Journal of Advanced Zoology

ISSN: 0253-7214

Volume 44 Issue 03 Year 2023 Page 93:106

Plant Disease Detection using Deep Learning in Banana and Sunflower

Chetan H R[1], Rajanna G S[2*], Sreenivasa B R[3*], Ganesh N Yallappa[4]

[1]Research Scholar, Dept of ECE, Srinivasa University, Mukka Mangalore, India-Dept. of EEE, Jain Institute of Technology, Davangere-577003, Karnataka, India.chetan04hr@gmail.com

[2]Professor, Dept of ECE, Srinivasa University, Mukka, Mangalore, India kgsrajannar@gmail.com

[3]Associate Professor, Dept, of ISE, BIET, Davangere, India.

4Assistant Professor, Dept. of Medicinal Chemistry, Jain Institute of Technology, Davangere-577003, Karnataka, India. gindi.ny1988@gmail.com

**Corresponding author's E-mail: br.seenu@gmail.com*

Article History

Received: 06 June 2023
Revised: 05 Sept 2023
Accepted:01Oct 2023

Abstract

In recent years plant disease detection and classification is finding a lot of scope in the field of agriculture. The use of image pre-processing along with deep learning techniques is making the role of farmers easy in the process of plant leaf disease detection. In this paper we propose a deep learning technique, ResNet-50 for the identification and classification of leaf diseases mainly in banana and sunflower. Images for the training and testing purpose are collected by visiting the farms and from village dataset for normal, leaf spot, leaf blight, powdery mildew, bunchy top, sigatoka, panama wilt. Pre-processing is done to remove eliminate the noise in the image by converting the RGB input to HSV image. Binary pictures are retrieved to separate the diseased and unaffected portions based on the hue and saturation components. A clustering method is utilized to separate the diseased region from the normal portion and the background. Classification of the disease is carried out using ResNet-50 algorithm. The experimental results obtained are compared with CNN, machine learning algorithms like SVM, KNN, DT and Ensemble algorithm like RF and XG booster. The proposed algorithm provided maximum efficiency compared to other algorithms.

Keywords: *Leaf diseases, ResNet-50, machine learning, deep learning, features*

1. Introduction

In most of the developing countries like India agriculture is considered to be important source of income (Zhou et al., 2021). Farmers harvest a variety of food crops depending on the environmental factors affecting the area and the demand for food (Ramesh & Vydeki, 2020). But the farmers face a lot of issues due environmental conditions and diseases that occur in the crops. The majority of issues are reduced by offering some technical facilities (Sharma et al., 2018). Finding specialists is not necessary if disease prevention is implemented in a timely manner to increase food yield. The field of agriculture provides an area for research in plant leaf disease detection. Hence many researches are being carried out in the agriculture domain.

In order to minimize the losses and to improve the quality of crop yield, recognition of plant diseases plays a vital role (Chakraborty et al., 2021). Disease detection and plant health monitoring are extremely beneficial for sustainable agriculture. According to studies connected to identifying plant diseases, the illnesses are represented by outward patterns seen on the plants (Shrestha et al., 2020). The manual method of monitoring the plants to identify the disease is more challenging task as it involves processing time, a significant amount of labor and knowledge about the plant disease is required (Mukti & Biswas, 2019; Paymode & Malode, 2022). Hence image processing techniques are used to identify the plant leaf disease (Pooja et al., 2017).

The processes in the image processing method for diagnosing diseases include image acquisition, pre-processing, segmentation, feature extraction, and classification (Lahza et al., 2023; Kumar et al., 2023). Only the outsides of the affected plants can be treated using these procedures. The most common source for identifying plant illnesses in the majority of plants is their leaves (Prajapati et al., 2016). The diseases that commonly occur in banana and sunflower like leaf spot, leaf blight, powdery mildew, bunchy top, sigatoka, Panama wilt is considered in our work. The symptoms of these diseases vary from each other depending on size, color and texture.

