

Plant Disease Identification Using Discrete Wavelet Transforms and SVM

Kiran S M¹, Dr. Chandrappa D N²

¹Research scholar, Dept of ECE, SJBIT, Assistant Professor, Dept of ECE, BNMIT, Bangalore, VTU Belagavi India

²Professor, Dept. of ECE, SJBIT, Bangalore, VTU Belagavi India

Abstract Disease detection in plants is one of the essential step in the field of agriculture to improve the quality and yield of crops. Applications of image processing plays major role in early detection of diseases and also in terms of accuracy and time consumption. In many plant health monitoring systems, Fourier and wavelet transforms are applied for feature extraction from plant images and then they are classified using different classifiers. In this study, tomato leaf images are collected from PlantVillage database, image are preprocessed to improve the contrast and then image segmentation is performed using k-means clustering technique. Texture features are extracted from the region of interest using Discrete Wavelet Transforms (DWT). Fourteen image features obtained from Daubechies (db3), Symlet (sym3) and biorthogonal (Bior 3.3, Bior 3.5, Bior 3.7) wavelets. These features are used to classify the images as healthy and unhealthy with the help of Support Vector Machine (SVM) classifier. Performance of the system is measured in terms of Sensitivity, Specificity and Accuracy. The proposed system classify the images with sensitivity of 92%, specificity of 84% and accuracy of 88%.

Keywords Discrete wavelet transforms, Fourier Transforms, PlantVillage, k-means clustering, plant diseases.

1. Introduction

After potato and onion tomato is the world's 3rd largest vegetable crop. Tomato is grown in outdoor fields, net houses and green houses in almost all the countries. China, India Egypt, USA, Italy Brazil and Spain are the leading tomato producing countries. In India tomato is most widely cultivating crop and India exports tomato to Pakistan, Bangladesh and many neighbouring countries. Predominantly tomatoes are summer crops, but can be cultivated in all the seasons.

A variety of insects, pests and diseases cause the problems on growth of the plant, leading to poor quality of fruits and also reduction of yield. As per the available statistical data, there are more than 20 types of tomato diseases. Most of the diseases that affect the tomato crop are due to various viruses, bacteria and fungi. Some important diseases of tomato include: early blight, anthracnose, bacterial canker, bacterial blight, fusarium, wilt bacterial spot, root rot, Late blight and powdery mildew [1]. These diseases affects affect any part of the plant such as stem, root and leaf. Most observable part of the symptoms are in leaves.

Around 30% of tomato crop losses due to the diseases [2]. Farmers can avoid losses due to diseases by following some basic steps, some of them include: rotation of crops, pinch off the leaves with disease symptom and disposing them to avoid infection spreading, avoiding working in farm when tomato foliage is wet, choosing disease resistance variety, providing adequate air circulation, following mulching techniques,

Generally farmers use naked eye observation for disease identification [3]. This is not only time consuming, but also inaccurate. In many cases farmer need to take help of an experts for disease detection. Detecting diseases in the early stages helps to reduce the plant damage and to improve the quality as well as yield of crops. Therefore an automatic system for health monitoring in tomato farms plays crucial role in the field of agriculture.

Disease detection in plants using image processing techniques plays a crucial role in the recent developments in technology. Analysis of affected area of plant parts is the best way to detect disease using object detection techniques in image processing. In the field of computer vision, pattern recognition in images is very important subject. The main aim of pattern recognition techniques is to accurately locate the region of interest (RoI) in the images. In this study, healthy and infected tomato leaf images were taken for experiment, k-means clustering technique is used to segment the region of interest from tomato leaf images, image features are extracted after applying discrete wavelet transforms and then images are classified as healthy and unhealthy using thresholding method.

