International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-8, June 2020

Tomato Leaf Disease Detection using Back Propagation Neural Network

Roopali Gupta, ToranVerma

***Abstract:** Most of the Indian economy rely on agriculture, so identifying any diseases crop in early stages is very crucial as these diseases in plants causes a large drop in the production and economy of the farmers and therefore, degradation of the crop which emphasize on the early detection of the plant disease. These days, detection of plant diseases has become a hot topic in the area of interest of the researchers. Farmers followed a traditional approach for identifying and detecting diseases in plants with naked eyes, which didn't help much as the disease may have caused much damage to the plant. Tomato crop shares a huge portion of Indian cuisine and can be prone to various Air-Bourne and Soil-Bourne diseases. In this paper, we tried to automate the Tomato Plant Leaf disease detection by studying the various features of diseased and healthy leaves. The technique used is pattern recognition using Back-Propagation Neural network and comparing the results of this neural network on different features set. Several steps included are image acquisition, image pre-processing, features extraction, subset creation and BPNN classification.*

***Key Words—**Feature Extraction; Image Processing; Tomato Disease Detection; GLCM; BPNN.*

I. INTRODUCTION

India is well recognized for an agricultural economic system. About 70% of the population relies on agriculture. Farmers have wide variety of variation to prefer compatible vegetation for their farm. The continuous demand for food with an increasing population, discount rates in agricultural land, local weather change the agriculture enterprise continues to search for new methods to increase production. Consequently researchers from more than one discipline are browsing for ways to include new applied sciences and precision into the agronomic methods. There is a need for effective and special procedures of farming, enabling farmers to put minimal inputs for high production. Tomato is the main ingredient of Indian cuisine like Ketchups, puree, sausages, jam etc. This vegetable has high medicinal value and is highly rich in Vitamin A and Vitamin C and also contains minerals like iron and phosphorus. Orissa, Andhra Pradesh, Madhya Pradesh, Karnataka, Gujarat, Orissa, West Bengal, Chhattisgarh, Maharashtra, Bihar, Haryana, Uttar Pradesh, Telangana and Tamil Nadu[15].

Revised Manuscript Received on June 30, 2020.
* Correspondence Author

Roopali Gupta*, Department of Computer Science Engineering, Rungta College of Engineering and Technology, Raipur, C.G. India. Email: rplgpt07@gmail.com

Toran Verma, Department of Computer Science Engineering, Rungta College of Engineering and Technology, Raipur, C.G., India. Email: toran.verma@rungta.ac.in

The most affected phase for the disorder is the leaves of the plant. Around 80 to 90% of turmoil on the plant is on its leaves. The infection on the leaves may occur because of the air or soil having Bacteria, Fungus or viruses. It may be predicted with the help of modern technology. The Image processing technique is utilized in agricultural purposes for predicting plant disease from pictures of plants. Tomato crop faces many problems due to diseases which affect it a lot and it is not possible to identify easily. We require the system which can precisely separate the variety of characteristics present on leaf image and rely on that variety for further contrast and database for storing those picture characteristics and features. [15]

This work focuses on various features of the diseased image of the plant and compares which combination gives better result. In the early time detection of Plant Disease is of significant concern. Thus, there is a need to consider speedy, gradually moderate and careful method to subsequently perceive the strikingly reasonable features of the disease that show up on the plant leaf area. This connects with machine vision that is to give picture based distinctive verification. The target of this task is to see the plant leaf infection recognizing verification. Overall, there is a noticeable increase in the demand and production of this crop, indicating the continuously growing global market of Tomato. The point to be considered here is that a large amount of crop is affected by different types of diseases because of carelessness and untimely disease detection, which in turn results in the downfall of quality, quantity and thus, productivity. An infected plant if left untreated can make its neighboring plant suffer also. Pattern recognition technique has attracted various researchers in this field. On a whole, it can be realized that bare human eye simply isn't good enough to the point that it can see the infected piece of picture with variation in its features.

A. Types of Tomato Leaf Spot Diseases

The infections on a plant can be seen on its stem, leaves or its fruits which can be detected from the outer surface of the plant. Around 80 to 90 percent of diseases on the plant are on its leaves. Fig.1shows some examples of diseased and non-diseased Tomato leaves. So, importance of the leaf of the crop is of much importance than entire crop. The various diseases in tomato plant are caused by Bacteria, Fungus and Viruses [4]. These diseases can be categorized as:-

- *Bacterial Diseasess* :- Bacterial Spot (Xanthomonascampestrispv. vesicatoria), Bacterial Canker (Clavibactermichiganensis sub sp. michiganensis), etc. [4]

- *Fungal Diseases* :- Early Blight (Alternariasolani), Late Blight (Phytophthorainfestans), septoria leaf spot (Septorialycopersici), etc. [4]
- *Viral Diseases* :- Tomato Mosaic (TMV), Tomato Leaf Curl (TLCV) etc.[4]

(a)Normal Leaf (b)Bacterial Spot (c) Early Blight (d) Late Blight

Fig. 1 Tomato Leaves (a) Normal (b), (c), (d) Diseased.

Hence the Goalsand pursuits of this research work are:

i. Collection of Dataset containing diseased and healthy images.
ii. Converting all the images to same size of 256*256.
iii. Extracting various texture features from the gray scale variant of the above images.
iv. Creating various combinations of these features.
v. Applying Back Propagation Neural Network to train and classify these images to be healthy or diseased.
vi. Conclude depending on the results which combination works best.

Further, this paper is organized as follows. Section 2 presents numerous methods and details of the proposed scheme, Section 3 presents literature survey of the previous scheme, and Section4 presents propose work and Section 5 presents experiment result analysis and conclusionsof the study are presented in section 6.

II. TECHNIQUES USED

A. Feature Extraction

To distinguish one class of objects from other, extraction of features provides valuable data of significance. Features of image normally consist of color, texture and shape. In this work, texture features are taken into consideration. A total of 200 sample images of 5 different classes of diseased leaves i.e. Bacterial Spot, Early Blight, Late Blight, Leaf Mold including healthy leaves are observed and extracted to make classification model. Following 11 Features are extracted through image processing tool.

