International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 07 | July 2022 www.irjet.net p-ISSN: 2395-0072

Plant Disease Detection and Severity Classification using Support Vector Machine and Convolutional Neural Networks

R V Sahiti Gupta[1], Sandhya H S[2], Geetha L S[3]

[1,2]*VIII semester, BE, Department of Computer Science and Engineering, BNMIT*
[3] *Assistant Professor, Department of Computer Science and Engineering, BNMIT, Bangalore, Karnataka, India*

---***---

Abstract - *Agriculture, is a well-known fact that it is the major source of food, income, and employment for rural populations. Various seasonal conditions, however, the crops get infected by many kinds of diseases. These diseases threaten farmers' income and food security. Identification of plant diseases by using the recent technologies can help farmers with further diagnosis and treatment. This paper introduces the modeling of plant disease detection along with severity classification using Support Vector Machine (SVM) and Convolutional Neural Networks (CNN). The dataset contains images of four different plant species of various diseases. The accuracy obtained by this model varies with each disease of the plant which is around 80 to 90%.*

Key Words: Plant Disease, Severity Classification, Support Vector Machine, Convolutional Neural Networks

1. INTRODUCTION

Being the major portion of the population is from rural segments, Agriculture plays a major role as of important sector of the Indian Economy. The maintenance of its health, quality & quantity is of utmost concern for the country. The disease is a major concern in every country, as the demand for food is rapidly increasing due to an increase in the population

As per the latest estimates by the Food and Agriculture Organization of the United Nations (FAO), around 40% of the crops are lost worldwide every year due to plant diseases. The crops get infected by various kinds of diseases, due to various seasonal conditions, animate (pests and weeds), and inanimate (weather, rainfall, wind, moisture) matters. Identification of plant diseases can help farmers with further diagnosis and treatment which in turn increases their income and food security. The influence of disease plays one of the factors in the quality and quantity of agricultural production.

The recent developments in technology, efficient data storage, and powerful hardware provide an opportunity for image classification in agriculture. Studying visually observable patterns of plant leaves can help identification of folic diseases and monitor the health of plants. Thus, it provides a way to reduce loss in yield substantially and increase plant production.

In order to identify the different diseases of plants, various machine learning algorithms are significantly been used. Hence, this paper focuses on detecting plant diseases using Support Vector Machine and classifying their severity by deep learning models such as DenseNet and EfficientNet which helps to improve the accuracy. Knowing the severity of the disease makes the farmers take necessary mitigations at the earliest.

The rest of the paper is organized as follows: Section 2 provides the related works on the identification of plant diseases. Section 3 discusses the proposed methodology. Section 4 presents the experimental results and analysis. Finally, section 5 summarizes the model.

2. RELATED WORKS

This section describes various approaches for detecting the disease in plant leaves using different techniques.

The plant diseases are detected based on the images provided as input. It is known that there is noise in the images, that are used for training will cause incorrect results. In order to obtain a better efficacy, methods like background subtraction and segmentation algorithms were used that help in cleaning the noisy background images. This approach is proposed in paper [1] which outlines the various deep models which were tested and applied on image sets with different backgrounds.

For the classification of plant diseases, techniques like Deep Learning (DL) algorithms are significantly used by most researchers. The following papers provide a comparative study on using DL methods. Paper [2] explains DL models that were in from 2012 to 2018 to visualize plant diseases. It presents the current trends and challenges for the detection of plant leaf disease using advanced imaging techniques and deep learning along with the problems that need to be resolved. Similarly, the authors in the paper [3] give an insight into the evaluation of various deep learning architectures. The best-performing model was further fine-tuned by various optimization algorithms. This resulted that the Xception model with Adam optimizer has achieved the highest F1 score of 0.9978. Paper [4] that was referred here presented a survey on the exemplary comparison, frameworks, Convolutional Neural Network (CNN) models and optimization techniques to detect plant diseases using leaf images as a data set. It highlighted merits & demerits easing the task of developers while applying DL techniques.

