Multivariant Disease Detection from Different Plant Leaves and Classification using Multiclass Support Vector Machine

Yangala Raghavendra[a], Dr. G.A.E. Sathish Kumar[b]

[a] scholar, ECE Dept., Vardhaman College of Engineering, raghuyangala1@gmail.com
[b] professor, ECE Dept., Vardhaman College of Engineering, gaesathi@gmail.com

Abstract: Organic farming is becoming smarter. Unfavourable conditions, such as drought or lack of nutrients, or excessive water, will negatively affect crop production. We will grow a good crop if we have a sensor that watches over the garden. A solution is being implemented by computer-aided image processing in short, we'da first approximation, our best approach is to treat the disturbance in the leaf is to make sophisticated machinery that picks up on it. In the current issue of "Science and Cooking", find the disease in tomato, potato, grape, apple and corn plant leaves. Plant germination is usually affecting early blight, Bacterial spot, leaf mould, common rust etc. Manually monitoring plant diseases is quite tough. It necessitates a great deal of effort, knowledge of plant diseases, and an inordinate length of processing time. As a result, image processing is utilized to detect plant illnesses. Image acquisition, image processing, and image analysis are all processes in the disease detection process. image segmentation, feature extraction, pre-processing and in addition, Multi class support vector machine classifiers (MCSVMs) are used in this work.

Keywords: Image processing, k-means clustering, Multiclass SVM

1. Introduction

Food-stuffs such as several varieties of wheat, rye, oat, barley, and rye or wheat and barley are especially vulnerable to various diseases. Unclothed vision is the most prevalent mode of human sight.

For growers, there's a big difference between success and failure with any crop Visual screening will tend to miss the disease in the short term, so precautions must be taken to check further in. in the majority of the farms are located in the countryside, and it is hard for them to find a nearby location where they can get help with their crops the effects are impressive when you use image processing instead of a lot of time and effort. There is a tremendous sense of inferiority felt by members of the theatrical profession about theatre critics. Theater has its share of egotists, of course, but the difference is that critics work with ideas. There are plenty of arrogant people in business, but business people work with numbers**.**

2. Related Work

For automatically detecting tomato leaf disease primary task is to identify the diseased or injured tomato leaf region. To obtain the best texture properties possible, the Haralick algorithm is used. To identify the features extracted from the tomato leaf disease database, various machine learning techniques such as logistic regression, random forest, and SVM were used [1]. Fungi, soil-borne pathogens, and viruses infect groundnut plants.

The study focuses on early and late leaf spot, rust, early and late spot, and other groundnut crop diseases. The machine determination to robotically diagnose and categorize groundnut leaf diseases is presented in this paper [2]. This method will boost crop production.

Many diseases wreak havoc on the growth of tea leaves, causing them to shrink and hampering tea production. An automated system for identifying 3 main kinds of tea-leaf diseases has been developed using a Support Vector Machine (SVM) with a limited set of features. [3].

In the farming sector, image processing has been used by creative farmers. This segment presents the new solution. Pushed, Chaudharyan and Push created separate layers of red, green, and blue images. These three-colour sets are included: Blue, Red, and Yellow. Yellow, Magenta, and Cyan are colours on the CIELAB Colors colour scale. The filter approaches the centre of the signal so as to exclude the noise. Dot-and-and-and-veckleak tests were given to the subjects. Kirsch and Rothe generated neural network colour representations of differentiating outcomes. A lot of testing and experimentation Instead of testing people in an old-fashioned, no diagnostic manner, a disease-detection system had to test someone at the brink of life. In the other hand, various colours, shapes, and textures of skin lesions generated the desired results There were 5 medical conditions in the GLC (Neural Network) and SVM beat out, as shown. It occurred to Zhang; Zhang started thinking about the whole leaf. And to better performance, the experiments were conducted using the RBF function.

Creative thinking, an original classification device Selam, Priya, and Balasaran [12] discovered 12 qualities and made a distinction between k-Nearest neighbour analysis and support vector machine learning. Omrani and El examined the ailments in the apple tree.