- *Contrast* is a measure of intensity of a pixel with respect to its neighbor over the entire image. In the visual observation of the real world, contrast is determined by the variation in the color and brightness of an object with respect to the other objects within the frame of reference. For a constant image contrast value is 0. Mathematical Equation of this feature is given by Equation 1:-

$$Contrast = \sum_{n=0}^{G-1} n^2 \left\{ \sum_{i=1}^{G} \sum_{j=1}^{G} P(i,j) \right\}, |i-j| = n \tag{1}$$

- *Correlation* is defined as the linear dependency of a gray level pixel with its neighboring pixel. For a perfectly positive and negative correlated image the correlation value is +1 and -1 and for a constant image the value is NaN. For Range = [-1, 1], the mathematical equation is given by Equation 2:-

$$Correlation = \frac{\sum_{i=1}^{G} \sum_{j=1}^{G} (i, \bar{x})(i, \bar{y}) P(i,j)}{\sigma_x \sigma_y}$$

Where

$$\bar{x} = \sum_{i}^{G} i \sum_{j}^{G} P(i,j), \bar{y} = \sum_{j}^{G} j \sum_{i}^{G} P(i,j), \tag{2}$$

- *Energy* in Texture Features is defined as the orderliness of the image. It is the summation of the square elements in GLCM. When the window is proficient orderly, energy value is high. It ranges from 0 to 1 and for a constant image energy value is 1. The Energy is given by Equation 3:-

$$Energy = \sum_{i,j=0}^{G-1} P^2{}_{i,j} \tag{3}$$

- *Homogeneity* is the value of tightness of distribution of the elements in the GLCM to its diagonal. For diagonal elements its value is 1 and it ranges from 0 to 1. The weights of Homogeneity elements are opposite as that of Contrast weights. These weights decreases exponentially as the elements go far from diagonal. Homogeneity is given by the Equation 4:-

$$Homogeneity = \sum_{i,j=0}^{G-1} \mathrm{P_{i,j}} / (1 + (\mathrm{i} - \mathrm{j})^2 \tag{4}$$

- *Mean* is the average value of the given data and is given by the Equation 5 below:-

$$Mean\ (\mu_i) = \sum_{i,j=1}^{G} i\,(P_{i,j}), (\mu_j) = \sum_{i,j=1}^{G} j\,(P_{i,j}) \tag{5}$$

- *Variance* is the sum of squares of differences between all numbers and means. It puts relatively high weights on the elements that differ from the average values.

Retrieval Number: H6531069820/2020©BEIESP
DOI: 10.35940/ijitee.H6531.069820
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The Equation 6 for the same is given below:-

$$Variance, \sigma_x^2 = \sum_{i=0}^{G}(i-\bar{x})^2 \sum_{j=0}^{G} P(i,j),\ \sigma_y^2 = \sum_{i=0}^{G}(i-\bar{y})^2 \sum_{j=0}^{G} P(i,j) \quad (6)$$

- *Standard Deviation* is square root of variance. It gives the extent to which the data varies from the mean. The Equation 7 for the same is given below:-

$$Standard\ Devaition, \sigma_i = \sqrt{\sigma_i^2}, \sigma_j = \sqrt{\sigma_j^2} \quad (7)$$

- *Entropy* is the measure of the randomness to characterize an image. Inconsistent image have low first order entropy, while a consistent image has high entropy. Entropy of an image is given the following equation 8:-

$$Entropy = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \times log\,(P(i,j)) \quad (8)$$

- *Root Mean Square* is defined as the square root of the mean of squares for a given set of numbers. RMS is given by the following Equation 9:-

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|x_n^2|} \quad (9)$$

- *Kurtosis* is a measure which determines whether the data is heavy-tailed or light-tailed with respect to the normal distribution. That is, data sets with high kurtosis incline towards heavy tails, or outliers and data sets with low kurtosis incline towards light tails, or lack of outliers. A uniform distribution would be the extreme case. For Normal Distribution the value of Kurtosis is 3. Kurtosis is given by the following Equation 10:-

$$Kurtosis(k) = \frac{E(x-\mu)^2}{\sigma^4} \quad (10)$$

- *Skewness* gives the measure of asymmetry. A distribution, or data set, is symmetric if it looks the same to the left and right from the center point. The Skewness of a normal Distribution is 0.Skewness is given by the following Equation 11:-

$$Skewness\ (s) = \frac{E(x-\mu)^3}{\sigma^3} \quad (11)$$

B. Back Propagation Neural network

Back Propagation Neural Network (BPNN) is a multilayer neural network consisting of an input layer, minimum one hidden layer and an output layer. As the name of the algorithm implies, the errors (hence, learning) propagate backwards from the output layer to the input layer. Therefore, the process of back propagation is used to estimate weights which minimize the errors in reference with the network's modifiable weights. As shown in the Fig. 2, the architecture of BPNN has three interconnected layers having weights on them. The working of BPNN is in two phases. One phase sends the signal from the input layer to the output layer, and the other phase back propagates the error from the output layer to the input layer.

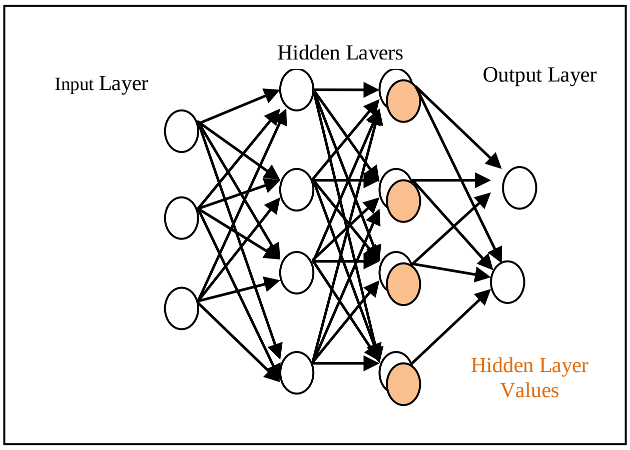

Fig. 2 Architecture of Back Propagation Neural Network

The steps involved in Back Propagation Neural Network are as follows:-

a. Initialize weights (to random values which are small) and biases in the network.
b. Propagate the inputs forward (by applying activation function)
c. Back Propagate the error (by uploading weights and biases)
d. Stopping condition when the error is decreased.

All of the above steps will be concluded in the algorithm as follows:-

Step 1. Initialization of weights to small random values.
Step 2. While stopping condition is false, then repeat steps 3 to 10.
Step 3. For each training pair repeat steps 4 to 9.
Step 4. Each input unit receives the input signal xi and transmits it to all the units in the hidden layer.
Step 5. Each hidden unit (Z_h, h= 1, 2 …p) sums its input signal.

$$Z_{inj} = V_{oj} + \sum x_i v_{ij}$$

Applying the activation function

$$Z_j = f(Z_{inj})$$

Now these signals are sent to the units in the output layer.

