Applied Computer Systems
ISSN 2255-8691 (online)
ISSN 2255-8683 (print)
June 2023, vol. 28, no. 1, pp. 92–99
https://doi.org/10.2478/acss-2023-0009
https://content.sciendo.com

Detection and Classification of Banana Leaf Disease Using Novel Segmentation and Ensemble Machine Learning Approach

Vandana Chaudhari[1*], Manoj P. Patil[2]
[1]*Smt.G.G. Khadse College, Muktainagar, India*
[2]*School of Computer Sciences, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, India*

***Abstract* – Plant diseases are a primary hazard to the productiveness of crops, which impacts food protection and decreases the profitability of farmers. Consequently, identification of plant diseases becomes a crucial task. By taking the right nurturing measures to remediate these diseases in the early stages can drastically help in fending off the reduction in productivity/profit. Providing an intelligent and automated solution becomes a necessity. This can be achieved with the help of machine learning techniques. It involves a number of steps like image acquisition, image pre-processing using filtering and contrast enhancement techniques. Image segmentation, which is a crucial part in disease detection system, is done by applying genetic algorithm and the colour, texture features extracted using a local binary pattern. The novelty of this approach is applying the genetic algorithm for image segmentation and combining a set of propositions from all the learning classifiers with an ensemble method and calculating the results. This obeys the optimistic features of all the learning classifiers. System accuracy is evaluated using precision, recall, and accuracy measures. After analysing the results, it clearly shows that the ensemble models deliver very good accuracy of over 92 % as compared to an individual SVM, Naïve Bayes, and KNN classifiers.**

***Keywords* – Banana, diseases, ensemble learning, local binary pattern.**

I. Introduction

Our daily lives depend primarily on agriculture, which also improves country's economy by providing occupation to many people. The geographical and climatic conditions decide the type of plant cultivation in a particular area. Banana is one of the major fruit crops in India [1]. Farmers cannot enhance their banana yield due to the plant diseases. The diseases on banana crops significantly impact the production. The plant diseases are mainly categorised into fungus, bacterial and viral diseases [2]. According to the review, environmental factors such as temperature, humidity, wounds and rainfall cause diseases to arise. The increased humidity and temperature are the growing situation for some bacterial diseases. The disease develops in hot, humid climate conditions with a plenty of rainfall and high temperature and soil moisture. Detection of correct disease by the farmers can help in improving the quality and yield of banana farming. Various disease types are found on the banana plant depending on the country and the region. Some of the major diseases found on the banana plants in Maharashtra state of India are banana bacterial wilt, yellow and black sigatoka, bunchy top and cucumber mosaic virus. In present circumstances, the farmers need help from experts for identification of various diseases. Disease detection in the traditional way is time-consuming and costly. The outcome of the disease detection depends on the experience and expertise. Accuracy and early detection of plant diseases are necessary to reduce loss in the crop yield and control disease severity [3], [4]. An automated system can help in reducing the overload of a manual system with the help of banana plant image analysis and identification by classification using machine learning techniques. The primary impact of this study is the development of an automated system using image processing through an artificial intelligent approach for the early identification of viral, bacterial and fungal diseases on the banana crop caused by ecological changes. The suggested method is quick, automatic, and low-cost in terms of computation. A variety of images of diseases and healthy leaves are used to train the proposed model. The ensemble classification approach can make the banana leaf disease classification effective and reliable. Hence, the use of automated system is useful to increase the yield and to reduce the loss occurred due to ecological changes.

In this study, the detection of banana plant disease was carried out using ensemble machine learning classification techniques. Visual symptoms appearing on the banana plant were used as inputs which were captured through the images. For analysis of the suggested task, a majority of disease symptoms that manifest on the leaf were taken into account. An automated approach for disease diagnosis was suggested for the common banana diseases that appear in the Maharashtra area. Corresponding symptoms are provided in Section 2. Section 3 introduces to the most promising work carried out in the area of disease identification on various plants. The proposed system architecture is described in Section 4 along with database creation by acquisition of images, image pre-processing,

*Corresponding author's e-mail: vandu2108@gmail.com , mpp145@gmail.com

©2023 Vandana Chaudhari, Manoj P. Patil.
This is an open access article licensed under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

segmentation, feature extraction from pre-processed and segmented images and the role of ensemble approach for classification of images. Section 5 describes the experimental results that are discussed thoroughly to draw fruitful conclusions. Finally, the conclusions are made in Section 6.

II. Banana Plant Disease

The banana disease identification using an automated system is carried out for common diseases appearing on banana leaves viz. Bacterial Wilt, Cucumber Mosaic Virus, Yellow Sigatoka, Bunchy Top, Black Sigatoka, Pitting and Streak Virus disease.

A. Bacterial Wilt

Bacterial wilt is a disease caused by bacteria *Xanthomonas vasicola* pv. *musacearum* (Xvm). Xanthomonas wilt is mostly found in India and Africa. In Ethiopia, the bacterial wilt was reported for the first time on banana in 1968 [5]. Bacterial wilt is more harmful to banana plants because of its easy spreading and easy in-plant development. The lack of resistive cultivators in plants is a major reason for this disease [6]. The infection is seen on a leaf by a change in its colour, initially it becomes yellow, later turns brown and finally it dies. The petiole breaks and leaves hangs. Older leaves become chlorotic, wilted and starts collapsing [7]. This bacterial infection is generally transmitted by farm equipment, animals, infected rhizomes and insects. High soil moisture and high temperatures usually favour the wilt disease [8].

