



## **Leaf Diseases Detection of Medicinal Plants Based on Support Vector Machine Classification Algorithm**

**Payal Bose<sup>1</sup>, Shawni Dutta<sup>2</sup>, Vishal Goyal<sup>3</sup> and Samir K. Bandyopadhyay<sup>4\*</sup>**

<sup>1</sup>Lincoln University College, Kota Bharu, Kelantan, Malaysia.

<sup>2</sup>Department of Computer Science, The Bhawanipur Education Society College, Kolkata, India.

<sup>3</sup>GLA University, Mathura-Delhi Road Mathura, Chaumuhan, Uttar Pradesh, India.

<sup>4</sup>Lincoln University College, Kota Bharu, Kelantan, Malaysia.

### **Authors' contributions**

*This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.*

### **Article Information**

DOI: 10.9734/JPRI/2021/v33i42A32391

**Editor(s):**

(1) Dr. Prem K. Ramasamy, Brandeis University, USA.

**Reviewers:**

(1) A Venkata Srinivasa Rao, Sasi Institute of Technology & Engineering, India.

(2) Inass Shahadha Hussein, Middle Technical University, Iraq.

Complete Peer review History: <https://www.sdiarticle4.com/review-history/73008>

**Original Research Article**

**Received 18 June 2021**

**Accepted 23 August 2021**

**Published 26 August 2021**

### **ABSTRACT**

On earth, plants play the most important part. Every organ of a plant plays a vital role in the ecological field as well as the medicinal field. But on the whole earth there are several species of plants are available. The different species of plants have different diseases. Therefore, it is required to identify the plants as well as their diseases correctly. It is difficult and also time consuming to identify the plants and their diseases manually. In this research an automatic disease detection system of plant is proposed. High-quality leaf images are used for training and testing. For detecting the healthy area and diseased area in a leaf, region-based and color-based region thresholding techniques are used. For feature selection Histogram Oriented Gradient (HOG) and Local Binary Pattern (LBP) method were applied. Finally, for classification two-class and multi-class Support Vector Machine (SVM) were used. It is found that both feature selection processes with SVM give 99% accuracy. An user oriented graphical user interface is created for understanding the automated system.

**Keywords:** *Image processing; automated plant diseases detection; histogram oriented gradient (HOG); local binary pattern (LBP); support vector machine (SVM).*

\*Corresponding author: E-mail: 1954samir@gmail.com;

## 1. INTRODUCTION

Among all the living things in this world, plants are one of the most important things. Horticulture, a part of agriculture, is used to deal with gardening of useful plants. Horticulture plants like flowers, fruits, vegetables are used to fulfil daily life necessities. So it is necessary to have complete knowledge about these plants as well as the diseases affecting these plants. Detection of plant diseases in the early stages can help to prevent leaf losses and poor production. Detecting of plant diseases manually is a time-consuming process. With the developing of technology, today is quite easy to monitor and detect the plant diseases. Computer Vision technology is much more suitable and gives more accurate results in these cases.

In this paper, an automatic plant diseases detection system is created. This plant disease detection system is based on image segmentation using region-based thresholding and color based thresholding techniques. These techniques are used to detect the areas of disease-affected and healthy leaf region. Finally, a classification algorithm is used to classify all the diseased and healthy leaf categories.

Section II describes the literatures survey. Section III presented proposed methodology for detecting diseases. Experimental results and discussion are in Section IV. The conclusions of the research is given in Section V. Section VI has the reference works that are consulted.

## 2. LITERATURES SURVEY

Different researchers have done researches to detect various plant diseases based on various techniques. The researchers used spectroscopic method and imaging techniques to detect the diseases. With the development of artificial intelligence, machine learning and deep learning based on computer vision study gives more satisfying results. These techniques are widely used to detect plant and leaf diseases automatically.

In a research paper, the authors proposed a real-time grape disease detection process based on an improved convolution neural network [1]. First, they used digital image processing techniques to create a grape leaf disease database. Second, they apply a faster-RCNN and a deep learning-based faster DR-IACNN model to detect the disease. They have also shown that the faster

DR-IACNN model achieved an 81.1% precision rate and detection speed reaches up to 15.01 FPS.

In another research paper, the authors used a convolution neural network with an image processing technique to detect plant diseases [2]. They mostly used common plants that are available all over the world specifically in Iraq. They classified 15 classes, among them 12 classes were diseased and 3 were healthy. They obtained more than 98% accuracy for both the training and testing dataset.