Manual detection of plant diseases is carried out through the naked eye, which is time consuming when taken on larger farms. Such manual detection some time provides error in prediction of the diseases (Nandish & Pushparajesh, 2021; Joshi & Jadhav, 2016). Due to the miss identification of the diseases, the yield of the plant decreases thereby causing loss to the farmers (Ramesh et al., 2018). Hence in order to overcome this limitation a suitable algorithm is required for the faster and accurate prediction of the diseases (Waghmare et al., 2016). Hence our work proposes a novel method for the disease identification and classification in banana and sunflower. Hence in order to carry out the proposed work images of diseased leaf of banana and sunflower were captured from the fields and some of the images were collected from the village plant website. Since the collected images consist of the noises, these noises are reduced using the filters (Govardhan & Veena, 2019). Then the background elimination is done to further reduce the noises. From the pre-processed image the segmentation is done to eliminate the unwanted portion of the image. Hence the diseased part of the image is segmented from the healthy part of the leaf. For further the classification of the disease, ResNet-50 is used.

The rest of the paper is structured as follows: The most recent research on the classification of plant leaf diseases is included in Section 2. The problem statement is described in Section 3. Proposed methodology is presented in Section 4. The experimental analysis is displayed in Section 5. The final section is conclusion.

2. Literature Review

Zhou et al. (2021) proposed a novel approach for the plant disease detection in grapes. The work consisted of identifying leaf spot disease in grapes. In order to increase the augmented images size, fine grained-GAN was used. The images obtained were preprocessed and desired features of leaf spot were obtained. Then the total number of images obtained was fed to deep learning model to train the algorithm. ResNet-50 was used for the identification and classification of the leaf spot disease. The proposed algorithm was used to identify one leaf spot disease in grapes.

Ramesh & Vydeki (2020) developed an optimized deep neural algorithm for the identification of diseases in paddy. Images were captured from the farm to form the database. The images were preprocessed in order to remove the noise by converting the RGB images to HSV. Image clustering was performed to separate the diseased part of the image from the healthy part. From the segmented image, features were extracted to train the proposed model. Deep neural network was used to classify and identify the diseased leaves. In order to improve the efficiency of the proposed algorithm a feedback loop was created in post processing step.

Sharma et al. (2018) proposed a machine learning algorithm for the identification of diseases in plant. Nearly eighty thousand images were collected to form the database. Tensor flow frame work was used for the implementation of CNN algorithm. The CNN algorithm consisted of multiple layers of filters. The proposed work provided a better accuracy in identifying the various diseases among the different plant leaf's.

Chakraborty et al. (2021) proposed an image processing and machine learning model for the identification of the diseases in plant leaf's. The proposed model was used to classify the healthy and diseased leaves. The captured images were preprocessed using Ostu thresholding and histogram equalization. The diseased parts of the image were segmented using Ostu thresholding by finding a threshold value. From the segmented part of the image GLCM features were extracted, obtained features were used to train the model for the identification of the diseases. Multi class SVM was used for the identification and classification of the diseases. The work aimed at identifying two diseases that normally occur in apple black rot and leaf rust.

Bose et al. (2020) proposed a machine learning model to detect the disease in Hemp plant, which is considered to be one of the commercial crops. Images related to Hemp diseases were collected to form the database. These images were preprocessed to reduce the noises and resized to 516*516. The features obtained after preprocessing were used to train SVM model for the identification of the diseases and features were also trained using three deep learning models. The results obtained from two models were compared with each other and work concluded that SVM has higher accuracy compared to deep learning algorithms.

Tomato is considered to be one of daily useable vegetable and it is subjected to different diseases. Hence Shrestha et al. (2020) proposed CNN algorithm that can is used to detect the diseases that occur in tomato plant. Images of diseased tomato leaves were collected from the farms. These images were converted in NumPy array and the images were labeled and stored to form the database. Further these images were split to form the as train and test images. The algorithm developed was trained to identify nearly fifteen different diseases.

Problem Statement

The conventional method of human vision is used for the identification of the diseases. This requires expert to identify the disease accurately, which is time consuming and costly approach. The accuracy of the identification of the disease in conventional method depends on the experience and vision of the person. While the machine learning techniques offers an accurate technique to identify the disease and also recommend accurate pesticides to control the identified diseases. One of the main advantages of using machine learning approach is accurate identification and less time consuming. Hence to overcome the limitations of conventional methods, there is requirement of developing deep learning techniques that can be used for the identification of the diseases.