B. Cucumber Mosaic Virus (Infectious Chlorosis)

Cucumber Mosaic Virus (CMV) is a viral disease. The leaves start becoming yellowish due to light yellow-white streaks appearing on the leaves. This creates a striped appearance on the leaf. The streaks generally span from midrib to blade edge. CMV is usually transmitted by aphids and also may appear through an infected seed [9].

C. Yellow Sigatoka

Yellow Sigatoka is a fungal disease caused by *Mycosphaerella musae*. The spread starts from the young leaves. This part of the leaf become greyish with a yellowish halo running parallel to veins on the third or fourth leaf from the top. Streaks are enlarged and turn brown with chlorotic halo. Yellow Sigatoka disease was first reported on banana in Java in 1902. Sigatoka transmits fungus that is conveyed by wind, rainwater, and old dried infected leaves, all of these contribute to the disease spread [10].

D. Bunchy Top

The banana bunchy top virus causes a viral infection as banana bunchy top. Leaf size reduction is the main symptom to discover this disease. The leaves grow short and narrow with overturned margins [10]. Leaf becomes stiff and brittle with short leaves becoming fragile. Young leaves develop dark green lies on the undersides [11].

E. Black Sigatoka

The fungus *Mycosphaerella gloeosporioides* causes Black Sigatoka, which is a fungal disease. Light yellow or brownish-green narrow streaks appear on a leaf and they extend over the entire leaf [7]. The disease spot enlarges further and the tissue around the lesions turns yellow and dies. Lesions merge, the whole leaf turns brown and leaf becomes dry rapidly, which causes defoliation [10].

F. Pitting

Banana pitting disease is caused by *Physalis angulata*. Primarily blast spots are found on emerging leaves of side suckers. Mature leaves show light brownish spots in the small oval shape or shape of eyes, sometimes with yellow border. During favourable weather conditions, the midrib petiole also worsens blast spots. Sporulating lesions are white grey in colour, similar to the spores [12].

G. Streak Virus

Banana streak virus is the member of the genus badnavirus [13]. Yellow vertical streaks appear on the leaves, which turn from yellow to brown and then to black. The disease is mainly spread by infected planting material and is likely to interrupt plant growth.

III. Related Work

Das et al. presented the model to detect tomato leaf diseases. Mean, entropy, skewness, and kurtosis of texture features were extracted using Haralick's algorithm. The accuracy of the extracted features was evaluated using the random forest, logistic regression, and support vector machine classifier. The experimental results show that the SVM classifier gives better results than the other classifiers [14]. Suganya Devi et al. proposed a system to identify groundnut leaf diseases. Images were segmented to find the diseased part by using the HSV segmentation, and features were extracted using Harris corner detector followed by a Histogram of Gradient method. Classification of disease was performed by using the K-Nearest Neighbour classifier. The results using the proposed H2K method were compared against the existing multiclass SVM method and demonstrated that the proposed method was robust and optimum to classify groundnut leaf diseases [4]. Hassanien et al. implemented an improved moth flame approach that automatically detected the disease on the tomato plants. In the proposed system, an algorithm for feature selection was implemented and evaluated (MFORSFS). In the work, the proposed algorithm of feature selection was compared against feature selection based on Particle Swarm Optimization (PSO), Genetic Algorithms. It was found that the proposed algorithm was robust, performed better and faster [15]. Mia et al. proposed mango leaf disease detection system. RGB images were converted to L*a*b images for processing. Image segmentation was performed using the K-means clustering method. GLCM and colour features were extracted from the segmented region. The extracted features were used to detect the disease using the SVM classifier. The proposed system showed an average of 80 % accuracy [14]. Anjna et al. presented the work to identify the capsicum disease. Image enhancement was performed using histogram equalization technique. GLCM features were extracted with tree, linear discriminant, SVM and KNN

techniques used for classification. Results showed that the tree and linear discriminant gave 98.4 % and SVM, KNN gave 100 % accuracy [16]. Bedi et al. proposed the hybrid model to detect peach leaf diseases. The model created a CAE network and reduced the dimension of images from 256×256 into 32×32. The CAE network was trained to reduce the reconstruction losses. The loss between the original and reconstructed images was computed using the Normalized Root Mean Squared Error (NRMSE) function. The CNN classifier was used to classify an image [17]. Morgan et al. classified and predicted mushroom and soybean diseases. Different classifiers (ANN, Naïve Bayes, Support Vector Machine, KNN, Decision Tree, and Random Forest) were used to predict diseases. Different parameters were calculated to check the correctness of the classifier, such as accuracy, MAE, TP-rate, FP-rate, Precision, Recall, and F-measure. Results showed that the naïve Bayes classifier gave better accuracy and other classifiers performed well on the mushroom dataset with almost 100 % accuracy [18]. Patil et al. presented the cotton disease detection system using machine learning techniques. Cotton images were segmented from background using the factorization based active contour method. The texture and colour features were extracted, and SVM, multilayer perceptron, Random Forest, Naïve Bayes, KNN and AdaBoost classifiers were used to categorize the diseases. Experiments were carried out with colour features, texture features and combination of both colour and texture features. The experiments were carried out using SVM, multilayer perceptron, Random Forest, Naïve Bayes, KNN and AdaBoost. The work concluded that only colour features were sufficient to improve the accuracy of classification [19].