For all nations, economic growth mostly depends on agricultural products. But due to the various plants' diseases, the growth and productivity are decreased from time to time. In the early stage of diagnosis of these diseases can help to prevent the spreading and helps to better productivity. The authors survey various kinds of classification models based on image processing and machine learning techniques to detect and recognized the diseases of agricultural fields. They also show the proposed method and its accuracy for detailed analysis [3].

In another research paper, the authors also survey a detailed analysis of different classification techniques to detect the diseases of various agricultural plants [4].

To detect the apple leaf disease authors proposed a deep neural base improved convolution neural network model. They create an apple leaf diseases dataset based on the laboratory and real-life complex images. Then to detect the diseases they introduce GoogleNet interception and Rainbow concatenation with a deep neural network. They used this model to train up to five apple leaf diseases. The detection performance is 78.80% with high detection speed is 23.13 FPS [5].

To identify the plant leaf diseases the author proposed a model based on image pre-processing and machine learning approaches. They use different image processing techniques to pre-process the image and the GLCM feature extraction method to extract the features of the leaf. Finally, apply the K Nearest Neighbour classifier to detect the diseases of the leaf. The proposed implementation predicts the 98% accuracy for disease detection [6].

To identify automatic crop diseases the authors survey 19 studies based on Convolution Neural

Network (CNN). In this study, the authors describe diseases profile, implementation techniques of CNN, and analysis of the performance of the techniques. Finally, they provide the guidelines for an improved CNN for future research [7]. In another paper, the authors' overview different plant leaf disease detection based on different machine learning classification techniques. They describe different algorithms and their performance [8].

### 3. PROPOSED METHODOLOGY

In this experiment to detect the leaf diseases from an input image, the following steps are involved: 1) Pre-Processing, 2) Segmentation, 3) Feature Selection and Extraction, and 4) Classification.

#### 3.1 Database Details

The complete set of images consists of 12 plants they are named as "Mango", "Arjun", "Saptarni (Alstonia Scholaris)", "Guava", "Bael", "Jamun", "Jatropha", "Pongamia Pinnata", "Basil", "Pomegranate", "Lemon", and "Chinar". All the plants have their own economical and environmental values. The entire database is divided between two main classes "Healthy" and "Diseased" and 22 subclasses. Figure 1 and 2 shows a detailed view of the database images. This database has a total of 4503 images, among them, 2278 images are for healthy leaf images and 2225 are for diseased images. All the images are collected from the Shri Mata Vaishno Devi University, Katra [9-10]. Table 1 shows the details about the experimental database.

#### 3.2 Segmentation Process

##### 3.2.1 Image pre-processing

Image pre-processing [11] means processing the images before the computation processes. This

process commonly involves eliminating the background from the objects, reduces the background noises, resize the input images, enhance the brightness and contrast level. In this experiment, the images in the leaf database are very high in quality [dimension  $6000 \times 4000$ ] [dimension  $\times \times \times$ ]. All the images are in RGB colour and taken with a high-quality digital camera. Therefore, for experimental usage, here the images reduce to dimensions  $600 \times 400 \times \times$  for display purposes and dimensions  $150 \times 100$  for feature extraction purposes. Fig. 3 shows the sample leaf example.

##### 3.2.2 Segmentation

Image segmentation [12-13] is one of the most important parts of image processing. These criteria used to divide the input image into the same type of area and extract the region of interest. Threshold [14] segmentation is a common and simple image segmentation technique. It is a region-based segmentation technique. This method can be divided into two categories 1) Global thresholding [15] and 2) Local thresholding [16]. In the Global thresholding technique, the input image is divided into two main regions 1) background and 2) the region of interest. The global thresholding method uses a single thresholding level to segment the input image. The Local thresholding technique uses multiple thresholding levels to divide the input region and background.

In this experiment, the global thresholding technique is used. For segmentation, a global optimum threshold value is selected. In the leaf image database, the background of the input images is much darker therefore the segmentation effect is more effective. This procedure is faster and the calculation is much simple. Figure 4 shows the segmentation result of the input image.