2. Related Work

Lots of research have been done on plant leaf disease detection with the use of digital image processing application. Arifur Rahman et al. [4], implemented a new approach for image segmentation, tomato leaf images are classified based on the diseases by using RGB images. In this approach noise from the images are removed using median filters and then RGB images are converted to LAB color space. Then color negation is done by removing the image background. Images are segmented using RGB thresholding technique. The image features are extracted using GLCM and finally leaf images are classified using deep learning. M Dhakate et al. [9] use the image processing and neural networks for disease identification and classification in pomegranate plants. In this work, three diseases were covered Leaf spot, Bacterial blight and Fruit rot. This system classify the disease with accuracy of 90%

MelikeSardogan et al. [5], worked on disease detection in tomato leaf using convolutional neural networks and learning vector quantization. This work focused on 4 diseases: bacterial spot, septoria spot, late blight and yellow curved, here, 400 image are used for training and 100 images for testing, the average accuracy obtained is 86%. Thair A. Salih et al. [6], implemented a system for tomato disease identification using image processing algorithms. Images are collected from plant village database and are preprocessed to maintain standard size and to improve the brightness. Images are classified using deep learning convolutional networks.

A neural network based system for mango leaf disease identification is implemented by Sujit Roy et al. [7], this method helps to detect diseases in mango plant very easily and correctly, this saves lot of time compared to traditional methods. Here, different symptoms of leaf are considered and the system is trained and tested using machine learning algorithms. Both healthy and diseased leafs are used for testing and training purpose so that the system detects diseases with good accuracy. The system detect the disease without agriculturist help, and able to examine the disease with 80% accuracy. This helps in improving the mango crop yields. Deep learning models are implemented in [8] for disease detection in mango plant. For their work database were collected from Konkan area in India. Here, the Convolutional Neural Network (CNN) is trained using transfer learning techniques. Disease is detected with accuracy of 91%.

In [10], tomato leaf images are classified as healthy or infected by late blight, features are extracted from each R, G and B components of pre-processed images using discrete wavelet transforms. The images are classified using Euclidean distance algorithm, Principle component analysis and independent component analysis. Experimental results shows the system detects late blight with 89.9% accuracy using principle component analysis.

3. Methodology

In this work, disease identification is done in tomato crop using leaf images. Images are collected from available standard database and then pre-processed to remove noise and to improve the contrast. Texture feature of an image are obtained discrete wavelet transforms. These texture features are feed to SVM classifier to classify the images as healthy or unhealthy. Fig 1 shows the various steps followed in the proposed algorithm.

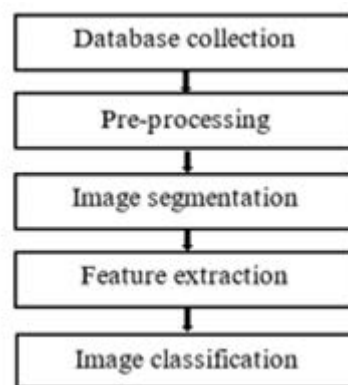


Fig 1: Steps for tomato disease identification

3.1. Image database: For this study, a database of 50 healthy images and 50 unhealthy images of tomato plant leaf are collected from PlantVillage dataset. Initially, features are extracted from 25 healthy and 25 unhealthy images, these features are used to classify the remaining 50 images. Fig 2 show the healthy and diseased tomato leaf images. The algorithm is implemented using MATLAB 2020b.



Fig 2: (a). Healthy Tomato Leaf (b),(c),(d). Unhealthy Tomato leaves

3.2. Image pre-processing: This is an important step in any image processing algorithms, importance of image pre-processing is to improve the image quality by removing background noise and suppressing undesired distortions. In this work, image is resized to 256*256 pixel. Using Gaussian filter, image noise is removed and finally contrast stretching is performed on the image.

3.3. Image segmentation: This involves the process of separation of a given image into homogenous parts based on certain features. Normally, image clustering techniques are used for segmentation purpose. Clustering is the technique used to group the large sets of data into clusters of smaller sets of similar data [2]. In this work, input leaf image is segmented in to three groups (Shown in fig 3) using k-means clustering technique. Before applying k-means clustering algorithm, input RGB image is converted to LAB colour space and then 'a' component is extracted to perform clustering.

K-means algorithm works as follow:

Step 1: Specify K value (number of clusters required)

Step 2: Randomly select cluster centres 'c'.

Step 3: Calculate Euclidean distance from cluster centre to each data point, Euclidean distance is the shortest distance between any two points.

