been produced. Our model's extension is H5, and its full name is Tomato base model. Then, in order to train our model for the prediction of plant diseases, we must load all of our data sets.

After our model has been successfully trained, it is tested. Images of diseases are loaded in that test. These photos were captured by our model, and they display the names of the diseases they depict.

The process involves the previous:

1. Acquiring images
2. Image Preparation
3. Segmentation
4. Features extraction
5. Classification

3.2.1 Acquiring Images

This process involves feeding our software images of plant leaves that will be examined for disease. As it is simpler to execute classification algorithms on 2-D black-and-white images, the images are transformed to grayscale versions. The system will retrieve the plant snapshot and load the image at this step. The actions that come after acquiring the image include image in JPG format. For visual analysis, higher standard resolutions will be used, and JPEG is the format in which these data are saved.

3.2.2 Image Preparation

The unsharp filter is a basic sharpening operator that gets its name from the fact that it actually strengthens edges (and other maximum frequency elements in an image) by removing the unsharp, or smoothed, versions of the original image.

3.2.3 Segmentation

Image segmentation is the process of breaking a digital image up into several pieces (sets of pixels, often known as super pixels). The output of picture segmentation is either a group of outlines or fragments that collectively protect the entire image. Each pixel in a zone is in close proximity to a few distinguishing or established characteristics including colour, shape, and texture.

3.2.4 Features Extraction

Solidity, extension, minor axis length, and eccentricity are the shape features that were extracted for this study. These characteristics are used to identify the sick area of the leaf under consideration. Extraction of texture features uses three types of texture features: contrast, correlation, and energy. These characteristics are used to identify the sick area of the leaf under consideration. The variation of pixels and their neighboring pixels will then be determined. Feature extraction for colours has a special manner of displaying image representation when it comes to translation, scaling, and rotation. Mean, skewness, and kurtosis are the characteristics utilized for color. Here, RGB is converted to LAB.

3.2.5 Classification

Data is divided into three genres: training sets, testing sets, and valuation sets for identification. For each instance or piece of data in the training set, there is one target value and a number of characteristics. We will train our model using training sets and validation sets. To test our model, we use testing sets. Which I have already described in diagrams of system architecture or model architecture.

4 Algorithm of the Model

Ø Step 1:

Take all the disease images data and process all the disease images data. Create two separate data one is train data another one is val data. Then load all disease images data to the train data and val data using Image Dara Generator.

Ø Step 2:

Using matplotlib.pyplot for preprocessing train data. Then try to view some skimage. color images (Fig. 3).

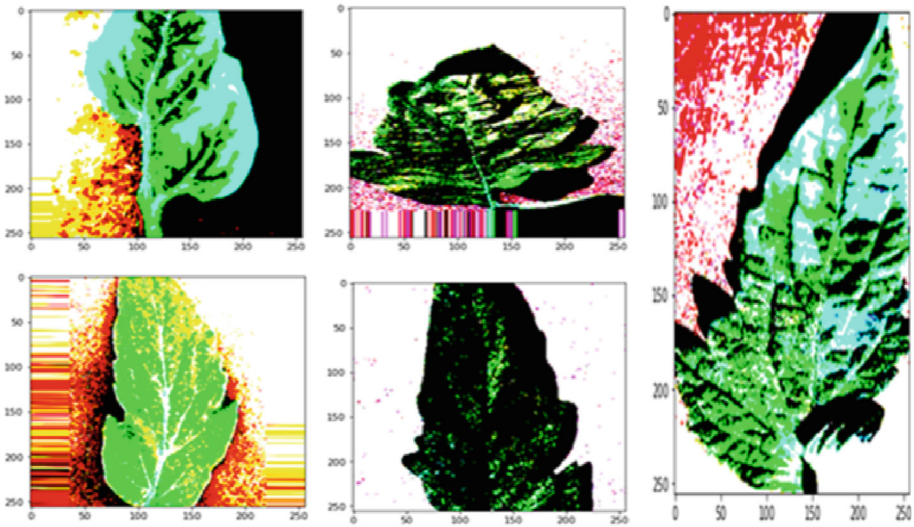

Fig. 3. Preprocessed images

Ø Step 3:

For printing the normalized data only it is divided by 255. Then try to view some skimage. color images (Fig. 4).

