Naïve Bayes and Textural Feature Analysis Technique for the Plant Disease Detection

Sarvjeet Kaur[1], Dr. Gaganpreet Kaur[2]
[1]Research Scholar, M.Tech (CSE), [2]Assistant Professor, Deptt. of CSE
Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India

Abstract: The plant disease detection is the technique which is applied to detect disease from the leaf images. In the agriculture, farmers face many problems due to the diseases of plants that affect the quality and quantity of crops. The plant disease detection technique consists of steps: image acquisition, image pre-processing, image segmentation, feature extraction and classification. In the proposed work, laws textural mask method is applied for the textural feature analysis and naïve bayes is applied for the classification of diseased and non-diseased portion of leaf. This classification method detects and classify the diseases into multiple classes and it enables to detect the infected part of the diseased leaf. The performance of this system is tested in terms of accuracy, precision and recall. The system has achieved high accuracy of 96.22%.

Keywords: *Feature Extraction, Naïve bayes, Segmentation, classification.*

I. INTRODUCTION

Agriculture is a very important part of our lives. Suitable crops are chosen by farmers and in order to ensure their proper growth, various fertilizers and pesticides are used. The quality and quantity of the products of agricultural products has to be maintained by ensuring that they suffer from no diseases. The observable patterns are to be found within the plants in order to detect any such diseases. In order to cultivate crops in a farm it is to be made sure that diseases are identified for which various studies have also been proposed. With the help of expertise person, the plant diseases were monitored and analyzed within the earlier days. Large amount of work is involved within these techniques along with proper processing times to handle such situations. Most of the techniques included various disease symptoms to be provided in order to diagnose the types of diseases [1]. The productivity of the crop is decreasing day by day because of diseases of plant as diseases affect the production quality of crops. The poor production of crops affects the health and wealth of the country. The farmers face various problems that affect the production of crops. These problems include environmental problems and plants diseases problems. There are some methods for disease control such as use of pesticides but they might be harmful for natural resources such as soil, water, air and food chain [2]. So advance methods are required to detect the diseases. There are various advance image processing methods available. Images of the plant leaves, fruit and stem are given as input to the processing system designed for detection of plant diseases. The disease detection system extracts the informational features from the images. According to the features extracted decision is generated that whether the plant is affected by disease or not. The diseases of plants can be easily detected through the disease detection system designed with the help of image processing tools [3]. Image processing tools can be used to recognize the plant diseases at early stage so that necessary steps should be taken to save the quality of product crops. Diseases interrupt the normal growth function of plants such as photosynthesis, transpiration, pollination, fertilization, germination etc. that affect the quality and quantity of crops. So it is very important to diagnose the disease at early stage [4]. The diseases may be present on leaves, stems and fruits of the plant. The diseases can be identified through the symptoms that causes change in the color and shape of leaves, fruit and stem. Some disease detection system also defines the name of the diseases as these systems are trained by defining the symptoms of various diseases. The symptoms help to recognize that the plant is affected by which disease [5].

1.1 Causes of Plant Diseases

The plant diseases are caused by mainly two factors. These factors are biotic and abiotic factors. The ecosystem is made up of biotic and abiotic factors. The biotic factors in ecosystem are dependent on abiotic factors. The minor change in these factors can have a great impact [6].

1.1.1 Biotic Factors: - Biotic factors include the diseases caused by living organisms such as pathogens, nematodes and parasitic, fungi, bacteria and virus. Biotic factors affect the plants of same age. The diseases caused by biotic factors can be detected earlier if the signs of fungus is present. But in some cases the signs are not present, in this case it is very difficult to recognize the diseases at early stage. So the disease can damage the quality of plant [6].