3. Materials And Methods

Our proposed work mainly consists of the following six phases like image acquisition, image preprocessing, image segmentation, feature extraction, disease identification and recommendation of pesticides. Images of sunflower and banana are collected by visiting the farms and by collecting the images from village dataset. Since the images captured are of different sizes, shapes and contain the noises, preprocessing is carried out in order to reduce the noises. Segmentation of the preprocessed images is done using k-means clustering in order to obtain area of interest from the input image (Ahmed et al., 2019). Features are extracted from the segmented part of the image and these features are used for the classification of the diseases (Ramesh & Vydeki, 2018). The identified diseases are further are recommended with pesticides in order to control the diseases. “Fig. 1” shows the methodology of the proposed work.

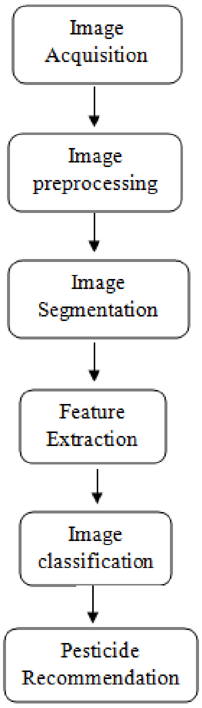

Figure1. Process flow graph

Image Acquization

Image acquisition is the process of collecting the images to form the database (Sarangdhar & Pawar, 2017). Images were collected by visiting the farms and few of the images were collected from the internet source. Nearly 200 images of normal leaf image of sunflower and banana, diseased leaf images consisting of leaf spot, leaf blight, powdery mildew, bunchy top, sigatoka, and Panama wilt were collected to form the database. “Fig. 2” shows sample images of diseased and normal leaf’s of sunflower and banana.

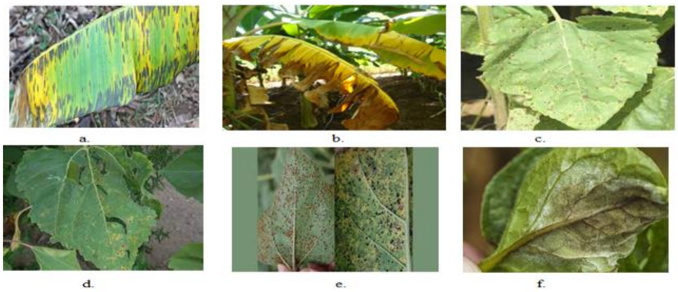

Figure2. Types of Leaf Diseases a. sigatoka b. Panama wilt c.alternia leaf spot d. bacterial wilt e. common rust f. late blight

Pre-processing

Since the images captured are of different size and shapes, images are resized to 250x250. This ensures reducing the memory size and increasing the computational speed (Pooja et al., 2017). The noise present in the images is eliminated by using Gaussian filter shown by the equation (Zhou et al., 2017). After the filtering of the input imaged the background of the image needs to be eliminated. From the original RGB model, the HSV model is first created, and from this created HSV model, the saturation value is chosen for the binary conversion since it covers the whiteness. When the threshold value equals 90, the binary image is converted from the original image. Then, for mask generation, both the binary and RGB original pictures are fused. By fusing with the pixel value 0, which in the RGB paradigm represents the color black, the backdrop is eliminated. Figure 3 depicts the preprocessing steps involved.

$$\varphi(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(z-\mu)^2}{2\sigma^2}} \qquad (1)$$

K-means segmentation

Segmentation is the process of separating the region of interest from rest of the image. K means clustering is used for separating the infected part of the leaf from the healthy part of the leaf. “Fig. 4” shows the clustered image using k-means clustering. The algorithm used for the segmentation in our work is explained below:

- Indicate K the number of clusters.
- The dataset is initially shuffled, and then K data points are randomly chosen and used as the initial values for the centroids.
- Till the centroids remain unchanged, keep iterating. In other words, the clustering of data points remains the same.
- Calculate the sum of the squared distances between all of the centroids and the data points.
- The closest cluster (centroid) should be used to assign each data point.
- By averaging all the data points that make up each cluster, calculate the centroids for the clusters.

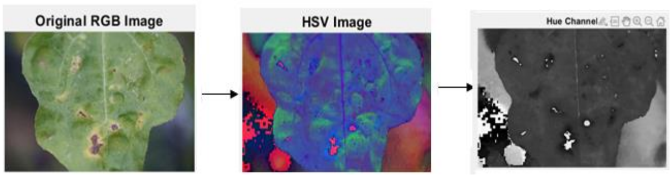

Figure 4: clustered image using k-means clustering

ResNet-50

Transfer learning consists of various models like AlexNet, GooLeNet, overfeat, VGNN. There were numerous convolutional layers overlaid. Deep CNN networks have some challenges, including degradation issues, vanishing gradient issues, and network optimization issues. Hence in order to over the issues related to vanishing gradient problem ResNet-50 is proposed in our work. The ResNet network introduces a novel concept. It helps to tackle challenging problems and improves detection precision. ResNet aims to address the saturation and accuracy degradation that occur during the deep CNN training process.