IV. System Architecture

The proposed system for automatic detection of banana plant diseases is described in this section.

The Proposed Approach

Five essential stages are performed to detect the plant diseases: Image acquisition, image pre-processing, image segmentation, feature extraction, and image classification. The architecture of the proposed system is shown in Fig. 1.

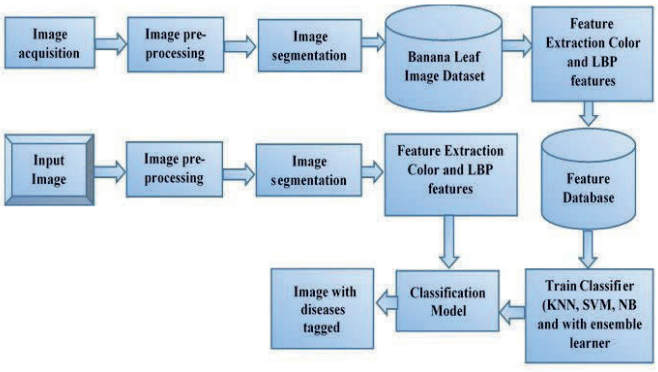


Fig. 1. The proposed banana leaf disease system architecture.

A. Image Acquisition

The first stage of the suggested system is image acquisition. To achieve accuracy in an automated system, an accurate image dataset must be constructed. Real images from several agricultural farms were collected to create the dataset needed for the proposed approach. For experimental work, Black Sigatoka, Yellow Sigatoka, Cucumber Mosaic Virus, Bunchy Top, Bacterial Wilt, Pitting, and Streak Virus with healthy leaves were taken into consideration. Disease images were captured using 48 MP resolution mobile camera in the present environment. Some images of Black Sigatoka and Bacterial Wilt were used from the predefined dataset [20].

Figure 2 illustrates the sample disease images used during the experiment.

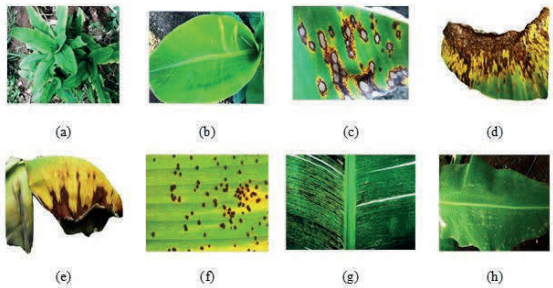


Fig. 2. Sample images: (a) Bunchy Top; (b) CMV; (c) Yellow Sigatoka; (d) Black Sigatoka; (e) Bacterial Wilt; (f) Pitting; (g) Streak Virus; (h) Healthy leaf.

B. Image Pre-processing

All the captured images were resized into uniform size 640×480. Noise was present in the images taken in the current environment. To eliminate the noise from the dataset, average filtering technique was applied. Equation (1) was used to calculate the average filtering with default filter size [3,3]. L*a*b colour space was explored because it closely matched with human perception. The RGB images were converted into L*a*b* colour space and the L* component of the image was used for image enhancement process. The adaptive contrast enhancement approach was used to execute the image enhancement procedure.

$$ones\left(n(1),n(2)\right)/\left(n(1)*n(2)\right). \quad (1)$$

C. Image Segmentation

Finding the area of interest requires image segmentation. The genetic algorithm was applied for component segmentation. The searching ability of genetic algorithms was used to perform clustering properly from the given image. The $r \times c$ colour image was used during the experiment. Each pixel in image consisted of three components: red, green, and blue. Each chromosome displays a solution that consists of the centre of K cluster. The population was initialised randomly and in each step of processing, the best chromosome was continued for the upcoming processing round.

In the initial phase of fitness function, the input pixels were clustered based on nearest corresponding cluster centres so that each colour pixel of the image was placed into its corresponding cluster, which had cluster centre CL_k for $k = 1, 2, \ldots K$ using (2).

$$\|IP_i - CL_k\| < \|IP_i - CL_l\|, \quad i = 1,2,\dots.,r \times c, \quad l = 1,2,\dots.K. \tag{2}$$

In next step, a mean of every pixel in the selected clusters was calculated in order to determine the centre of the new cluster CL_i.

$$CL_i(r,g,b) = \frac{1}{n_i}\sum_{IP_k \in C_i}\left(IP_k(r,g,b)\right), i = 1,2\dots k. \tag{3}$$

Finally, the fitness function was computed by finding the Euclidean distance between pixels and its associated cluster by (4).

$$F = \sum F_i. \tag{4}$$

$$F_i = \sum_{IP_k \in C_i}\left|\left(IP_k(r,g,b) - CL_i(r,g,b)\right)\right|. \tag{5}$$

The disease cluster was automatically determined by executing following steps.

Step 1: [check the green component of cluster image]
for i = 1: No of clusters
Select the ith cluster image as input image.
Choose the green component of the image.
Observe the values of green component greater than the specified value (0.5).
Calculate the amount of green pixels in the segmented image.
Select the cluster having a minimum amount of green pixels.
end for
Step 2: [Store the selected disease image for next process].