Many researchers also used Convolution Neural Networks in their model which gave a higher accuracy. Paper [5] deals with the usage of minimal computing resources over traditional models for better results such as Neutral network models by using feature extraction achieved an average accuracy of 94.8% indicating its efficacy even under unfavorable situations. In paper [6], the authors proposed a model which is based on the inception layer and residual connection. The technique was tested on 3 plant disease datasets resulting in accuracy percentage as appended, plant village is 99.39%, Rice is 99.66% and Casava is 76.59 %.

Other techniques include SVM, Artificial Neural Networks and GAN network. Paper [7] highlights the usage of Wasserstein generative adversarial network with gradient penalty (WGAN-GP) combined with Label Smoothing Regularization (LSR) which addressed the overfitting issues due to limiting training data. The model fine tunes the classification accuracy by 24.4% when compared to 22% using synthetic samples without LSR and 20.2% of classic data augmentation. Paper [8] dealt with an image processing framework for plant disease identification and classification which consists of 3 stages, image segmentation, feature extraction and classification. The tests were conducted on 4 different classes of Tomato leaf diseases by using GLCM, multithreshold, etc. This method achieved 98.3% overall accuracy with 10- fold cross-validation.

The identified disease of a leaf image should be shown to the users so that they can take necessary mitigations at the right time. There are various ways like a web interface, mobile application, or messaging tool developed for this purpose. Paper [9] focuses on providing users the name of the disease detected and directs them to an e-commerce website where they can purchase the medicine for the diseases and use them appropriately according to the directions given. In paper [10], two crops i.e., Corn and Tomato are taken for disease identification using SVM and ANN, and the detected disease is sent to users through GCM. The accuracy obtained by SVM is 60-70% and by ANN 80-85%. In case of Corn, by using SVM is 70-75% and by ANN is 55-65%. Table 1 provides a summary of the related works in the literature.

Table -1: Summary of related works

Source	Dataset/Crops	Methods and Results
[1]	Plant Leaf	Fine-tuned Densenet121 (Removed Background); 93%
[2]	Plant Village	DL models
[3]	Plant Village	Xception with Adam; 99.7%
[4]	Plant Village	DCNN
[5]	Plant Village	VGG16; 94.8%
[6]	PlantVillage, Rice, Cassava	99% for plantvillage and rice, 76% for cassava
[7]	Plant Village	GAN
[8]	Tomato	SVM; 98.3%
[9]	Apple	CNN; around 70-80% for different diseases
[10]	Corn, Tomato	SVM, ANN; 60-80%

3. PROPOSED METHODOLOGY

In this section, the approach is defined for the detection of plant diseases. First, we describe the dataset, highlight the segmentation steps, and discuss the methods for training the learning models. Figure 1 shows the data flow diagram of the proposed model.

3.1 Dataset Description

The proposed experiment utilized the Plant Village Dataset available on Kaggle which contains 20,639 images of high resolution of 38 different healthy and diseased leaves pertaining to 14 plant species. For the implementation of this model, segmented images of 4 crops with their diseases are considered. All the diseases of a crop are grouped into a single folder and stored in Google Drive plant-wise as shown in table 2.

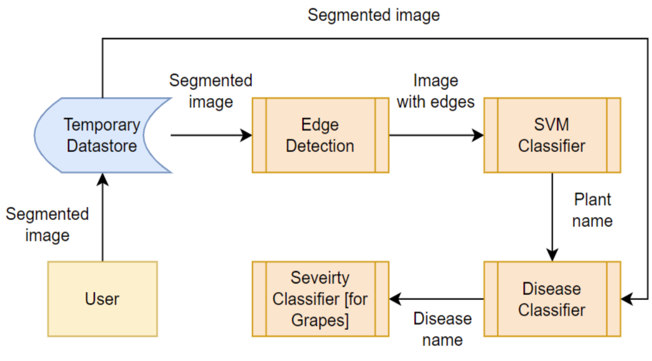

Fig – 1: Data Flow Diagram

3.2 Image Segmentation

Image segmentation is an important part of image processing. For the segmentation of images, various methods like region and edge-based methods, boundary and spot detection algorithm, K-means clustering, Otsu's method, etc. are available. Canny edge detection is one of the optimal edge detectors as it provides a good, reliable, and lowest error rate in real edge point detection. It is a technique that detects a wide range of edges in images using a multi-stage algorithm. The steps followed for edge detection using the canny edge detector are:

1. Smoothing: Gaussian filter is applied to smooth the image in order to reduce noise.

2. Finding intensity gradients: The edges are marked where the gradients of the image are having large magnitudes. The gradients of the images of large magnitudes are highlighted as edges.