Step 6. Each output unit (Yk, k=1, 2 ….m) sums its weighted input signals

$$Y_{ink} = W_{ok} + \sum Z_j W_{ij}$$

Now, applying the activation function to

Calculate the output signals

$$Y_k = f(Y_{ink})$$

Step 7. Each output unit (Y_k, k=1, 2 ….m) receives a target pattern corresponding to an input pattern, error information (δ) is then calculated as:-

$$\delta_k = (t_k - Y_k) f(Y_{ink})$$

Step 8. Each hidden unit (Z_j, j= 1,2 ….n) sums its delta input units as:

$$\delta_k = \sum \delta_j W_{jk}$$

And the error information is calculated as

$$\delta_k = \delta_j f\left(Z_{inj}\right)$$

Step 9. Each output unit (Y_k, k=1, 2 ….m) updates its bias and weights (j = 0 to n)

The weight correction is given by

$$\Delta W_{jk} = \alpha \delta_k Z_j$$

And bias correction is givenby:

$$\Delta W_{ok} = \alpha \delta_k$$

Therefore,

$$W_{jk(new)} = W_{jk(old)} + \Delta W_{jk}$$

$$W_{ok(new)} = W_{ok(old)} + \Delta W_{ok}$$

Each hidden unit ($Z_{i,}$ j = 1,2 …. p) updates the bias and weights (i = 0,1, ….n)

The weight correction is given by

$$\Delta V_{jk} = \alpha \delta_j x_j$$

And similarly the bias correction is given by

$$\Delta V_{jk} = \alpha \delta_j$$

$$V_{ij(new)} = V_{ij(old)} + \Delta V_{ij}$$

$$V_{oj(new)} = V_{oj(old)} + \Delta V_{oj}$$

Step 10. Testing for the stopping condition, i.e. minimization of the errors.

Back Propagation Neural Network some most prominent advantages which are listed below:

- It is fast, simple and easy to implement.
- Can reduce the computing time if weights chosen initially are small.
- It is a flexible method as it does not need previous knowledge about the network.
- It is a standard technique that in general works well.
- It does not need any particular mention of the features of the function which are to be learned.

III. LITERATURE REVIEW

A. Overview

Fuentes et al. (2016) analyses the importance of the detection of diseases in plants as a strategy to reduce production and economic losses in the agriculture sector. According to their study due to the advances in the technology era, recent works in the agricultural area have proposed the use of nondestructive methods to detect diseases in plants; therefore, Techniques based on images are used to detect diseases without causing any secondary impact in the plant. Their study describes the challenge of accuracy in the computer vision and the use of proper data as an important issue of each intelligent system. Several diseases that affect tomato plants and their characteristics are the main part of this study. Furthermore, some samples of the collected dataset which contains different diseases and pests of tomato plants are shown. [8] Gavhale and Gawande (2014) define diseases in plants as a major cause of production and economic losses which results to reduction in both quality and quantity of agricultural products. In their research, they reviewed the need of simple plant leaves disease detection system that would facilitate advancements in agriculture. Early information on crop health and disease detection can facilitate the control of diseases through proper management strategies. This technique will improves productivity of crops. This paper also compares the benefits and limitations of these potential methods. It includes several steps viz. image acquisition, image pre-processing, features extraction and neural network based classification. [9] Gurle et al. (2019) worked on various mechanisms of Tomato Plant Disease Detection and Pest Management as India is an agricultural country and most of people wherein about 70% depends on agriculture. So, disease detection is very important research topic. Their work defines many species of tomato diseases and pests, the pathology of which is complex. According to them, Crop diseases are a major threat to crop production, but their identification remains difficult in many parts of India due to the lack of the necessary infrastructure. Thus, it is difficult and error-prone to simply rely on manual identification. Recent advances in computer vision with the help of deep learning have made the way possible for automatic disease detection. In this article, the authors have analyzed a method of disease detection and pest management using convolution neural networks (CNN), K-means clustering, and acoustic emission. [11] Iqbal et al. (2018)worked onPlant Disease recognition system and classification which plays a vital role in agriculture field. According to their work, Product quality, quantity or productivity of plants is harshly affected by slight negligence in this domain; therefore, huge amount of human work load for crops intensive care in vast farms can be reduced by automatic system accomplished of perceiving and classifying plant disease at early phases. This paper presents image processing framework for plant disease identification and classification. The proposed workflow consists of three stages that are images segmentation, feature extraction and classification. For segmentation Multithresholding technique is used then that of other common segmentation techniques. GLCM for feature extraction for pixel level information and Support Vector Machines for classification due to its robustness and optimality are used. Experiments were conducted on Tomato leaf dataset comprising of 4 different classes. Proposed framework achieved 98.3% overall accuracy with 10-fold cross validation. [12] Karmarkar et al., (2018) discussed various types of diseases on Tomato Fruit. Here, images are captured through Raspberry PI camera and are interfaced with Raspberry PI 3B Model; Images are segmented by using k-means clustering algorithm, then the conversion of the images into HSV color model is done from RGB for getting the region of interest only. As future scope, the author plans to use texture features in this system and SVM Classifier for detecting the types of images. [13] Patil and Kumar, (2011)proposed a theory that parameters like rainfall, temperature; and other infections on crops are the essential points which influence production quality and the quantity. So, for managing the diseases in plant leaves this research focuses on extracting color features of the leaves. Mean Standard Deviation and Skewness of the images of the diseases leaves are calculated and evaluated. As future work they proposed to give a content based retrieval system using Indexing. [18]

