An Approach of Tomato Leaf Disease Detection Based on SVM Classifier

J. Keerthi, Suman Maloji, P. Gopi Krishna

Abstract: Disease Identification and management is a challenging task. Diseases of plants are commonly seen on the leaves of the plant. Precise identification of the disease by visually observing them is difficult because of the complexity of the patterns on the leaf. So the demand for identifying the diseases using computers has raised more in recent years. This work employs a machine learning technique to identify the diseases of a tomato plant and suggest appropriate control measures to handle the disease. The system is designed using python software programmed into raspberry pi modules. After the image is uploaded for the disease identification, images are pre-processed using histogram equalization, filtering, color transformation and segmentation then the images are taken to the classification using SVM classifier and the appropriate disease identified is displayed on the screen along with the corresponding control measures to be taken.

Index Terms: Precision Farming, Machine Learning, Plant diseases, Database, Feature Extraction, Support Vector Machine.

I. INTRODUCTION

Identification of the plant disease is one of the major issues to overwhelm the production losses and to increase crop growth. The basic idea is to monitor the Plant and their condition throughout the reap. At present, the identification of plant disease mostly depends on the manual recognition and this can cause some problems like farmers might mistakenly identify a disease because they judge them by their past experiences. Compared to a person’s vision, computer image processing is way better because it can identify characteristics like speediness huge information and distinguish even small diversity which cannot be done by human vision. Therefore image processing can help farmers to identify the disease accurately and suggest pests to handle them.

The tomato is one of the major crop production all over the world. According to the annual report 2017-2018 Ministry of Agriculture [17], India is the second largest producer of vegetables after china. India occupies the third position in the production of potato and tomato in the world. The scientific name of the tomato is known as Solanum lycopersicum comes under the family of Solanaceae with the energy 17.69 Cal (per 100g).

The diseases in the tomato plant are mainly caused by
1. Fungi.
2. Bacteria.

I. Fungi:

The diseases caused by fungi are further classified into different types such as Septoria leaf spot, Early blight, Anthracnose, Fusarium wilt, late blight, Verticillium wilt etc.

a. Septoria leaf spot

Caused by the fungus Septoria lycopersici.

Fig1: Septoria Leaf Spot

This is the most common fungus disease in Tomatoes It first appears as small, water-soaked spots that soon become circular spots. These are spread to new leaves through the rains, infected leaves turn yellow and gradually fall off. Lower leaves are infected first and the disease progress upward. This disease can occur at any stage of the plant, but it is mostly seen when the plant begins to set fruit. This disease can be controlled by taking measures like water at the base of the plant in the morning rather than in the evening to make sure leaves are wet for a minimum amount of time, rotation of crops, using disease-resistant plants etc.

b. Early blight

Caused by the fungus Alternaria solani.

Fig2: Early Blight

Premature loss of lower leaves is the symptom of this disease. Brown to black spots appears on lower leaves. Like the Septoria leaf spot these can also occur at any stage of the plant. Wet weather favors rapid spread of early blight.
An Approach of Tomato Leaf Disease Detection Based on SVM Classifier

2. Bacteria

a. Bacterial spot

Caused by the bacterium Xanthomonas,

b. Bacterial speck

Caused by the bacterium Pseudomonas syringae,

3. Fungi

c. Anthracnose

Caused by the fungus Colletotrichum,

![Fig3: Anthracnose](image)

This is the most common fruit attacking disease of the tomato plant. Symptoms of this disease are fruits will have small circular spots and as they expand they develop dark spots on the fruit and the fruits may rot completely. Harvesting frequently and picking up all the ripe fruits at each harvest can control this disease.

d. Fusarium Wilt

Caused by the Fungus Fusarium Oxysporum,

![Fig4: Fusarium Wilt](image)

Leaf yellowing and wilting are the symptoms of this disease and this spread from the base of the stem to upwards and the wilted leaves drop prematurely. Using resistant varieties of the plant can control this disease.

e. Verticillium Wilt

Caused by the Fungus Verticillium wilt,

![Fig5: Verticillium Wilt](image)

This fungus disease is common in most of the plants like potato, strawberry, watermelon, tomato plants, and radish plants. Lower Leaves initially turn yellow and rapidly turn completely turn yellow and wither and drop off, an infected plant may survive but the yield is greatly reduced. Growing resistant varieties of the plant can handle this disease.

f. Late Blight

Caused by the fungus Phytophthora infestans,

![Fig6: Verticillium Wilt](image)

Late curse may taint either youthful or old leaves. It initially shows up as water-drenched territories that grow quickly, shaping sporadic, greenish dark imprints, giving the plant harmed appearance. Control measures are the same as Septoria leaf spot and in addition, avoiding rotating with potato can be helpful.