ResNet is used to extract the features from the segmented images and the features obtained are used to identify the diseases. “Fig. 5” shows the architecture of ResNet model used for the identification and classification of the diseases.

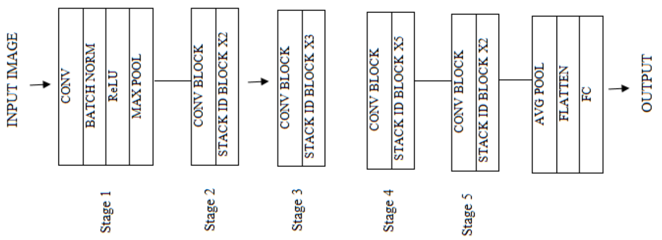

Figure 5: ResNet-50 architecture

ResNet-50 is divided into different stages, each containing multiple residual blocks. Here's an outline of its architecture:

- Input Layer: Accepts the input image.
- Initial Convolutional Layer: Applies a convolutional layer to extract basic features from the input image.
- Stage1: consists of convolutional layer, batch normalization, ReLU activation and Max Pooling.

- Stage 2: consists of 3 residual blocks like identity connection, convolutional layers, batch normalization, ReLu activation.
- Stage 3,4 and 5: consists of residual blocks.
- Global Average pooling: converts the output of the last residual block into a fixed size feature map.
- Fully connected layer: connects the feature map to the output classes for the classification.

Image classification and pesticide recommendation

“Fig. 6” shows the proposed algorithm, ResNet-50 used for the identification and classification of the disease. The proposed work also suggests organic and chemical pesticide that can be used to control the identified disease.

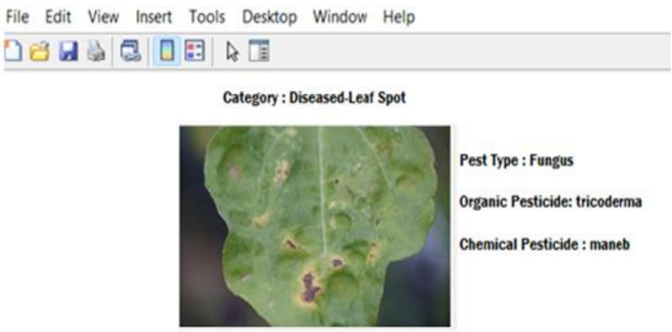

Figure 6. Disease identification by ResNet in cotton plant

3. Results and Discussion

Using the Matlab, 2020 version we put our suggested methods into practice. The performance of the proposed approach ResNet-50 is estimated and contrasted with the performance of currently used deep learning model like CNN, machine learning classifiers like KNN, SVM, DT and ensemble classifiers like XG-booster and RF, the outcomes are contrasted according on the disease classes, which include normal leaf spot, leaf blight, powdery mildew, bunchy top, sigatoka, Panama wilt. 70% of the images in the dataset are utilized for training, 30% are used for testing.

“Fig. 7” shows the confusion matrix of ResNet-50 for banana. The values of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are predicted from this confusion matrix. The values of TP, TN, FP and FN from the above confusion matrix are 185, 756, 18 and 41 respectively for the normal banana image; the values of TP, TN, FP and FN are 179, 785, 20 and 16 respectively for the Panama wilt; the values of TP, TN, FP and FN are 182, 772, 28 and 18 respectively for the sigatoka; the values of TP, TN, FP and FN are 175, 783, 21 and 21 respectively for the bunchy top; the values of TP, TN, FP and FN are 170, 785, 22 and 13 respectively for the leaf spot affected image. The values of

“Fig. 8” shows the confusion matrix of ResNet-50 for sunflower. The values of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are predicted from this confusion matrix. Table provides confusion matrix table for the banana leaf. The values of TP, TN, FP and FN from the above confusion matrix are 190, 757, 17 and 36 respectively for the normal sunflower image; the values of TP, TN, FP and FN are 182, 786, 19 and 13 respectively for the leaf blight; the values of TP, TN, FP and FN are 175, 781, 19 and 25 respectively for the powdery mildew; the values of TP, TN, FP and FN are 180, 781, 23 and 16 respectively for the leaf blight; the values of TP, TN, FP and FN are 172, 794, 23 and 11 respectively for the downy mildew affected image.