Rani and Nagaraj(2016) have proposed a technique to detect Bacterial Spot Disease in Tomato Plant using SOM based clustering. In their research work, Tomato plant pictures were acquired from field after which image preprocessing was done using four methods Gamma correction, Color space conversion, Down sampling and Cutting for center parts is done after which images were segmented using SOM based clustering method. To overcome the limitations of Centroid-based K-means clustering, Self-Organizing Maps (SOM) is introduced for achieving effective classification result and to improve the detection performance. [19] Sivagami and Mohanapriya(2019)have describedTomato Crop as one of the most important vegetable to use in the entire world. Disease easily affects the tomato plant due to insects and nutrient deficiency. Their main focus was to detect nutrient deficiency using image segmentation and classification. By detecting these deficiencies at an early stage the yields can be increased and the disease caused due to lack of nutrient deficiency can be also reduced. Their research work uses k-means and Expectation maximization segmentation algorithms for segmentation and SVM classifier for classification. Based on the results Expectation Maximization provides better result than K-means segmentation. [23] Vetal and R.S., (2017) worked on detecting Tomato plant Leaf Diseases using Image Processing Technique. Their work identifies four key diseases using image segmentation and Multi-class SVM algorithm. For parting of damaged area on leaves, image segmentation is used and for classification, Multi-class Support Vector Machine(SVM) algorithm is used, Texture Features like Energy, Entropy, Correlation and Homogeneity are determined on Tomato leaf diseases like Early Blight, Bacterial Spot, Septoria Leaf spot and Iron Chlorosis. The claimed accuracy is 93.75%. [24]

B. Outcomes of Literature Review

Developing countries like India the financial system is often relies on agriculture. Because of plant Diseases the quality and quantity of agriculture product is diminished. Disease to the plants prompted typically through the leaf, stem and fruit. The intent of agriculture isn't simplest to feed ever increasing populace but it is as significant basis of vigor and an option to remedy the crisis of worldwide warming. Plant Disease Detection is very important in early stage in order to treat and manage the disorder. Based on the various research works done in this field, many researchers have used Preprocessing, segmentation, training and classification for disease detection processes. Different procedures of color conversions were also taken into consideration or the same.Feature Extraction was of utmost importance for determining the diseases in leaves, stems or other parts of the plants and fruits. A lot of work is done to extract color, texture and shape features of the images of the diseased plants. For training purpose mainly Back Propagation Neural Networks and SOM are used and for classification also BPNN and SVM with multiple kernels were utilized. These methodsareuseful for determining any deterioration in the crops which can be useful for both farmers and good production. Relying on the purposes, many techniques were proposed to clear up or at the least to minimize the issues, by using utilizing Image Processing, Pattern Recognition and some automatic classification instruments. Within the next section this paper tries to present these proposed methods in meaningful approach.

IV. PROPOSE WORK

A. Problem statement

Agriculture is one of the most significant sectors of Indian Economy. Almost 50% of the total employment of the country's workforce is provided by the Agriculture sector. India stands amongst the world's largest producer of many crops like pulses, rice, wheat, spices, tomatoes and other crops. A Farmer's financial condition depends on the quality of the crops which they produce, which in turn relies on the plant's growth and their harvest. Therefore, in agriculture sector, Plant Disease detection plays an important role. Plants leaves, being exposed to environment are highly prone to diseases that affect the plant's growth, which in turn affects the productivity of the crop. In order to detect a plant leave's disease early and at initial stage, Automatic Disease Detection Techniques can prove to be very valuable. The symptoms of plant diseases are conspicuous in different parts of a plant such as leaves, stem etc.

Detection of plant disease with naked eyes is a very tough job and gives less accurate results. Hence, it is required to develop computational methods which will make the process of disease detection and classification using leaf images automatic. The primary purpose of thesis is to develop an efficient methodology for the Tomato Leaf disease detection by extracting various features of the images collected.

B. Proposed Methodology

Any type of diseases on the Tomato Plant diminishes the productivity of the crop. Thus, to increase the productivity, Image processing system is used for distinguishing diseases on Tomato Plant leaves. The leaves are categorized as healthy and unhealthy leaves. This Tomato Leaf Disease Recognition System is used to detect and identify the patterns on diseased leaf with the help of Feature Extraction and Back Propagation Neural Network. Various Texture Features of the image are extracted and analyzed for the studying which set of features are dominant enough to give accurate results. The steps of the Disease Detection Process are given below:-

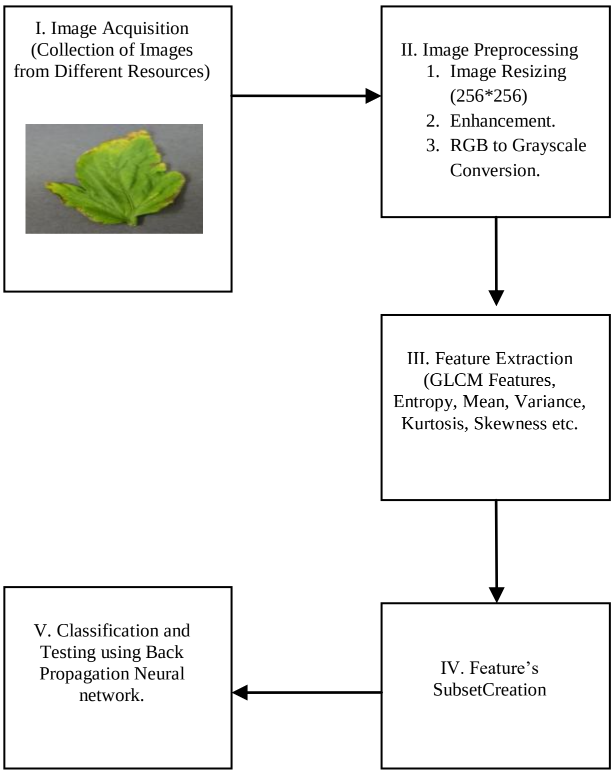

Fig. 3 Block Diagram of Tomato Leaf Disease Recognition System

The proposed work for Tomato leaf disease detection determines the best set of features to be used in the study. This methodology is mainly divided into 5 steps which are listed below:-

i. Image Acquisition.
ii. Image Preprocessing.
iii. Feature Extraction.
iv. Features' Sub-set creation.
v. Neural Network Training and Classification.

C. Proposed Algorithm

1. Firstly the images are collected from dataset.
2. All Collected images are resized to 256*256 pixels. Contrast Stretching is done to enhance the original image.
3. Then the image is converted from RGB toGray Scale.
4. Texture Features of the image are extracted and stored in the Database.
5. As the Features are extracted following 5 subsets of the features are created and stored separately.
   a. (Contrast, Correlation, Energy, Homogeneity, Mean, Variance, Standard Deviation, Entropy, Root Mean Square, Kurtosis and Skewness).
   b. (Contrast, Correlation, Energy and Homogeneity).
   c. (Entropy and Kurtosis).
   d. (Entropy and Skewness).
   e. (Entropy, Kurtosis and Skewness).
6. These vectors are trained and classified using Back Propagation Neural Network.
7. The accuracy of these feature subsets are then evaluated and compared against each other.
8. Best combination is then determined on the basis of the results obtained.

