INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2021, VOL. 67, NO. 3, PP. 517–522
Manuscript received June 9, 2020; revised July, 2021. DOI: 10.24425/ijet.2021.137841

IoT based Automated Plant Disease Classification using Support Vector Machine

Hiren Mewada, and Jignesh Patoliaya

***Abstract*—Leaf - a significant part of the plant, produces food using the process called photosynthesis. Leaf disease can cause damage to the entire plant and eventually lowers crop production. Machine learning algorithm for classifying five types of diseases, such as Alternaria leaf diseases, Bacterial Blight, Gray Mildew, Leaf Curl and Myrothecium leaf diseases, is proposed in the proposed study. The classification of diseases needs front face of leafs. This paper proposes an automated image acquisition process using a USB camera interfaced with Raspberry PI SoC. The image is transmitted to host PC for classification of diseases using online web server. Pre-processing of the acquired image by host PC to obtain full leaf, and later classification model based on SVM is used to detect type diseases. Results were checked with a 97% accuracy for the collection of acquired images.**

***Keywords*—Plant Disease classification; Support Vector Machine; Graph Cut; Gray-level Co-occurance Matrix**

I. Introduction

Progress in irrigation system and planting with adequate fertilization has increased the quality of mass foods. A plant diseases control system plays an important role to control diseases by preventing the growth of bacteria, fungi etc. Based on the type of diseases, it may be managed with proper pesticides.. However, if diseases are identified then this entire system can become more effective. Hence this paper proposed the diseases diagnoses and classification process.This paper classifies five types of diseases including bacterial bright, leaf curl, myrothecium leaf spot, grey mildew, alternaria leaf spot and normal leaf.

In addition, Internet of Things (IoT) - a sensor embedded network allows access of leaf images remotely and provide ease to classify the diseases in laboratory. In the proposed system, leaf images are accessed remotely using camera and image processing steps are used to extract full leaf and to identify the diseases region from the extracted leaf. The basic steps followed for leaf diseases detection were: image acquisition and pre-processing, segmentation of diseases region, feature extraction and diseases classification. Image acquisition is the process to capture the leaf image using digital camera interfaced with Raspberry PI SoC. The image pre-processing step enhances the image by removing the noise from the leaf image. Then region of leaf affected mostly due to diseases was identified using segmentation process. Features were extracted from the segmented region for further classification of the type of diseases. The proposed work uses multi-support vector machine (SVM) [1], a supervised machine learning technique to classify the leafdiseases. The Block diagram in figure 1 shows the complete working process of the proposed work.

The major constraints for the classification of the diseases are insufficient leaf region capturing, overlay of leaves, few numbers of images in the database. Due to advancement in neural network (NN), many NN algorithms were presented to detect the plant diseases [2]. However, overall detection accuracy is limited to 95% and in all these algorithms, plant images were pre-processed to extract the single leaf of the plant. Thus all these methods are semiautomatic requiring human intervention with high computation cost. **Overall contributions of the paper are:** (1) Leaf diseases symptoms classification require for identification is proposed. (2) Full leaf extraction from the plant image is used without any human intervention. (3) Machine learning proved to be excellent classification algorithms has been used to classify the disease. (4) Disease region is extracted and used in classification algorithm which increases overall accuracy of the classifier.

This paper is divided into five sections. Section II describes types of the leaf diseases. Section III presents work carried out for diseases classification. Proposed system and corresponding results are described in section IV. Finally conclusion is established in section V.

II. Types of leaf diseases

Fungal is the most common disease amongst all types of plants producing spores. It can be seen as spots over the leaf region [3]. Other diseases are due to viral and bacterial. Leaf region has brown necrotic laceration surrounded by a bright yellow circle of light. In the proposed mode, five types of diseases are considered: (1) Fungal (2) The bacterial bright is caused by the bacterial pathogen Xanthomonas Campestris (3) the leaf curl is caused by the fungus genus taphrinavirus (4) gray mildew: in this fungal disease, Initial infected region appear as triangular, square or irregularly size on the leaf. As the diseases increases small spot merge together and form bigger spot. High humidity and low temperature help in the spread of the diseases. (5) The Myrothecium leaf disease is caused by the fungal called myrotheciumroridum. It is causes small-dark brown circular lesion on leaves. (6) The alternaria leaf disease is caused by the fungal called alternariacucumerina. During the initial stage, a small brown (often yellow) spot forms on

Hiren Mewada is with Faculty of Electrical Engineering, Prince Mohammad Bin Fahd University, Al Kobhar, Kingdom of Saudi Arabai (e-mail: hmewada@pmu.edu.sa).