3. Non-maximum suppression: This is applied to get rid of spurious responses to edge detection. The edges marked are local maxima.

4. Double Threshold: Thresholding is taken as a criterion for determining potential and actual edges.

5. Edge tracking by hysteresis: Weak edges that are connected to strong edges are actual or real edges and those that are not connected to strong edges will be suppressed.

Table - 2: Dataset Specifications

Plant	Disease Name	Count
Apple	Healthy	1645
	Apple Scab	630
	Black Rot	621
Corn	Healthy	1162
	Common Rust	1192
	Gray leaf spot	513
Grapes	Healthy	423
	Black rot	1180
	Esca	1383
	Isariopsis	1076
Tomato	Healthy	1591
	Leaf mold	952
	Septoria	1771
	Early blight	1000

3.3 Classification

A. Support Vector Machine algorithm

For the identification of plant diseases, the SVM algorithm is used. An SVM is a supervised learning and vector space-based machine learning method where the goal is to find a maximum marginal hyperplane (MMH) that divides the training data into classes. This algorithm helps to analyse data used for clustering, classification and regression analysis. The following steps are used to search for the maximum marginal hyperplane:

1. Hyperplanes are recursively generated to segregate the classes in the best way.

2. The next step is to select the hyperplane with the maximum segregation from both nearest data points for accurate results.

B. Convolutional Neural Networks

Convolutional Neural Networks (CNN) is a class of Artificial Neural Networks (ANN) that works on the concept of hidden layers and is most widely used in image processing and recognition. The different methods used in this model are DenseNet and EfficientNet. They are done by using activation functions and an optimizer.

DenseNet: A DenseNet is one of the CNN that uses dense connections between layers where they are connected with each other, through Dense Blocks. This approach is suitable if there are similar kinds of plant images. For example, the leaves of Tomato and Apple.

EfficientNet: EfficientNet is a CNN architecture that uses a compound coefficient technique that uniformly scales each dimension with a certain fixed set of scaling coefficients. It achieves better efficiency and higher accuracy.

Activation functions: In neural networks, the activation function is responsible for transforming the weighted sum of the input into an output from nodes in a layer of the network. This model uses ReLU (Rectified Linear Unit) as the activation function in the hidden layer, that outputs the input directly if it is positive, otherwise, it will output zero:

$$f(x) = \max(0, x) \quad (1)$$

In the last layer, the SoftMax function is used, which returns values between 0 and 1 and determines the probabilities of data related to the class used for the multiclass problems:

$$\pi(z)j = \frac{e^{zj}}{\sum_{k=1}^{K} e^{zj}}, \quad \text{for } j = 1, 2, 3, 4, \ldots k. \quad (2)$$

Optimizers: Optimizers are methods used to adjust the parameters such as weights, and learning rate for a model to reduce losses. One of the optimizers used in this model is Adam Optimizer. This algorithm accelerates the gradient descent algorithm by considering the exponentially weighted average of the gradients.

4. RESULTS

The proposed model is divided into three modules namely, identification of the plant, classification of disease, and classification of severity. In the plant classifier module, a segmented image is given as the input which first detects the edges in the leaf image using Canny Edge Detection method. Then the image is processed through SVM classifier for identification of plant. Deep learning methods are used in disease classifier module and severity classifier module to provide better performance and accurate results. Figures 2, 3, 4 and 5 display the output of the model.