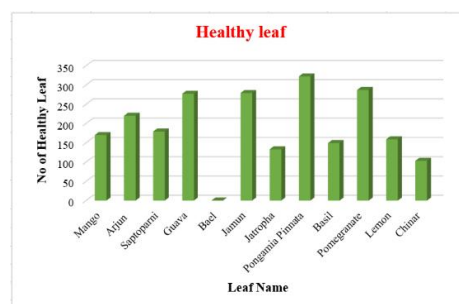


Fig. 1. Healthy leaf details in leaf database

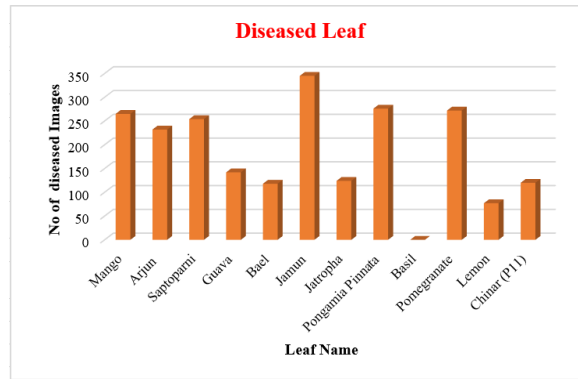


Fig. 2. Diseased leaf details in leaf database

Table 1. Database details

Name of the plant	Healthy Leaf (No of images)	Diseased Leaf (No of images)	Disease Name
Mango	170	265	Anthraco nose
Arjun	220	232	Leaf Spot
Saptoparni	179	254	Foliar Galls
Guava	277	142	Fungal Disease
Jamun	279	345	Fungal Disease
Jatropha	133	124	Leaf Spot
Pongamia Pinnata	322	276	Chlorotic lesions
Pomegranate	287	272	Cercospora spot
Lemon	159	77	Citrus Canker
Chinar	103	120	Leaf Spot

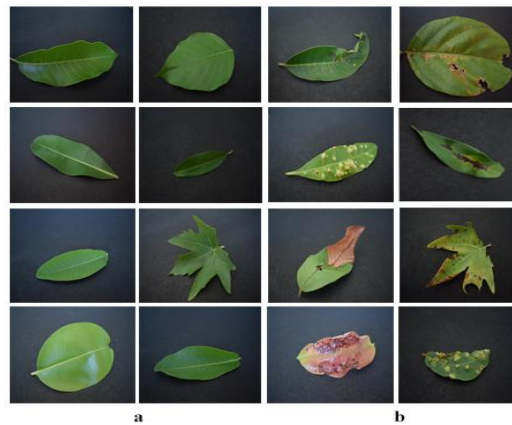


Fig. 3. The sample leaf (a) healthy leaf, (b) diseased leaf

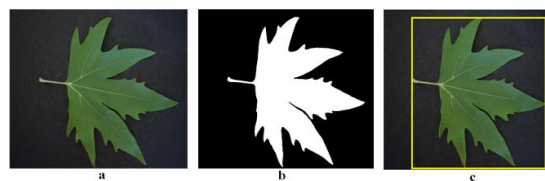
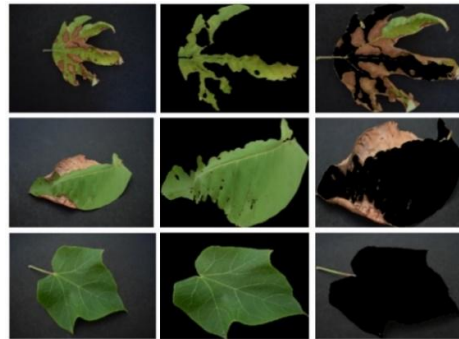


Fig. 4. a) Original Image, b) segmented Image, c) region of interest



**Fig.5. a) Original Image, b) green leaf area, c) disease affected area**

In this experiment, for detecting the diseased affected area a color-based region segmentation technique is used. To segment the green color region from the diseased defected regions first convert the input image into Lab color space. Lab color space is a 3-axis color channel. It split the RGB color image into one lightness (L) and two chromaticity layers (a, b). After converting into Lab color space image, equation 1 is applied to segment the green leaf and diseased affected area.

$$\text{RGB segmenation} = [(0 \leq R \leq 80) \& (-55 \leq G \leq -10) \& (5 \leq B \leq 100)] \quad (1)$$

Fig. 5 shows the segmentation of diseased and green leaf area.

#### **4. FEATURE SELECTION AND EXTRACTION**

Feature selection is a technique in image processing. This technique helps to reduce a large set of features by selecting the best features from the original dataset. It abandons the redundant data from the original dataset. In computer vision object classification or object recognition is one of the most popular subjects. The aim of the object classification is to extract features from the input image and use a proper classifier to classify the feature class label Histogram Oriented Gradient (HOG) and Linear Binary Pattern (LBP) are the two efficient gradient-based feature selection techniques [17]. Their performance is much better than any other feature selection technique.