Jignesh Patoliya is with Charotar University of Science and Technology, Changa, India (e-mail: jigneshpatoliay@charusat.ac.in).

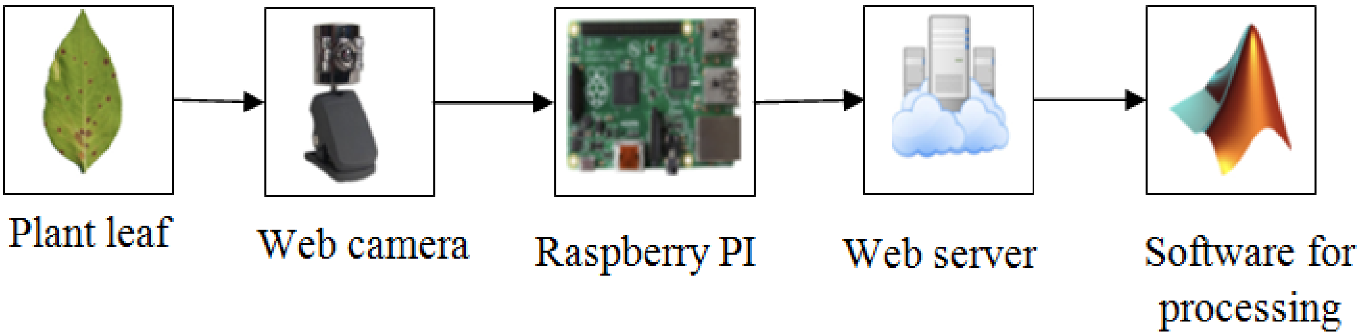


Fig. 1. IoT based leaf image capturing and diseases classification block diagram

TABLE I
SYMPTOMS OF LEAF DISEASES

Fungal [4]	Birds-eye spot on berris septoria brown spot (Leaf spot) yellowing of leaves (Chlorosis)
Bacterial [4]	Leaf spot with yellow halo Fruit spot Sheperd's crook stem ends on woody plants
Leaf curl	reddish areas on developing leaves
Gray mildew	White spot at initial stage grayish colored soft, mushy spots at later stage fuzzy spores
Myrothecium [5]	small and circular, evolve into irregularly shaped spots with a light brown color Oval spots with light-brown centre portions encircled by thick, dark-brown rings
Alternaria	yellow-dark brown to black circular leaf spots concentric rings with black color

leaf and later grows onto irregular brown spots. Few symptoms of the diseases are as expressed in table I and Figure 2 shows the these symptoms over different plants leaf.

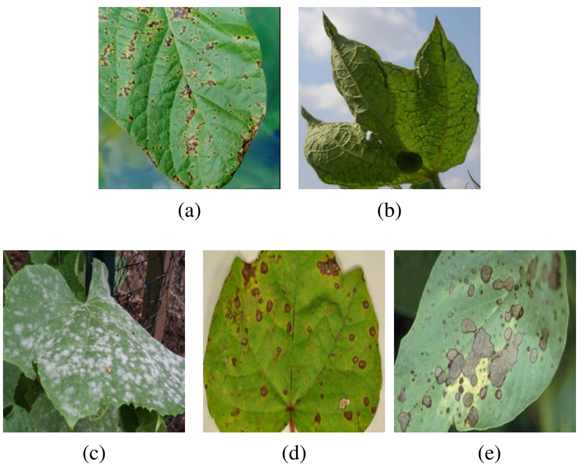


Fig. 2. Type of diseases on leaf (a) Bacterial Blight (b) Leaf curl (c) Gray mildew (d) Myrothecium (e) Alternaria

Overall classification of leaf diseases is proposed in figure 3.