Figure 3 shows the sample disease image and the corresponding pre-processed and segmented image.

D.Feature Extraction

Feature extraction plays a significant role in the process of image categorization. In this phase, colour and local binary pattern texture features were used to describe the textural pattern of banana leaves diseases.

Colour Feature

Colour is a very important aspect in categorizing the diseases. For the manual disease detection system colour is one of the most visible component. Similarly, anyone may simply manipulate the image colours using digital image processing techniques. Generally, the input image is given in RGB colour space, in which each pixel is addressed as the combination of RGB. Also other colour spaces, such as CIE, HIS, HSV, YCbCr, L*a*b*, are frequently used in various segmentation processes to analyse and examine the images [21].

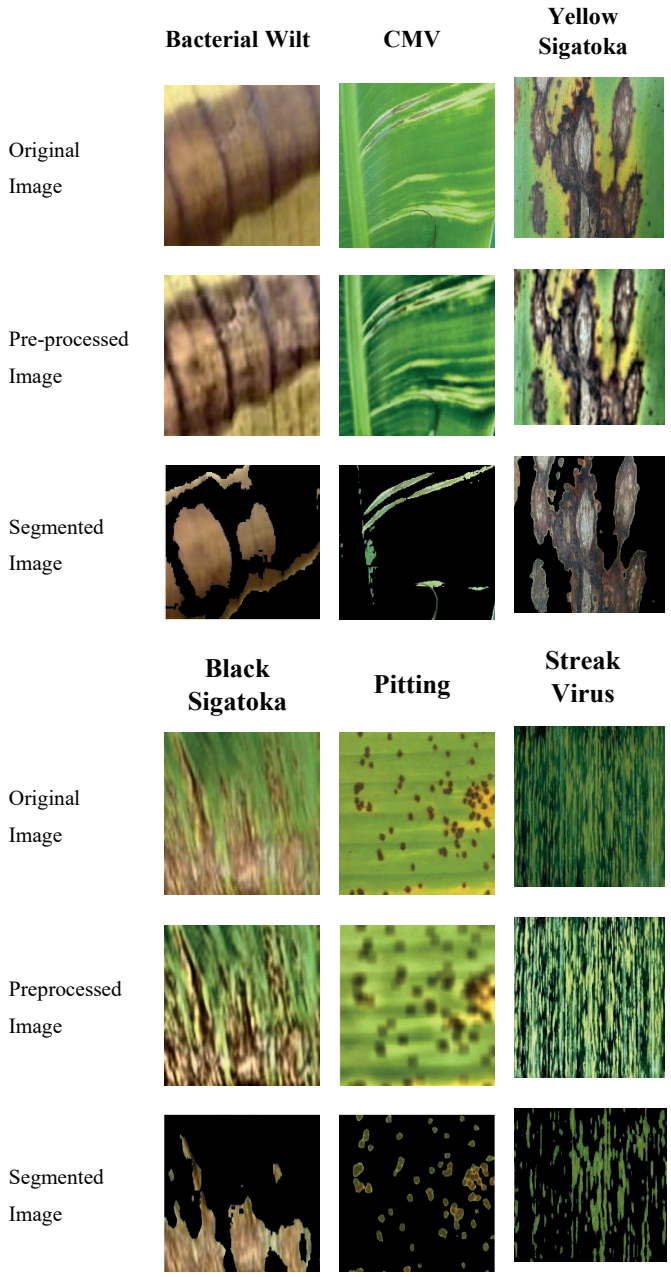


Fig. 3. Sample of disease image and the generated segmented image.

Different colour features are extracted such as mean, moment, standard deviation, skewness and kurtosis for the experimental work. Mean is the ratio of summation of all elements with the number of elements calculated using (6). Moment is the asymmetry in the image evaluated using (7). Equation 8 calculates the square root of the variance called standard deviation. Skewness describes the amount of data spread around the sample mean computed using (9). Equation (10) is used for kurtosis to measure in what way the outlier prone is distributed.

$$\text{Mean} = \frac{1}{N}\sum_{i=1}^{N} Ai. \tag{6}$$

$$m_k = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^k. \quad (7)$$

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{\sum(x_i - \mu)^2}{n}}. \quad (8)$$

$$s1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^3}. \quad (9)$$

$$k1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}. \quad (10)$$

Texture Feature

Different authors recommended various methods for extracting texture features, such as the extraction of Gabor features, local binary patterns, gradient histograms, and the co-occurrence matrix of grey levels. In the proposed study, a local binary pattern technique was applied to extract the local texture features in the images. Centre pixel and the adjacent pixels were compared using (11) having results zero or 1.

$$LBP_{P,R} = \sum_{P=0}^{P-1} S(g_P - g_c)2^P \quad ,S(x)$$

$$= \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}. \quad (11)$$

Here, centre pixel grey value is indicated by g_c and value of neighbourhood pixel is represented by g_p; $S(x)$ shows the function of thresholding stage [22].

E. Classification

In the classification phase, the KNN, Naïve Bayes, and SVM were employed to categorize the banana disease images.