Figure 1: Leaf spot disease

For classification support vector and associative neural networks (SVMs) can also be used. They used the support vector machine (SVM) and discovered that Phadar, Santanu, and Jaya attempted to identify diseases in rice plants. Isolating the contaminated regions is done using optical imaging software. The neural network classifier is now being applied in this case. PNN was attempted on the data, but it failed.

3. Diseases Present in leaves

CMV and aleriac yellows, as well as bacterial wilt are often present in the foliage. An overview of these diseases is given in this section. Cucumber has the highest numbers of CMV, by far. On the basis of these results, real-time data is applied to cucumber mosaic and CMV disease.

A. Leaf Spot

Rot is caused by Pseudomonas syringae bacteria Every portion of the plant leaves, including leaves, can be colonised by this particular type of bacteria. If there are leaves on the route, the impacted area would be waterlogged and look like a muddy. During the latter stage, it eventually takes on a brownish hue. As a consequence, the tree will die. It is possible for bacteria to move from part to part after heavy rainfall. Because the disease is often found between around 13.9° and 35.6°C, 24 to 28 degrees Celsius is the appropriate temperature and relative humidity is best for it. The contaminated leaf seen in Figure 1.

B. Leaf Miner

An insect that feasts on foliage it will decrease the output of the plant by eating away at the leaves. Generally, large-leaved plants are found to have tannins. Many species of this insect feed on leaves, too. In terms of percentage, this is a small percentage of the total impact. But when cultivated on a large scale, this could lead to low yield. propermanent degradation is unlikely.

Figure 2: Leaf Miner

Figure 3: CMV

Figure 4: Healthy leaf

It can be expected that they would not have problems from the leaf miner. Fig. 2 indicates the presence of leaf mines in the cucumber.

C. CMV

The Mosaic virus leaves artwork can leave attractive designs on the leaves. Also, sometimes, because of this fungus, the plant can wither and die. Green and dark-green pickles. If CMV-infected grains may have a sour taste, they are unlikely to germinate (unless given the right moisture and conditions). The ideal temperature for this condition is in the range of 79 to 84 degrees and above. To stop this, we must first discover the problem and then eliminate it. CMLV is very evident in this photograph. Isb disjointed and for lack of unity.4.

D. Alternaria Leaf Blight

Spots are commonly occurring on the oldest leaves first. For the most part, the spots would be yellow-brown.

If the disease gets worse, the patches develop into large, tumours occur. The details follow the same formula. The leaves finally start to drop.
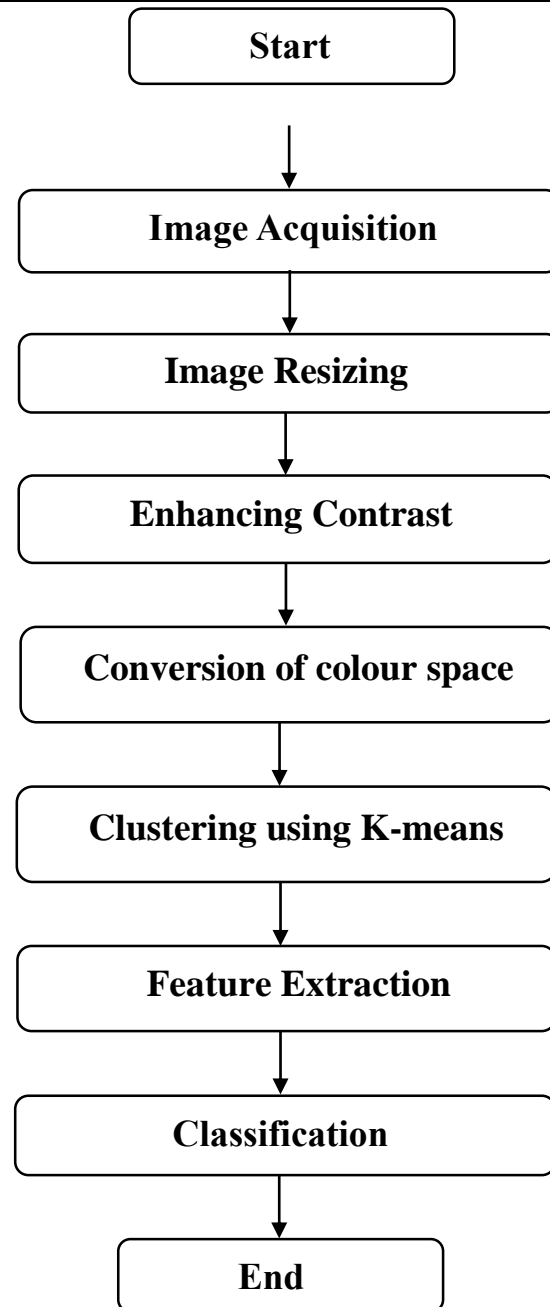
```
┌─────────────────────┐
│        Start        │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│  Image Acquisition  │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│   Image Resizing    │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│ Enhancing Contrast  │
└─────────────────────┘
           │
           ▼
┌──────────────────────────┐
│ Conversion of colour space │
└──────────────────────────┘
           │
           ▼
┌──────────────────────────┐
│  Clustering using K-means │
└──────────────────────────┘
           │
           ▼
┌─────────────────────┐
│ Feature Extraction  │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│   Classification    │
└─────────────────────┘
           │
           ▼
┌─────────────────────┐
│        End          │
└─────────────────────┘
```

