

1 Abstract

Support Vector Machines

(SVMs) along with neural network functions in order to detect and identify type of disease that infects tomato plant. In this project, we used dataset consisting of images of tomato leaves. Supported Vector Machines technique has been employed to identify an optimal feature subset. Finally, a support vector machine classifier with different neural network algorithms was employed to evaluate the ability of this approach to detect and identify where tomato leaf infected with Bacterial spot, early blight, Septoria leaf spot, etc. To evaluate the performance of presented approach, we tested our proposed model on dataset consisted of 10000 images comprised of 10 types of tomato leaf diseases. Extensive experimental results demonstrate that the proposed approach provides excellent annotation with accuracy more than 80%. Efficient result obtained from the proposed model can lead to tighter connection between agriculture researchers and computer technology, yielding more effective and reliable results.

2 Introduction

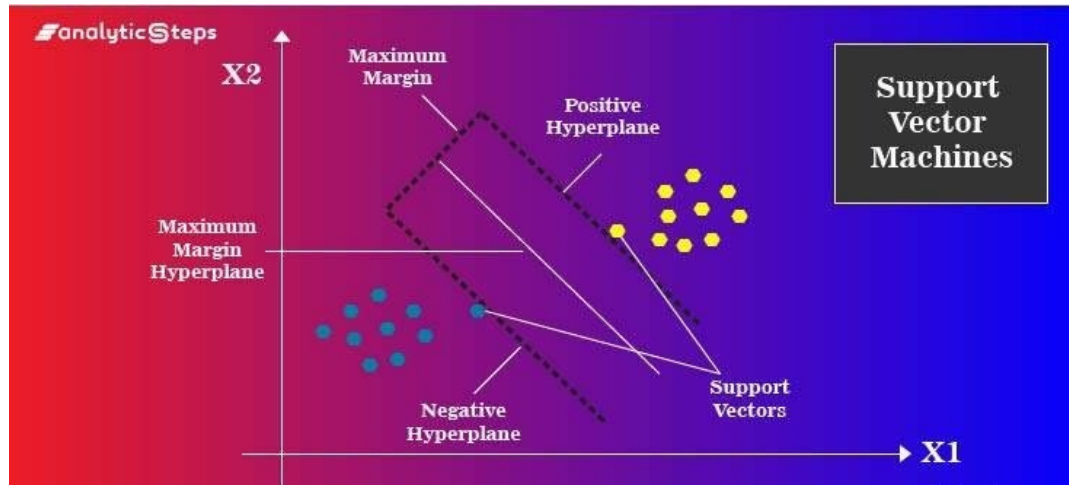


Figure 1: SVM

Tomatoes are one of the most widely cultivated food crops throughout the world due to its high nutritive value. It contains a lot of vitamins and nutrients such that vitamin C. It occupies the fourth level between word vegetables. Although, the naked eye observation of experts is the main approach adopted in practice for detection and identification of plant diseases, it requires continuous monitoring of experts, which might be expensive and difficult especially in large farms. So, it is necessary to help farmers in automatically detect symptoms of disease as soon as they appear by analyzing the digital images. Aiming at minimizing economic losses, ensuring both quality and quantity of agricultural products and minimizing agrochemicals use, computer based applications have been developed and revealed high efficacy. In this paper, a combination of various machine intelligence techniques has been used to detect diseased tomato leaf for the purpose of diagnosing pathogen types that infect tomato. The proposed SVM-based tomato diseases detection approach utilizes some preprocessing techniques such that image cropping and image filtering for improving the quality of leaf images.