Traditional Classifier

KNN is one of the simplest classifiers used for input image classification. This classifier categories the testing sample based on nearby training samples in the feature space. Euclidean distance measure is used to calculate distance between the two scenarios. SVM is one of the widely used supervised techniques generally used in pattern recognition and classification. This technique converts the features into high dimensional space and forms a hyperplane to categorize the data [23]. Figure 4 shows the binary classification of SVM classifier. For real life categorization, multiclass classification approach is used.

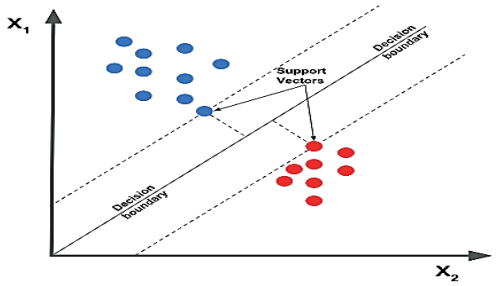


Fig. 4. Binary classification of SVM.

The disease categorization is also performed using Naïve Bayes classification techniques. Naïve Bayes classification technique is based on Bayes theorem. It assumes independence among predictors [24]. Equation (12) shows the Bayes Theorem:

$$P\left(\frac{c}{x}\right) = \frac{P(x/c)P(c)}{P(x)}, \quad (12)$$

$P(c/X) = P(x_1/c) \times P(x_2/c) \times \ldots\ldots \times P(x_n/c) \times P(c)$, where
$P(c/x)$ – class posterior probability;
$P(c)$ – class prior probability;
$P(x/c)$ – probability of predictor given class;
$P(x)$ – predictor prior probability.

Ensemble Learning Approach

The ensemble learning approach uses a number of classifiers to categorize data by using the weighted vote of classifiers [25]. The regular machine learning models attempt to learn a single hypothesis from the given data. Ensemble learning combines the set of hypotheses from all learners and predicts the result. According to the previous study, ensemble learning predicts good accuracy as compared to single classifier results.

The following algorithmic steps are applied to build the basic ensemble approach.

Algorithm Ensemble_method
Input: Training data $TD = \{x_i, y_i\}_{i=1}^{n}$
Output: Ensemble learning classifier E
Step 1: Learn the base level classifiers
for c_l=1 to CL do
learn e_{cl} based on TD
end for
Step 2: Generate new dataset of predictions
for i=1 to n do
$TD_E = \{x_i', y_i\}, where\ x_i' \{e_1(x_i), \ldots, e_{CL}(x_i)\}$
end for
Step 3: Learn the meta classifier
Learn E based on TD_E
return

To calculate the performance of a system, k-fold cross-validation process was used. This method divides the data into two sets: training and testing. The data were divided into k-fold and one subsample among the k-fold was used to test the system and other k-1 subsamples were used to train the system. Applying the k-fold method, all the sample images were used in the training and testing process.

V. Results and Discussion

For the purpose of identifying diseases on banana plants, KNN, SVM, Naïve Bayes, and ensemble classification methods were applied. The ensemble classifier is stacked with KNN, Naïve Bayes, and SVM with Gaussian and polynomial kernel functions. The performance of the proposed work is to calculate, using confusion matrix, three measures: accuracy, precision and recall.

Accuracy: It measures the proportion of properly categorized as disease affected and calculated using (13).

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP}. \tag{13}$$

Precision: This measures the proportion of truly positives sample among the positives using (14).

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{14}$$

Recall: Equation (15) is used to measure the proportion of truly positives sample among the retrieved samples.

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{15}$$

Figure 5 presents the confusion matrix of ensemble learner and different traditional classifiers, such as SVM, NB and KNN. The number from 1 to 8 refers to Bacterial Wilt, Cucumber Mosaic Virus, Yellow Sigatoka, Bunchy Top, Black Sigatoka, Healthy leaf, Banana Pitting, and Banana Streak Virus Disease, respectively. Graphical representation of accuracy measure is shown in Fig. 6.

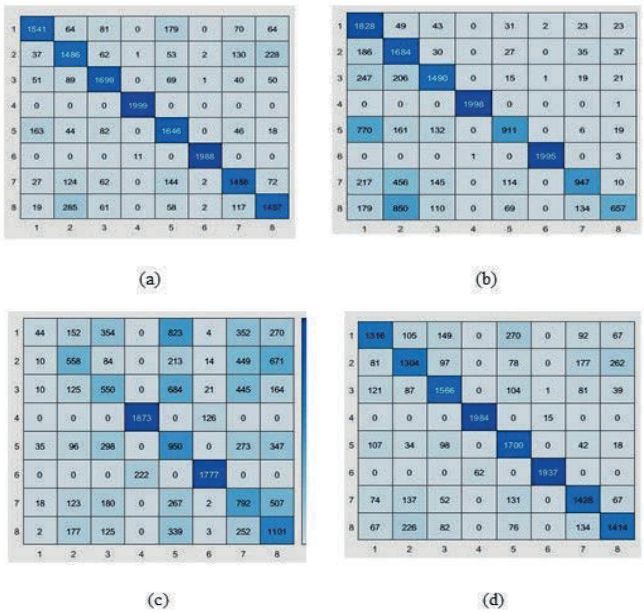


Fig. 5. (a) Confusion matrix of the proposed ensemble learner classifier; (b) SVM Classifier; (c) Naïve Bayes Classifier; (d) KNN Classifier.