Figure 5: Flow chart

E. Anthracnose

The areas would be yellow and brown coloured by the growth of the lesions. Any of these findings is shown in leaves, roots, and fruits. Some types will be completely unresponsive to treatment, so the lesions will look green on the darker side.

F. Aster Yellows

The stems will expand vertically and this disease will display a peculiar characteristic of development perpendicular to the ground. In general, the leaves are thin, but they can appear thicker. When either or more of the plant's flower and leaf bracts get disfigured, they are owned by that flower or that leaf. When they're ripened, the fruits can lose their colour and quality.

4. Methodology

An unwell leaf will be one of the first things you will investigate. Photometric processing is absolutely necessary. Post processing must occur, which will separate, remove, and classify the components It is a technique that follows step-by-by-step rules.

I) Image acquisition.

II) Image Preprocessing.

➢   Image Resizing

➢   Enhancing Contrast

III) Conversion of colour spaces

IV) Clustering using K-means

V) Feature Extraction

➢   Using GLCM

VI) Classification.

➢   Multiclass SVM

On the fly pictures, where new pictures are captured and updated as needed. The therapy commences as soon as possible after the disease has been diagnosed. In Fig. 5 is demonstrated. Follow these step-by-by-by-step instructions to the letter.
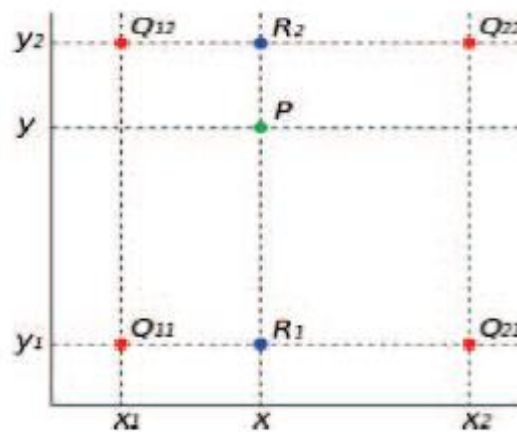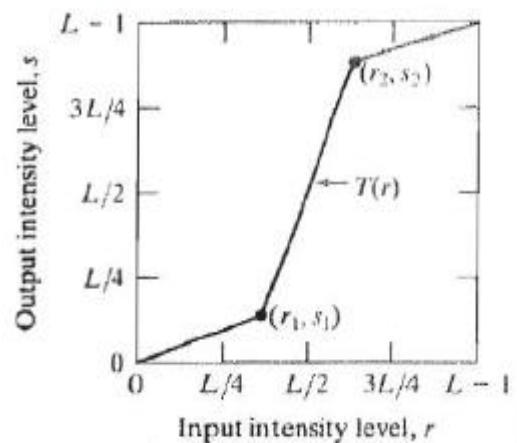

Fig. 6. Bilinear interpolation [14].