a. Bacteria: Bacteria do not cause harm to the plant immediately. It needs favorable environment conditions such as high humidity, poor soil health, nutrient imbalance to harm the crop [6]. Some of the bacterial diseases are Black rot, Bacterial leaf spot, Bacterial Blight, Bacterial Brown spot etc. [7]

b. Virus: Virus diseases spread slowly on plant. The identification process of viral diseases is very difficult. The virus of infected plant is easily transmitted to healthy plant. In this way virus slowly infect the other healthy plants that are near to the infected plant. They are ultra-microscopic organism. They are so tiny that they can only be seen with a very powerful microscope called an electron microscope. These kinds of diseases are difficult to detect and diagnose. [6]. Common viral diseases are Leaf Curl, Leaf Crumple, Leaf Roll etc. [8]. The plant which is infected by virus has frizzed and farrowed leaves [9].

c. Fungi: Fungi affect large number of plant species. Fungal diseases are reproduced at very high rate. Fungi diseases can be controlled with the use of fungicides [6]. There are some basic fungal diseases such as Wilt, Powdery Mildew, Downy Mildew, Anthracanose, Alternia, Leaf Spot, Grey Mildew, Rots, Cankers, Mold etc.[9].

1.1.2 Abiotic Factors: Abiotic factors can affect the plant of various ages. Diseases caused by biotic factors are not associated with the pests and insects. Abiotic factor diseases are caused by environmental factors such as weather conditions and nutritional deficiencies such as deterioration in nitrogen, phosphorus, magnesium etc. Weather conditions such as high moisture due to the floods or excess amount of water can turn the plant into yellowish and low moisture can lead to death of plant [8].

The paper is organized as follows: First section of the paper briefly discusses the introduction to the plant disease detection. Second section discusses the review of the previous work related to the various plant disease detection techniques. Third section explains the proposed methodology. Fourth section briefly defines the results and comparative analysis is performed on the basis of accuracy, precision and recall. Fifth section consists of conclusion and future scope.

II. LITERATURE REVIEW

This section provide the review of literature. The various image processing techniques used by various researchers in their papers are discussed.

Karthik G. et al. [10] monitored banana leaf diseases. The disease detected is a viral disease also known as a Banana Streak virus. The symptoms of banana streak disease can be seen on the banana leaf when the color changes into reddish yellow. This symptom can be identified from fifth month of plant. An embedded linux development board interfaced with camera is introduced for the detection and prevention of banana streak, in which the role of camera is to capture image. An algorithm ca lied Economic Threshold Level (ETL) is utilized to process the captured image. A threshold value is defined to detect the disease. If the pixel values of captured image surpasses the threshold value then the leaf is near to the infected of banana streak.

Dhaware C.G. et al. [11] discussed plant disease detection and identification stages like image acquisition, image pre-processing, image segmentation, feature extraction and classification. Some of the image segmentation algorithms are discussed to automatically recognizing the plant leaf diseases. For classifying the leaf diseases, classification methods are described which can be used by researchers.

Islam M et al. [12] proposed computer vision based phenotyping that gives auspicious step towards the agricultural security. Image processing techniques are combined with machine learning to detect the disease. Potato leaf diseases are detected in this work. Gray Level Co-occurence Matrix (GLCM) is applied for feature extraction and the features extracted are contrast, energy, correlation and homogeneity. Multiclass SVM is used as classifier. Potato leaf diseases are classified under class labels late blight affected, early blight affected and healthy leaf.

Patrick A. et al. [13] presented a system which monitor the health of plant through the amount of visible and close infrared light that plants reflect. The multispectral image of tomato spot shrivel disease were acquired by a quadcopter having high throughput.

Rajan P. et al. [14] introduced a system for the identification of pests at early stages to reduce the use of pesticides. Two parameters time and accuracy are considered for checking the system results. SVM classifier is used to classify the pest pixels and leaf pixels. For removing the annoying elements Morphological operators are used. SVM is trained by providing the color features.

Tichkule S.K et al. [15] presented a summary of detecting different plant diseases caused by bacteria, fungus and virus. The diseases and their causes and symptoms are described. The Soybean leaves diseases are detected using SIFT algorithm and SVM classifier, Cotton leaves using PCA/KNN, Wheat leaves using PCA and morphological features, Grape leaves using BPNN and K-means algorithm.