#### **4.1 Histogram Oriented Gradient (HOG)**

It is a gradient-based feature descriptor [18]. It calculates the existing gradient and orientation of the input image. Basically, this method broke the input image into a small piece of regions, then calculate the vertical and horizontal gradients of those blocks locally. Finally calculate the magnitude and orientations of those gradients.

#### **4.2 Linear Binary Pattern (LBP)**

It is a gradient-based effective texture descriptor [18]. This method divided the input image pixels into  $3 \times 3$  matrices. Then it considers the central value of  $3 \times 3$  matrices as a threshold value and set a binary value based on the threshold value. This process is continued for all pixel matrix and finally, it converts all the binary numbers into decimal and represents the input image in a better way.

#### **5. CLASSIFICATION**

In machine learning classification [19-20] is a process, used to understanding, recognizing, and grouping the same type of data based on their categories. The classification algorithms are used a pre-trained training dataset to predict the category of an unknown sample whose data fall into the predetermined categories.

For classification of an image support vector machine is one of the most popular and best classification methods. It is a supervised learning method. This algorithm is divided the whole data space with a maximum margin to predict the class of the unknown data sample. Fig. 6 shows the classification model of this experiment.

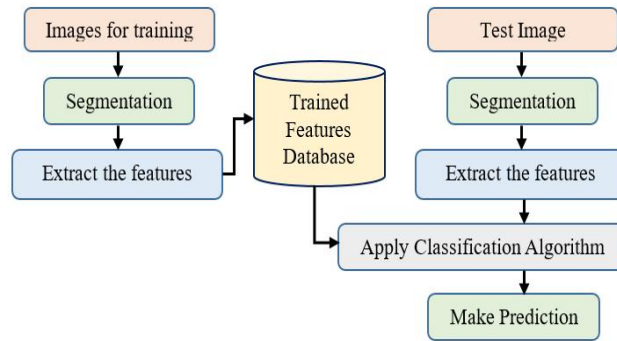


Fig. 6. Classification model

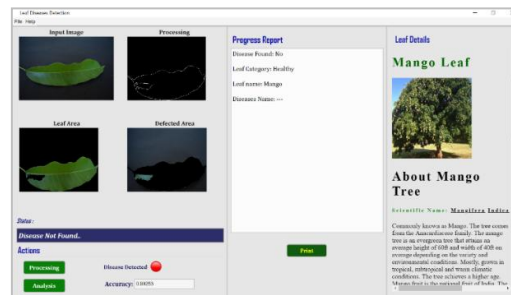


Fig. 7a. Leaf diseases detection and classification (healthy leaf)

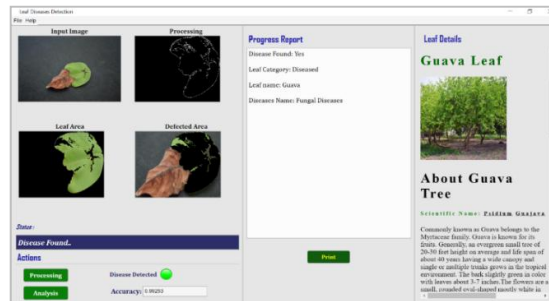


Fig. 7b. Leaf diseases detection and classification (diseased leaf)

Table 2a. Performance analysis using HOG features

Methods	Accuracy	Specificity	Sensitivity
SVM (Proposed Method)	0.99175	0.99447	0.98905
Decision Tree	0.98113	0.98419	0.97811
Linear Analysis	0.89623	0.91621	0.87647
Discriminant Analysis	0.98192	0.97866	0.98514

Table 2b. Performance analysis using LBP features

Methods	Accuracy	Specificity	Sensitivity
SVM (Proposed Method)	0.98146	0.97695	0.99592
Decision Tree	0.98045	0.97695	0.98391
Linear Analysis	0.85305	0.95525	0.75201
Discriminant Analysis	0.96562	0.95932	0.97185