Figure 7: Revision/opposite of construction

A. Image Acquisition

It is defined in image processing as the action of retrieving an image from some source, typically a hardware-based source, for processing. It is the first stage in the workflow sequence because no processing can take place without an image.

To begin, all leaf photos are obtained from various sources. Both healthy and sick leaf photos can be found in the datasets. The dataset contains 20638 samples organised into 25 folders. Twenty folders contain exclusively diseased leaf images, while the remaining five folders have healthy leaf images.

B. Image Preprocessing

1) Image Resizing: If a new picture is useful, it would be collected at a different scale Resizing the images is advised. It shrinks in size to $256 \times 256 \times 3 \times$ the default size Bilinocular interpolation is utilized in this process when the new value is calculated using the values of four adjacent pixels.

$$P = \frac{Q1+Q2+Q3+Q4}{4} \qquad (1)$$

This pixel value is called P Four pixels are the neighbours of Q1,Q2,Q3,Q3, and Q4. This can be seen in figure 6.

2) Enhancing Contrast: Focusing on the foreground and backdrop (of an image enhancement) Linear contrast enhancement and non-linear contrast have two distinctive visual characteristics. Linear curvature in this work is defined as a piecewise function. As seen in Figure 7, contrast stretching is performed to get the picture to the maximum intensity spectrum of the medium and to the full dynamic range. The distribution of the input values (r1, r, s) as well as the strength levels of the output values (s1, s) determines the overall shape of the transformation matrix [15].

C. Color Space Conversion

This is a device-dependent conversion since RGB colour space is heavily tied to specific devices. This is also an issue that relies on the light intensity. In L*a*b space, colour processing is much easier. Also, the space is computer-driven. As a colour composition is created, it is then applied to the picture to produce a visual change.

Human colour vision was incorporated into the RGB colour model in order to provide a convincing simulation of how the eye sees. One significant difference is that it has colour.

D. K Means Clustering

It is the most commonly used for natural looking clustering in an un-directed, L*a* b-color space (a learning algorithm that uses both the primary and secondary colours). We know how many clusters there would be, as we had foreseen. We referred to k = 3 cases and the healthy leaf, the second cluster, and the diseased third leaf. Each centroid remains the same. Placing the reference point is a difficult one in the simplest terms, the only solution is to hold them apart. One hundred data points are related to the correlated with the centroid. No analysis of the data has taken place. All data has been tabulated, and the first step is complete. New centroids are derived from the previous centroid and the furthest centroid from the original centroid. Finally, the matching process is underway with the latest centroids in place. Before you cannot adjust the position of the k centroids the objective of the algorithm is to minimize the square of the error function.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \qquad (2)$$

What is the formula: J = error function data value of point centred on: c is roughly equivalent to (x.i)^n The data point to cluster squared distance Using the basic algorithm below, it is simple to explain:

1) Initial artefacts are located at the center of the cluster.

2) The point with the nearest to each target coordinates are selected, and it is connected with all the other points.

3) As the new centroid point is found, the process is restarted.

4) After [performing] these two moves, no new   points can be located.

After this image has been rendered, three separate renderings can be created. Function extraction is reserved for the relevant disease cluster in which it is detected.

E. Feature Extraction

In the goal of data compression, we are looking to focus on minimizing the amount of data which are essential for broad sets of features. When the classifier has a wide sample size, it underperforms. By careful study of this texture analysis, some special features and variations in personality are uncovered. Grey degree pixel co-occurrence (GLCM). Four aspects of image processing can be selected by using GLCM: Contrast, Correlation, Symmetry, Coherence, and Symmetry Besides entropy, RMS, Kurtosis, and skewness are part of the calculation in general. Figuring out the right picture parameter co-element of the p! Or Consider the co-occurrence matrix p(i, j). Moving from pixel I to pixel j has a high likelihood. is represented by the grey degree. For the input of the GLC models, we have the equations below.