Singh V et al. [16] proposed an algorithm for image recognition and segmentation. Banana, beans, rose leaf diseases such as early scorch, bacterial leaf spot are considered for recognition. Banana, beans, rose leaf diseases such as early scorch, bacterial leaf spot are considered for recognition. Leaf image is clipped to remove the undesired portion and soothed using smoothening filters. Threshold value is computed to mask the green pixels. Color Co-occurrence Method is applied for feature extraction. Three color co-occurrence matrixes are generated for each H,S,I channel of leaf image. SVM classifier is used to classify the leaf image.

Padol P.B. et al. [17] proposed a system for classification of grape leaf disease using SVM. Downy mildew, Powdery mildew fungal diseases of grape leaf are recognized. Firstly segmentation is performed using K-means then features like contrast, Uniformity, Maximum probability, Homogeneity, Inverse difference, Difference variation, diagonal variance, Entropy, Correlation are extracted. Classification is done using Linear Support Vector Machine (LSVM).

Gowthami K. et al. [18] discussed various segmentation techniques, feature extraction techniques and classification techniques. Artificial Neural Network (ANN), Back Propagation Neural Network, Self-Organizing Feature Map and Support Vector Machine (SVM) can be used efficiently to classify the diseased image. Various features such as Correlation, Smoothness, Contrast, Energy are extracted that helps in identifying the disease.

Sethy P et al. [19] proposed a novel approach to recognize the defected part of rice leaf using 3-means clustering algorithm. The RGB image is converted into L*a*b color space. Colors of leaf image are classified using k-means segmentation. Euclidean distance is calculated to get the difference between colors. Pixels of image are assigned to cluster to which they belong.

III. RESEARCH METHODOLOGY

In this section the basic steps that are followed in the proposed system for detecting potato leaf diseases are discussed. The steps are as follows:

Step 1. Image Acquisition: It is the first step in which leaf images are collected from which the disease is to be detected. In this research work potato leaf diseases are detected. The potato leaf images affected by various diseases are taken from plantsvillage**.**

Step 2. Pre-processing Phase: In the first phase, the input leaf image is converted to greyscale.

C. Feature Extraction: The Law's textural method algorithm is applied which will extract the textural features of the input image. Following are the equations are used for calculating the features by law's texture mask method.

$$\text{Mean} = \frac{\sum_{W} neighboring\ pixels}{W} \quad \ldots(1)$$

$$\text{Absolute mean} = \frac{\sum_{W} abs\ (neighbouring\ pixels)}{\text{W}} \quad \ldots(2)$$

$$\text{Standard deviation} = \sqrt{\frac{\sum_{W} (\text{neighbouring pixels} - \text{mean})^2}{W}} \quad \ldots(3)$$

Step 3. Image Segmentation: The K-mean segmentation technique is applied which will segment the image on the basis of their properties. The image segmentation techniques are generally applied categorized into region based segmentation and threshold based segmentation.

Step 4. Classification: In this work naïve bayes classifier is used to classifying the diseases. Naïve bayes classifier classifies the similar and dissimilar data into more than one classes. In this work naïve bayes classifier classifies the potato leaf diseases into multiple classes and also detects the infected part of the leaf. Following equation is used for the classification .

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad \ldots(4)$$

where, P(c|x) is the posterior probability of class (target) given predictor (attribute).

P(c) is the prior probability of class. P(x|c) is the likelihood which is the probability of predictor given class. P(x) is the prior probability of predictor.

Fig. 1 shows the flowchart of proposed work.

The basic steps of flowchart are:
1. Input the image to the system to test.
2. Apply k-mean algorithm for image segmentation.
3. Apply law's textural mask method for feature extraction.
4. Apply Naïve bayes classifier for the classification of the diseases into multiple classes.