III. LITERATURE SURVEYS

The diseases classification algorithm works with the pattern and/or features extracted from the leaf region. Thus leaf to be processed must have sufficient region to make classification algorithm work efficiently. And this can be achieved of complete isolated leaf image is used as an input to the classification algorithm. It is challenging task in automated system due to varying background, over-placed leaf composition, illumination problem etc. Basic apporach required is leaf sementation or its boundary extraction as suggested in [6]. M Rzanny et al [7] utilized smartphone to capture leaf images. They recorded the leaf front side and back side under flash in and flash off illumination condition with bright background. Then grab cut method is used to segment the leaf. They applied this process for leaf species identification and expressed the importance and effect of wisely image acquisition process over image classification algorithm. This method is appropriate with human intervention. An automatic approach is required which is able to capture leaf image for further process. D Vukadinovic and G Polder [8] proposed fully automated leaf segmentation process. The method was partitioned into two steps: plant segmentation and leaf segmentation. A neural network model was used to segment plant from soil with consideration that all plant images are captured from top. Later leaf was segmented using color histogram and watershed technique. C Niu et al [9] extracted the green component to obtain the leaf region from the background. Then leaf region was enhanced using wavelet transform and canny edge detector. They proposed the improved watershed marking by integrating morphological operation to segment the leaf region. Due to success of machine learning algorithm in all classification application, many authors proposed the use of machine learning in diseases classification. Pinto et al [10] used k-means clustering to extract the region of disease from the leaf image. Later, various machine learning algorithms were tested to classify the diseases.

To develop an automatic plant monitoring systems, many researchers started the use of IoT [11], [12]. Similarly use of IoT is agriculture is also proposed in Yun Shi et al [13]. They presented the structure of IoT which can be used to monitor various information for the diseases identification and insect pest control information. A Kapoor et al [14] used of Ardiuno based IoT and image processing algorithms for smart agriculture. They sensed the temperature and humidity to study effect on pesticides. In the proposed approach, image acquired only once leaf is closest to the camera. This avoid

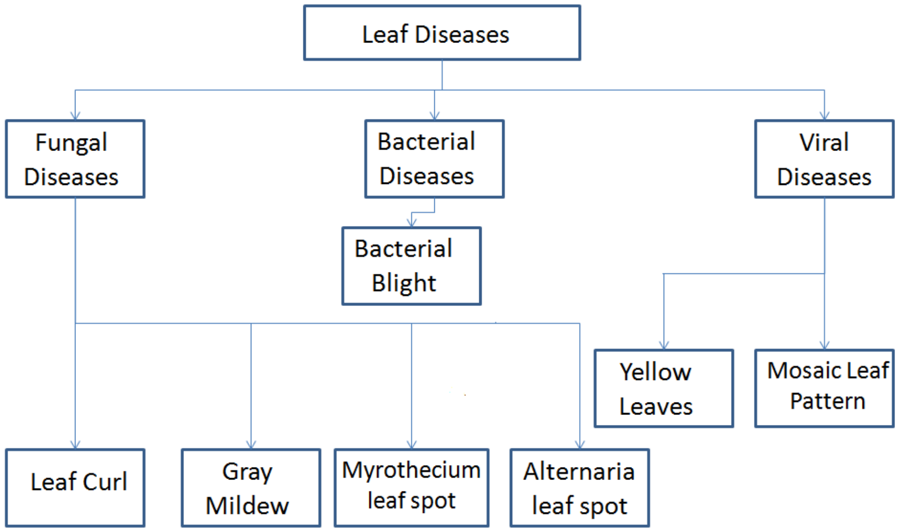


Fig. 3. Classification of leaf disease

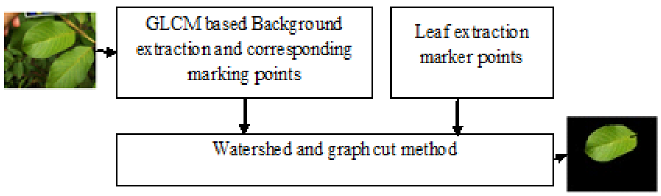


Fig. 4. Leaf segmentation process from acquired image

the chances of over-placed by other leaf and allows capturing of maximum region. Next section describes the proposed leaf image acquisition process using IoT and discusses the leaf disease classification algorithm.