D. Flow Chart

The Flowchart of the proposed Algorithm is given by Figure 4:-

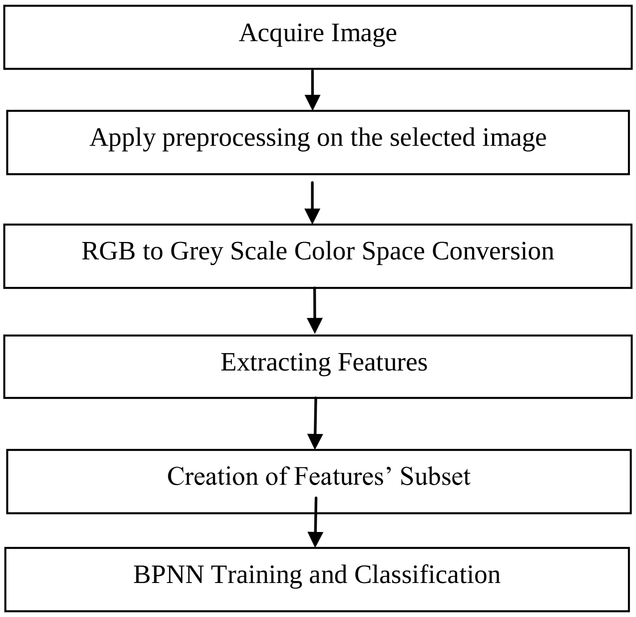

Fig. 4 Flow chart of proposed work

V. EXPERIMENT & RESULT ANALYSIS

For our system, the hardware configuration is a PC with an Intel Pentium Dual CPU T3400 @2.16 (GHz) CPU with 3GB RAM. The Operating system used Windows 7 with MATLAB installed. .

Here, we provide the outcomes of different experiments conducted on different features set of the images database. In order to be assured that it is resistant to different situations, we test them on different datasets. Afterward, we provide the quantitative and qualitative results of our experiments. Colored pictures are considered for estimation process. The algorithm is designed on MATLABR2017a using Image Processing toolbox.

Some Sample images are shown in Fig. 5.1

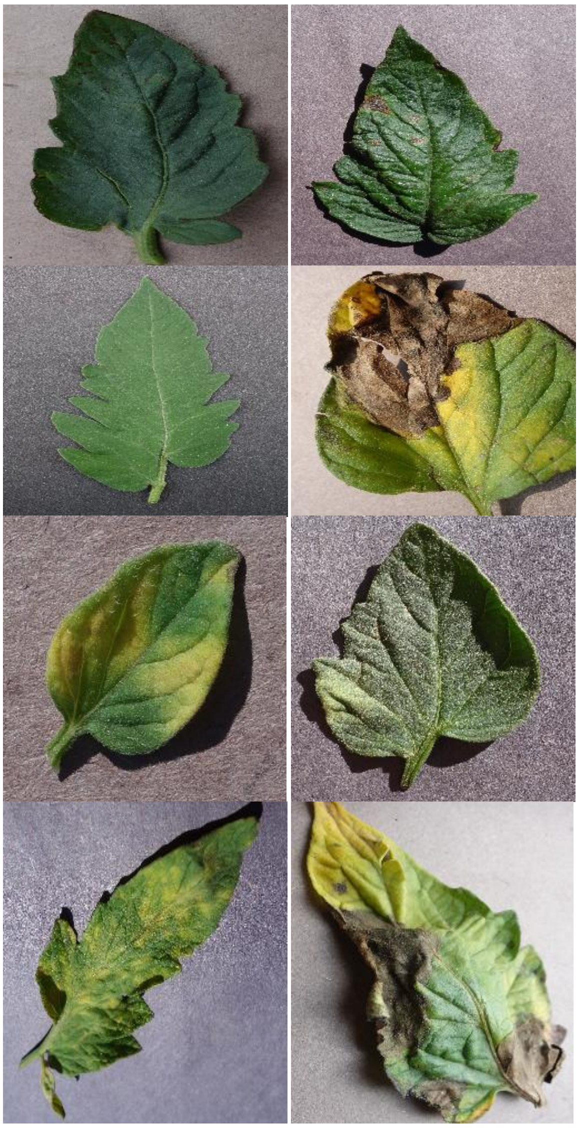

Fig. 5.1 Sample images from the Database.

In this operation, different features of 5 distinguished sets, each of 200 images are analyzed. The output process is as follows:-

- Image from the database is selected via browsing window. The selected images are shown in Fig 5.2(a) and Fig 5.2(b).

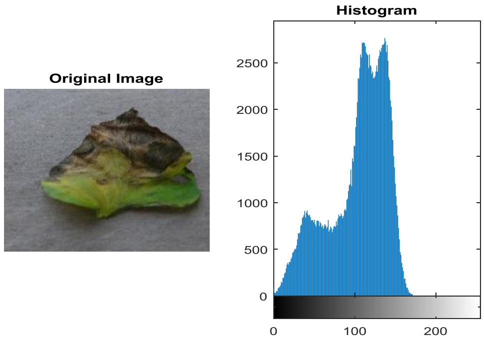

Fig. 5.2(a) Original Image 1

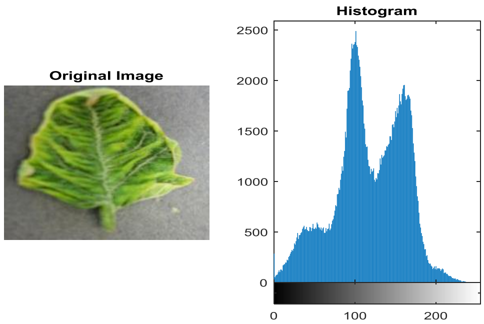

Fig. 5.2(b)Original Image 2

- Browsed image is then preprocessed for Contrast Stretching, Enhanced images are shown in Fig 5.3(a) and Fig 5.3(b).

