ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 3, January 2018

IDENTIFICATION OF DISEASES IN TOMATO LEAVES USING SIFT

RAVIKIRAN.R
Dept. of Electronics and communication
S.J.C.I.T
Chickballapur, India
rksravikiran@gmail.com

SHASHIKIRAN.R
Dept. of Electronics and communication
Nagarjuna College of Engineering
Banglore, India
rkrshashikiran@gmail.com

RASHMI.K.A
Dept. of Computer Science and Engineering
S.J.C.I.T
Chickballapur, India
rashmiravikiran.rr@gmail.com

RAVI.M.V
Dept. of Electronics and communication
S.J.C.I.T
Chickballapur, India
ravimvreddy1986@gmail.com

ABSTRACT

Due to various health benefits, tomato is one of the major horticulture crop consumed by majority of the population in all over the world. Because of large number of tomato consumers and tomato being a high yield and a multi-time crop, large numbers of farmers cultivate this horticulture crop in a large scale. But the major problem is that the variations in climatic conditions makes the horticulture crops and various parts of the plants such as roots, stem, leaf and seeds susceptible to the virus attacks. Of all these parts, leaf is most susceptible to the virus attacks. Furthermore, the horticulture crop diseases spread at a faster rate compared to the other category of crops. In addition to these problems, the manual identification of leaf diseases is challenging and tedious. To address these problems, we have automated the process of identification of the health condition of the tomato leaf, with the assumption that a proper image acquisition process exists. We have proposed a algorithm to identify whether the tomato leaf is healthy or unhealthy is based on the Scale Invariant Feature Transform (SIFT) To train the algorithm, a total of 80 tomato leaf images (healthy and unhealthy) are captured using the camera. Using the proposed algorithms, the features are extracted from these images and stored in the database. When a query tomato leaf image is given as an input, the algorithm extracts the features and compares these features with the stored features and gives the output of the algorithm as health or unhealthy leaf.

Keywords— *SIFT, Tomato leaf disease*

I. INTRODUCTION

Due to variations in climatic conditions, the horticulture crops and various parts of the plants such as roots, stem, leaf and seeds are attacked by the viruses [1]. Also, the horticulture crop diseases spread at a faster rate compared to the other category of crops. This will result in low crop yield and cause financial loss to the farmers. The plant's and the crop's health can be maintained by applying the necessary medicine on the plants, but this indirectly affects the health of the consumers [2]. Further, the continuous increase in population demands more crops. So, necessary steps have to be taken to produce more and healthy crops.

The horticulture crops consist of fruits and vegetables. According to the Indian government, in the year 2015, about 500 million people belong to middle class and below poverty line [3]. For these categories of the people, in addition to the food grains, vegetables are of higher priority and forms the necessary and important part of the meal compared to the fruits.

The most frequently used vegetables in the day to day life are given in Table 1[4],Experts (Experienced Farmers) have suggested that, compared to the roots and stem parts of the plants, leaves are more susceptible for virus attacks. They have suggested that, identifying the health condition of the plant from the leaves is important and as well as easy compared to other parts of the plant. Hence, leaf is considered in this work for identifying the health condition of the plants. Experts have provided the number of days taken by the virus to attack the plants completely and this data is summarized in Table 1. From the table it can be observed

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 3, January 2018

that the following three vegetables are having high water content, namely: tomato, potato and beans. Potato is a one-time crop for each plant, as the yield is obtained after rooting out the plant. The life span of beans plant is around three months and in this life span it yields four to five times. The life span of tomato plant is around six months and in this life span it yields seven to eight times [6], [7]. Moreover, tomato has more health benefits compared to potato and beans [7], [8], [9]. Some of the health benefits of consuming tomato are as follows: protects vision and degenerative eye disease, reduces cardiovascular diseases, reduces the risk of prostate cancer and breast cancer, prevents kidney and gall bladder stones, reduces the risk of blood clot, increases fat burning capacity, prevents stroke, restores biochemical balance in diabetics, improves digestion and prevents constipation, provides healthy and glowing skin, and nourishes the hairs [10], [11]. Because of these health benefits and high yield from the tomato plant, we have considered tomato in our work [12].