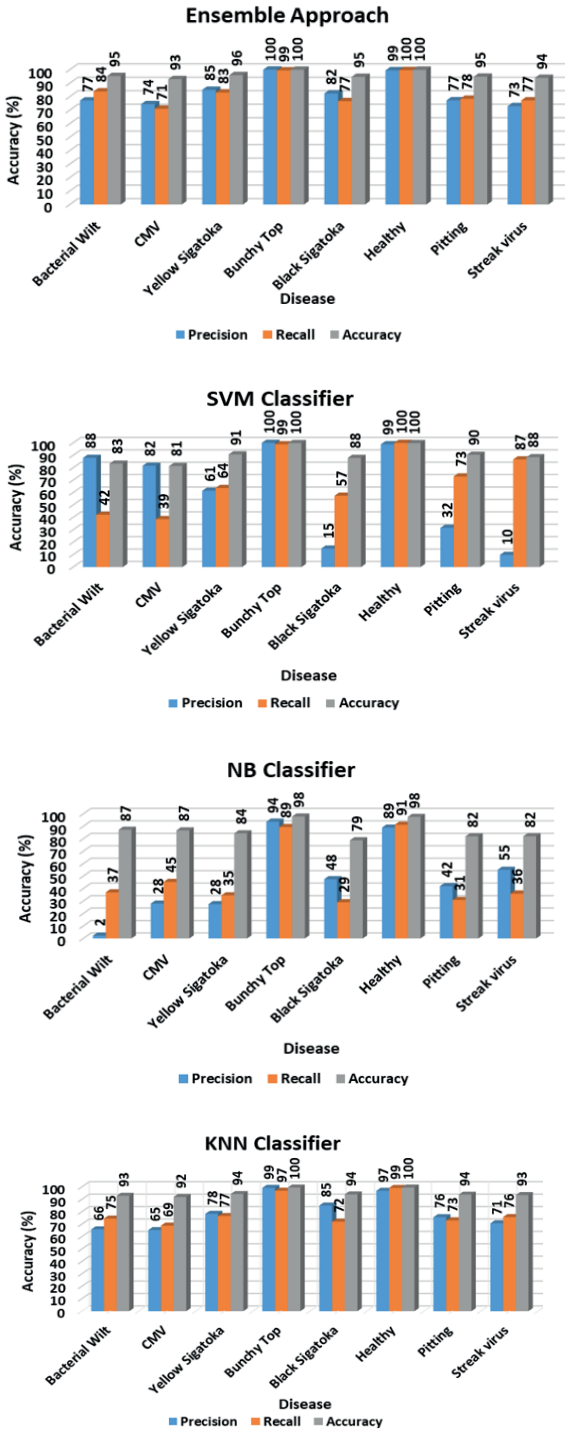


Fig. 6. Graphical representation of accuracy measure.

Ecological factors are responsible for the occurrence of diseases on the plants. Due to the colour change and texture of the leaf, it is presumed that it is affected by disease. In the automated disease identification system, the sample images are classified according to their visual symptoms. However, there is always some disparity in colour, shape, and sizes of the symptoms. It was observed that the visual symptoms of Bacterial Wilt were identical between the primary and the final stages of Black Sigatoka. As a result, the retrieved colour and texture features were essentially identical, which led to the

incorrect classification of both samples. During the segmentation process, it was found that the brown spots were more obvious than white yellow streaks which meant Yellow Sigatoka samples were classified correctly as compared to CVM. Different machine learning techniques were to applied on extracted colour and local binary pattern texture features, such as Support Vector Machine, Naïve Bayes, K Nearest Neighbour. The ensemble approach was also used to check the accuracy of extracted features. It was observed that the confusing sample having the same visualized symptoms of diseases was classified correctly using KNN classifier, so the overall accuracy of KNN was good as compared to SVM and NB classifier.

TABLE I
COMPARISON OF DIFFERENT CLASSIFIER

Disease	Classifier	Precision	Recall	Accuracy (%)
Bacterial Wilt	SVM	88	42	83
	NB	2	37	87
	KNN	66	75	93
	Ens	77	84	**95**
CMV	SVM	82	39	81
	NB	28	45	87
	KNN	65	69	92
	Ens	74	71	**93**
Yellow Sig	SVM	61	64	91
	NB	28	35	84
	KNN	78	77	94
	Ens	85	83	**96**
Bunchy Top	SVM	100	99	100
	NB	94	89	98
	KNN	99	97	100
	Ens	100	99	**100**
Black Sigatoka	SVM	15	57	88
	NB	48	29	79
	KNN	85	72	94
	Ens	82	77	**95**
Healthy	SVM	99	100	100
	NB	89	91	98
	KNN	97	99	100
	Ens	99	100	**100**
Pitting	SVM	32	73	90
	NB	42	31	82
	KNN	76	73	94
	Ens	77	78	**95**
Streak Virus	SVM	10	87	88
	NB	55	36	82
	KNN	71	76	93
	Ens	73	77	**94**

In the NB classifier, every feature contributes individually, irrespective of some possible correlations between them. The value of colour features is independent of the value of texture features. Accordingly, the accuracy result is low using the NB classifier. The colour or texture feature contributes individually during the working of Naïve Bayes; thus, positive predictive value (precision) and sensitivity (recall) of Naïve Bayes classifier is lower. To utilize the efficiency of each classifier, the ensemble approach was used. KNN, NB and SVM with Gaussian and polynomial kernel function were used in the ensemble learning process. As compared to any single classifier, an ensemble model produced higher accuracy.