Fig6: Verticillium Wilt

![Fig7: Bacterial Spot](image)

The infected plants will have spots at leaves and stem in circular to irregular in shape and have a greasy feel but unlike the Septoria leaf spots do not develop grayish brown color and the fruits first appear as black and they become gray-brown. The control measures are the same as the Septoria leaf spot and using disease-resistant plants could help. Fixed copper spray can reduce the disease.

Fig7: Bacterial Spot

![Fig8: Bacterial Speck](image)

Symptoms are dark spots appear on the leaves surrounded by yellow halos. Specs can be seen in the fruits which are smaller than that of bacterial spot and these specs do not penetrate the fruit deeply. Planting only disease-free transplants and
destroying weeds around a tomato field or garden will reduce the survival of the bacteria.

c. **Bacterial Canker**

Caused by the *bacterium Clavibacter*.

Leaflets turn down brown at the edges and initially one side of the plant develop symptoms and these eventually spread. The spots on the fruits are white and slightly raised at first and turns brown as they become older. Control measures are the same as bacterial spot except that it will have minimal effect for copper sprays.

3. **Viruses**

The most common virus disease is *wilt virus*. Leaves stop growing and will have purplish discoloration. Planting only virus-free plants are the best way of managing viruses. There is no way to cure a virus, but removing the infected plant at the initial stage itself can avoid spreading of the disease.

**Current practices and challenges faced by farmers**

According to the surveys, farmers are well aware of the most common diseases and they seek the help of other farmers and pests dealers in case of doubt. Identifying if the plant is infected is easy, but identifying the type of disease that is infected is a challenging task for the farmers. Therefore the monitoring of the plants continuously by the experts is necessary, but it is expensive and a difficult task. Therefore the need for a quick automatic less expensive source for identifying the plant disease is very much needed.

**Image-based solutions**

Studies show that machine learning can be applied to detect the type of the disease infected. Employing machine learning methods such as artificial Neural Networks (ANNs), Decision Trees, K-means, k-nearest neighbors, and Support Vector Machines (SVMs) can be helpful in accurate detection.

II. **LITERATURE SURVEY**

 Priyanka G. Shinde, Ajay K. Shinde, Ajinkya A. Shinde, Borate S. P [1] have studied the detection and prevention of the plant diseases using a raspberry pi. K-means clustering algorithm is used for the image analysis, which automatically detects the symptoms by using an image. In the first step RGB images are captured using a camera then, color space transformation is formed and segmentation is done by using k-means clustering and then the masking of green pixels is done and the infected cluster is converted from RGB to HSI format and after that SGDM matrix are created for each pixel. The features are calculated for the pixel present inside the edge of the infected part of the leaf and then disease detection is done and the result is transferred to GSM module.

Halil Durmus, Ece Olcay Guneu, Murvet Korco have studied [2] ailment identification in plants and profound learning techniques were utilized to identify the tomato plant illnesses. They pointed that profound learning techniques ought to be kept running progressively on the robots. The engineering determination was the key issue for the usage. Two distinctive profound learning system models were tried AlexNet and after that SqueezeNet. Preparing and approval were done on the Nvidia Jetson TX1. Tomato leaf pictures from the PlantVillage dataset was utilized for the preparation. From the comparison AlexNet produces the accuracy of 0.9565 and model size is about 227.6 Mbyte with an inference time of ~150ms and SqueezeNet produces 0.943 accuracies with a model size of 2.9Mbyte and an inference time of ~50ms. Exactness results are the aftereffects of the Caffe tests. AlexNet performed marginally superior to SqueezeNet. The examination appears that the SqueezeNet display is very nearly multiple times littler than AlexNet. What's more, it is a decent contender for the portable profound learning grouping because of its light weight and low computational necessities.

Vyshnavi.G.K.P Sirpa.M.R Chandramoorthy.M, Padmapriya,B [3] have proposed a method on Hierarchical Clustering. As it takes fewer iterations for segmentation and is efficient. And it consumes less time for the identification of the disease and the classifier used is the support vector machine which is used to identify the type of disease. Feature extraction has depended upon the 1. The color of the leaf 2. The texture of the leaf and 3. The shape of the leaf.