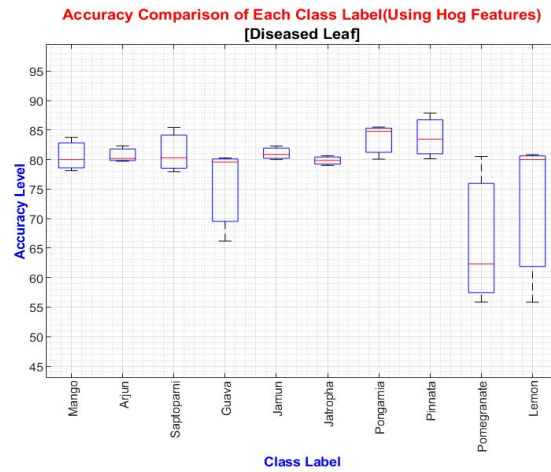


Fig. 8a. Performance details each diseased leaf classes (Using HOG Features)

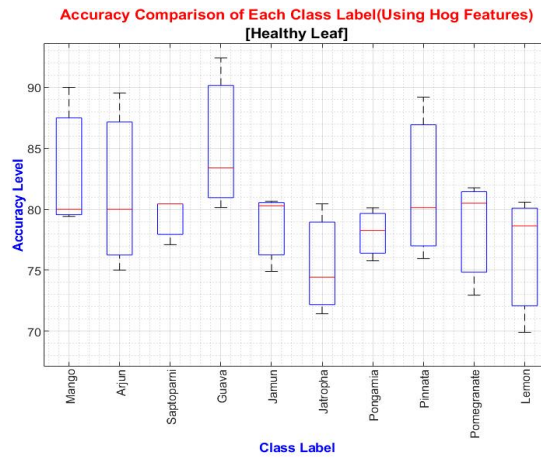


Fig. 8b. Performance details each healthy leaf classes (Using HOG Features)

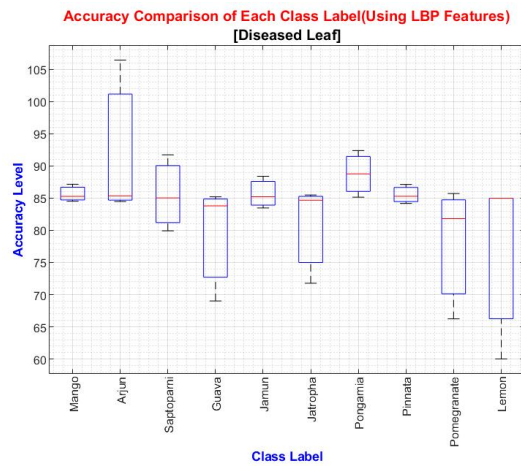
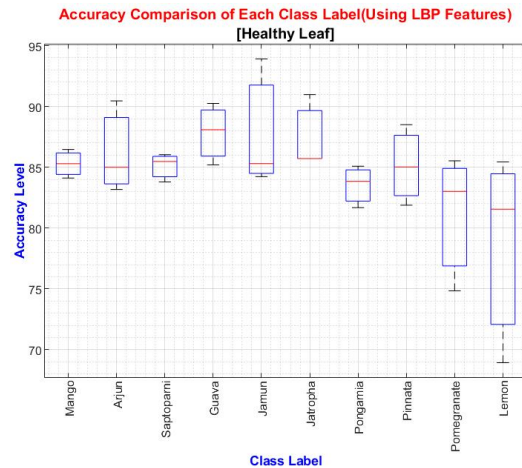


Fig. 9a. Performance details each diseased leaf classes (Using LBP features)



**Fig. 9b. Performance details each healthy leaf classes (Using LBP features)**

## 6. EXPERIMENTAL RESULT AND DISCUSSIONS

The experimental database has 10 leaf classes and two main classes 'Healthy' and 'Diseased'. Figs. 7a and 7b show the leaf disease detection and classification process.

Table 2a and 2b shows the performance analysis of different classifiers with HOG and LBP features extraction.

Figs. 8a and 8b show the performance details of each main class and the subclasses using the HOG feature selection method. Figs. 9a and 9 b show the performance details of each main class and the subclasses using the LBP feature selection method.

## 7. CONCLUSIONS

In this paper, an automated leaf detection system is proposed. Here, region-based thresholding technique is used to detect the healthy and diseased leaf image. Again, to detect the particular diseased area of a diseased leaf the color-based region thresholding method is used. For feature selection from the input images both HOG and LBP feature selection technique is used. To classify the category healthy and diseased leaf with subclasses, leaf name, and disease name the two-class, and multi-class SVM classifier is used. A detailed performance analysis was done between main classes and subclasses by using different classifiers. Finally, a graphical user interface is created for all users.

