Accurate Detection and Quantization of Leaf-Diseases through Soft Computing

Divya Srivastava¹, Rajitha B.¹ and Suneeta Agarwal¹
¹Motilal Nehru National Institute of Technology, Allahabad, India

Email: divyalknw@gmail.com

Abstract: Diseases in leaves can cause the significant reduction in both quality and quantity of agricultural production. If early and accurate detection of disease/diseases in leaves can be automated, then the proper remedy can be taken timely. A simple and computationally efficient approach is presented in this paper for disease/diseases detection on leaves. Only detecting the disease is not beneficial without knowing the stage of disease thus the paper also determine the stage of disease/diseases by quantizing the affected of the leaves by using digital image processing and machine learning. Though there exists a variety of diseases on leaves, but the bacterial and fungal spots (Early Scorch, Late Scorch, and Leaf Spot) are the most prominent diseases found on leaves. Keeping this in mind the paper deals with the detection of Bacterial Blight and Fungal Spot both at an early stage (Early Scorch) and late stage (Late Scorch) on the variety of leaves. The proposed approach is divided into two phases, in the first phase, it identifies one or more disease/diseases existing on leaves. In the second phase, amount of area affected by the disease/diseases is calculated. The experimental results obtained showed 97% accuracy using the proposed approach.

Keywords: Bacterial Blight, Fungal Spot, 3-means clustering, Texture, Morphology

1 Introduction

Initially, agriculture, used to be a term for feeding the country's population but now it has increased its scope in various areas. Plants now play a vital role in economic development, global warming, ecological balance, etc. Researchers from various countries have proved that the diseases on plants may cause a great loss to the country's economy as it is estimated that in the United States the crop losses due to plant pathogens result in about 33 billion dollars every year as investigated by Roberts, M.J., Schimmelpfennig, et al.[1]. The bacterial, fungal and viral infections are the most common diseases of plants causing severe damage to them. A survey by Pimentel, D., Zuniga, R., et al.[2] showed that there are about 50,000 parasitic and non-parasitic plant diseases around the world. Symptoms of infection can be seen in various parts of the plant-stem, roots, leaves, buds, etc. Visually it is more prominent on leaves as compared to stems or roots. The naked eye investigation by the experts has been the main approach in practice for the detection and identification of plant disease. Manual investigation or diseases have a lot many challenges like: 1) A large team of experts is required. 2) Continuous monitoring of plants is required. 3) Illiterate farmers do not have the proper knowledge to contact suitable experts within time. 4) Consulting experts are expensive and also time-consuming.

Diagnosing the diseases accurately and timely is utmost important. For this, Digital Image processing can be used: 1) To detect and identify diseases in plants, 2) To quantify the diseases by calculating the area affected which further helps in taking the suitable remedy.

The diseases in plants can be classified on various bases like: 1) By host plants (cereal diseases, vegetable diseases, etc). 2) By crops (Wheat, Rice, Grapes, etc). 3) By location
(along the edges of leaf, aerial part of leaf, stem, pods, etc). 4) By color change (Scorch, Bacterial Blight, Fungal Spot, Rust). 5) By climatic conditions. 6) By pathogens, etc.

This paper deals with the disease/diseases mainly found on leaves based on the color change. Early scorch, late scorch, fungal spot, bacterial blight are the four types of diseases that are taken into consideration. The detected diseases are quantified and also arranged concerning their levels.

Remaining part of the paper is arranged in the following manner, Section 2 gives a background of the related work done. Section 3 gives an overview of the symptoms of diseases classified. Section 4 deals with the proposed approach. Section 5 discusses the experimental results. Section 6 is the concluding one.

2 Literature Review

Researchers have used different classifiers to automate the system to detect diseases in plants. Artificial Neural Networks, Decision Tree, K-means, K-Nearest s and Support Vector Machine are the most wide machine learning methods used in this. Since plants have a very wide range of variety and so the diseases, most of the researchers have limited themselves either to diseases found on specific plants/crops or specific diseases (bacterial, fungal, etc) on particular parts of the plant: leaves, stems, roots, etc. Most of the work has been done on particular crops like tomatoes, potatoes, rice, etc. Gutiérrez, P.A., López-Granados, et al. [3] and Wang, X., et al. [4] predicted Phytophthora infection on tomatoes by using ANNs, work has also been done on diseases in potatoes by Vleeshouwers, Vivianne GAA, et al. [5] and on soybeans by Dou, Daolong, et al.[6]. Camargo, A., Smith, J. Et al. [7] identified visual symptoms of cotton diseases. Automatic disease detection in grapevine leaves and comparison with symptoms detected with plant pathologist is done by Oberti, Roberto, et al. [8]. Diseases are also sometimes categorized according to the nature of the plant like diseases found common in citrus plants for such diseases work is discussed by Oberti, Roberto, et al. [8] and Guo, Y., P. W. Woods et al. [9].