There are two factors that cause diseases to the leaves, namely: biotic factors and abiotic factors. Biotic factors means living things such as virus, bacteria and fungi causes diseases to the plants. Abiotic factors constitutes of environmental conditions, nutritional deficiencies and pollutants. These abiotic factors make the leaves prone. [5] discussed about an eye blinking sensor. Nowadays heart attack patients are increasing day by day."Though it is tough to save the heart attack patients, we can increase the statistics of saving the life of patients & the life of others whom they are responsible for. The main design of this project is to track the heart attack of patients who are suffering from any attacks during driving and send them a medical need & thereby to stop the vehicle to ensure that the persons along them are safe from accident. Here, an eye blinking sensor is used to sense the blinking of the eye. spO2 sensor checks the pulse rate of the patient. Both are connected to micro controller.If eye blinking gets stopped then the signal is sent to the controller to make an alarm through the buffer. If spO2 sensor senses a variation in pulse or low oxygen content in blood, it may results in heart failure and therefore the controller stops the motor of the vehicle. Then Tarang F4 transmitter is used to send the vehicle number & the mobile number of the patient to a nearest medical station within 25 km for medical aid. The pulse rate monitored via LCD .The Tarang F4 receiver receives the signal and passes through controller and the number gets displayed in the LCD screen and an alarm is produced through a buzzer as soon the signal is received.

Vegetables Names	Tomato	Potato	Capsicum	Chilly	Pumpkin	Beans
Number of days	**2**	**2**	**4**	**7**	**3**	**2**

Table 1 The vegetable names and number of days taken by the biotic factors to attack the leaves.

In [ref], Harini and Lalitha have proposed an algorithm for identifying the leaf diseases in tomato plant based on the Wavelets and the PCA. Here, the Haar Wavelet is used for the extraction of the important features from the preprocessed input image. Further, the dimensionality is reduced by using the PCA and thus retaining only the most important features. They have tested their algorithm on 45 images and obtained 90% efficiency based on the similarity score.

The remaining part of the paper is organized as follows. In Section 1, the mathematical framework for the proposed algorithm is provided. The proposed algorithm using SIFT methodology is presented in Section 2. The simulation results and the corresponding discussion are given in Section 3. Section 4 ends with the conclusion and the future scope of work.

II. MATHEMATICAL FRAMEWORK

One of the various challenges faced by the farmers is to find a particular plant's leaves that have been diseased or not, out of large number of tomato plants cultivated in the farm. Assuming that a good system exists for the purpose of image acquisition, this work automates the process of identifying whether a particular plant (leaf) is healthy or not. Here, we are using Scale Invariant Feature Transform (SIFT) .A brief review of the SIFT is as follows.

A. SIFT: The SIFT extracts the fixed local feature points from an input image, that are invariant to basic image transformations such as scale and rotation. The steps are summarized as follows:

Step 1: ***Finding the extreme points in the scale space*****.**

The input image is subjected to Gaussian filters with different mask size using different values of σ, to get a set of images $L(x, y, \sigma)$ with different scales of blurring. The extreme points are extracted by comparing every point of the blurred image with the points in the same space and the points in the neighborhood space. The Gaussian scale transform operator used by Lowe was the difference of Gaussian (DoG) and the scale space is

Gaussian difference scale space is {G(x; y; σ)}, Lowe proved that the point extracted by this operator is invariant in scale transformation of original image *I(x,y)*. The scale space can be formulated as follows.

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \quad \ldots (1)$$

Gaussian different scale space is given blow:

$$D(x,y,\sigma)=\left(G(x,y,k\sigma)-G(x,y,\sigma)\right)*I(x,y)$$

$$=L(x,y,\sigma)-L(x,y,\sigma) \quad \ldots (2)$$

Where * is convolution operator

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-\left(x^2+y^2\right)/2\sigma^2} \quad \ldots (3)$$

Each point in the scale space is compared with 8 points in the same space and 9 points above and below the current scale space, summing to 26 points. The points with extreme values (maximum or minimum) are retained and called as extreme points that are invariant to changes in the scale space.