Ø Step 4:

Then create one model using keras.applications.vgg19. In Vgg19 pass input shape and include the top. Then print the created model summary.

Ø Step 5:

In the model summary add two more layers that are Flatten and Dense. In Dense use "Softmax" activation in the model.

Ø Step 6:

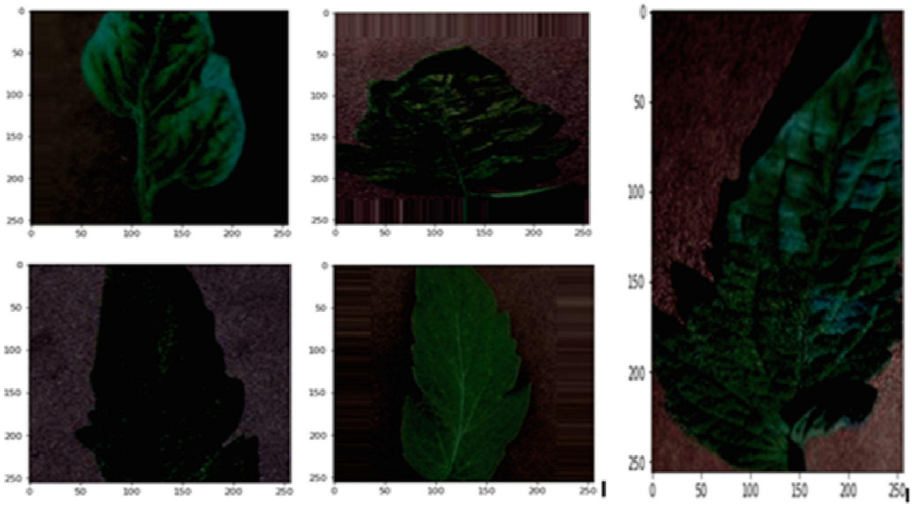

Fig. 4. Normalized images

Then compile the model using keras.losses.categorical_crossentropy. In this use adam for optimization and matrices are used for accuracy.

Ø Step 7:

Before training the model first do Early Stopping and Model Check Point using keras.callbacks. Here Early Stopping and Model Check Point is used to find val_accuracy of the model.

Ø Step 8:

Then train the model using model.fit generator. In this pass our train disease images data set, epochs, steps of epochs, verbose, callbacks and our validating disease images data set, is called validation data.

Ø Step 9:

Then find the accuracy of the model as well as validation accuracy of the model in a graphical way. Also evaluate the loss and validation loss of the model in a graphical way using MATPLOTLIB PYPLOT.

Ø Step 10:

Then load the model and try to find the final accuracy of our created model. In this scenario final accuracy is evaluated as 96.92307829856873.

Ø Step 11:

Now the model is ready to detect the disease of the plant. For testing purpose, load disease plant image then disease name and pictures of the loaded disease plant image is displayed (Table 1).