3 Motivation

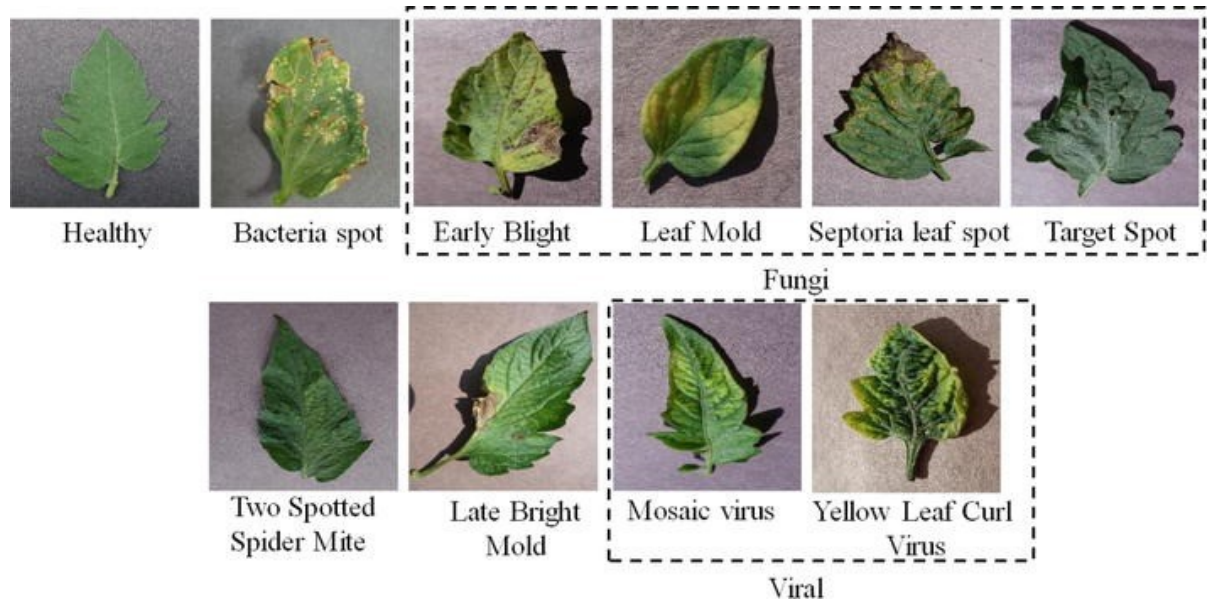


Figure 2: Tomato leaves with various diseases

Tomato, worldwide, is infected by several viral diseases which cause stunting, leaf curl, yellowing, mosaic, mottling, necrosis and shoe-string symptoms on plants, leaves or fruits. Among them, bud necrosis disease caused by an orthospovirus is emerging as a major constraint to the cultivation of tomato for resource-poor farmers. In the Indo-Gangetic eastern plains, various disease incidence on tomato ranged from 0 to 45% under field conditions during 2015 and 2016, along with other diseases such as leaf curl and mosaic (0–35%) and target spot (0–18%) respectively. Thirty four viral infected samples collected from 11 different villages were screened for different viruses using serological and PCR-based methods. The result revealed that most samples were positive for RNA (leaf curl virus, tomato mosaic virus). So, in order to stop the spread of such type of harmful diseases, our model accurately predicts which type of disease crop is going through with and help farmer take appropriate precautions.