Step 2: ***Localizing the key points***

The extreme points must be filtered as some of the points are unstable and sensitive to noise. The unstable points that are sensitive to noise may not be detected as they are dominated by noise and have smaller values. The points on the edge of the texture are sensitive to image transformations and leads to extraction of different extreme values from the same location. The remaining points are considered as key points that are invariant to noise and affine transformation.

Step 3: ***Assigning orientations to each key points***

Compute the Gradient for each blurred image using the following equations.

$$m(x,y)=\sqrt{\left(L(x+1,y)-L(x-1,y)\right)^2+\left(L(x,y+1)-L(x,y-1)\right)^2} \quad \ldots (4)$$

$$\theta(x,y)=\tan^{-1}\left(\left(L(x,y+1)-L(x,y-1)\right)/\left(L(x+1,y)-L(x-1,y)\right)\right) \quad \ldots (5)$$

An orientation histogram is created with 36 bins covering 360° for the region around the key point. Each point is then weighted with a Gaussian window of size 1.5σ. The orientations are assigned to key points by considering the highest peak in the histogram and all other local peaks within 80% of the highest peak.

Step 4: ***Building the description for key points***

The key points assigned with orientation are grouped and summarized into 16×16 sample array. A sub-region of size 4×4 is considered to create a histogram with 8 bins. The key points are weighted with normalized Gaussian widow of size 1.5σ. The trilinear interpolation is used to place in the histogram bins.
The SIFT algorithm uses the above technique to extract key points that are invariant to scale changes, affine transformation, rotation, noise and illumination variations.

III. PROPOSED ALGORITHM

To identify whether a tomato leaf is healthy or not, two of the algorithms are proposed in this paper. In both the algorithms, the execution is carried out in two phase. The first phase is the database creation phase, where in the features are extracted from the input training set of tomato leaf images. The training set consists of both healthy and unhealthy tomato leaves. The second phase is the identification phase, where in the features are extracted from the query image and compared with the features stored in the database. If there is a match above some particular threshold then, the relevant metadata (healthy or unhealthy) is produced as the output. One of the algorithm is based on the SIFT and the other algorithm is based on the color histogram and the PCA. Both the algorithms are explained in the following subsections.

A. Proposed algorithm based on the sift

The block diagram of the proposed algorithm based on the SIFT is shown in Figure 1. From the block diagram it can be observed that the sequence of steps followed in the database creation phase and the identification phase are same. The sequence of steps is as follows: input image (training image or the query image), normalization and feature extraction.

In the database creation phase, the input training image is normalized, by transforming the RGB image in to the PGM format and then resizing to a predefined size. After normalization, the feature is extracted

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 3, January 2018

using the SIFT (the procedure was explained in Section 2.1) and the feature is stored in the database along with the metadata of the corresponding input training image. The metadata is nothing but whether the tomato leaf is healthy or not. In other words, if the input tomato leaf is unhealthy, then the corresponding extracted features are stored under the unhealthy class and if the input tomato leaf is healthy, then the corresponding extracted features are stored under the healthy class. In the identification phase, a query image whose health condition is yet to be identified by the algorithm is given as an input. Following the similar procedure in the database creation phase, the feature is extracted. This extracted feature is compared with the features stored in the database and the relevant metadata is produced as the output of the algorithm.

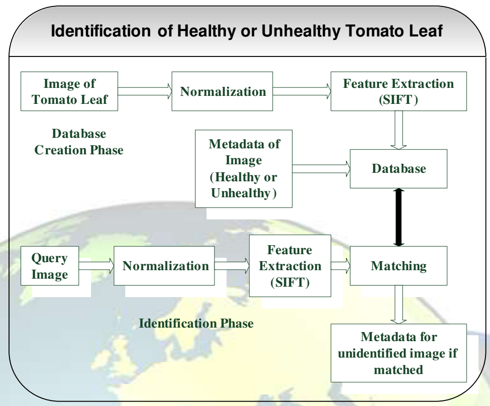

Fig 1 Block diagram of the proposed algorithm based on the SIFT.