Table. 1. Comparative Study

Plant Disease Detection using ANN	Plant Disease Detection using CNN	Plant Disease Detection using Image Processing	Plant Disease Detection using Proposed approach
1. Proposed a "Classification of Pomegranate Diseases Based on Back Propagation Neural Network" which mainly works on the method of Segment the defected area and color and texture are used as the features 2. Here they used neural network classifier for the classification proposed a "Classification of omegranate Diseases Based on Back Propagation Neural Network" which mainly works on the method of Segment the defected area and color and texture are used as the features 3. Using image segmentation, the background is removed from the image. The KNN, ANN, and SVM methods are used to carry out the classification process. Using the closest distance between trained and tested subjects, KNN classifies samples [15, 17].Varun et al. developed models for the extraction noise removal and features extraction. So after that, a multiclass SVM classifier is used. For segmentation, L*A*B* colour spaces are used, which is based on a collection of marks produced by analysing the palette and luminescence components of various sections of the image. The GLCM is applied to extract features. Vijai Singh et al. investigated plant leaf samples taken with a digital camera, including rose/beans (bacterial disorder), lemon (sun burn disorder), banana (early scorch), and beans (fungal).Finally, the genetic algorithm is implemented to obtain the segmentation results. The colour co-occurrence is altered in order to extract useful features from segmented images. The Minimum Distance Criterion is used first for identification, abided by the SVM classifier	1. The logical next step in automating the identification of plant species, as well as solving many other machine learning challenges, was to eliminate the need for explicit feature selection entirely 2. Deep learning CNNs have recently experienced a significant breakthrough in computer vision because of the accessibility of effective and hugely parallel computing on graphics processing units (GPUs) and the huge amounts of image data required for training deep CNNs with millions of parameters 3. With the exception of prototype and design approaches, CNNs do not require direct and rigorous feature extraction and detection stages. Instead, both are incorporated into an iterative training procedure that incredibly quickly finds an image representation that is statistically appropriate for a given issue 4. Nature is compositional in a similar fashion; small units assemble to create large ones, and each level of clustering increases the structure's diversification. Whether they include or exclude manually created features, such hierarchical representations produce classification results that are largely unattainable using shallow learning techniques	1. Despite the fact that machine learning methods have many applications, classification algorithm remains the most difficult challenge. Deep neural networks have empowered promising plant pathology solutions to be built without the need for time-consuming feature extraction 2. Deep neural networks significantly improve image classification accuracy. This section describes the various deep learning methods used by scientists to identify plant diseases 3. Mohanty et al. [10, 11] used AlexNet [9] to train and characterise completely undiscovered plant diseases. Prediction accuracy was drastically decreased when testing image conditions varied widely from training image conditions. Disease is rarely visible on the top and bottom sides of the leaves. Rangarajan et al. [14] trained both AlexNet and VGG19net using the hyper-parameters of least sample size, weight, and bias learning rate. Accuracy and minimal level batch size are inversely related in the case of VGG19net For data preprocessing, pooling layers and convolution layer were blended and distributed as Inception V4 to the GoogleNet architecture [16]. Too et al. fine-tuned this architecture's average pooling layer of 8x8 by adding pre-trained weights from Image Net	1. In this proposed work, image processing is used along with VGG 19 model. Using this model, images can be uploaded and trained very fast and efficiently 2. This model increases image classification process. It will be able to identify diseases that appear in different sections of the leaf 3. Accuracy of the model has been increased

5 Results and Discussion

The accuracy of the model as well as validation accuracy of the model in a graphical way and it is shown in Fig. 5

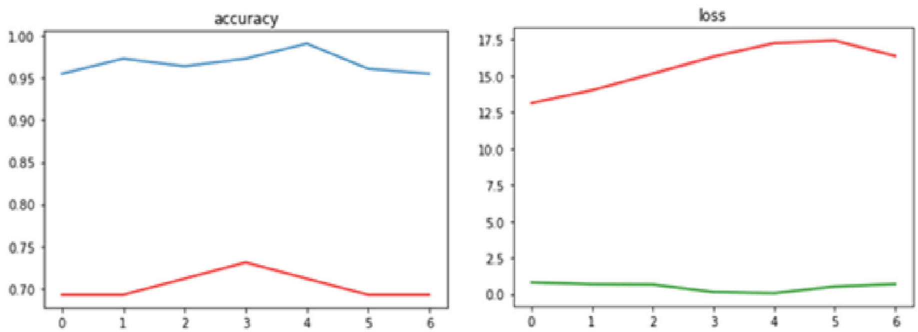

Fig. 5. Accuracy and loss of the model

The model loaded and the final accuracy of our created model is obtained. Now the model is ready to detect the disease of the plant. The disease plant image is loaded for testing and then disease name and pictures of the loaded disease plant image is displayed. It is shown in Fig. 6.