	Normal	Panama Wilt	Sigatoka	Bunchy Top	Leaf Spot
Normal	185	8	12	10	11
Panama Wilt	5	179	5	2	4
Sigatoka	4	5	182	5	4
Bunchy Top	6	4	8	175	3
Leaf Spot	3	3	3	4	170

Figure 7. Confusion Matrix of ResNet-50 for banana

	Normal	Leaf Spot	Powdery Mildew	Leaf Blight	Downy Mildew
Normal	190	7	9	10	10
Leaf Spot	5	182	3	2	3
Powdery Mildew	5	6	175	7	7
Leaf Blight	5	3	5	180	3
Downy Mildew	2	33	2	4	172

Figure 8. Confusion Matrix of ResNet-50 for Sunflower

confusion matrix table for resnet-50 in banana leaf

Leaf Disease	TN	TP	FN	FP
Leaf spot	795	170	13	22
Bunchy top	789	175	21	21
Sigatoka	772	182	18	28
Panama wilt	785	179	16	20
Normal	756	185	41	18

Table 2 and Table 3 show the performance of ResNet50 classifier in the identification of diseases in banana and sunflower respectively. The proposed algorithm provided a high accuracy of 96.5 for leaf spot in banana and 96.8 leaf spot in sunflower.

confusion matrix table for resnet-50 in sunflower leaf

Performance parameter\ Leaf Disease	**Leaf Spot**	**Bunchy Top**	**Sigatoka**	**Panama Wilt**	**Normal Leaf**
Accuracy	96.5	95.8	95.4	96.4	94.1
Sensitivity	92.8	89.2	92	91.7	81.8
Specificity	97.3	97.3	98.1	97.5	97.6
Precision	88.5	89.2	86.6	89.9	91.1
F1 Score	90.59	89.2	89.2	90.79	86.19
NPV	98.3	97.3	97.7	98	94.8
FRP	2.7	2.7	1.9	2.5	2.4
FDR	11.4	10.7	13.3	10	8.8

classification performance for resnet-50 for diseased and normal leaf images of sunflower

Performance Parameter / Plant disease	Downy Mildew	Leaf Blight	Powdery Mildew	Leaf Blight	Normal
Accuracy	95.6	96.1	95.6	96.8	94.7
Sensitivity	93.9	91.8	87.5	93.3	84.07
Specificity	97.1	97.1	97.6	97.6	97.8
Precision	88.2	88.6	90.2	90.5	91.7
F1 Score	90.9	90.1	88.8	91.87	87.7
NPV	98.6	97.9	96.8	98.3	95.4
FRP	2.9	2.9	2.4	2.4	2.2
FDR	11.7	11.3	9.7	9.4	8.2

The graphs plotted shows the performance of ResNet-50 with respect to accuracy, Sensitivity, specificity, precision, F1 score, NPV, FRP and FDR with respect to banana normal and diseased leafs. “Fig.9” shows the comparison of accuracy with respect to different classifiers used for the identification of the diseases in banana. Equation 2 is used for the purpose of calculating the accuracy form confusion matrix table1.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (2)$$

The proposed ResNet-50 when used for the identification of diseases in banana provides an accuracy of 96.5% for leaf spot, 95.8% for bunchy top, 95.4 for sigatoka, 96.4% for panama wilt and 94.1% for normal image of banana. Similarly, when CNN used, accuracy of 94% for leaf spot, 93.9% for bunchy top, 92.4% for sigatoka, 94% for panama wilt and 92.3% for normal image of banana is obtained. When RF is used, accuracy of 92.5% for leaf spot, 92% for bunchy top, 91.6% for sigatoka, 91.6% for panama wilt and 89.6% for normal image of banana. For XG-booster accuracy of 90.2% for leaf spot, 90% for bunchy top, 91.5% for sigatoka, 88.7% for panama wilt and 87.7% for normal image of banana. KNN provides accuracy of 88.5% for leaf spot, 88.3% for bunchy top, 84.8% for sigatoka, 88.3% for panama wilt and 85.5% for normal image of banana. DT provided accuracy of 80.3% for leaf spot, 80.4% for bunchy top, 77.8% for sigatoka, 81.2% for panama wilt and 78.3% for normal image of banana. Accuracy of 79% for leaf spot, 80% for bunchy top, 76.2% for sigatoka, 79.9% for panama wilt and 76.2% for normal image of banana using SVM classifier.