4 Literature Review

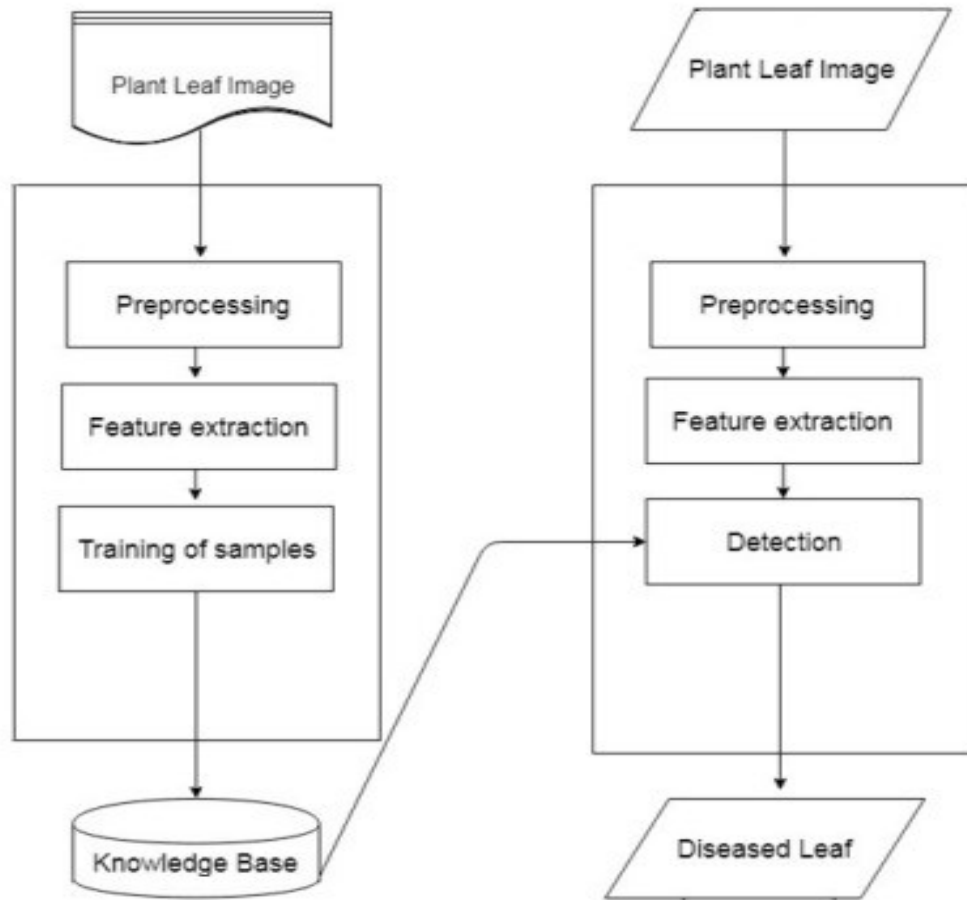


Figure 3: Flow chart for model

Support vector machine (SVM) is a classification technique originally developed by Vapnik. It was recently used in many applications including pattern recognition, bioinformatics, cancer diagnosis. It tries to find the trade-off between minimizing the training set error and maximizing the margin, to achieve the best generalization ability and remains resistant to over fitting. SVM works by mapping data to a high-dimensional feature space, so that data points can be categorized, even when the data are not otherwise linearly separable. SVM uses a nonlinear mapping to transform the original training data into a higher dimension, then it searches to find a linear optimal hyperplane, so that the margin of separation between the two classes is maximized.

5 Research Gaps

1. The main motive behind the project is to detect disease in tomato leaf. This further can extended to enhancing the model which will be capable of answering the cause and solution for the particular disease.
2. This can be achieved by suggesting the name of pest that caused particular disease and name of pesticide that is used for the remedy of the disease.
3. It will require large dataset consisting of pests, pesticides along with leaf images for making appropriate predictions of the disease.
4. Efficient result obtained from the proposed approach can lead to tighter connection between agriculture specialists and computer system, yielding more effective and reliable results.

6 Problem Statement and Objectives

1. Our primary objective of this project is to develop a model to categorize input plant leaf pictures as healthy or unhealthy. Moreover the objective of the proposed method is to develop a technique to identify leaf disease in tomato plant by improving the classification accuracy and reducing computational time.
2. In this project, we are trying to help farmers by solving the issue regarding the early detection of tomato leaf diseases so that they can start preparing for prevention of diseases. Thus minimizing their economic losses.
3. This proposed model is a self-learning model. We aim to improve the accuracy of prediction made by model as much as possible. It can be achieved by testing more number of images ensuring an accurate prediction.

7 Methodology

Support Vector Machine (SVM) classifier :

To classify the various diseases in plants any of the machine learning techniques can be used. The classification phase suggested deciding if the input image is healthy or diseased. In this paper Support Vector Machine (SVM) classifier has been used because it has some advantages over other classifiers such as effective in high dimensional spaces also in cases where the number of dimensions is greater than the number of samples. It is memory efficient as it uses a subset of training points in the **decision function** (called **support vectors**).

SVM is a supervised machine learning algorithm used for both classification and regression. SVM is a discriminative classifier. In this approach, for classification, the SVM technique has been applied. In SVM, each data item is plotted as a point in n-dimensional space; the number of dimensions corresponding to the number of features being classified. The classification is obtained by discovering the hyper-plane that uniquely distinguishes between different groups of scattered data points. By finding the hyper-plane the classification is performed. Hyper-plane differentiates two classes very well. This is shown in Figure 4.