IV. Proposed Methodology

Leaf image acquisition is the first process for leaf diseases classification. Tradition method is capture leaf image placed on plain surface using standard camera. Then apply the segmentation method which includes thresholding, edge detection using gradient operators, region extraction using morphological process. As explained in section III, image acquisition plays important role leaf diseases classification. This paper presents an IoT based image acquisition. In IoT, sensors are connected to backbone processor and it can transmit the sensed data to internet by wireless network. To adopt this method, Raspberry PI SoC capable to capture the leaf images using interfaced USB web camera is used. This web camera captures leaf images. The captured images are transmitted over the internet using Raspberry PI. The host PC accesses the leaf images from the web server and processes it for diseases classification.

A. Leaf region segmentation: This paper makes use of watershed and graph cut segmentation model to segment the individual leaf as suggested in [15]. Overall leaf segmentation process used is shown in figure 4.

There are two possibilities in the acquired leaf image's background: non-green background and texture background. The non-green background is processed using color histogram components. As suggested in [15], [16], two indices with excessive green and red are computed as follows.

$$Ex_R = 2G - R - B, \qquad Ex_G = 1.4R - G - B \quad (1)$$

Otsu algorithm is used to divide the difference between these two indices into two groups i.e background and leaf. In addition, the region having the blue value larger than green value is also considered as background region. To segment the leaf from the textured region, local entropy over the window of $3x3$ is calculated. The region is identified as background if local entropy in the image domain is greater than optimum threshold. The optimum threshold is calculated using gray level co-occurrence matrix (GLCM) which was used to classify the blood vessels from the background in [17]. GLCM provides information about gray level distribution. Co-occurrence matrix are calculated for the direction of $0^o, 45^o, 90^o, 135^o$for each matrix fourteen features to be calculated [3]. With the consideration that the captured image $f(r, c)$ in spatial domain is $n-$ bit image with maximum intensity values of $P = 2^n - 1$, Then GLCM matrix can be represented as G = $[g_{ij}]_{PxP}$, where

$$g_{ij} = \sum_{r=1}^{R} \sum_{c=1}^{C} \delta \quad (2)$$

$$\delta = \begin{cases} 1, & f(r,c) = i, \quad \text{and} \\ & (f(r,c+1) = j \,||\, f(r+1,c) = j) \\ 0, & otherwise \end{cases} \quad (3)$$

Using this co-occurance probability between the two intensity values i and j can be obtained as

$$p_{ij} = \frac{t_{ij}}{\sum_i \sum_j t_{ij}} \quad (4)$$

This GLCM matrix is partitioned in four quadrants using the certain threshold value Th. The First quadrant is used to calculate the local entropy. Finally optimum threshold is obtained by the intensity values corresponding to the maximum entropy of first quadrants. The optimum threshold is used to identify the background of leaf.

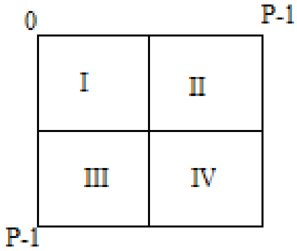
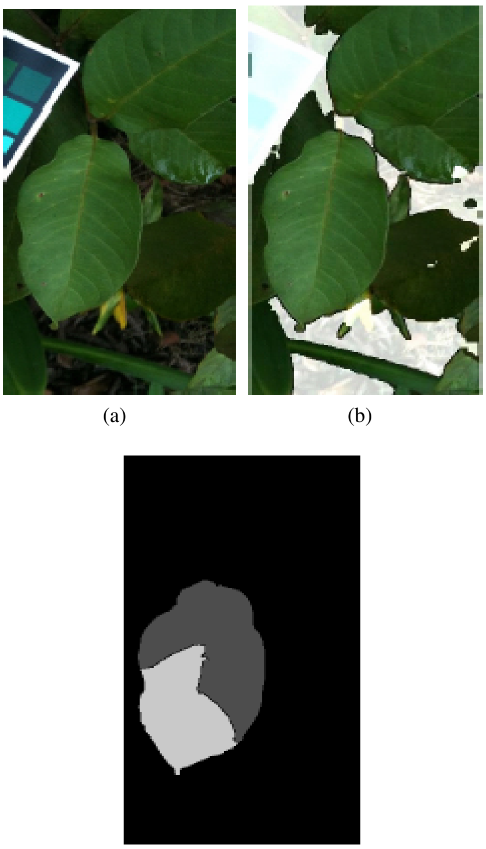