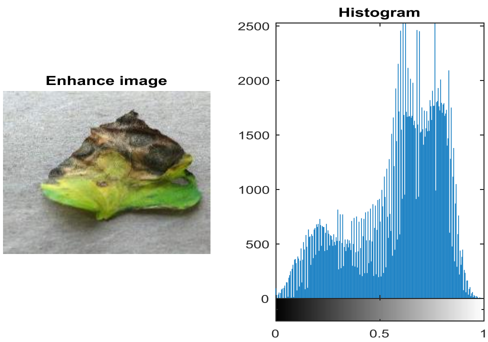

Fig. 5.3(a) Preprocessed Image 1

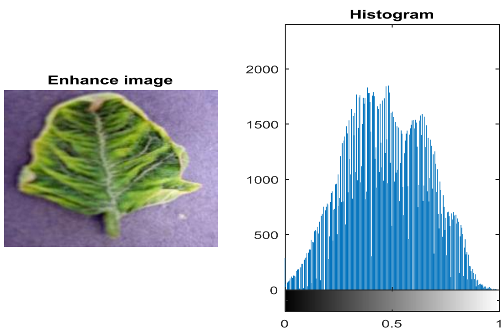

Fig. 5.3(b) Preprocessed Image 2

- The images in RGB are then converted to Grayscale color space. The converted images are shown in Fig 5.4(a) and Fig 5.4(b)

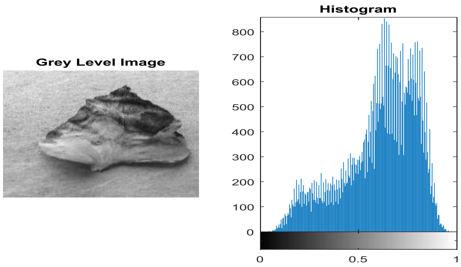

Fig. 5.4(a) Grey Scale Image 1

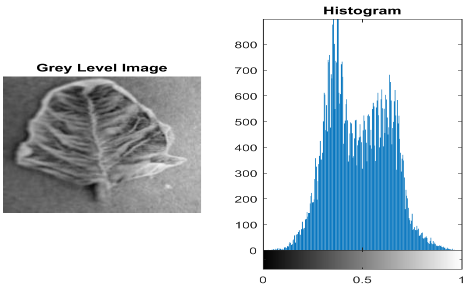

Fig. 5.4(b) Grey Scale Image 2

- This grayscale image is used to calculate 11 different features i.e. Contrast, Correlation, Energy, Homogeneity, Mean, Variance, Standard Deviation, Entropy, Root Mean Square, Kurtosis and Skewness from 5 categories of 1000 images containing 1 set of Healthy and 4 sets of diseased leaves. All theseTexture Features are extractedusing Image Processing Toolbox in Matlab. Sample features of 50 images are taken and their 11 Features are depicted in Graph Form with the help of Fig 5.5.

- 5 Subset of these features are created, then correct and Incorrect Classification is calculated using confusion matrix and then compared against each other. The correct classification of Total (11 Features set), (GLCM Features set), (Entropy and Kurtosis), (Entropy and Skewness set) & (Entropy, Kurtosis and Skewness set) is 91%, 76.5%, 60.8%, 59.9% & 70.9 % respectively. This comparison is given by Table I and the comparative graph is shown in Fig 5.6 as below:-

Table I Comparison of Different Features' Subset

S. N.	Feature Subset	Correct Classifi - cation %	Incorrect Classifi- cation%
1	11 Features	91.0	9.0
2	GLCM Features	76.5	23.5
3	Entropy and Kurtosis	60.8	39.2
4	Entropy and Skewness	59.9	40.1
5	Entropy, Kurtosis and Skewness	70.9	29.1

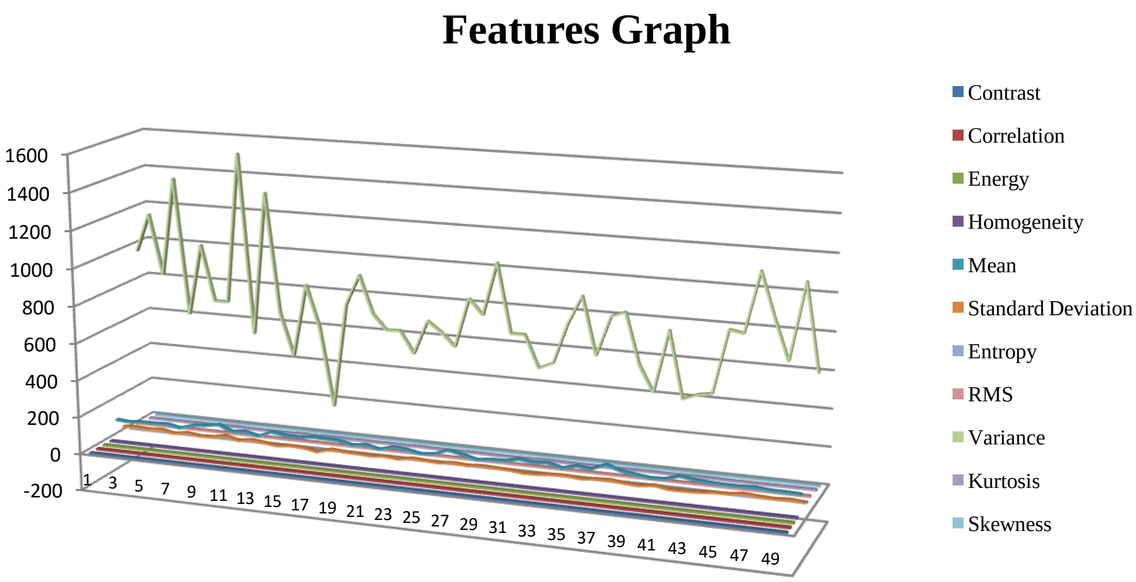

Fig. 5.5 Graph Showing 11 Features

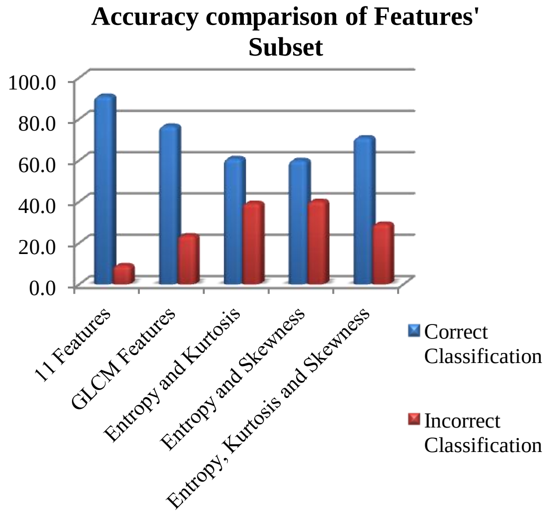

Fig. 5.6 Comparison Chart for Accuracy of 5 Subsets

- As we can see, the best results are obtained by extracting all the 11 features of the image with correct classification of 91%. Thus, the Back Propagation Neural Networktool, Confusion Matrix, Performance Plot, Training plot and Error Histogram of Feature Extraction of 11 Features of the images is represented by Fig 5.7 (a), 5.7 (b), 5.7 (c), 5.7 (d) & 5.7 (e) respectively.