A. Input image: The input image is captured using digital camera. In common, images can be in different formats like white and black or RGB colour leaf image. An image is a picture that has been copied or created and put in to storage in the electronic form. An image can be defined in terms of vector graphics. An image stored in vector form is sometimes called a bitmap. A leaf image map is a file having data that associates different locations on a specified leaf image with hypertext links. Image formats like JPEG, PNG, GIF, SVG, and TIFF. as shown in fig (2).

Fig 2 input leaf image.

B. Normalization: The input image is converted into the form of PGM formats. The Portable Pix Map format (PPM) and Portable Gray Map format (PGM) are image file formats designed to be used on different platforms with ease.

The PGM format is lower most common denominator grayscale file format. It is designed in extremely easy ways to write programs and learn. The PGM is a transparency mask which consists of sequence of one or more PGM images.

There are no delimiters, data or padding after, before or between images.

C. Feature Extraction using SIFT Algorithm: The SIFT algorithm input leaf image is in PGM format. The SIFT algorithm uses the Gaussian blur to create the scale space to original image and processed continuously, whose output is blurred leaf image and the original image is resized to half its size. And SIFT creates blurred out image again. And these processes keep repeating. Blurring is referred to as the convolution of Gaussian operator as shown in fig(2) , Locate Maxima and Minima in DoG leaf image, Getting rid of low contrast key points, Key points orientations and Generating features. The fig (3) is feature extraction of input leaf, fig (4) feature extraction of query leaf and fig (5) shows the resultant output of leaf.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 3, January 2018

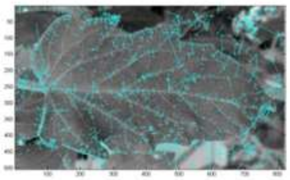
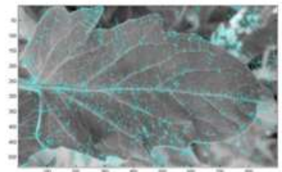

Fig 3 input image key point extracting Fig 4 input query image key point extracting.

D. Database: A set of data collected together in one location like file, folder, record and system data. The database is used for storing group of SIFT result leaf images, which is arranged for ease, speed of search and retrieval of leaf images those are converted into PGM formats and also for extracting the key points. Most data contains multiples leaf images.

E. Query image: This is inquiring about leaf images. The query leaf image waits for a reply.

F. Check: The check block is used to compare the database images and as well as SIFT Query image. The output obtained will indicate that the key points matching correspond to a healthy leaf or not.

IV. EXPERIMENTAL RESULTS:

This paper includes the proposed methodology which is implemented using SIFT algorithm and colour histogram-PCA. In this project, the database contains 80 images collected from tomato plant which is captured using Sony camera. The experimental result has two methodologies as follows.

SIFT: The sift algorithm is used to find the key points generated from the PGM image. In this work, database collected consists of 80 images which are either healthy or unhealthy leaves. Each image is in different angle, rotation and position. SIFT algorithm is used for comparing two different leaf images, one from database and other is the query image. This algorithm extracts the key points of two different images. After extraction, if more number of key points matches it is considered as healthy else it is considered to be unhealthy.

Comparing the two healthy leaves with almost same similarities as shown in below figure.

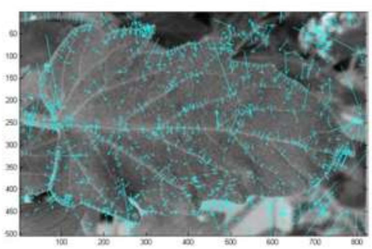
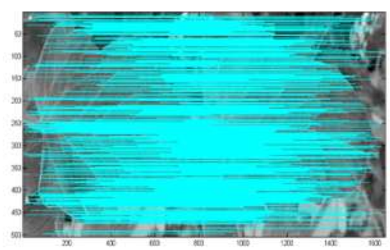

Fig 6 Healthy leaf key points found **Fig7** Extracting key points matching of healthy.

Comparing the two healthy leaves with almost same similarities as shown in below figures.