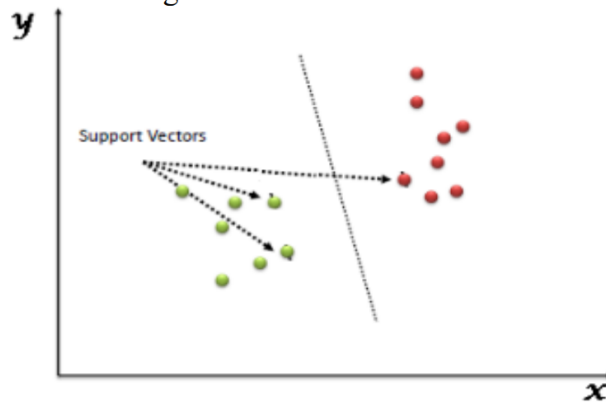


Figure 4: Classification of healthy (in green) and diseased part of leaf (in red)

The proposed SVM-based tomato diseases detection approach is comprised of the following four fundamental building phases :

1. **Image acquisition phase :** The proposed SVMbased tomato diseases detection approach begins with acquiring and collecting digital images from suitable environment.
2. **Pre-processing phase :** To improve the quality of the images and to make the feature extraction phase more reliable prepare, some image-processing techniques are applied including leaf image isolation, image resizing and background removing.
3. **Feature extraction phase :** 402 texture-based features have been extracted, and represented in a database as vector values.

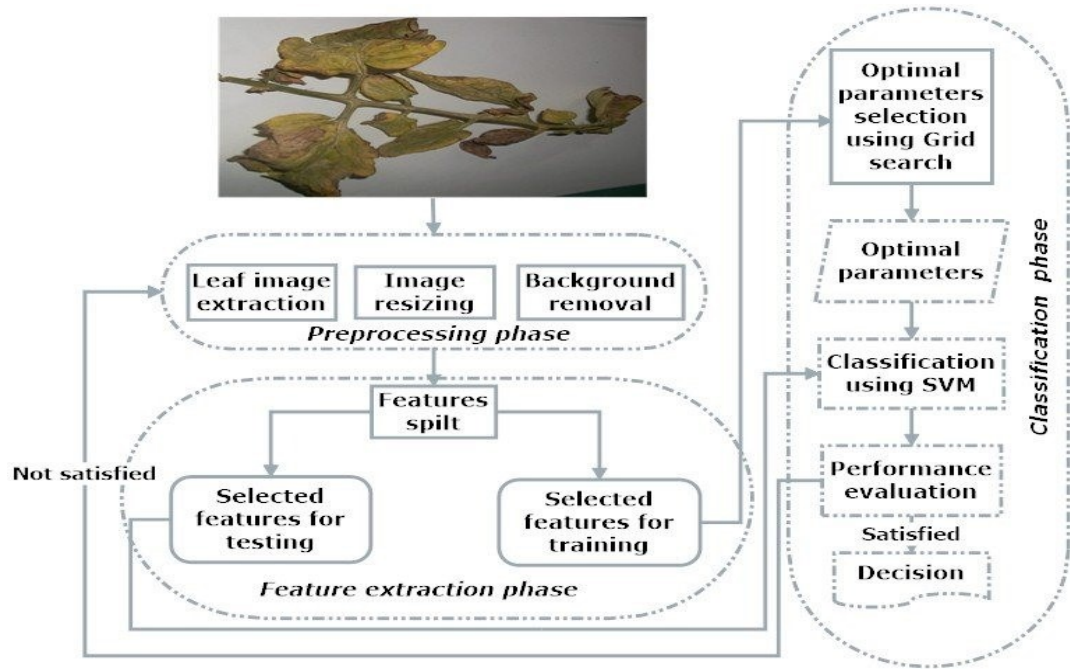


Figure 5: Prediction Process of disease

4. **Classification phase :** The last phase is the classification and prediction of new objects, it is dependent on the SVM classifier. These four phases are described in detail in the following subsections and depicted in Figure 5.

8 Hardware and Software Requirements

The proposed model must be able to detect disease. The performance will depend upon GPU power. Users can follow up the interface step by step for their purpose. The proposed model must be available for use to the user as and when needed provided that the user's system meets the specified requirements. The proposed system must be able to recover from failure in case of the application crashing abruptly and become ready-to-use after recovery. Along with the necessary peripheral hardware, and Python 3 will be used to implement the software functionality of drowsiness detection.

- **Software Requirements:** python3, Tensorflow 2.10.0, Keras 2.10.0, NumPy 1.23.4, Panda 1.5.1
- **Hardware Requirements:** NVIDIA-SMI 472.19, Driver Version : 472.19, CUDA Version : 11.4, GPU Name : TCC/WDDM.