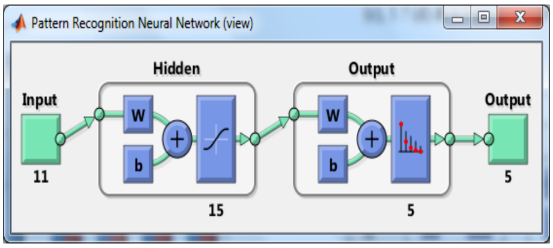

Fig. 5.7 (a) Neural Network Tool

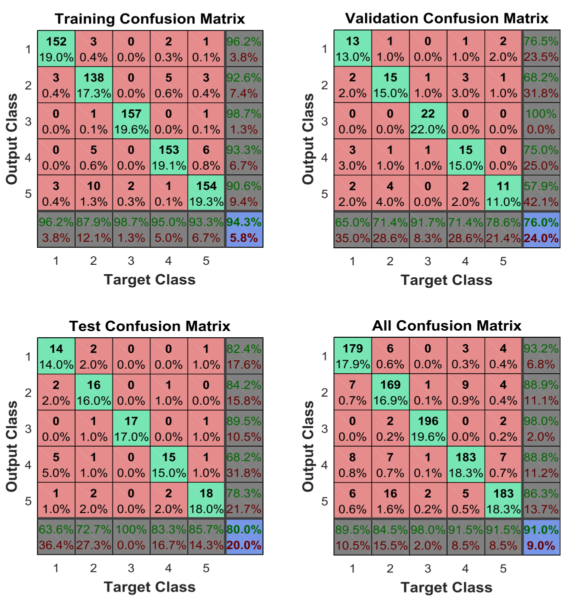

Fig. 5.7 (b) Confusion Matrix

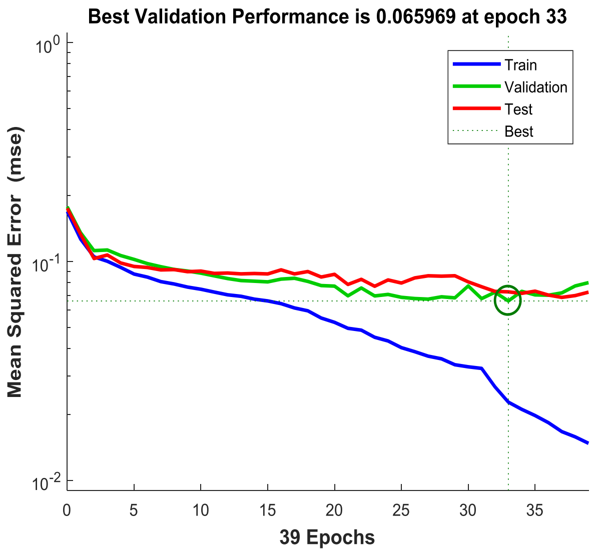

Fig. 5.7 (c) Performance Plot

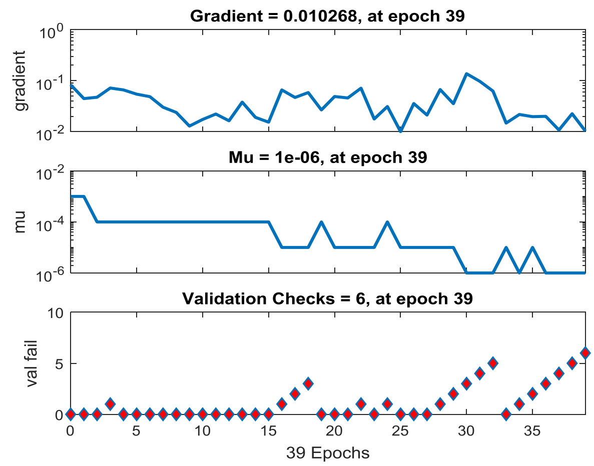

Fig. 5.7 (d) Training Plot

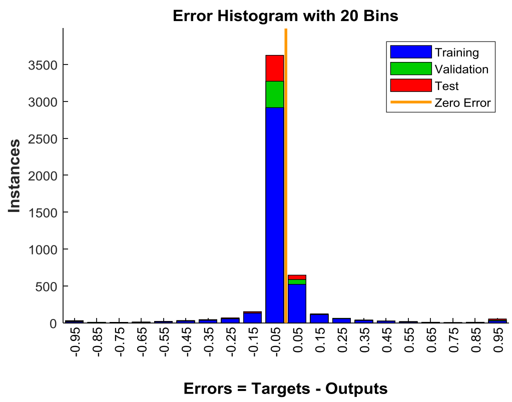

Fig. 5.7 (e) Error Histogram

VI. CONCLUSION AND FUTURE WORK

This work consists of identifying the affected part of the diseased leaf. 5 Types of Tomato leaf Diseases are taken into consideration with 200 images each. The 11 Features: - Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, RMS, Variance, Kurtosis and Skewness are extracted for each image.

5 Subsets of these features are created and stored in the database. We used Back Propagation Neural network training and classification in which Extracted features set is given as input and a unique target set is used for validation, 80% of the total images are used for training, 10% for validation and 10% for testing. The results obtained from 5 types of Features set are compared and analyzed using confusion matrix. In this paper, we tended to know which set of features gives the more accurate results and thus help in identifying the leaves with affected disease. Here, Feature set of all 11 Features gives more accurate results with accuracy of 91%. Therefore, this set of features can be more prominently used for detecting the diseased part of the leaves and reducing computational time and efforts.

Some future works are to be done in order for improvement and extension of our work. For various challenging scenarios, the proposed system still needs to be improved to allow better performance. It can be done in terms of analyzing the results with different classification approaches like Support Vector Machines, Deep Neural Network etc. Thus, for future work, we intend to address the above-mentioned problems.