S. Askraba, A. Paap, K. Alameh, Senior, IEEE, J. Rowe, and C. Mill operator [4] have considered that the unearthly reflectance-based plant separation sensor utilized in the specific herbicide showering frameworks is planned. Its dynamic execution was tentatively surveyed for the two plants. The unearthly reflectance based plant separation sensor which includes two 3-laser-diode mix modules with new laser drivers, two multi-spot shaft generators, a line examine picture sensor, and a control unit is created. The tasks are synchronized by a calculation executed in a microcontroller. At the point when an objective plant is identified a flag is produced which is utilized to trigger a shower unit or record arranges from a Differential Global Positioning System (DGPS) which gives a way to apply herbicide just to regions plagued by weeds.

Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang and Qiao-Liang Xiang [5] have utilized Probabilistic Neural Network (PNN) with picture and information preparing methods for leaf acknowledgment. They have executed a broadly useful robotized leaf acknowledgment for plant classification. Around 1800 leaves are prepared to arrange 32 sorts of plants with a precision more prominent than 90% for every sort of plant, 10 bits of leaves of every sort of plant from testing sets are utilized to test the exactness of the calculation and the normal exactness found was of 90.312%. The accuracy comparison between the other algorithm is shown where PNN gives 90%. Here 12 features are extracted and processed by PCA (Principal Component Analysis) to form the input vector of PNN. Experimental results indicate that the algorithm is
An Approach of Tomato Leaf Disease Detection Based on SVM Classifier

workable with an accuracy greater than 90% on 32 kinds of plants. Where, the algorithm is fast in execution, compared to other methods which are efficient in recognition and easy in implementation. Peifeng xu, gangshan wu, yijia guo, xiaoyin chen, hetong yang, rongbiao zhang [11] has studied a wheat leaf disease detection system based on embedded image recognition technology. The system utilized the ARM9 processor with the embedded Linux platform. The program is developed in a Qt integrated environment. It first captures the image by means of web camera which is connected to the ARM microcontroller through USB. Diagnosis and classification are done on the 30 images of wheat leaf rust by artificial intelligence and image processing. The results are displayed with the help of LCD. The results that are verified shown that the recognition rate reaches 96.2% and accuracy rate reaches 92.3% with the method of the image processing.

Mr. Girish Athanikar, Ms. Priti Badar [12] describes a neural network base detection and classification of Potato leaf samples using Segmentation of K-Means Clustering. Neural Network, one of the Machine Learning Techniques, is used to categorize the query leaf image as either healthy or diseased sample. A total of 150 images of potato leaf samples are used in the experiment. BPNN feed-forward Neural Network is used for training, learning, testing, detecting and classifying the leaf types. Implementation is carried out using the MATLAB. Studies prove that BPNN could effectively detect the disease spots and classify the particular disease type by mentioning the name of the disease, cause of disease and pesticides to be to prevent the diseases with an accuracy of 92%.

Prof Dr. D. A. Godse, Nalini tripathi, and Co-authors [13] developed a system which can operate by using a mobile application on Android phones. In order to detect the diseases of jute plants through image processing, farmers have to capture the image of the stems of their affected jute plant and send them to the system, server where the image will undergo several levels of processing to detect and identify the disease. Support Vector Machines (SVM) are used to classify the disease. The result shows the condition of the plant and any necessary suggestions to be taken to improve the plant growth.

Suneeta Budihal, Sandhya R. And Co-authors [14] develops a user-friendly automated system for the farmers that will help them in determining the diseases of tomato leaves without bringing an expert in the field. For these, an android application is developed that takes an image of the leaf via accelerometer sensor and detects the disease of the leaf using image processing techniques that produce a proper solution for the detected disease. The proposed algorithm was implemented by creating a JAVA library in android for leaf disease detection of tomato. According to Otsu's threshold method, the number of pixels is calculated and based on the calculated value it is identified whether the leaf is diseased with fungal screen or the leaf is diseased with a spot or not or the leaf is healthy.

Vidyaraj K, Priya S [15] proposed four structure modules in particular picture pre-handling. Here the RGB pictures of leaves are changed over into HSI shading space portrayal. At the preprocessing step tainted bit of the leaf is removed. From that point forward it is fragmented into various patches of equivalent size utilizing K-medoid bunching. For removing the list of capabilities shading co-event technique or CCM is utilized. The CCM surface investigation technique is created using spatial dark dimension reliance lattices (SGDM). From that removed capabilities are encouraged into the LS-SVM classifier. Least Squares Support Vector Machine (LS-SVM) is utilized for finding an ideal hyperplane, which isolates different classes where Kernel-based SVM are utilized for the LS-SVM classifier.