In image processing, most of the researchers have worked based upon the visual symptoms of diseases in plants either seen on leaves, roots, stems, etc. Most of the algorithms have used the common steps for determining the diseases in plants.

A survey paper by Bertolini, E., et al. [10] presents different classification techniques used for detection of disease in leaves. This paper suggests that k-nearest neighbour method is simple and suitable for segmentation and Support Vector Machine (SVM) is most suitable for the prediction of disease.

The paper by Ghaiwat, Savita N. et al. [11], proposed the detection of disease on the basis of color by masking and removing the green pixels and then the pixels of the pre-computed image is compared with the threshold value for classification of disease. Dhaygude, Sanjay B. et al. in [12] have used the color co-occurrence matrix for the classification of diseases in leaves. Arivazhagan, S., et al. [13], used texture feature for the detection of unhealthy region in leaves. Anand. H. Kulkarni and Ashwin Patil [14] and Bashir, Sabah, et al. [15] have compared different classifiers for the classification of diseases in leaves using image processing techniques.

The basic difference in state-of-art approaches is the different methods used for segmentation either by using k-means clustering, nearest-neighbor, ostu' method, etc. Most of the researchers have preferred SVM over other classifiers as it showed the better accuracy in comparison to others.
3 Symptoms of the Leaves Diseases
There exists a huge variety of plant pathogenic, and so a variety of bacterial and fungal diseases. The proposed approach focuses mainly on four types of diseases commonly seen on leaves: Fungal Spot both at early and late stage, Bacterial Blight and Leaf Spot as shown in figure 1. The symptoms of each disease are discussed in the following subsections.

![Symptoms of diseases in leaves](image)

**Figure 1.** Symptoms of diseases in leaves

3.1 Bacterial Blight
Bacteria [16] are single-celled and very small (microscopic) organisms. Their life cycle is rapid and is very short. They attack the plants at all the stages of growth and are generally seen on leaves, pods, and fruits. Beans, cotton, tomato, rice, pepper, cassava, etc are it’s host plants. As the bacteria need moisture to spread, so they are most common in wet climates. Initially bacterial infection begins as a small water-soaked spots shown in figure 2 which later turns yellow and then brown as the tissue dies.

![Bacterial Blight](image)

**Figure 2.** Bacterial Blight

3.2 Fungal Diseases
Fungi are very common parasites on plants causing Fungal Diseases [17]. They are very small in size or say microscopic in nature. Fungi produce spores which may be carried from plant to plant by wind, water, and insects. These fungal spores are spread through moisture and are common during wet, humid seasons. Leaves are the most common area of the plant affected by fungal disease. They appear as small spots better known as frog eye, rust, alternaria alternata, early scorch, late scorch, anthracnose, powdery mildew, etc.

There are various fungi organisms that cause fungal diseases in plants. The fungal diseases cause the destruction and death of the plant tissues. Different fungi organisms cause different types of diseases like leaf spot, scorch, anthracnose, cankers, blasts, etc. All the diseases have
different symptoms. This paper discusses the scorch both at the early and late stages, and the fungal spot i.e. Leaf spot.

At first, early scorch generally appears on leaves as small and irregular spots starting from color yellow then brown and later on dark-brown, or black spots. Along with the color change the spots also expands in size and tends to cover the whole affected area. They generally grow along the edges of the leaves [18] as shown in figure 3 (a) Early Scorch and (b) Late Scorch. Early scorch and late scorch can appear simultaneously on the same leaf as shown in figure 4.

Figure 3. Early and late Scorch

Figure 4. Early and Late Scorch in same leaf

4 Proposed Approach
The flowchart of the proposed approach is shown in figure 5. The detail description is given in the subsections. The Flowchart is divided into two blocks, training and testing.

4.1. Training Block
4.1.1 Image Pre-processing
All the images in the dataset are first pre-processed by enhancing their contrast. Each image in the dataset is pre-processed by enhancing its color using imadjust function in Matlab as shown in Eq. 1.

\[ I = \text{imadjust}(I, \text{stretchlim}(I)) \]  

(1)

Where stretchlim returns a two-element vector of pixel intensity values that specify lower and upper limits that can be used for contrast stretching of the input image I as shown in figure 6.
Figure 5. Training and Testing Block

Figure 6. (a) Input image  (b) Color Enhanced of Input image
4.1.2. Image Segmentation

The enhanced image is then segmented by using 3-means clustering based on the color intensity of the pixels. An image is segmented into three clusters having green pixels, yellow pixels and the brownish gray pixels.