REFERENCES

1. Adhikari, S., Kumar, S., Shrestha, B., Baiju, B. and Kumar, S. (2018). Tomato Plant Diseases Detection System Using Image Processing. Kantipur Engineering College (KEC) Conference Proceedings, 1(1), pp. 81–86.
2. Ashok, V. and Vinod, D.S. (2014). Automatic Quality Evaluation of Fruits Using Probabilistic Neural Network Approach. In: International Conference on Contemporary Computing and Informatics (IC3I). Mysore, India: IEEE, pp. 308–311.
3. Bhimte, N.R. and Thool, V.R. (2018). Diseases Detection of Cotton Leaf Spot Using Image Processing and SVM Classifier. In: Second International Conference on Intelligent Computing and Control Systems (ICICCS). Madurai, India: IEEE, pp.340–344.
4. Blake, J.H., Kluepfel, M. and P. Keinath, A. (2018). Tomato Diseases & Disorders. [online] Home & Garden Information Center | Clemson University, South Carolina. Available at: https://hgic.clemson.edu/factsheet/tomato-diseases-disorders/ [Accessed Nov. 2019].
5. Brahimi, M., Boukhalfa, K. and Moussaoui, A. (2017). Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. Applied Artificial Intelligence, 31(4), pp. 299–315.
6. Dubey, Y.K., Mushrif, M.M. and Tiple, S. (2018). Superpixel Based Roughness Measure for Cotton Leaf Diseases Detection and Classification. In: 4th International Conference on Recent Advances in Information Technology (RAIT). Dhanbad, India: IEEE, pp.1–5.
7. Ferentinos, K.P. (2018). Deep Learning Models For Plant Disease Detection and Diagnosis. Computers and Electronics in Agriculture, 145, pp. 311–318.
8. Fuentes, A., Youngki, H., Park, D.S., Yoon, S. and Lee, Y. (2016). Characteristics of Tomato Plant Diseases - A Study for Tomato Plant Disease Identification. International Symposium on Information Technology Convergence (ISITC). pp. 226–230.
9. Gavhale, K.R. and Gawande, U. (2014). An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques. IOSR Journal of Computer Engineering (IOSR-JCE), 16(1), pp. 10–16.
10. Gavhale, K.R., Gawande, U. and Hajari, K.O. (2014). Unhealthy Region of Citrus Leaf Detection Using Image Processing techniques. In: International Conference for Convergence for Technology (ICCT). Pune, India: IEEE, pp.1–6.
11. Gurle, A.S., Barathe, S.N., Gangule, R.S., Jagtap, S.D. and Patankar, T. (2019). Survey Paper on Tomato Crop Disease Detection and Pest Management. International Journal of Applied Evolutionary Computation (IJAEC), 10(3), pp. 10–18.
12. Iqbal, J., Adnan, S.M., Ahmad, R.W., Aziz, S., Ismail, W. and Din, Z. ud (2018). Classification of Tomato Plants’ Leaf Diseases Using Image Segmentation and SVM. Technical Journal, 23(02), pp. 81–88.
13. Karmarkar, S., Jadhav, G., Parkhe, M., Kadam, A. and Patel, H. (2018). Tomato Disease Detection using Image Processing. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE), 7(12), pp .4212–4220.
14. Krishnaveni, M. and Radha, V. (2013). View-based Signer Independent Approach for Indian Sign Language Recognition Using Extreme Learning Machine. [Online] pp.107–121. Available at: http://shodhganga.inflibnet.ac.in:8080/jspui/handle/10603/13306 [Accessed Nov. 2019].
15. Monthly Report Tomato. (2018). [online] agricoop.gov.in, Ministry of Agriculture and Farmers welfare, Govt. of India, pp.1–7. Available at: http://agricoop.gov.in/horticulture-reports?page=6 [Accessed Nov. 2019].
16. Narmadha, R.P. and Arulvadivu, G. (2017). Detection and Measurement of Paddy Leaf Disease Symptoms Using Image Processing. In: International Conference on Computer Communication and Informatics (ICCCI). Coimbatore, India: IEEE, pp. 1–4.
17. Narvekar, P. and Patil, S.N. (2015). Novel Algorithm for Grape Leaf Diseases Detection. International Journal of Engineering Research and General Science, 3(1), pp. 1240–1244.
18. Patil, J.K. and Kumar, R. (2011). Color Feature Extraction of Tomato Leaf Diseases. International Journal of Engineering Trends and Technology (IJETT), 2(2), pp. 72–74.
19. Rani, X.A.K. and Nagaraj, R. (2016). SOM Based Clustering for Detecting Bacterial Spot Disease in Tomato Field. Indian Journal of Innovations and Developments (IJID), 5(7), pp. 1–8.
20. Sarangdhar, A.A. and Pawar, V.R. (2017). Machine Learning Regression Technique for Cotton Leaf Disease Detection and Controlling Using IoT. In: International conference of Electronics, Communication and Aerospace Technology (ICECA). Coimbatore, India: IEEE, pp.449–454.
21. Shah, N. and Jain, S. (2019). Detection of Disease in Cotton Leaf using Artificial Neural Network. In: Amity International Conference on Artificial Intelligence (AICAI). Dubai, United Arab Emirates: IEEE, pp.473–476.
22. Sharma, K., Priyanka, Kalash, A. and Saini, K. (2015). GLCM and its Features. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), 4(8), pp. 2180–2182.
23. Sivagami, S. and Mohanapriya, S. (2019). Automatic Detection of Tomato Leaf Deficiency and its Result of Disease Occurrence through Image Processing. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(11), pp. 4165–4172.
24. Vetal, S. and R.S., K. (2017). Tomato Plant Disease Detection using Image Processing. International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 6(6), pp. 293–297.
25. Zhang, S., Wang, H., Huang, W. and You, Z. (2018). Plant Diseased Leaf Segmentation and Recognition by Fusion of Superpixel, K-means and PHOG. Optik, 157, pp. 866–872.

AUTHORS PROFILE

Roopali Gupta is a M. Tech. (C.S.E.), Research Scholar of Rungta College of Engineering and Technology and has received her Bachelor of Engineering (B.E.) Degree in Information technology in the year 2012 from C.S.V.T.U., Bhilai, India Currently working in the area of Image Processing and Data Mining.

Toran Verma is working as a Profesoor and Head of Computer Science and Engineering Dept., Rungta College of Engineering and Technology, C.S.V.T.U., Bhilai, India. He has completed his graduation in Computer Science & Engineering from Barkatullah University-Bhopal, India. His area of research includes soft computing; image processing and computer linguistics.