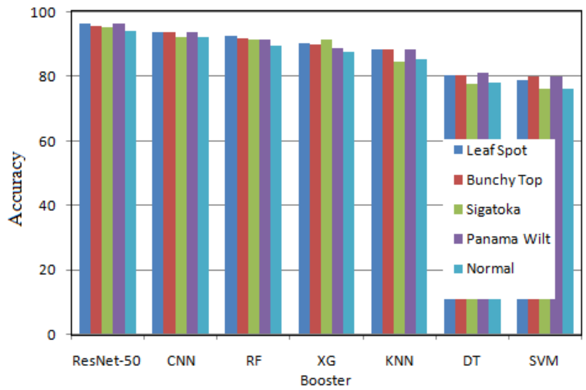

Figure 9: Comparison graph of accuracy

The sensitivity is a metric used to assess how well a classification model can identify the positive cases. It’s often referred to as the true positive rate or recall. The formula used for the calculation is given by equation 3. “Fig. 10” shows the comparison graph of sensitivity between different classifiers used in our study. The proposed ResNet-50 provided sensitivity of 92.8 for leaf spot, 89.2 for bunchy top, 92 for sigatoka, 91.7 for panama wilt and 81.8 for normal leafs. While CNN provided 87.4 for leaf spot, 86.3 for bunchy top, 83 for sigatoka, 86.1 for panama wilt and 77.4 for normal leafs. Similarly, RF provided 83 for leaf spot, 81.4 for bunchy top, 77.4 for sigatoka, 82 for panama wilt and 73 for normal leafs. XG booster provided 77.5 for leaf spot, 75 for bunchy top, 78 for sigatoka, 74.3 for panama wilt and 67.6 for normal leafs. While KNN provided 75 for leaf spot, 63 for bunchy top, 67.6 for sigatoka, 75.2 for panama wilt and 60.7 for normal leafs. DT provided 53 for leaf spot, 43 for bunchy top, 50.2 for sigatoka, 55.3 for panama wilt and 45.3 for normal leafs. SVM provided 51.1 for leaf spot, 45.8 for bunchy top, 44 for sigatoka, 50.5 for panama wilt and 40.4 for normal leafs.