In the proposed approach the image is segmented on the basis of color by using 3-means clustering. Yellowish green, green and brown-tan are the most prominent colors seen on the leaves. Thus the segmentation is done on the basis of these three colors. The image is first converted from RGB to L*a*b* color space. It is a 3 axis color system where L represents the Luminance (Lightness), a and b are used to represent the color dimensions. All of the colors information are in the layers a* and b*. RGB is converted to L*a*b* by using rgb2lab function as in Eq.2. Initially taking a small sample region for each color and then calculating each sample region's average color in 'a*b*'. The approach aims at choosing a small sample region for each color and then calculate each sample region's average color in 'a*b*' space. The resultant color markers are then used to classify each pixel. This is done by calculating the Euclidean Distance between the pixel and each color marker. Smaller the distance more closely the pixel matches the color marker for instance, if the distance between a pixel and the green color marker is smallest, then the pixel is labelled as a green pixel. The clusters formed by using 3-means clustering is shown in figure 7. Figure 7(a) is the image of the defected leaf, (b, c, d) are the images of the clusters formed on the basis of color- green, yellow and brown-tan respectively.

\[ I_1 = \text{rgb2lab}(I) \]  

(2)

Figure 7. (a) Image of defected leaf. (b)cluster formed on basis of green pixels, (c)cluster formed on basis of yellow pixels, (d)cluster formed on basis of brown pixels.

4.1.3. Feature Extraction

Texture and Morphology features (Entropy, Correlation, Homogeneity, Contrast, Kurtosis, Skewness and Smoothness) are computed for the texture and shape analysis of each segment of size M*N.
Gray-Level Co-occurrence matrix [19] is used to calculate the following set of texture-features of the input image.

**Entropy**: It measures the textural uniformity of the image, entropy is large if the image is texturally uniform.

\[
\text{Entropy} = -\sum_{i} \sum_{j} P[i,j] \log_2 P[i,j]
\]

where \( P[i,j] \) is the Gray level co-occurrence matrix having gray levels \( i \) and \( j \).

**Energy**: It measures the pixel pair repetitions, and is also called as an Angular second moment.

\[
\text{Energy} = P[i,j]^2
\]

**Homogeneity**: It is also called as Inverse Difference Moment and measures the homogeneity of the image.

\[
\text{Homogeneity} = \sum_{i} \sum_{j} \frac{P[i,j]}{1+(i-j)^2}
\]

**Correlation**: It measures the gray tone linear dependencies in an image.

\[
\text{Contrast} = \sum_{i} \sum_{j} (i - j)^2 P[i,j]
\]

4.1.4. **Morphology Feature Extraction**

Morphology features are used as they provide the physical and external structure information of the leaves. Skewness [20] and Kurtosis are the two morphology features extracted in the proposed approach.

**Skewness**: The shape of the leaf is described by its skewness, as it gives the direction of the tail of the leaves.

\[
\text{Skewness} = \frac{\sum_{i,j} (P[i,j] - \mu)^3}{MN\sigma^3}
\]

**Kurtosis**: It determines that whether the data values are peaked or flat as compared to the normal distribution.

\[
\text{Kurtosis} = \sum_{i} \sum_{j} (P[i,j] - \mu)^4 \frac{1}{MN\sigma^4} - 3
\]

4.1.5. **Classify**

The data obtained is passed to the SVM to train it.

4.2. **Testing Block**

First of all the contrast of the query image is enhanced. The image is now segmented on the basis of color of the pixels by using 3-means clustering. Features extraction (**Entropy, Correlation, Homogeneity, Contrast, Kurtosis, Skewness and Smoothness**) are computed. Already trained SVM is now used to identify the disease.

5 **Experimental Results and Observations**

All the experiments were performed in Matlab 2015Rb. For input training samples of various disease leaves affected with early scorch, bacterial blight, leaf spot and late scorch are collected from various sources like, Google images, papers: Singh, Vijai, and A. K. Misra
[21], Al-Hiary, H., et al. [22] and Patil, Jayamala K., et al.[23]. From figure 8 to figure 11, show the classification of diseases of four different input leaves. For instance, figure 8(a) shows the image of an original leaf and figure 8(b) is the disease identified on leaf in figure 8(a) as Early Scorch. Similarly, for the leaf in figure 9(a), the disease identified is shown in figure 9(b) as Late Scorch. For the other two leaves in figure 10(a) and 11(a) the diseases identified are Bacterial Blight (figure 9(b)) and Leaf Spot (figure 11(b)) respectively. 30 leaves affected with Bacterial Blight and Early Scorch were taken and 25 for Late Scorch and Leaf Spot were taken for experiment.