Step 4: Assign data point to the cluster whose distance from cluster centre is closest of all other cluster centres.

Step 5: Repeat the steps for all the data points and assign them to the clusters with closest distance to cluster centre.

Step 6: Stop the process ones the clusters become stable.

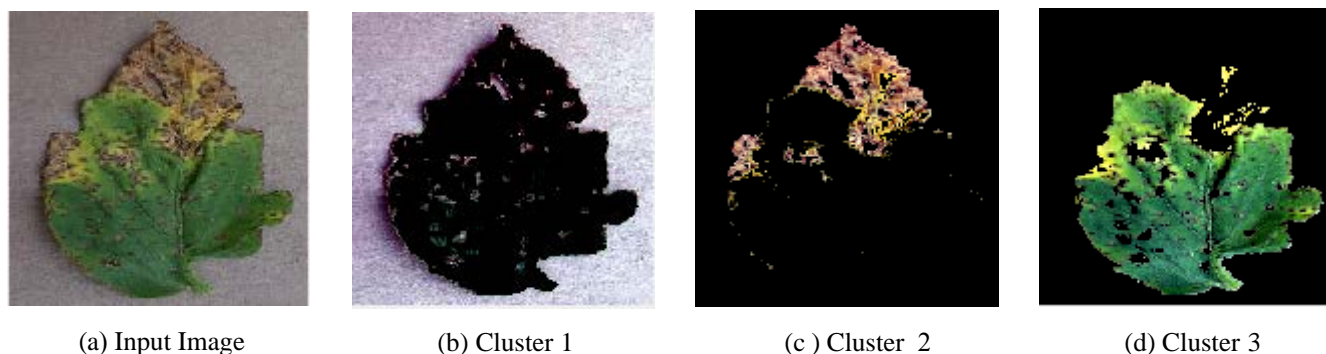


Fig 3: Segments obtained from k-means clustering

3.4. Feature Extraction: Image classification is mainly depends on the number and type of features extracted from an image. In images, feature is a significant representative which can be used to distinguish images. There exists two kinds of image features such as texture and structural features. Structural features depends on size and shape of image parts where a texture features depends on the pixel intensity and location of that pixel [11].

In this proposed method, 2-D discrete wavelet transforms are used to obtain texture features from the image segment having region of interest using. Discrete wavelet transforms consider both frequency as well as spatial information of a signal. This property is useful for texture feature extraction with good accuracy. 2-D DWT have the capacity to store information of a signal in less number of coefficients [12]. In image processing, DWT stores the image as two dimensional signal with rows and columns. Wavelet transforms analyse the details of image like horizontal, vertical and diagonal sub-bands of the image [13]. Use of 2-D DWT for image decomposition is shown in fig 4

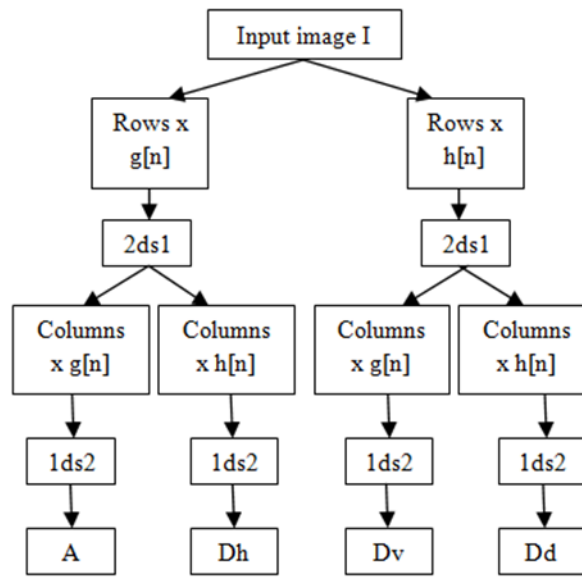


Fig 4: Fig 4: 2-D DWT Decomposition

In the fig 4, I indicates an input image, h[n] is high pass filter and g[n] is low pass filter. 2ds1 means rows are down sampled by 2 and columns by 1, 1ds2 means rows are down sampled by 1 and columns by 2. In this work, significant results are obtained from level one itself. A represents an approximation coefficient. Dh represents a horizontal coefficient, Dv is vertical and Dd is a diagonal coefficient. In this work, detailed coefficients obtained from five different wavelets are used to extract energy features. The wavelets used here are daubechies wavelet (db3), symlet wavelet (sym3) and biorthogonal wavelets (Bior 3.3, Bior 3.5 and Bior 3.7). These wavelets are defined by different scaling and wavelet functions.