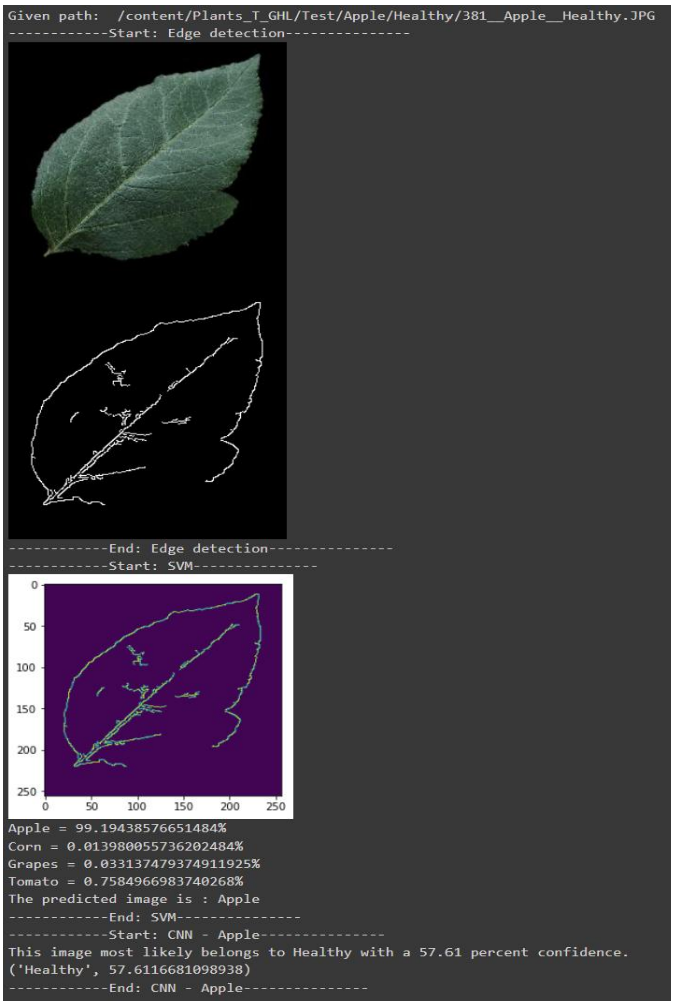

Fig – 2: Output of Apple disease detection

Taking the various images of plant disease datasets from Kaggle namely, apple, corn, tomato and grapes, the test scores obtained for the disease and its severity classifier are shown in table 3.

Table – 3: Test scores of plant disease classification

Plant	Disease Name		Accuracy (%)
Apple	Healthy		92.5
	Apple Scab		100
	Black Rot		100
Corn	Healthy		100
	Common Rust		100
	Gray leaf spot		57.5
Grapes	Healthy		90
	Black rot	High	100
		Low	85
	Esca	High	90
		Low	100
	Isariopsis	High	95
		Low	100
Tomato	Healthy		100
	Leaf mold		100
	Septoria		100
	Early blight		95