Fig. 5. Co-occurrence matrix partitions

Second step is to find the full leaf from the foreground. The inverse of background is used as input for this step. Then watershed and graph cut algorithm are used as explained in [15] to extract full leaf. Figure 6 presents the intermediate results for the full leaf segmentation.



(a) (b)

(c)

Fig. 6. (a) Original image (b) background extraction (c) Full leaf segmentation

This leaf region is extracted from the original image and further processed form disease classification.

B. Leaf disease classification: The leaf classification algorithms can be classified in three categories: color index based algorithm, threshold based algorithm and machine learning algorithms [18]. Color index methodologies are easy to compute. These algorithms fail when light source is not proper. Threshold based algorithms has fair accuracy if optimum threshold are identified. Machine Learning algorithms offered high accuracy at the cost of complexity [18]. This paper uses the machine learning technique to classify leaf disease. As suggested in section II, classification amongst five diseases is proposed. The diseases classification accuracy depends on the types of features extracted. Machine learning algorithms are proved to be efficient for such feature classification applications. This paper uses support vector machine (SVM) for lulti feature classification. The common stepsare to train SVM using different sets of features obtained from the set of manually labeled samples (i.e. training samples) and supply the query features (i.e. test sample) to classify using the training. The objective of SVM is to cluster the samples having similar features by splitting them with hyperplane w and b. Let $\mathbf{x}_i, y_i \in R^i \times \{\pm 1\}, i = 1......N$ be the training images of leaf with input $\mathbf{x}_i$ and output $y_i \in R^i \times \{\pm 1\}$. After the training,the decision function of the classifier can be written as [19] :

$$f_{w,b}(\mathbf{x}) = \mathrm{sgn}(\mathbf{wx} + b) \tag{5}$$

Where $\mathbf{w}$ presents coefficient vector and b is the bias value, signum (sgn) function provides the sign of number. The classification between the two sets of features can be obtained by satisfying the hyper plane condition

$$y_i\left[(w)(x_i) + b\right] \geq 1, i = 1, 2, ..., K \tag{6}$$

The extension of this binary classification is used to classify multiple features. For the K features, K- SVM classifier can be used as

$$y_i(\mathbf{x}) = f_i(\mathbf{x}) = \mathbf{w}_j \phi_j(\mathbf{x} + b_j), j = 1, 2....K \tag{7}$$

And output of multi-SVM classifier is written as

$$f(\mathbf{x}) = \sum_{j=1}^{K} \mathbf{w}_j \phi(\mathbf{x}) + \sum_{j=1}^{K} b_j \tag{8}$$

Now basic steps involved to classify the disease are as follow (i.e 7):

The full leaf is extracted using the watershed and graph cut segmentation method as discussed in section IV (A). For better accuracy, it is desirable to extract the features from the region infected by the diseases. To extract this region, image was converted to binary using defined threshold and morphological operations i.e. dilation was used to remove the misleading dots. After that, again Otsu's algorithm is used to convert the dilated image into binary image. It split the image domain into white and black region $T0 = \{0, 1, 2, ...t\}$ and $T1 = \{t+1, t+2, ..., L-1\}$ respectively. Where t represents the threshold value and L is number of bits used per pixel in image.