I) Contrast: checks for grey levels in a picture

$$Contract = \sum_{i,j} |i-j|^2 p(i,j) \qquad (3)$$

II) Correlation: Analyses a pixel to establish its association with its neighbours.

$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \qquad (4)$$

III) Energy: In G-th, this function returns the value of the squared magnitudes.

$$Energy = \sum_{i,j} p(i,j)^2 \qquad (5)$$

IV) Homogeneity: The overlap of the distribution of variables is calculated by 2GLCM.

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad (6)$$

V) Entropy: The quantity of image complexity represents the amount of information contained in the image is measured by the relative entropy of an image provides.

$$Entropy = -\sum P_i \log_2 P_i \qquad (7)$$

VI) Average: Consider an image to be a certain portion of a map to be a particular brightness in the grayscale. In other words, mean pixel ma would take the value of this across all the samples in the area.

$$m_a = \frac{1}{\Lambda} \sum_{(m,n)\epsilon R} a[m,n] \qquad (8)$$

VII) Standard Deviation (S.D): The ability to perceive the intensity of pixels in the area R in a picture the estimated brightness is the sa, which is equal to this brightness:

$$s_a = \sqrt{\frac{1}{\Lambda-1} \sum (a[m,n] - m_a)^2} \qquad (9)$$

The other variables were drawn from the section of the diseased leaf.

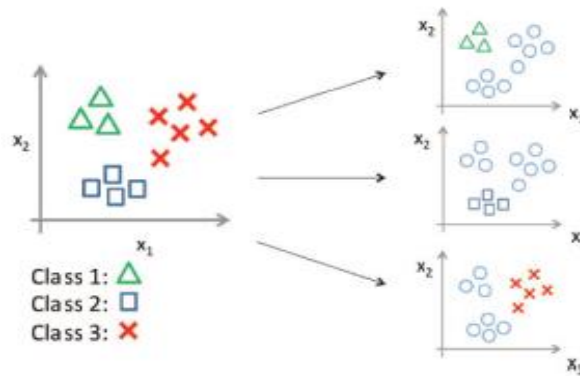**F. Classification: Multiclass SVM**

Figure 8: SVM 1AA approach.

Enthusi-mentary vector machines are a kind of classification method. In supervised classification, labeled data is gathered from sample images and new classifications are generated from it Classify each disorder and classifying the classifications are two different techniques. This data is being used to help build a database for the training process in supervised grouping, one group is compared to each other and the label is formulated for each comparison. When testing data are provided, it is appropriate to assign the class names. Also, when the training data is limited, SVMs can be powerful and reliable. According to earlier research, SVM had higher classification accuracy than neural networks. Thus, in this project, the technique of Support Vector Machines is used to do supervised classification. Although a binary classifier would only be able to place things into one category, SVMs can be trained to recognize several classifications. In this case, there are two methods to apply for this system: One-against-One (OK) and One-against-all (OKA).

A class system is born by subdividing N classes into N classifications. For example, in Figure 8, if three classes are numbered 1, 2, and 3, classification can be done with the following tables: 1 versus non class 1, 2, 2 versus 3, and 3, compared to class 1 (3). N-1 classifications were a variation on the methodology A2. Since the diseases have been described as of lower degrees of severity, it has taken the place of 20A ICD Leaf miner spotting is included. Good leaves are picked too.

A single classifier will only predict one type of object class, whereas a multi-classifier system is able to recognize more than one type of object. If the Machine Learning classifier learns to separate the fifteen diseases in our dataset, it will know more about them. The fifteen diseases are listed in one disease class in this dataset.

5. Training the Detection System

In the above picture, we can see that we have 25 classifications for each disease. Multi-vector SVMs are built in the following the CODE sheet In the Multiclass SVM panel, we can choose the 'Calculate the Multiclass Accuracy' and then press to see the accuracy of the model on individual sample c.