$$Sensitivity\ (SN) = \frac{TP}{TP+FN} \quad (3)$$

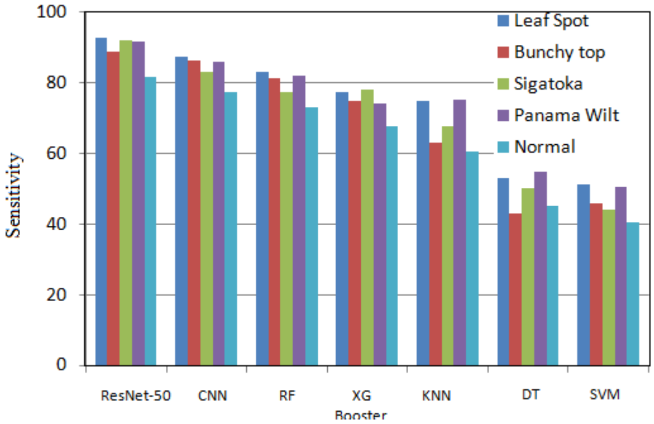

Figure 10: Comparison graph of sensitivity

A model's ability to correctly distinguish negative samples from all the other negative samples is measured by the specificity. Specificity of a given matrix from the confusion matrix is calculated by the formula given by equation 4. "Fig. 11" depicts the comparison of specificity between different classifiers used in our study. The proposed ResNet-50 provided specificity of 97.3 for leaf spot, 97.3 for bunchy top, 98.1 for sigatoka, 97.5 for panama wilt and 97.6 for normal leafs. While CNN provided 95.4 for leaf spot, 96.2 for bunchy top, 85 for sigatoka, 95.9 for panama wilt and 96.6 for normal leafs. Similarly, RF provided 94.8 for leaf spot, 94.7 for bunchy top, 95.3 for sigatoka, 94.1 for panama wilt and 94.8 for normal leafs. XG booster provided 93 for leaf spot, 93.6 for bunchy top, 94.8 for sigatoka, 92.7 for panama wilt and 93.5 for normal leafs. While KNN provided 91 for leaf spot, 94 for bunchy top, 81.2 for sigatoka, 91 for panama wilt and 93 for normal leafs. DT provided 86 for leaf spot, 89 for bunchy top, 84.9 for sigatoka, 87.1 for panama wilt and 89.1 for normal leafs. SVM provided 85 for leaf spot, 88.1 for bunchy top, 84.4 for sigatoka, 86.6 for panama wilt and 87.9 for normal leafs.

$$Specificity\ (SP) = \frac{TN}{TN+FP} \quad (4)$$

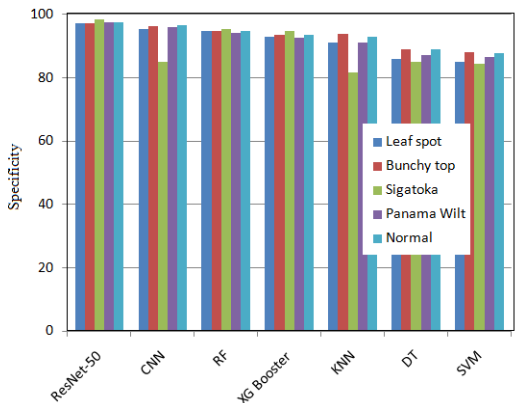

Figure 11: Comparison graph of specificity

One of the measures used in a confusion matrix to assessa classification models' effectiveness is precision. The accuracy of the models' favorable predictions is measured by precision. It reveals the percentage of projected positive things that were true. In other words, it measures how well the model can avoid making false positive errors. The formula used for calculating the precision is given by the equation 5. "Fig. 12" depicts the comparison of precision between different classifiers used in our study. The proposed ResNet-50 provided precision of 88.5 for leaf spot, 89.2 for bunchy top, 86.6 for sigatoka, 89.9 for panama wilt and 91.1 for normal leafs. While CNN provided 81.2 for leaf spot, 84.6 for bunchy top, 80.5 for sigatoka, 83.5 for panama wilt and 87.6 for normal leafs. Similarly, RF provided 78.3 for leaf spot, 79 for bunchy top, 80.7 for sigatoka, 77.2 for panama wilt and 80.4 for normal leafs. XG booster provided 71.3 for leaf spot, 74.2 for bunchy top, 79.1 for sigatoka, 71.4 for panama wilt and 75.3 for normal leafs. While KNN provided 63 for leaf spot,72 for bunchy top, 62.2 for sigatoka, 66.3 for panama wilt and 75.7 for normal leafs. DT provided 44.1 for leaf spot, 48.8 for bunchy top, 46.6 for sigatoka, 49.5 for panama wilt and 57.7 for normal leafs. SVM provided 41.1 for leaf spot, 47.8 for bunchy top, 42.7 for sigatoka, 46.3 for panama wilt and 52.3 for normal leafs.