9 Conclusions

In this paper, an effective SVM-based tomato leaf diseases detection approach is proposed to recognize two types of tomatoes leaf diseases. The dataset consists of in total 11000 (10000 images-training dataset and 1000 images-testing dataset) infected tomato leaf images with Bacterial spot, early blight, etc were used for both training and testing phase. Overall validation accuracy (after 10th epoch) is 82.93% whereas training accuracy is 89.30%.

10 Timeline Chart

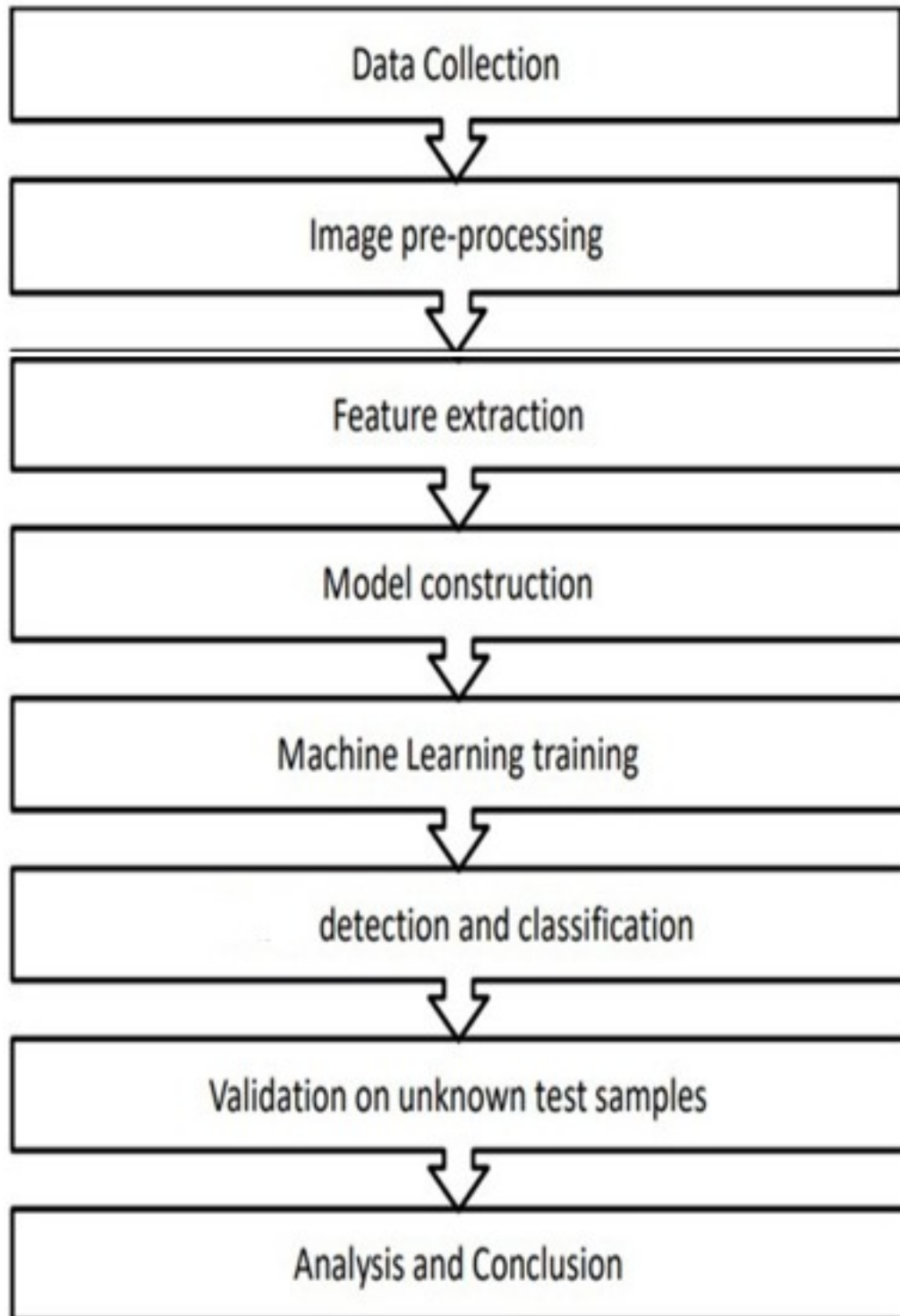


Figure 6: Flowchart for design and development of proposed model