Let $P()$ gives the probability of each pixel in the original image. And $W_{T0}(t), \mu_{T0}(t)$ and $\sigma_{T0}(t)$ are the weight, mean and variance of class $T0$ with intensity value ranging from 0 to t respectively. $W_{T1}(t), \mu_{T1}(t)$ and $\sigma_{T1}(t)$as the weight , mean and variance of class $T1$ with intensity ranging from $t+1$ to $L-1$ respectively. Then best value of threshold t is the value when weighted variance is minimum. This is expressed as follow:

$$\sigma_W^2 = W_{T0}(t)\sigma_{T0}^2(t) + W_{T1}\sigma_{T1}^2(t) \tag{9}$$

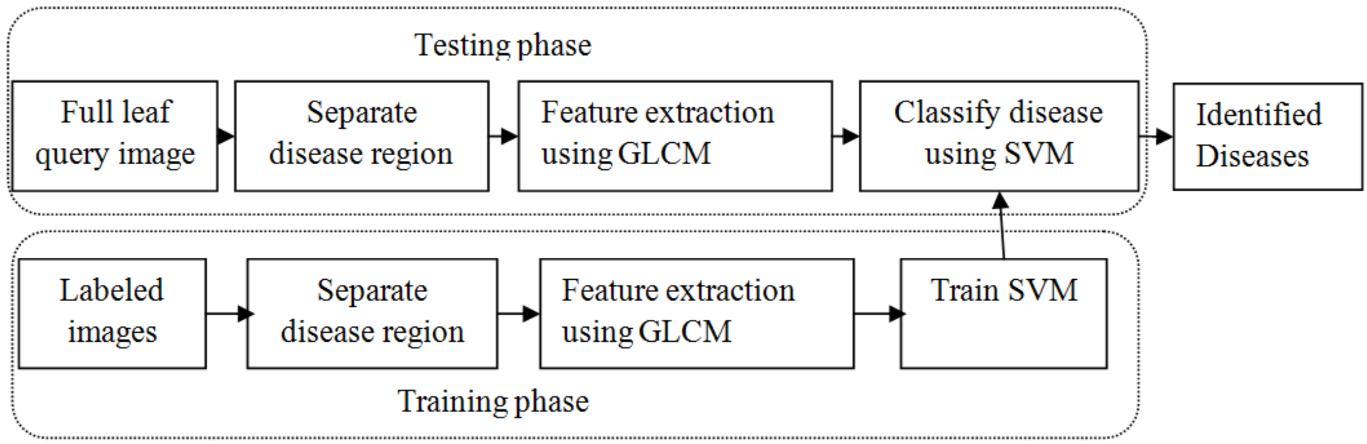


Fig. 7. Full leaf disease classification steps using SVM

Where,

$$W_{T0}(t) = \sum_{i=0}^{t} P(i), \qquad W_{T1}(t) = \sum_{i=t+1}^{L-1} P(i) \tag{10}$$

$$\mu_{T0}(t) = \sum_{i=0}^{t} iP(i), \qquad \mu_{T1}(t) = \sum_{i=t+1}^{L-1} iP(i) \tag{11}$$

$$\begin{aligned} \sigma_{T0}(t) &= \sum_{i=0}^{t}(i - \mu_{T0}(t))^2 P(i), \\ \sigma_{T1}(t) &= \sum_{i=t+1}^{L-1}(i - \mu_{T1}(t))^2 P(i) \end{aligned} \tag{12}$$

With minimization of these variances, best threshold value is selected to obtain binary image. The resultant binary image is masked over the original color image to obtain the diseases region. Figure 8 represents the extracted diseases region from the crop leaf image.