Different methodologies were used in plant disease detection. Average accuracy for each disease in the proposed work is provided in Table II

TABLE II
COMPARISON WITH EXISTING METHODOLOGIES

Author	Plants	Methodology used	Accuracy (%)
Das et al., 2020 [14]	Tomato	SVM logistic regression random forest	87.60 67.30 70.05
K. Suganya Devi et al., 2020 [4]	Groundnut	KNN H2K (Harris, HOG and K-Nearest Neighbour)	97.67
Das et al., 2020 [14]	Mango	SVM	80.00
Anjna et al., 2019 [16]	Capsicum	linear discriminant SVM KNN	98.4 100 100
Punam Bedi and Pushkar Gole, 2021 [17]	Peach	CNN	98.38
Maria Morgan et al., 2021 [18]	Mushroom and soybean	Naïve Bayes ANN, Support Vector Machine, KNN, decision tree and random forest	95.83
Patil and Burkpalli, 2021 [19]	Cotton	SVM multilayer perceptron Random Forest Naïve Bayes KNN AdaBoost	93.38 96.69 95.86 90.91 94.21 92.56
Proposed work	Banana	Ensemble Approach SVM NB KNN	96 90.13 87.13 95

VI. CONCLUSION

Plants may be harmed by changing environment. Due to ecological changes, several viral and bacterial diseases can arise on plants. Diseases on the banana crop reduce the amount of production harvested. Identification of correct disease is a

crucial step to treat the diseases. The visible changes in plants are mostly symptomatic of the ecological impact. The work focused on the development of automated and optimum model to accurately identify and categorize diseases on the banana leaf. The similarity in disease symptoms at the onset increases the complexity of disease detection. The image segmentation process carried in L*a*b* colour space closely matched with human perception and became easier to find the region of interest. The colour and texture features extracted using Local Binary Pattern were more eloquent to classify the images accurately. The present study demonstrated that the ensemble approach provided good accuracy as compared to KNN, SVM, NB. The benefit of this approach is by combining the set of propositions from all the learning classifiers and calculating its result, which obeys the optimistic features of all the learning classifiers.

References

[1] N. B. A. Mustafa et al., "Image processing of an agriculture produce: Determination of size and ripeness of a banana," in *2008 International Symposium on Information Technology*, Kuala Lumpur, Malaysia, Aug. 2008, pp. 1–7. https://doi.org/10.1109/ITSIM.2008.4631636

[2] M.J. Vipinadas and A. Thamizharasi, "Banana leaf disease identification technique," *International Journal of Advanced Engineering Research and Science (IJAERS)*, vol. 3, no. 6, pp. 120–124, Nov. 2016. https://doi.org/10.22161/ijaers/3.11.21

[3] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik, and Z. AL-Rahamneh, "Fast and accurate detection and classification of plant diseases," *International Journal of Computer Applications*, vol. 17, no. 1, pp. 31–38, Mar. 2011. https://doi.org/10.5120/2183-2754

[4] K. Suganya Devi, P. Srinivasan, and Sivaji Bandhopadhyay, " H2K – A robust and optimum approach for detection and classification of groundnut leaf diseases," *Computers and Electronics in Agriculture*, vol. 178, Nov. 2020, Art. no. 105749. https://doi.org/10.1016/j.compag.2020.105749

[5] D. Yirgou and J. F. Bradbury, "Bacterial wilt of enset (*Ensete ventricosum*) incited by *Xanthomonas musacearum sp. n*," *Phytopathology*, vol. 58, pp. 111–112, 1968. https://www.musalit.org/seeMore.php?id=10016

[6] L. Tripathi, M. Mwangi, V. Aritua, W.K. Tushemereirwe, S. Abele, and R. Bandyopadhyay, "Xanthomonas wilt: a threat to banana production in East and Central Africa," *Plant Disease*, vol. 93, no. 5, May 2009. https://doi.org/10.1094/PDIS-93-5-0440

[7] D. Surya Prabha and J. Satheesh Kumar, "Study on banana leaf disease identification using image processing methods," *International Journal of Research in Computer Science and Information Technology (IJRCSIT)*, vol. 2, no. 2(A), pp. 89–94, Mar. 2014. https://www.researchgate.net/publication/299486544_Study_on_Banana_Leaf_Disease_Identification_Using_Image_Processing_Methods

[8] V. Devappa and T. C. Archith, "Wilt diseases of ornamental crops and their management," in *Wilt Diseases of Crops*. Today and Tomorrow Printers and Publisher, New Delhi, India, 2019, pp. 141–164. https://www.researchgate.net/publication/331917069_Wilt_diseases_of_ornamental_crops_and_their_management

[9] R. Dheepa and S. Paranjothi, "Transmission of cucumber mosaic virus (CMV) infecting banana by aphid and mechanical methods," *Emirates Journal of Food and Agriculture*, vol. 22, no. 2, pp. 117–129, Oct. 2010. https://doi.org/10.9755/ejfa.v22i2.4899

[10] H. R. Almadhoun and S. S. Abu-Naser, "Banana knowledge based system diagnosis and treatment," *International Journal of Academic Pedagogical Research (IJAPR)*, vol. 2, no. 7, pp. 1–11, Jul. 2018. https://hal.science/hal-01847731v1/file/IJAPR180701.pdf