The data on the screen shows that there are 25 disease types, but validation SVM is only looking at 4128 of the photos in the dataset. To see how accurate our SVM predictions are after these forecasted numbers, check out these test images. Since the SVM dataset contains over two million items, it takes about five to ten minutes to load.
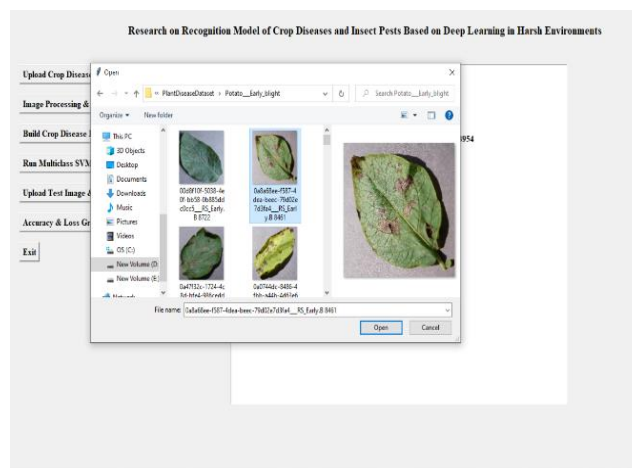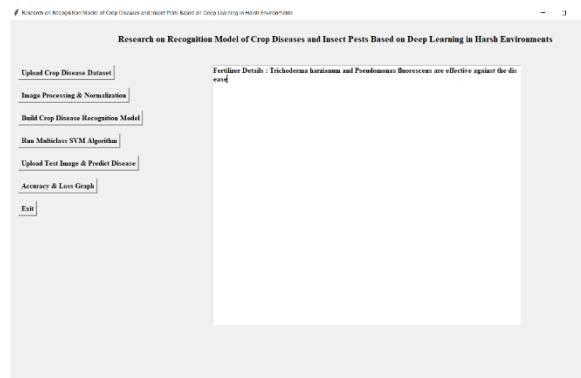

6. Results

In this project after adding fertilizers, we will get disease prediction name as well as fertilizers details.

In above screen we can select and upload one image and then click on 'Open' button to get below result.
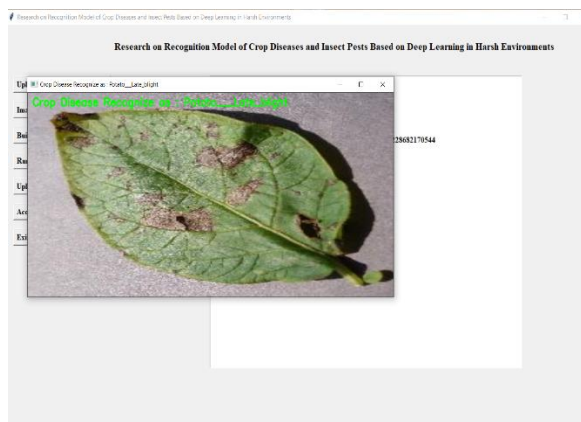
In above screen in green text, we can see predicted disease and now close above image to get below fertilizers details.

The fertilizer information for the predicted disease is shown in the text region of the above screen.

7. Conclusion

Everything of our life is rooted in nature and the economy of the soil. Additionally, diseases damage the flora in varying ways, advanced treatment methods are used to treat the various types of illnesses that they cause. When GDP goes down, that means there are fewer crops to go around. The only way to deal with plant disease is to recognize it. Primarily intended to aid in the accurate diagnosis of disease in plants. This paper discussed various techniques to segment the disease part of the plant. This paper also discussed some Feature extraction and In order to retrieve the features of an infected leaf, classification techniques were used.as well as plant disease classification can be put to better use. We can accurately identify and classify a wide range of plant diseases techniques for image processing and also provide fertilizer details of disease leaves.