Average and energy feature are calculated from horizontal and vertical components using:

$$\text{Average } Dh = \frac{1}{R * C} \sum_{x=M} \sum_{y=N} |Dh(x, y)| \quad (1)$$

$$\text{Average } Dv = \frac{1}{R * C} \sum_{x=M} \sum_{y=N} |Dv(x, y)| \quad (2)$$

$$\text{Energy} = \frac{1}{R^2 * C^2} \sum_{x=M} \sum_{y=N} (Dv(x, y))^2 \quad (3)$$

Where R represents number of rows and C represents columns in an input image. Dv and Dh are vertical components and horizontal of first level of decomposition. These average and energy values are the good indication of pixel intensities distributed in an image.

3.5. Classification: Tomato leaf image are classified as healthy or unhealthy using the extracted texture features. Sometimes all the features may not be important for classification, therefore the features having little or no predictive information are eliminated and only important features having enough information for classification will be selected.

In this work, fourteen texture features are used as input for Support Vector Machine (SVM) classifier. SVM is a supervised machine learning algorithm useful for classification. Basically, SVM model represents the different classes of data in hyperplane, by generating hyperplane in iterative manner errors can be reduced. SVM optimizer finds an hyper plane using labeled training vectors, that will then maximizes the margin of separation among the two classes as shown in Fig 5.

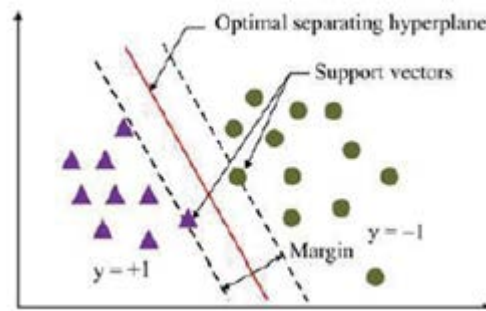


Fig 5: SVM in Linearly Separable Condition

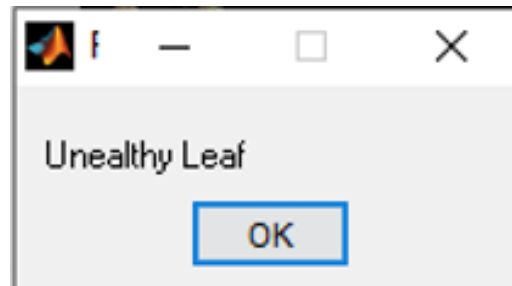


Fig 6: classification result

4. Experimental results:

In this work totally 100 images are used for experimentation, 50 images for training and 50 images for testing purpose. The image shown in fig 3 is an unhealthy image. Fig 6 shows the classification result as unhealthy leaf image.

The performance of the system is analysed using 3 evaluation metrics such as accuracy, specificity and sensitivity. True negative, true positive, false positive and false negative are used to evaluate the performance of the system.

Table 1: Definition of TN, TP, FP and FN

Definition	Input Image	Output obtained
True Negative (TN)	Healthy	Healthy
True Positive (TP)	Unhealthy	Unhealthy
False Positive (FP)	Healthy	Unhealthy
False Negative (FN)	Unhealthy	Healthy

Accuracy: The ability of the system to give correct results is given by accuracy. It is defined by Eq. (4)

$$\text{Accuracy} = \left(\frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \right) * 100 \quad (4)$$

Specificity: The ability of a system to classify healthy images as healthy is defined by specificity. This gives the ability of the system to detect normal conditions. It is calculated using Eq. (5).

$$\text{Specificity} = \left(\frac{\text{TN}}{\text{TN} + \text{FP}} \right) * 100 \quad (5)$$