[11] S. C. Nelson, "Banana bunchy top: Detailed signs and symptoms," 2004. [Online]. Available: https://www.ctahr.hawaii.edu/bbtd/downloads/bbtv-details.pdf

[12] S. Ganesan, H. Shankar Singh, S. Petikam, and D. Biswal, "Pathological status of *Pyricularia angulata* causing blast and pitting disease of banana in Eastern India," *Plant Pathol. J.*, vol. 33, no. 1, pp. 9–20, Feb. 2017. https://doi.org/10.5423/PPJ.OA.08.2016.0162

[13] S. M. Muturi, F. N. Wachira, L. S. Karanja, and L. K. Njeru, "The mode of transmission of banana streak virus by *Paracoccus burnerae* (*Homiptera; Planococcidae*) vector is non-circulative," *British Microbiology Research Journal*, vol. 12, no. 6, pp. 1–10, 2016. https://doi.org/10.9734/BMRJ/2016/21574

[14] S. K. Das, Md. R. Mia, S. Roy, and Md. A. Rahman, "Mango leaf disease recognition using neural network and support vector machine," *Iran Journal of Computer Science*, vol. 3, pp. 185–193, Apr. 2020. https://doi.org/10.1007/s42044-020-00057-z

[15] A. E. Hassanien, T. Gaber, U. Mokhtar, and H. Hefny, „An improved moth flame optimization algorithm based on rough sets for tomato diseases detection," *Computers and Electronics in Agriculture*, vol. 136, pp. 86–96, Apr. 2017. http://doi.org/10.1016/j.compag.2017.02.026

[16] Anjna, M. Sood, and P. K. Singh, "Hybrid system for detection and classification of plant disease using qualitative texture features analysis," *Procedia Computer Science*, vol. 167, pp. 1056–1065, 2020. https://doi.org/10.1016/j.procs.2020.03.404

[17] P. Bedi and P. Gole, "Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network," *Artificial Intelligence in Agriculture*, vol. 5, pp. 90–101, 2021. https://doi.org/10.1016/j.aiia.2021.05.002

[18] M. Morgan, C. Blank, and R. Seetan, "Plant disease prediction using classification algorithms," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 1, pp. 257-264, Mar. 2021. https://doi.org/10.11591/ijai.v10.i1.pp257-264

[19] B. M. Patil and V. Burkpalli, "A perspective view of cotton leaf image classification using machine learning algorithms using WEKA," *Advances in Human-Computer Interaction*, 2021, Art. no. 9367778. https://doi.org/10.1155/2021/9367778

[20] "Godliver Owomugisha Implementation-BBW and BBS Diseases". [Online]. Available: https://github.com/godliver/source-code-BBW-BBS (accessed on July 2019).

[21] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Image processing techniques for diagnosing rice plant disease: A survey," *Procedia Computer Science*, vol. 167, pp. 516–530, 2020. https://doi.org/10.1016/j.procs.2020.03.308

[22] V. N. T. Le, B. Apopei, and K. Alameh, "Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods," *Information Processing in Agriculture*, vol. 6, no. 1, pp. 116–131, Mar. 2019. https://doi.org/10.1016/j.inpa.2018.08.002

[23] S. Iniyan, R. Jebakumar, P. Mangalraj, M. Mohit, and A. Nanda, "Plant disease identification and detection using Support Vector Machines and Artificial Neural Networks," in *Artificial Intelligence and Evolutionary Computations in Engineering Systems. Advances in Intelligent Systems and Computing*, S. Dash, C. Lakshmi, S. Das, and B. Panigrahi, Eds., vol. 1056, Springer, Singapore, Feb. 2020, pp. 15–27. https://doi.org/10.1007/978-981-15-0199-9_2

[24] L. J. Muhammad, E. A. Algehyne, S. S. Usman, A. Ahmad, C. Chakraborty, and I. A. Mohammed, "Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset," *SN Computer Science*, vol. 2, 2021, Art. no. 11. https://doi.org/10.1007/s42979-020-00394-7

[25] G. M. James and S. C. Punitha, "Tomato disease classification using ensemble learning approach," *IJRET: International Journal of Research in Engineering and Technology*, vol. 5, no. 10, pp. 104–108, Oct. 2016. https://ijret.org/volumes/2016v05/i10/IJRET20160510019.pdf

Vandana Chaudhari received B.Sc. in Computer Science from North Maharashtra University, Jalgaon, India in 2005. She obtained her Master degree in Computer Science at North Maharashtra University, Jalgaon, India in 2007. She is currently working as an Assistant Professor at Smt. G. G. Khadse College, Muktainagar, India. Her current research interests include natural language processing, artificial intelligence, deep learning and image processing.
E-mail: vandu2108@gmail.com
ORCID iD: https://orcid.org/0000-0002-5526-0761

Manoj Patil received his PhD degree in Computer Science from North Maharashtra University, Jalgaon, India in 2014. He works as an Associate Professor at KBC North Maharashtra University, Jalgaon, India. His current research interests include image processing, machine learning and algorithms.
E-mail: mpp145@gmail.com
ORCID iD: https://orcid.org/0000-0001-7579-5694