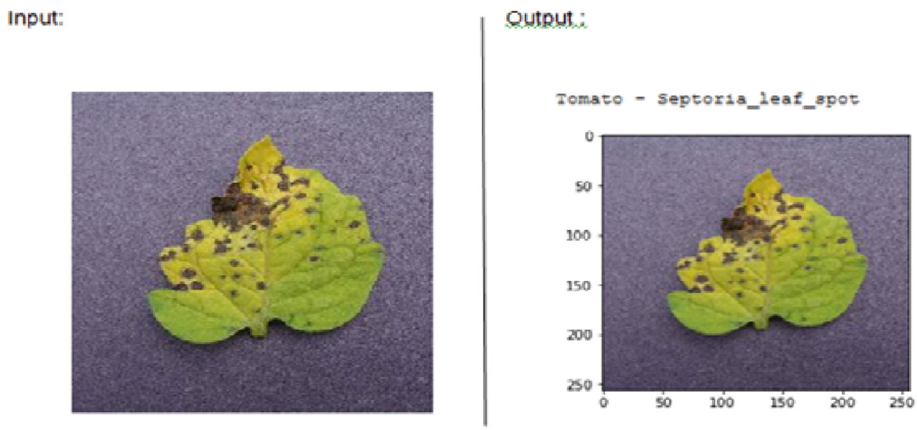

Fig. 6. Plant diseased image with disease name

6 Conclusion

Through the use of image processing techniques, this work provides effective and precise methods for classifying and detecting plant diseases. The manpower cost and detection time are decreased by this automated method. The farmers can use it to diagnose the illness and take appropriate corrective action. This database will expand in subsequent work to identify more leaf diseases. We have developed a model that is completely original from all others since it makes it simple to examine all varieties of plant diseases. Also system software is developed that allows users to enter an image of their sickness and quickly obtain both the name of the ailment and a cure. So that users' problems can be easily solved and they can produce different good products. Although there are many different versions on the market, this one is superior to all the others. Because numerous other plants, including potato, tomato, banana, and orange ones, have also undergone testing, it gives the disease an image of a disease plant.

References

1. Mohanty, S.P., Hughes, D.P., Salathé, M.: Using deep learning for image-based plant disease detection. Front. Plant Sci. **7**, 14199 (2016)
2. Rangarajan, A.K., Purushothaman, R., Ramesh, A.: Tomato crop disease classification using pre-trained deep learning algorithm. Procedia Comput. Sci. **133**, 1040–1047 (2018)
3. Too, J., Abdullah, A.R., Mohd Saad, N.: A new co-evolution binary particle swarm optimization with multiple inertia weight strategy for feature selection. Informatics **6**(2), 21. MDPI (2019)
4. Barbedo, J.G.A.: A review on the main challenges in automatic plant disease identification based on visible range images. Biosys. Eng. **144**, 52–60 (2016)
5. Ferentinos, K.P.: Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. **145**, 311–318 (2018)
6. Chen, J., Chen, J., Zhang, D., Sun, Y., Nanehkaran, Y.A.: Using deep transfer learning for image-based plant disease identification. Comput. Electron. Agric. **173**, 105393 (2020)
7. Madiwalar, S.C., Wyawahare, M.V.: Plant disease identification: a comparative study. In: 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), pp. 13–18. IEEE (2017)
8. Martinelli, F., et al.: Advanced methods of plant disease detection. A review. Agron. Sustain. Dev. **35**(1), 1–25 (2015)
9. Wallelign, S., Polceanu, M., Buche, C.: Soybean plant disease identification using convolutional neural network. In: FLAIRS Conference, pp. 146–151 (2018)
10. Shirahatti, J., Patil, R., Akulwar, P.: A survey paper on plant disease identification using machine learning approach. In: International Conference on Communication and Electronics Systems (ICCES), pp. 1171–1174. IEEE (2018)
11. Nagasubramanian, K., Jones, S., Singh, A.K., Sarkar, S., Singh, A., Ganapathysubramanian, B.: Plant disease identification using explainable 3D deep learning on hyperspectral images. Plant Methods **15**(1), 1–10 (2019)
12. UR Rahman, H., et al.: A comparative analysis of machine learning approaches for plant disease identification. Adv. Life Sci. **4**(4), 120–126 (2017)
13. Kothari, J.D.: Plant disease identification using artificial intelligence: machine learning approach. Int. J. Innov. Res. Comput. Commun. Eng. **7**(11), 11082–11085 (2018)
14. Vamsidhar, E., Rani, P.J., Babu, K.R.: Plant disease identification and classification using image processing. Int. J. Eng. Adv. Technol. **8**(3), 442–446 (2019)
15. Pertot, I., Kuflik, T., Gordon, I., Freeman, S., Elad, Y.: Identificator: a web-based tool for visual plant disease identification, a proof of concept with a case study on strawberry. Comput. Electron. Agric. **84**, 144–154 (2012)

16. Guo, Y., et al.: Plant disease identification based on deep learning algorithm in smart farming. Discrete Dyn. Nat. Soc. (2020)
17. Kiani, E., Mamedov, T.: Identification of plant disease infection using soft-computing: application to modern botany. Procedia Comput. Sci. **120**, 893–900 (2017)