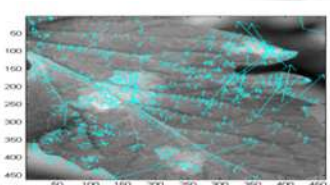
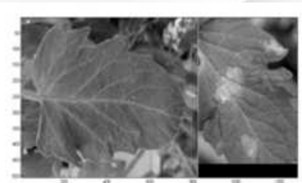

Fig 8 Unhealthy leaf extracting key points**Fig9** Extracting key points matches

V. CONCLUSION:

The farmers are facing huge losses due to the leaf diseases. Leaf diseases are caused due to the uneven environmental conditions. Due to this the entire crop yield has been reduced. Therefore, price of the harvested yield will be sold at very high prices. The consumers are indirectly affected by paying huge sum. Therefore, a precaution has to be taken in order to prevent this disease by designing a system which monitors the real time data of the crops.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 3, January 2018

The proposed paper uses SIFT and Histogram-PCA algorithm, which is used to identify the leaf disease, by monitoring the leaf, is healthy or not. The SIFT algorithm is used to find the features of each dataset and compare the query image. In this dataset, the input image and query image key points are considered for matching. If the key points used are not matched then it is considered as unhealthy.

REFERENCES:

[1]"Explore Cornell - Home Gardening - Minimizing Diseases in Vegetable Gardens", *Gardening.cornell.edu*, 2016. [Online]. Available: http://www.gardening.cornell.edu/homegardening/scene7af3.html. [Accessed: 15- Jan- 2016].

[2]"Medicines from Plants", *Webcache.googleusercontent.com*, 2016. [Online]. Available: http://webcache.googleusercontent.com/search?q=cache:http://www.rainforesteducation.com/medicines/PlantMedicines/rfmedicines.htm&gws_rd=cr&ei=Zqw6V5LYI4rRvgSvrq7gCw. [Accessed: 08- Feb- 2016].

[3]"Agriculture in India", *Wikipedia*, 2016. [Online]. Available: https://en.wikipedia.org/wiki/Agriculture_in_India. [Accessed: 17- Jan- 2016].

[4]"The Most Popular Vegetables", *Ranker*, 2016. [Online].Available: http://www.ranker.com/crowdranked-list/the-most-delicious-vegetables-v1. [Accessed: 01- Feb- 2016].

[5] Christo Ananth, S.Shafiqa Shalaysha, M.Vaishnavi, J.Sasi Rabiyathul Sabena, A.P.L.Sangeetha, M.Santhi, "Realtime Monitoring Of Cardiac Patients At Distance Using Tarang Communication", International Journal of Innovative Research in Engineering & Science (IJIRES), Volume 9, Issue 3,September 2014,pp-15-20

[6]"3 Vegetable Seeds That Have a Large Yield Per Seed", *Gardening | Pioneer Thinking*, 2015. [Online]. Available: http://pioneerthinking.com/gardening/3-vegetable-seeds-that-have-a-large-yield-per-seed. [Accessed: 17- Feb- 2016].

[7]"AZ Master Gardener Manual: Potatoes", *Ag.arizona.edu*, 2016. [Online]. Available: https://ag.arizona.edu/pubs/garden/mg/vegetable/potatoes.html. [Accessed: 17- Feb- 2016].

[8]"Health Benefits of White Beans", *Healwithfood.org*, 2016. [Online]. Available: http://www.healwithfood.org/health-benefits/white-beans-navy.php. [Accessed: 24- Feb- 2016].

[9]"10 Healthy Reasons to Dig Into Red Potatoes", *Blackgoldfarms.com*, 2016. [Online]. Available: http://blackgoldfarms.com/community/the-dirt/2015/april/26/10-healthy-reasons-to-dig-into-red-potatoes/. [Accessed: 24- Feb- 2016].

[10]"10 Reasons Why You Should Be Eating More Tomatoes | Florida Tomato Committee", *Floridatomatoes.org*, 2016. [Online]. Available: http://www.floridatomatoes.org/news-events/10-reasons-why-you-should-be-eating-more-tomatoes/. [Accessed: 28- Feb- 2016].