Second step is extraction of features. Each disease has different symptoms. Color characteristic can not be used to classify the disease as expressed in table 1 (i.e. yellow color is common symptom amongst all diseases). The features which can present geometrical shape along with color is the best suitable for such classification. Therefore, texture features extracted using Gray level co-occurrence matrixes were used for the classification. The co-occurrence matrices are calculated for the direction of $0^o, 45^o, 90^o and 135^o$. The joint probability matrix is calculated between the pair of pixels separated by the distance d in the give direction. GLCM calculates the statistical features from these joint probability matrices. Total thirteen statistical features including contrast, correlation, energy, entropy, homogeneity, maximum probability, sum variance, sum entropy, Difference Entropy, IDM Normalize, Cluster shade and Cluster Performance were calculated for each matrix [3]. These statistical features were used to train the SVM classifier and for further classification of query disease. Initially USB camera attached with Raspberry PI was used to capture images. Later, the datasets was prepared by acquiring plant images from various web sources. Due to limited set of datasets, 10 training images were used to train the SVM classifier and 10 testing images were used to calculate the accuracy as suggested in [20], [21].

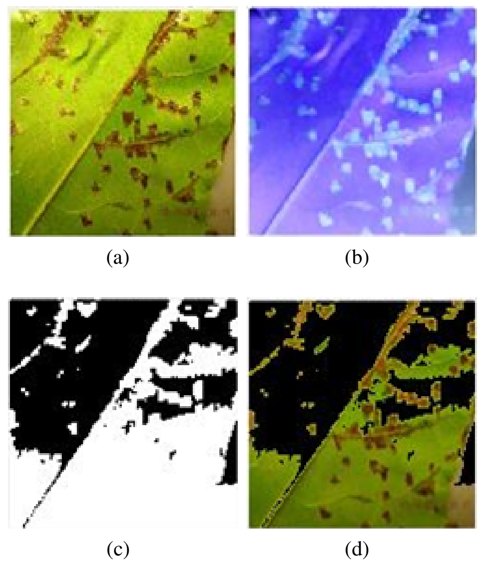


Fig. 8. (a) Original resized image (b) Filtered image (c) image after color inversion and dilation operation (d) disease region extraction

Following table II presents comparative analysis of various SVM based leaf disease classification. For the comparison, average accuracy is used. In [16], the genetic algorithm was used to obtain appropriate clustering. Hence accuracy of this model is better than that obtained in [10]. However, the initial pre-processing like leaf extraction, cropping of desired region was done manually. Hence it is not fully automated system. Whereas, in the proposed model, color-indices were used to obtain marker point required for graph cut segmentation and watershed algorithm. Hence it is fully automated system in comparison with [16].

TABLE II
AVERAGE ACCURACY COMPARISON FOR SVM BASED DISEASES CLASSIFICATION

Algorithm	Features used in classifier	Average accuracy
Morphological based segmentation + SVM [1]	Discrete cosine transform and discrete wavelet transform	94.45 %
Morphological based segmentation + SVM [1]	Discrete wavelet + texture features	83%
Morphological based segmentation + SVM [1]	Discrete cosine transform and + texture features	75%
K means clustering + SVM [10]	Texture features	92.15%
K means clustering + Nearest Neighborhood classifier [10]	Texture features	89.32
K means clustering + Logistic regression model [10]	Texture features	92.57
hline K-means clustering using genetic algorithm + SVM [16]	Texture features	97.6%
Proposed model	**Texture features**	**97%**

V. CONCLUSION

This paper proposed automated leaf extraction and disease classification system using IoT. A raspberry Pi based SoC was used for leaf image acquisition and to transmit over internet. At host PC, automatic full leaf extraction model was used to extract leaf from captured plant image. Later textured features based SVM classifier was used to classify the leaf diseases. Overall five types of diseases were classified with classification of accuracy of 97%. Further testing of the model is required for large sets of data. The experiment can be extended by calculating the accuracy of different sets of plants.

REFERENCES

[1] A. Akhtar, A. Khanum, S. A. Khan, and A. Shaukat, “Automated plant disease analysis (apda): performance comparison of machine learning techniques,” in *2013 11th International Conference on Frontiers of Information Technology*. IEEE, 2013, pp. 60–65.

[2] M. H. Saleem, J. Potgieter, and K. M. Arif, “Plant disease detection and classification by deep learning,” *Plants*, vol. 8, no. 11, p. 468, 2019.

[3] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, “Textural features for image classification,” *IEEE Transactions on systems, man, and cybernetics*, no. 6, pp. 610–621, 1973.