Figure 8. (a) Original image  
(b) Disease Identified: Early Scorch

Figure 9. (a) Original image  
(b) Disease Identified: Late Scorch

Figure 10. (a) Original image  
(b) Disease Identified: Bacterial Blight

Figure 11. (a) Original image  
(b) Disease Identified: Leaf Spot
The leaves when classified using the proposed approach 97% accuracy was attained. Out of 30 leaves affected with bacterial blight 28 were classified correctly and two were classified wrong one with Early Scorch and the other with the Leaf Spot as shown in Table 1. In case of Early Scorch, out of 30 test images, 29 were classified correctly and one was classified in wrong class as Leaf Spot. The best results were seen in Late Scorch where all the 25 leaves affected were classified in the correct class achieving 100% accuracy. Leaf Spot result had the least accuracy of 92% as out of 25 test images, 23 were correctly classified whereas 2 were classified in the wrong class as Early Scorch. The overall accuracy obtained was 97%.

Table 1. Classification result per class for proposed method

| Disease       | Bacterial Blight | Early Scorch | Late Scorch | Leaf Spot | Accuracy |
|---------------|------------------|--------------|-------------|-----------|----------|
| Bacterial Blight | 28               | 1            | 0           | 1         | 93       |
| Early Scorch   | 0                | 29           | 0           | 1         | 97       |
| Late Scorch    | 0                | 0            | 25          | 0         | 100      |
| Leaf Spot      | 0                | 2            | 0           | 23        | 92       |
| Average        |                  |              |             |           | 97       |

Table 2, shows comparison of proposed approach with other state-of-art approaches. The main difference between the state-of-art approaches and the proposed approach lies in the fact, that in state-of-art approaches only one disease is detected even if there exists more than one disease. In the proposed approach, more than one disease are taken into consideration. Also the state-of-art approaches have not quantified the amount of area affected, while the proposed approach does it. As tabulated in table 2, Image001, suffered from 2 types of diseases- Early Scorch and Leaf Spot, both the state-of-art approaches, i.e Approach 1 and Approach 2 had detected only the Early Scorch whereas the proposed approach detected both the diseases and quantified that 45% area is affected with Early Scorch and 14% is affected with Leaf Spot with the accuracy of 98% and 96% respectively. Similarly for the second test image i.e. Image002, all the three approaches gave the correct results but with an edge that the proposed approach also quantified the amount of area of leaf affected with Bacterial Blight. By having the quantification of disease one can analyze at what level the disease has affected the leaves and is a great help to take any further course of action.
### Table 2. Comparison with state-of-art Approaches

| Image     | Disease/Diseases | Approach 1 [21] | Approach 2 [22] | Proposed Approach |
|-----------|------------------|-----------------|-----------------|-------------------|
|           |                  | Disease Detected | Disease Detected | Disease Detected | Disease Detected | Acc. | Area Affected |
| Image 001 | Early Scorch     | Early Scorch    | 92              | Early Scorch     | Early Scorch     | 92   |             |
|           | Leaf Spot        |                 |                 |                  |                  |      |             |
|           |                  |                  |                 |                  |                  |      |             |
| Image 002 | Bacterial Blight | Bacterial Blight | 88              | Bacterial Blight | Bacterial Blight | 91   |             |
| Image 003 | Late Scorch      | Early Scorch    | 89              | Late Scorch      | Late Scorch      | 93   |             |
|           | Bacterial Blight |                 |                 |                  |                  |      |             |
| Image 004 | Leaf Spot        | Bacterial Blight | 90              | Leaf Spot        | Leaf Spot        | 90   |             |
|           |                  |                 |                 |                  |                  |      |             |
|           |                  |                  |                 |                  |                  |      |             |
| Image 005 | Leaf Spot        | Leaf Spot       | 94              | Early Scorch     | Early Scorch     | 89   |             |
|           | Bacterial Blight |                 |                 |                  |                  |      |             |

#### Conclusion

The paper presents an approach for automatic detection of disease on leaves using Digital image processing. The proposed approach considers four types of diseases mainly found on leaves. The experimental results confirm that it can be used for the fast detection, identification along with quantification of diseases in leaves. Only detecting the disease is not beneficial in practical terms unless the person has the idea that to what extent the leaves are affected so that suitable course of action can be taken. The proposed approach thus, quantifies the amount of area affected so that proper remedy can be used.
There is a lot of future scope as the present paper considered only four types of diseases, rather there exist a lot many more other diseases to explore. Also instead of checking diseases only on leaves another part of the plants may also be used for detecting the diseases.