IV. EXPERIMENTAL RESULTS

The proposed work consists of four steps image acquisition, image segmentation, feature extraction and classification. In the first step image is selected from the database as an input, in the second step the k-mean algorithm is applied for the segmentation. Segmentation technique divides the input image into number of clusters. Then the user can select the one region of interest from the different clusters. In the third step the textural features are extracted by the law's textural mask method. In the last step the naïve bayes classifier is used to detect and for classifying the diseases. Figure 2 shows an input image, Figure 3, 4, 5 and 6 shows the different clusters. Fig. 6 represents the infected part of the leaf. Then classifier classifies the leaf disease according to the extracted features which are calculated in the previous step which is feature extraction phase. Naïve bayes classifier detects the infected part of the diseased leaf in an efficient manner.

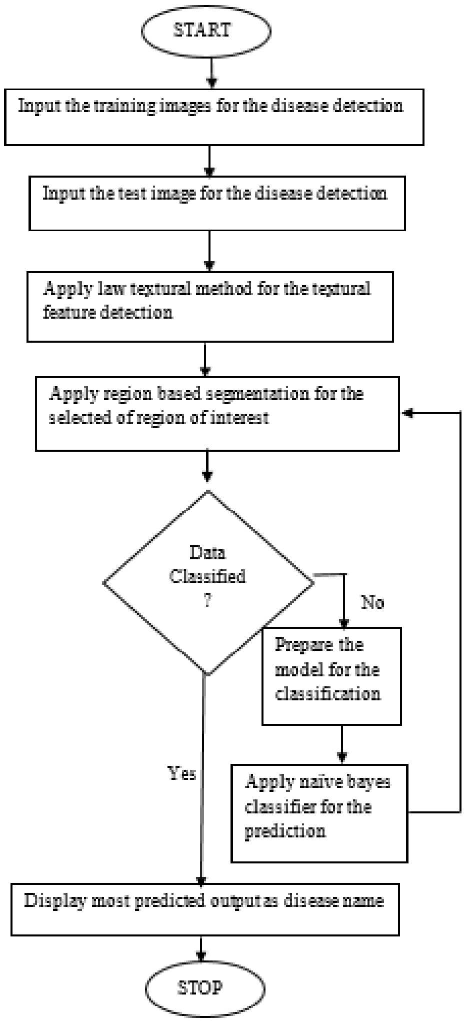

Fig.1: Flowchart of Proposed Work

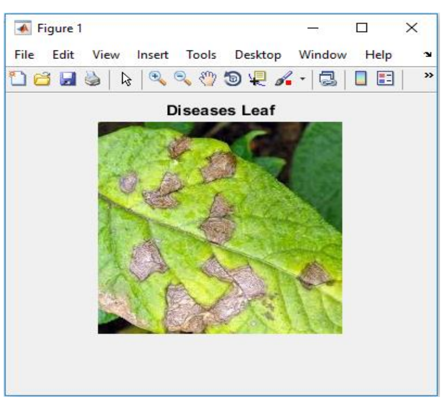

Fig.2: Input Image

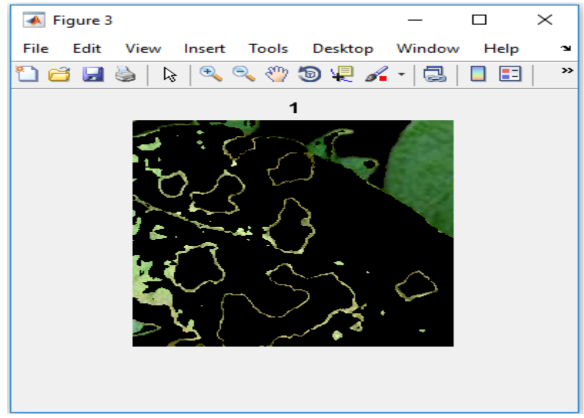

Fig. 3 Cluster 1

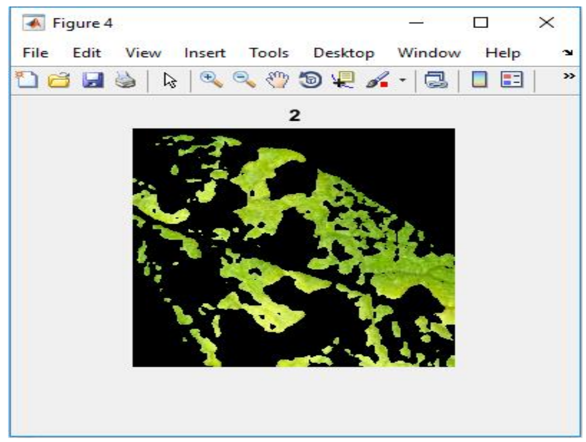

Fig. 4 Cluster 2

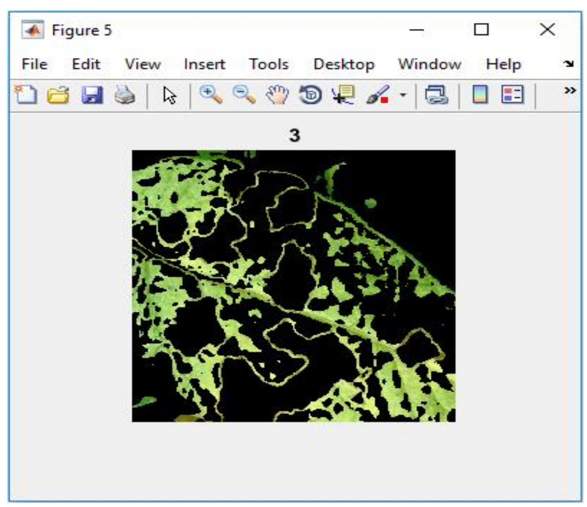

Fig. 5 Cluster 3

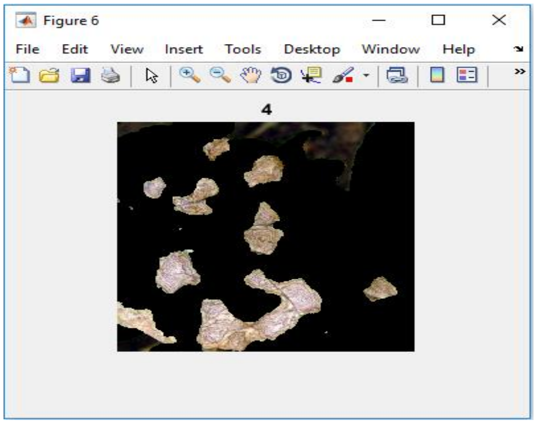

Figure 6: Cluster 4 (Infected Part)