[4] J. Isleib, *Signs and symptoms of plant disease*, 2019 (accessed February 3, 2019). [Online]. Available: https://www.canr.msu.edu/news/signs_and_symptoms_of_plant_disease_is_it_fungal_viral_or_bacterial

[5] R. Borges, M. Rossato, M. Santos, M. Ferreira, M. Fonseca, A. Reis, and L. Boiteux, “First report of a leaf spot caused by paramyrothecium roridum on tectona grandis in brazil,” *Plant Disease*, vol. 102, no. 8, pp. 1661–1661, 2018.

[6] H. K. Mewada, A. V. Patel, and K. K. Mahant, “Concurrent design of active contour for image segmentation using zynq zc702,” *Computers & Electrical Engineering*, vol. 72, pp. 631–643, 2018.

[7] M. Rzanny, M. Seeland, J. Wäldchen, and P. Mäder, “Acquiring and preprocessing leaf images for automated plant identification: understanding the tradeoff between effort and information gain,” *Plant methods*, vol. 13, no. 1, pp. 1–11, 2017.

[8] D. Vukadinovic and G. Polder, “Watershed and supervised classification based fully automated method for separate leaf segmentation,” in *The Netherland Congress on Computer Vision*, 2015, pp. 1–2.

[9] C. Niu, H. Li, Y. Niu, Z. Zhou, Y. Bu, and W. Zheng, “Segmentation of cotton leaves based on improved watershed algorithm,” in *International Conference on Computer and Computing Technologies in Agriculture*. Springer, 2015, pp. 425–436.

[10] L. S. Pinto, A. Ray, M. U. Reddy, P. Perumal, and P. Aishwarya, “Crop disease classification using texture analysis,” in *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. IEEE, 2016, pp. 825–828.

[11] A. Abraham, R. Falcon, and M. Koeppen, *Computational Intelligence in Wireless Sensor Networks: Recent Advances and Future Challenges*. Springer, 2017, vol. 676.

[12] S. Hu, H. Wang, C. She, and J. Wang, “Agont: ontology for agriculture internet of things,” in *International Conference on Computer and Computing Technologies in Agriculture*. Springer, 2010, pp. 131–137.

[13] Y. Shi, Z. Wang, X. Wang, and S. Zhang, “Internet of things application to monitoring plant disease and insect pests,” in *2015 International conference on Applied Science and Engineering Innovation*. Atlantis Press, 2015.

[14] A. Kapoor, S. I. Bhat, S. Shidnal, and A. Mehra, “Implementation of iot (internet of things) and image processing in smart agriculture,” in *2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*. IEEE, 2016, pp. 21–26.

[15] N. Anantrasirichai, S. Hannuna, and N. Canagarajah, “Automatic leaf extraction from outdoor images,” *arXiv preprint arXiv:1709.06437*, 2017.

[16] V. Singh and A. K. Misra, “Detection of plant leaf diseases using image segmentation and soft computing techniques,” *Information processing in Agriculture*, vol. 4, no. 1, pp. 41–49, 2017.

[17] N. P. Singh, R. Kumar, and R. Srivastava, “Local entropy thresholding based fast retinal vessels segmentation by modifying matched filter,” in *International Conference on Computing, Communication & Automation*. IEEE, 2015, pp. 1166–1170.

[18] E. Hamuda, M. Glavin, and E. Jones, “A survey of image processing techniques for plant extraction and segmentation in the field,” *Computers and Electronics in Agriculture*, vol. 125, pp. 184–199, 2016.

[19] B.-y. Sun and M.-c. Lee, “Support vector machine for multiple feature classifcation,” in *2006 IEEE International Conference on Multimedia and Expo*. IEEE, 2006, pp. 501–504.

[20] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini, “Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features,” *Agricultural Engineering International: CIGR Journal*, vol. 15, no. 1, pp. 211–217, 2013.

[21] S. Arivazhagan, R. Shebiah, S. Ananthi, and S. Varthini, “Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features,” *Agricultural Engineering International: CIGR Journal*, vol. 15, no. 1, pp. 211–217, 2013.