References

[1] Roberts, M.J., Schimmelpfennig, et al., 2006. The value of plant disease early warning systems. Economic Research Service No. 18, United States Department of Agriculture.

[2] Pimentel, D., Zuniga, R., Morrison, D., 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics 52 (3), 273–288.

[3] Gutiérrez, P.A., López-Granados, F., Pena-Barragán, J.M., Jurado-Expósito, M., Gómez Casero, M.T., Hervás-Martínez, C., 2008. Mapping sunflower yield as affected by Riddolfia segetum patches and elevation by applying evolutionary product unit neural networks to remote sensed data. Computers and Electronics in Agriculture 60 (2), 122–132.

[4] Wang, X., et al., 2008. Spectral prediction of Phytophthora infestans infection on tomatoes using artificial neural network (ANN). International Journal of Remote Sensing 29.6: 1693-1706.

[5] Vleeshouwers, Vivianne GAA, et al., 2008. Effector genomics accelerates discovery and functional profiling of potato disease resistance and Phytophthora infestans avirulence genes. PLoS one 3.8 : e2875. network (ANN)." International Journal of Remote Sensing 29.6: 1693-1706.

[6] Dou, Daolong, et al., 2008. RXLR-mediated entry of Phytophthora sojae effector Avr1b into soybean cells does not require pathogen-encoded machinery. The Plant Cell 20.7: 1930-1947.

[7] Camargo, A., Smith, J., 2009. Image pattern classification for the identification of disease causing agents in plants. Computers and Electronics in Agriculture 66 (2), 121–125.

[8] Oberti, Roberto, et al., 2014. Automatic detection of powdery mildew on grapevine leaves by image analysis: Optimal view-angle range to increase the sensitivity. Computers and electronics in agriculture 104: 1-8.

[9] Guo, Y., P. W. Woods, and S. Rooney-Latham, 2014. First report of Candidatus Liberibacter asiaticus associated with citrus huanglongbing in California.

[10] Bertolini, E., et al., 2014. Tissue print and squash real time PCR for direct detection of Candidatus Liberibacter species in citrus plants and psyllid vectors. Plant Pathology 63.5: 1149-1158.

[11] Ghaiwat, Savita N., and Parul Arora., 2014. Detection and classification of plant leaf diseases using image processing techniques: A review. International Journal of Recent Advances in Engineering and Technology (IJRAET). ISSN (Online): 2347-2812.

[12] Dhaygude, Sanjay B. and Nitin P. Kumbhar., 2013. Agricultural plant leaf disease detection using image processing. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering 2.1: 599-602.
[13] Arivazhagan, S., et al., 2013. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agricultural Engineering International: CIGR Journal 15.1 (2013): 211-217.

[14] Anand. H. Kulkarni and Ashwin Patil R. K., 2012. Applying image processing technique to detect plant diseases. International Journal of Modern Engineering Research, vol. 2, no. 5, pp. 3661-3664.

[15] Bashir, Sabah, and Navdeep Sharma., 2012. Remote Area Plant disease detection using image processing. IOSR Journal of Electronics and Communication Engineering 2.6: 31-34.

[16] R.Hare Kondaiah and A. Sreeramulu, 2014. Survey on Bacterial Diseases of crop and Non-crops of Cuddapah District, Andhra Pradesh, India. International Journal of Microbiology and Applied Sciences. ISSN: 2319-7706 Volume 3 Number 8 pp. 1095 1100.

[17] Fungal Disease: http://www.uky.edu/Ag/PAT/cat1/leafdis.htm

[18] Fungal Symptoms: http://www.oisat.org/pests/diseases/fungal/anthracnose.html

[19] GLCM: http://www.cse.usf.edu/r1k/MachineVisionBook/MachineVision Chapter7.pdf

[20] Skewness: https://www.google.co.in/search?q=skewness+in+plant & source=lnms & tbm.

[21] Singh, Vijai, and A. K. Misra, 2017. Detection of plant leaf diseases using image segmentation and soft computing techniques. Information Processing in Agriculture 4.1: 41-49.

[22] Al-Hiary, H., et al., 2011. Fast and accurate detection and classification of plant diseases. Machine learning 14.5.

[23] Patil, Jayamala K., and Raj Kumar., 2011. Advances in image processing for detection of plant diseases. Journal of Advanced Bioinformatics Applications and Research 2.2: 135-141.