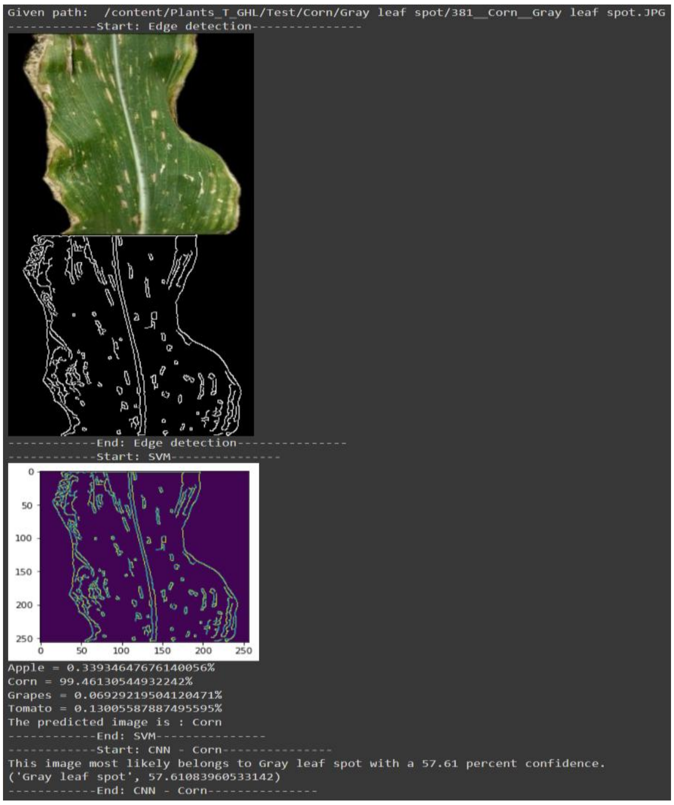

Fig – 3: Output of Corn disease detection

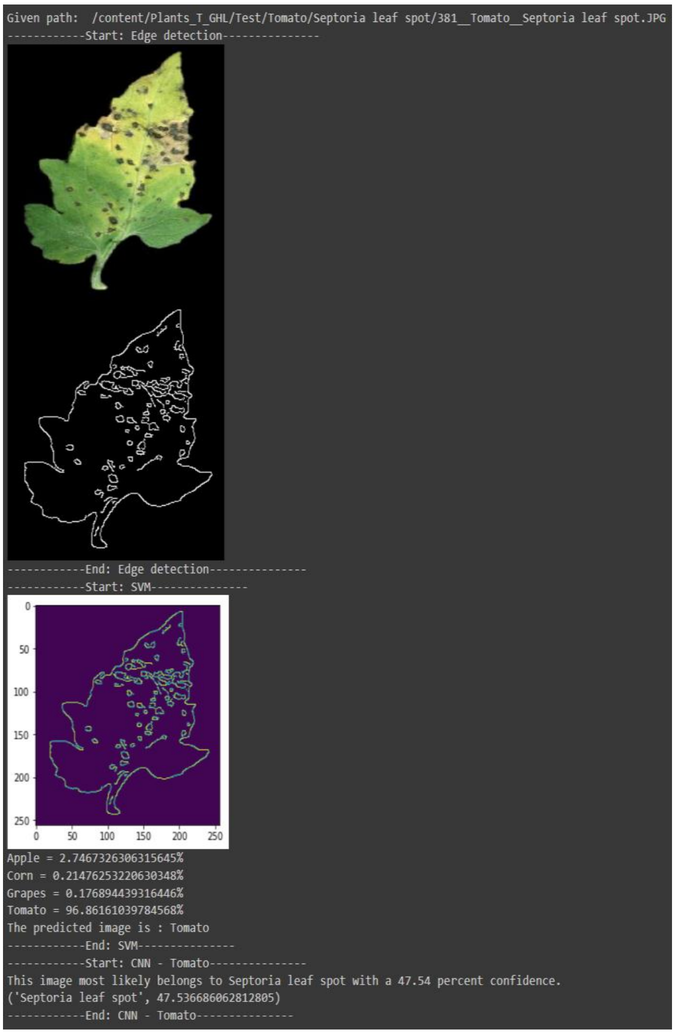

Fig – 4: Output of Tomato disease detection

5. CONCLUSION AND FUTURE SCOPE

The primary objective of the model "Plant Disease Detection and Severity Classification" is to detect plant diseases with severity using images of diseased plant species. This model highlights the identification of severity of disease through image processing and extracts information which helps in the classification task and supports disease detection in plants. The model is tested on four plants namely, Apple, Corn, Grapes, and Tomato having two to three diseases each. The technologies used here are SVM, CNN methods like DenseNet and EfficientNet that predict the diseases and their severity early and in a simple manner to make the growers take disease measures at right time. The experimental results indicate that both approaches significantly improve the accuracy of leaf diseases.

The future research work includes the following:

- The use of other algorithms can be explored to enable them to detect and classify diseases during their complete cycle of occurrence and enhance the efficiency of the system.
- Development of a user-friendly interface or mobile application which is handy for the growers to carry in their pockets which may prove to be a great asset to the agricultural sector.