It is not applicable.

## ETHICAL APPROVAL

It is not applicable.

## COMPETING INTERESTS

Authors have declared that no competing interests exist.

## REFERENCES

- Xie X, Ma Y, Liu B, He J, Li S, Wang H. A Deep-Learning-Based Real-Time Detector for Grape Leaf Diseases Using Improved Convolutional Neural Networks. *Front Plant Sci.* 2020;11:1–14.
- Jasim MA, Al-Tuwaijari JM. Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques. *Proc 2020 Int Conf Comput Sci Softw Eng CSASE 2020.* 2020;259–65.
- Swain S. A Review on Plant Leaf Diseases Detection and Classification Based on Machine Learning Models Page No : 5195. 2020;IX(Vi):5195–205.
- Santhosh Kumar S, Raghavendra BK. Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review. 2019 5th Int Conf Adv Comput Commun Syst ICACCS 2019. 2019;313–6.
- Jiang P, Chen Y, Liu B, He D, Liang C. Real-Time Detection of Apple Leaf



- Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks. IEEE Access. 2019;7:59069–80.
6. Tulshan AS, Raul N. Plant Leaf Disease Detection using Machine Learning. 2019 10th Int Conf ComputCommunNetwTechnol ICCCNT 2019. 2019;1–6.
  7. Boulent J, Foucher S, Théau J, St-Charles PL. Convolutional Neural Networks for the Automatic Identification of Plant Diseases. Front Plant Sci. 2019;10.
  8. Sherly Puspha Annabel L, Annapoorani T, Deepalakshmi P. Machine learning for plant leaf disease detection and classification - A review. Proc 2019 IEEE Int Conf Commun Signal Process ICCSP 2019. 2019;538–42.
  9. Chouhan SS, Singh UP, Kaul A, Jain S. A Data Repository of Leaf Images: Practice towards Plant Conservation with Plant Pathology. 2019 4th Int Conf Inf SystComput Networks, ISCON 2019. 2019;700–7.
  10. CHOUHAN, Siddharth Singh; Kaul, Ajay; SINGH, UDAY PRATAP; & Science, Madhav Institute of Technology (2020), "A Database of Leaf Images: Practice towards Plant Conservation with Plant Pathology", Mendeley Data, V4, DOI: 10.17632/hb74ynkjc.4
  11. Koprowski R. Image pre-processing. Stud ComputIntell. 2017;682:21–38.
  12. Zaitoun NM, Aqel MJ. Survey on Image Segmentation Techniques. Procedia Comput Sci [Internet]. 2015;65:797–806. Available: <http://dx.doi.org/10.1016/j.procs.2015.09.027>
  13. Singh V, Misra AK. Detection of plant leaf diseases using image segmentation and soft computing techniques. Inf Process Agric [Internet]. 2017;4(1):41–9. Available: <http://dx.doi.org/10.1016/j.inpa.2016.10.005>
  14. N S, S V. Image Segmentation By Using Thresholding Techniques For Medical Images. Comput Sci Eng An Int J. 2016;6(1):1–13.
  15. All Answers Ltd. Global Thresholding Image Segmentation Techniques [Internet]. November. [Accessed 19 August 2021]; 2018. Available: <https://ukdiss.com/examples/global-thresholding-image-segmentation-techniques.php?vref=1>.
  16. Guruprasad P. Overview of Different Thresholding Methods In Image Processing. TEQIP Spons 3rd Natl Conf ETACC; 2020.
  17. Greeshma K V, College C, Gripsy JV. Image Classification using HOG and LBP Feature Descriptors with SVM and CNN. 2020;8(04):4–7.
  18. Islam MA, Billah M, Yousuf SI. Automatic Plant Detection Using HOG and LBP Features With SVM. Int J Comput [Internet]. 2019;33(1): 26–38. Available: <http://ijcjournal.org/>
  19. Oo YM, Htun NC. Plant Leaf Disease Detection and Classification using Image Processing. Int J Res Eng. 2018;5(9):516–23.
  20. Shruthi U, Nagaveni V, Raghavendra BK. A Review on Machine Learning Classification Techniques for Plant Disease Detection. 2019 5th Int Conf Adv Comput CommunSyst ICACCS 2019. 2019;281–4.

© 2021 Bose et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*

*The peer review history for this paper can be accessed here:*  
<https://www.sdiarticle4.com/review-history/73008>