Shima Ramesh, Mr. Ramachandra Hebbar and co-creators [16] proposed whether the leaf is infected or solid, for that specific advances must be pursued, i.e., Preprocessing, Feature extraction, Training of classifier and Classification. Preprocessing of a picture conveys every one of the pictures size to a decreased uniform size. At that point comes extricating highlights of a preprocessed picture which is finished with the assistance of HOG(Histogram of Oriented Gradients) utilized for article discovery. The calculation was executed utilizing irregular woods classifier which is adaptable in nature and can be utilized for both arrangement and relapse strategies.

Using HOG the feature extraction is implemented. It considers shape, size, color, for extracting the features and it trained with the SVM classifier to get the accurate results.

III. PROPOSED SYSTEM

Fig10: Block Diagram

A database is created in the server. Depend upon the shape, color, feature is extracted using HOG (Histogram of Oriented Gradients) used in object detection and also it is trained by using the SVM classifier.

Whenever we have given an image to test it compared with the database and display the result with its control measures and sends back to the user through E-Mail.

IV. FLOW CHART

4.1 Collection of Database:

In the first step, the sample images are collected from the datasets of tomato using different digital cameras with different resolutions. Which are used to train the system the sample images are stored in the form of the JPG format. All
the sample images are in RGB (Red, Green, Blue) form. The obtained images include the healthy images and also diseased images like bacterial spot, tomato mosaic virus, etc. Various methods of pre-processing can be applied to the image to get better results.

4.2 Image Pre-processing:
In the second step, the image should be read, for that, we need to create the directory for both positive and negative images if it doesn't exist in our image dataset into a variable then we create a function to load folders containing images into arrays. And also the image contains some of the unwanted noise as well as redundancy. So pre-processing techniques are used. To remove the background noise and also to suppress the undesired distortion which is present in the image which occurs due to many reasons such as camera settings, variations in the light. To overcome these basic problems, The input RGB image is to be converted into a grayscale intensity image to provide accurate results.

4.3 Image Segmentation:
Image Segmentation is a technique of partitioning an image into the number of pixels with respect to their intensity levels. Assigning a label to every pixel in an image such that pixels with the same labels will share some characteristics. Masking of an image used for edge detection to increase the sharpness of an image. Thresholding operation is done with the pixels whose value is greater than the specified threshold value, are assigned with the standard value. Edge detection is a basic problem in image processing and machine visions. Therefore, Sobel edge detection is used in image segmentation. It calculates the gradient of image intensities at each pixel within the image.

4.4 Feature Extraction:
The process of extracting the relevant information from the input image and transferring data into a set of features with their labels is known as feature extraction. In the fourth step, based upon the features like color, size, shape, texture the effective features are extracted. The HOG (Histogram of oriented gradients) is a dense feature descriptor that simplifies the image by extracting the useful information by sorting out the extraneous information and it is widely used for image recognition, object detection. We can neglect lightening conditions by using HOG which is more accurate.

4.5 SVM Classifier:
The generated feature sets are aided into Linear SVM (Support vector machine) which is a supervised machine learning algorithm which can be used for classification and regression challenges. The classifier will assign the label to the image and it specifies which class it belongs to from where the classifier is pre-defined based upon the feature. This classification is used for both training and the testing phase. SVM uses the technique called the kernel.

4.6 Training:
The labeled datasets are set into training and testing steps. Using SVM classifier the dataset is being trained and after successful training. The training features are providing efficient results by analyzing the different number of features.

4.7 Testing:
In the testing stage, the tomato leaf images are pre-processed and then the infected regions are segmented based on edge detection algorithm by using a Sobel operator according to pre-defined methods in training steps. At that point the divided districts are encouraged into highlight extraction for extricating the distinctive kind of highlight. The removed element is coordinated with the preparation dataset for identifying the infection present in the tomato leaf picture dataset.

4.8 Email:
The obtained result will be sent to the user's or farmers e-mail id to know the details of their leaf and to get alerted if it is diseased and then according to the disease the control measures should be done. The e-mail is done by using SMTP (Simple Mail Transfer Protocol) protocol is a part of the application layer of the TCP/IP protocol. A protocol for sending e-mail messages between servers. Here raspberry pi acts as a server. It sends the result to the appropriate user.