Following are the five sample images on which the testing is performed.

Image 1

Image 2

Image 3

Image 4

Image 5

Comparison of the existing and proposed system is done on the basis of three parameters which are precision, recall and accuracy. Comparative analysis shows that the proposed system has the higher precision, recall and accuracy than the existing system. In this paper the comparison is shown in the form of table and graphs. Table 1, Table 2 and Table 3 shows the values of precision, recall and accuracy of both existing and proposed system. Fig. 6, Fig. 7 and Fig. 8 shows the graphical representation of the comparison of precision recall and accuracy of proposed and existing system. Table 1 shows the comparative results of precision.

Table 1. Comparison of Precision

Test Images	Precision	
	Existing System	Proposed System
Image 1	93.77	99.20
Image 2	93.66	98.77
Image 3	93.05	99.04
Image 4	92.40	97.76
Image 5	93.13	98.99

From Table 1, it is analyzed that the proposed system achieved higher precision rate as compared to the existing system. Table 2 shows the comparative results of recall.

Table 2. Comparison of Recall

Test Images	Recall	
	Existing System	Proposed System
Image 1	73.87	83.07
Image 2	74.04	83.85
Image 3	75.09	83.36
Image 4	76.22	85.72
Image 5	74.95	83.46

From Table 2, it is analyzed that the proposed system achieved higher recall rate as compared to the existing system. Table 3 shows the comparative results of Accuracy.