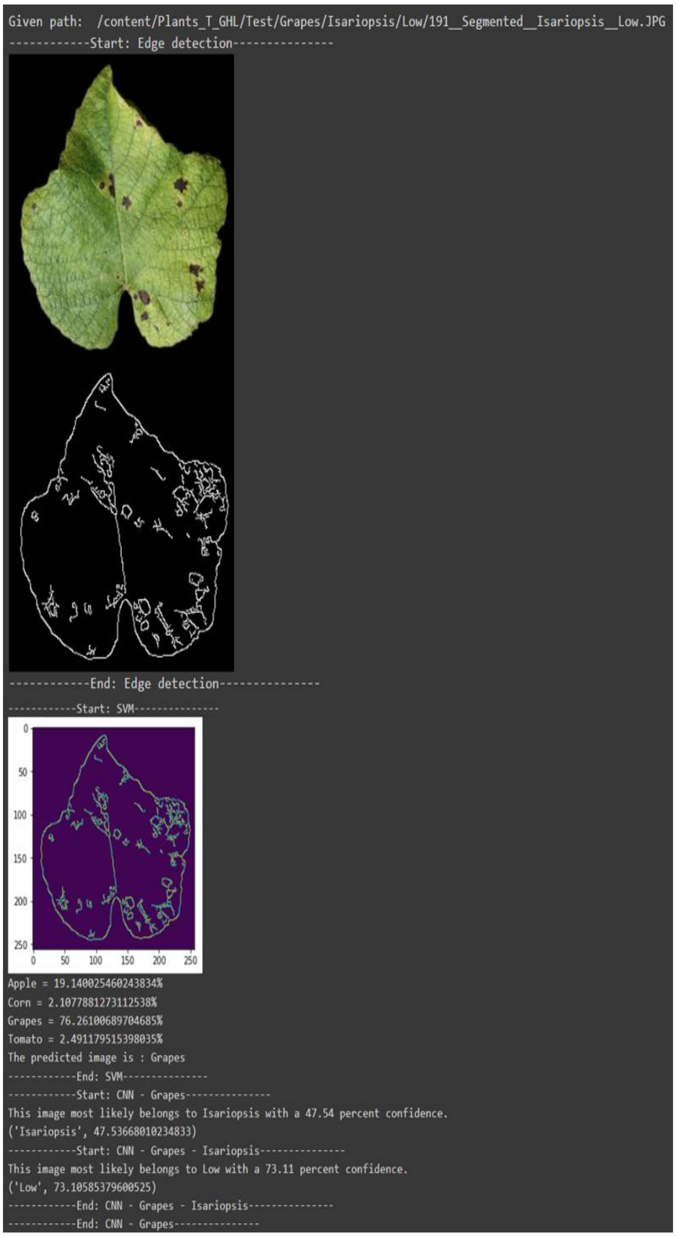

Fig – 5: Output of Grapes disease detection with the severity level

REFERENCES

[1] KC, K.; Yin, Z.; Li, D.; Wu, Z., "Impacts of Background Removal on Convolutional Neural Networks for Plant Disease Classification In-Situ", MDPI Plants, Agriculture 2021, 11, 827

[2] L. Li, S. Zhang and B. Wang, "Plant Disease Detection and Classification by Deep Learning—A Review," in IEEE Access, vol. 9, pp. 56683-56698, 2021

[3] Saleem, M.H.; Potgieter, J.; Arif, K.M. Plant Disease Classification: A Comparative Evaluation of Convolutional Neural Networks and Deep Learning Optimizers. Plants 2020, *9*, 1319

[4] Dhaka, V.S.; Meena, S.V.; Rani, G.; Sinwar, D.; K.; Ijaz, M.F.; Wo´zniak, M. "A Survey of Deep Convolutional Neural Networks Applied for Prediction of Plant Leaf Diseases", MDPI Sensors 2021, 21, 4749

[5] Rinu R, Manjula S H, "Plant Disease Detection and Classification using CNN", IJRTE, Volume-10 Issue-3, 2021

[6] S. M. Hassan and A. K. Maji, "Plant Disease Identification Using a Novel Convolutional Neural Network," in IEEE Access, vol. 10, pp. 5390-5401, 2022

[7] Bi L and Guiping Hu, "Improving Image-Based Plant Disease Classification With Generative Adversarial Network Under Limited Training Set", Front. Plant Sci. 11:583438, 2020

[8] M. Z. Din, S. M. Adnan, W. Ahmad, S. Aziz, J. Rashid, W. Ismail, M. J. Iqbal,"Classification of Disease in Tomato Plants' Leaf Using Image Segmentation and SVM", Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan Vol. 23 No. 2-2018

[9] Plant Disease Detection using Image Processing, Mr. V Suresh, D Gopinath, M Hemavarthini, K Jayanthan, Mohana Krishnan, IJERT, ISSN: 2278-0181, Vol. 9 Issue 03, March-2020

[10] Plant Disease Identification Using SVM and ANN Algorithms, N. Kanaka Durga, G. Anuradha, IJRTE, ISSN: 2277-3878, Volume-7, Issue-5S4, February 2019