V. UML DIAGRAMS

i) Class Diagram
In programming designing, a class chart in the Unified Modeling Language (UML) is a sort of static structure graph that depicts the structure of a framework by demonstrating the framework's classes, their properties, tasks (or strategies), and the connections among articles. In the chart, classes are spoken to with boxes which contain three sections:
The best part contains the name of the class
The center part contains the characteristics of the class.
The base part gives the strategies or activities the class can take.

ii) Activity Diagram
Action Diagram demonstrates the arrangement of steps that make up the mind boggling process. It demonstrates the stream of control, like grouping, however centers around activity as opposed to on articles. The parts utilized in this are
An Approach of Tomato Leaf Disease Detection Based on SVM Classifier

as per the following: Rounded Rectangle – It indicates the process.

Arrow – It indicates the transition line.
Rhombus – It indicates the decision.
Bars – It represents the start or end of concurrent activities.
Solid Circle – It represents the initial state of the workflow.
Encircled Black Circle – It represents the final state of the workflow.

Fig12: Class Diagram

Fig13: Activity Diagram

Fig14: Sequence Diagram

VI. HARDWARE AND SOFTWARE ENVIRONMENT

6.1 Hardware Environment:-
Raspberry Pi 3 Model B: Raspberry Pi 3 model B is the third generation Raspberry Pi with 1.2 GHz, 64-bit Quad-Core ARM Cortex processor with in-built Wi-Fi, wireless LAN and Bluetooth operations and also 4 x USB 2.0 Connector capabilities. This powerful single board can be used for many applications. It operates as a standard PC, requires a keyboard, mouse for command entry, a display unit, and a power supply. It has improved power management to handle more powerful external USB devices. It boots from micro SD card, running a version of the Raspbian operating system.

6.2 Software Environment:-
a. Python: Python is an interpreted based high-level, general-purpose programming language. We can use the programming language for evolving both desktop and web claims. It is designed with many features to aid data investigation and visualization. It takes a little amount of time to process with less effort. The data visualization libraries and APIs provided by python help to visualize and present information in a more appealing and efficient manner.

b. Open CV: OpenCV is an open source library of python ties intended to take care of PC vision issues. It makes utilization of Numpy, which is a profoundly enhanced library for numerical tasks. All the OpenCV cluster structures are changed over forward and backward numpy exhibits. Which makes it simpler to incorporate with different libraries that utilization Numpy, for example, Matplotlib.

VII. RESULT
In this work total, 191 tomato leaf images were taken. From which 50 are the healthy tomato leaf images and 107 are the diseased tomato leaf images for training phase and testing phase. Out of which 157 tomato leaf images are used as training phase, similarly, 34 leaf images are used for the testing phase.

The below figure 1 shows the database creation for both tomato healthy leaf images and tomato diseased leaf images dataset.

The below figure shows the feature extraction technique based upon size, shape, color, texture from the images of both healthy and the diseased leaf images.

The below figure 17 shows the original image and the masked image. Image masking is a non-destructive process which is used to remove the background unwanted noise and also to suppress the undesired distortion present in the image which is to get the accurate results.

The below figure 18 shows the SVM classifier in which the images are trained and tested when we presented the label image name in that classifier the result first shows the original image and masked image and then it indicates whether the image is healthy or diseased if it is diseased it shows whether the leaf is effected with bacterial spot or tomato mosaic virus based upon image classification the test indicated that the accuracy is more and also it acquires less time to receive the result.
An Approach of Tomato Leaf Disease Detection Based on SVM Classifier

Fig18: SVM Classifier Testing

The test result will send through the receiver's e-mail to get the alertness to the farmer or the user to get the awareness of their diseased leaf and gets sprayed.

Fig19: Result obtained with its Control Measures.

Finally, the below figure 20 shows the obtained result that which it sends to the user e-mail.

Fig 20: E-Mail

VIII. CONCLUSION

The proposed methodology has been implemented successfully and performance tests on python software. To get the accurate result, we have taken SVM classifier in the training phase and also tested. The results are getting accurately. It has been taken two commonly tomato diseases, namely bacterial spot, tomato mosaic virus which is mainly significant for yield loss. Hence, plant disease detection using Machine Learning Technique has been achieved to increase the speed and its accuracy. And also the obtained result is sent to the users using SMTP protocol through the mail. Thus the SVM classifier is getting an accurate result.

IX. FUTURE SCOPE

As for the future work increase in the database can produce better performance using the SVM classifier. Further, we can include automatic detection of tomato plant diseases. Whenever we capture the image for testing it undergo several processing and produces the result for future analysis.

REFERENCES

8. Prof. Bhavana Patil and Mr. Hemant Panchal, “Plant Monitoring Using Image Processing, Raspberry Pi & IoT,”