Table 3. Comparison of Accuracy

Test Images	Accuracy	
	Existing System	Proposed System
Test Image 1	92.26	96.07
Test Image 2	92.46	96.79
Test Image 3	94.20	95.98
Test Image 4	92.37	96.22
Test Image 5	92.90	95.48

From Table 3, it is analyzed that the proposed system achieved higher accuracy as compared to the existing system. Fig. 7 shows the comparative results of precision.

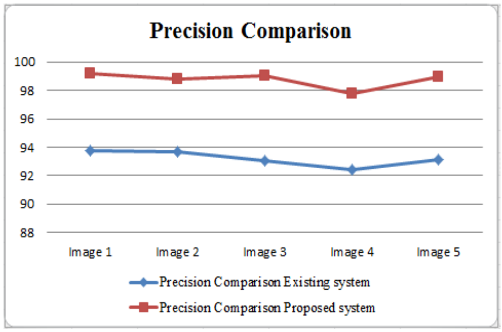

Fig. 7: Precision Comparison

From Fig. 7 it is analyzed that the value of precision generated by the proposed system is high as compared to the existing system. Fig. 8 shows the comparative results of recall.

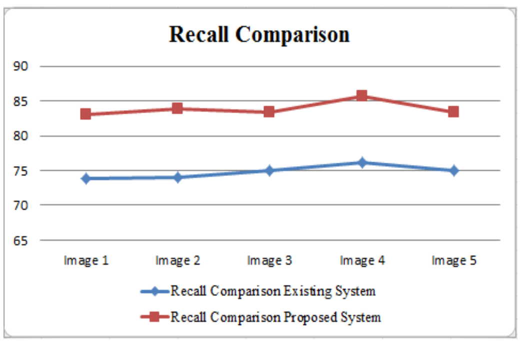

Fig. 8: Recall Comparison

From Fig. 8 it is analyzed that the value of recall generated by the proposed system is high as compared to the existing system. Fig. 9 shows the comparative results of accuracy.

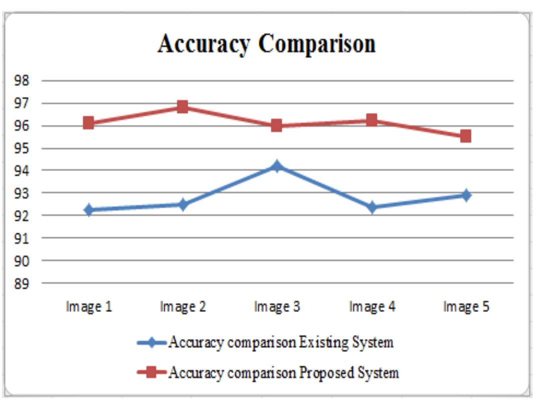

Fig. 9: Accuracy Comparison

From Fig. 9, it is analyzed that the proposed system achieved high accuracy.

V. CONCLUSION

In this work, it is been concluded that plant disease detection required three main steps which are feature extraction, segmentation and classification. Features are extracted through the feature extraction algorithm; the feature can be color features and textural features. From the comparative study of existing and proposed system it is clear that proposed system is better than the existing system in terms of precision, recall and accuracy. The naïve bayes classifier classifies the diseased and non-diseased portion of the leaves. The Bayesian Classification represents a supervised learning method as well as a statistical method for classification. The proposed algorithm is implemented in the MATLAB. The performance of proposed algorithm is tested in terms of accuracy, precision and recall and the results shows that proposed system has high accuracy.

VI. REFERENCES

[1]. Rothe P.R., and Kshirsagar R.V., (2014), “SVM-based Classifier System for Recognition of Cotton Leaf Diseases”, International Journal of Emerging Technologies in Computational and Applied Sciences, Vol. 7, No. 4, pp. 427-432.

[2]. Mainkar P.M, Ghorpade S., and Adawadkar M., (2015), “Plant Leaf Disease Detection and Classification Using Image Processing Techniques”, International Journal of Innovative and Emerging Research in Engineering, Vol. 2, No. 4, pp. 139-144.