References

1. Debasish Das, Mahinderpal Singh, Sarthak Swaroop Mohanty and S. Chakravarty (2020) "Leaf Disease Detection using Support Vector Machine", International Conference on Communication and Signal Processing: IEEE, 2020.
2. Vaishnnave, M. P., et al. (2019) "Detection and
3. Classification of Groundnut Leaf Diseases using KNN classifier", IEEE International Conference on System, Computation, Automation and Networking (ICSCAN): IEEE, 2019.
4. Hossain, Md Selim, et al. (2018) "Recognition and detection of tea leaf's diseases using support vector machine", 14th International Colloquium on Signal Processing & Its Applications (CSPA): IEEE, 2018.
5. P. R. Rothe and R. V. Kshirsagar, "Cotton leaf disease identification using pattern recognition
6. techniques," in Pervasive Computing (ICPC), 2015 International Conference on. IEEE, 2015.
7. P. Revathi and M. Hemalatha. "Advance computing enrichment evaluation of cotton leaf spot disease detection using Image Edge detection," in Computing Communication & Networking Technologies (ICCCNT),2012 Third International Conference on. IEEE, 2012.
8. D. S. Guru, P. B. Mallikarjuna, and S.Manjunath, Segmentation and classification of tobacco seedling disease, IEEE, COMPUTE11, (2010),2011.
9. Marek Schikora and Adam Schikora, "Image-based analysis to study plant infection with human pathogens,"Computational and Structural Biotechnology Journal, vol. 12, no. 20, pp. 1–6, 2014.
10. Zhang Jian and Zhang Wei, "Support vector machine for recognition of leaf diseases," in Advanced Computer Control (ICACC), 20102nd International Conference on, 2010, vol. 5. IEEE.
11. C. Arun Priya, T. Balasaravanan, and Antony Selvadoss Thanamani, "An efficient leaf recognition algorithm for plant classification using support vector machine," in Pattern Recognition, Informatics and Medical Engineering (PRIME), 2012 International Conference on. IEEE, 2012.
12. Elham Omrani, et al. "Potential of radial basis function-based support vector regression for apple disease detection," Measurement, vol. 55, pp. 512–519, 2014.
13. Santanu Phadikar and Jaya Sil, "Rice disease identification using pattern recognition techniques," in Computer and Information Technology, 2008.ICCIT 2008. 11th International Conference on. IEEE, 2008.
14. Ravi Shankar, Seema Harsha, and Raj Bhandary, "A practical guide to identification and control of cucumber diseases,"
15. Mary Ann Hansen, "Angular leaf spot of cucumber," (2009).
16. Viktor Knutsson, "Terra mechanics based wheel-soil model in a computer game enviroment," (2016).
17. Rafael C. Gonzalez and Richard E. Woods,
18. "Image processing," DigitalImage Processing, vol. 2, 2007.
19. Shitala Prasad, Sateesh K. Peddoju, and Debashis Ghosh. "Multi-resolution mobile vision system for plant leaf disease diagnosis,"Signal, Image and Video Processing, vol. 10, no. 2, pp. 379–388, 2016.

20. R. Anand, S. Veni, and J. Aravinth, "An application of image processing techniques for detection of diseases on brinjal leaves using k-means clustering method," in Recent Trends in Information Technology (ICRTIT),2016 International Conference on. IEEE, 2016.

21. P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar "Image texture feature extraction using GLCM," International Journal of Scientific and Research Publications, vol. 3, no. 5, May 2013, ISSN 2250–3153.

22. Vijay Kumar and Priyanka Gupta (Aug,2012) Importance of statistical measures in digital image processing, ISSN 2250–2459, vol 2, no. 8, International Journal of Emerging Technology and Advanced Engineering.

23. K. P. Soman, R. Loganathan, and V. Ajay. Machine learning with SVM and other kernel methods. PHI Learning Pvt. Ltd., 2009.1281

24. Dheeb AI BAshish, Malik Braik, Sulieman Bani-Ahmad (2010) A frame work for detection and classification of plant leaf and stem diseases, signal and image processing (ICSIP): IEEE international confererence.

25. Zainon Radhiah, "Paddy Disease Detection System Using Image Processing," (2012).

26. Piyush Chaudhary,et al (2012) Color transform based approach for diseases pot detection on plant leaf, vol. 3, no. 6, pp. 65-70, International journal of computer science and telecommunications: 2012.