Sensitivity: The ability of a system to classify unhealthy images as unhealthy is defined by Sensitivity. This gives the ability of the system to detect the disease. It is calculated using Eq. (6).

$$\text{Sensitivity} = \left(\frac{\text{TP}}{\text{TP} + \text{FN}} \right) * 100 \quad (6)$$

Table 2 shows the experimental results. As mentioned in table 2 proposed system gives the results with accuracy of 88%, sensitivity 92% and specificity of 84%

Table 2: Experimental results

Parameter	Value
Total images used	100
Trained images	50
Tested images	50
Failed cases	6
Successful cases	44
True Negative (TN)	21
True Positive (TP)	23
False Positive (FP)	4
False Negative (FN)	2
Sensitivity	92%
Specificity	84%
Accuracy	88%

5. Conclusion

An automatic system for tomato disease identification is developed using various image processing algorithms. Tomato leaf images are resized, enhanced, filtered and segmented using 'a' component of LAB colour space. From the segmented results image having RoI is used for feature extraction with 2-D DWT. Classification is performed using the features extracted from images with the help of SVM classifier. The proposed systems identify the disease with accuracy of 88%. The future expansion of this work can be focused on detecting the type of disease affected and also focus on multiple crops is necessary.

REFERENCES

- [1] Jones, J. B., Zitter, T. A., Momol, T. M., and Miller, S. A. "Compendium of Tomato Diseases and Pests", 2nd edition. American Phytopathological Society Press, St. Paul, 2014.
- [2] Pranjali B. Padol, Anjali A. Yadav, "SVM Classifier Based Grape Leaf Disease Detection", 2016 Conference on Advances in Signal Processing (CASP), Cummins College of Engineering for Women, Pune. Jun 9-11, 2016.
- [3] Weizheng, S., Yachun W., Zhanliang C., and Hongda W., "Grading Method of Leaf Spot Disease Based on Image Processing" International Conference on Computer Science and Software Engineering – Volume 06 ,PP. 491-494, December 2008.
- [4] Md. Arifur Rahman, Md. Mukitul Islam, G M ShahirMahdee, Md. WasiUKabir, "Improved Segmentation Approach for Plant Disease Detection", 1st International Conference on Advances in Science, Engineering and Robotics Technology 2019 (ICASERT 2019).
- [5] MelikeSardogan, AdemTuncer, YunusOzen, "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm", 3rd international conference on computer science and engineering 2018.
- [6] Thair A. Salih, Ahmed J. Ali, Mohammed N. Ahmed, "Deep Learning Convolution Neural Network to Detect and Classify Tomato Plant Leaf Diseases", Open Access Library Journal, May 2020.
- [7] Sujit Roy, Md. Atikur Rahman, Md. Rasel Mia, Subrata Kumar Das ., "Mango leaf disease recognition using neural network and support vector machine", Iran Journal of Computer Science · September 2020.
- [8] SampadaGulavnai, RajashriPatil, "Deep Learning for Image Based Mango Leaf Disease Detection", International Journal of Recent Technology and Engineering, Nov 2019.
- [9] Mrunmayee Dhakate and A. B. Ingole, "Diagnosis of pomegranate plant diseases using neural network," in Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2015 Fifth National Conference on, IEEE, pp. 1–4, 2015.
- [10] HiteshwariSabrol, Satish Kumar, "Recognition of Tomato Late Blight by using DWT and Component Analysis", International

Journal of Electrical and Computer Engineering (IJECE) Vol. 7, No. 1, February 2017, pp. 194~199.

- [11] Gaurav Kumar, Pradeep Kumar Bhatia, “A Detailed Review of Feature Extraction in Image Processing Systems”, Fourth International Conference on Advanced Computing & Communication Technologies, 2014.
- [12] Kiran S M, Dr. Chandrappa D N, “Automatic Detection of Glaucoma Using 2-D DWT”, International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 06, June-2016
- [13] R. Polikar, “The Wavelet Tutorial Part I,” <http://users.rowan.edu/~polikar/WAVELETS/WTpart1.html>.