[3]. Ugale V., and Gupta D., (2016), “A Comprehensive Survey on Agricultural Image Processing”, International Journal of Science and Research, Vol. 5, No. 1, pp. 133-135.

[4]. 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), Vol. 16, No. 1, pp. 10-16.

[5]. Khirade D. S., and Patil A.B., (2015), “Plant Disease Detection Using Image Processing”, International Conference on Computing Communication Control and Automation, pp. 768-771.

[6]. Thakur P., Aggarwal P., and Juneja M., (2017), “Plant Disease Detection and Classification using Image Processing: A Review”, pp. 22-27.

[7]. Husin Z.B., Shakaff A.Y., Aziz A.H., and Farook R.B., (2012), “Feasibility study on plant chili disease detection using image processing techniques”, Intelligent Systems, Modelling and Simulation (ISMS), Third International Conference, IEEE, pp. 291-296.

[8]. Manikrao N.D., and Dr. Prof. Vyavahare A.J., (2015), “Disease Detection of Cotton crop using Image Processing Technique: A Survey”, International Journal for Research in Applied Science & Engineering Technology, Vol. 3, No. 4, pp. 204-210.

[9]. Priya P., and Dony A. D'souza., (2015), "Study of Feature Extraction Techniques for the Detection of Diseases of Agricultural Products", International Journal Of Research In Electrical, Electronics, Instrumentation and Control Engineering. Vol. 3, No. 1, pp. 4-8.

[10]. Karthik G., and Praburam N., (2016), "Detection and prevention of Banana leaf diseases from banana plant using Embeeded Linux board", International Conference on Green Engineering and Technologies (IC-GET), Online, IEEE, pp. 1-5.

[11]. Dhaware C.G., and Wanjale K.H., (2017), "A modern approach for plant leaf disease classification which depends on leaf image processing", International Conference on Computer Communication and Informatics (ICCCI), IEEE, pp. 1-4.

[12]. Islam M., Dinh A., Wahid K., and Bhowmik P., (2017), "Detection of potato diseases using image segmentation and multiclass support vector machine", 30th Canadian Conference on Electrical and Computer Engineering (CCEC), IEEE, pp. 1-4.

[13]. Patrick A., Pelham S., Culbreath A., Holbrook C.C., De Godoy I.J. and, Li C., (2017), "High throughput phenotyping of tomato spot wilt disease in peanuts using unmanned aerial systems and multispectral imaging", Instrumentation & Measurement Magazine, Vol. 20, No. 3, pp. 4-12.

[14]. Rajan P., Radhakrishnan B., and Suresh L.P., (2016), "Detection and classification of pests from crop images using Support Vector Machine", International Conference on Emerging Technological Trends (ICETT), IEEE, pp. 1-6.

[15]. Tichkule S.K., and Gawali D.H., (2016), "Plant diseases detection using image processing techniques", Online International Conference on Green Engineering and Technologies (IC-GET), IEEE, pp. 1-6.

[16]. Singh V., and Misra A.K., (2017), "Detection of plant leaf diseases using image segmentation and soft computing techniques", Information Processing in Agriculture, Vol. 4, No. 1, pp. 41-49.

[17]. Padol P.B., and Yadav A.A., (2016), "SVM Classifier Based Grape Leaf Disease Detection", Conference on Advances in Signal Processing (CASP), IEEE, pp. 175-179.

[18]. Gowthami K., Pratyusha M., and Somasekhar B., (2017), "Detection of Diseases in Different Plants Using Digital Image Processing", International Journal of Scientific Research in Engineering, Vol. 2, No. 2, pp. 18-23.

[19]. Sethy P., Negi B., and Bhoi N., (2017), "Detection of Healthy and Defected Diseased Leaf of Rice Crop Using K-Means Clustering Technique", International Journal of Computer Applications, Vol. 157, No. 1, pp. 24-27.