$$Precision\ (prec) = \frac{TP}{TP+FP} \tag{6}$$

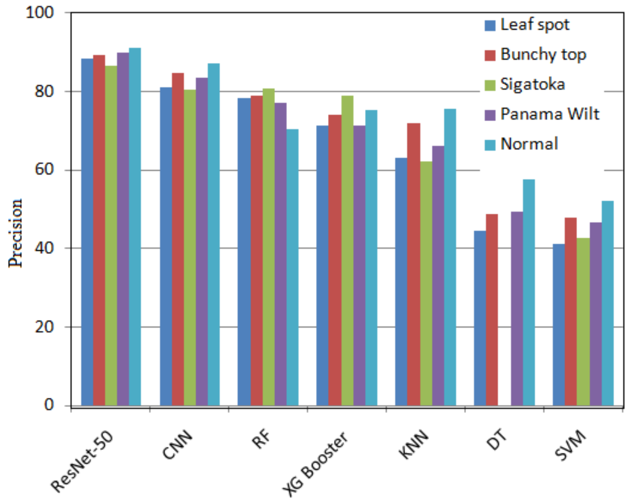

Figure 12: Comparison graph of Precision

A statistic for assessing a classification model performance, especially when working with unbalanced datasets is F1 score. It is determined using the values included in a confusion matrix which condenses the outcomes of a binary classification. The F1 score provides a balanced evaluation of a model's performance by combining precision and recall into one metric. When trying to balance reducing false positives and false negatives in our classification model. The formula used for the calculation of F1 score is given by equation 6. "Fig. 13" depicts the comparison of F1 score between different classifiers used in our study. The proposed ResNet-50 provided F1 score of 90.59 for leaf spot, 89.2 for bunchy top, 89.2 for sigatoka, 90.79 for panama wilt and 86.19 for normal leafs. While CNN provided 84.18 for leaf spot, 85.4 for bunchy top, 81.73 for sigatoka, 84.7 for panama wilt and 81.94 for normal leafs. Similarly, RF provided 80.6 for leaf spot, 80.1 for bunchy top, 79.2 for sigatoka, 79.5 for panama wilt and 76.5 for normal leafs. XG booster provided 74.27 for leaf spot, 74.59 for bunchy top, 78.5 for sigatoka, 72.8 for panama wilt and 71.2 for normal leafs. While KNN provided 68.4 for leaf spot, 67.2 for bunchy top, 64.2 for sigatoka, 70.4 for panama wilt and 67.3 for normal leafs. DT provided 48.1 for leaf spot, 45.7 for bunchy top, 48.3 for sigatoka, 52.2 for panama wilt and 50.7 for normal leafs. SVM provided 45.5 for leaf spot, 46.7 for bunchy top, 43.3 for sigatoka, 48.3 for panama wilt and 45.5 for normal leafs.

$$F1\ Score = \frac{2*recall*precission}{recall+precision} \tag{7}$$

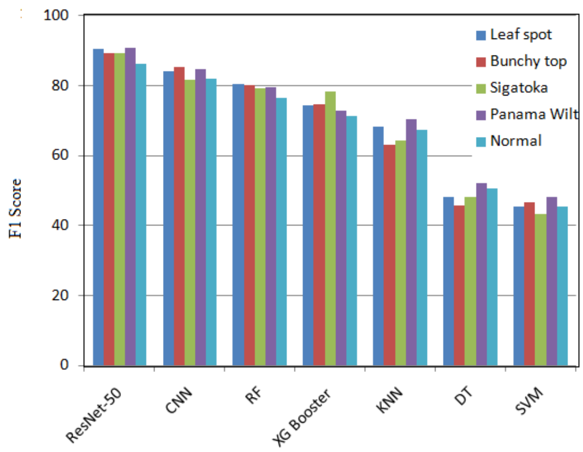

Figure 13: Comparison graph of F1 score

Negative predictive value is a way to gauge how well a model can identify true negatives or how accurately a negative forecast is made. A high NPV shows that the model is good at spotting the actual negatives, which is crucial in situation were failing to spot a negative case has serious repercussions. Equation 7 is used to calculate the NPV of a classifier model. "Fig. 14" shows the comparison of NPV between different classifiers used in our study. The proposed ResNet-50 provided NPV of 98.3 for leaf spot, 97.3 for bunchy top, 97.7 for sigatoka, 98 for panama wilt and 94.8 for normal leafs. While CNN provided 97.1 for leaf spot, 96.1 for bunchy top, 95.7 for sigatoka, 96.6 for panama wilt and 93.6 for normal leafs. Similarly, RF provided 96.1 for leaf spot, 95.4 for bunchy top, 94.4 for sigatoka, 95.5 for panama wilt and 92.2 for normal leafs. XG booster provided 94.8 for leaf spot, 93.8 for bunchy top, 94.5 for sigatoka, 93.6 for panama wilt and 90.8 for normal leafs. While KNN provided 94.7 for leaf spot, 91.4 for bunchy top, 91.3 for sigatoka, 94.1 for panama wilt and 97.9 for normal leafs. DT provided 86.9 for leaf spot, 86.8 for bunchy top, 86.7 for sigatoka, 89.5 for panama wilt and 83.2 for normal leafs. SVM provided 89.5 for leaf spot, 89.6 for bunchy top, 85.1 for sigatoka, 88.4 for panama wilt and 81.8 for normal leafs.

$$NPV = \frac{TN}{TN+FP} \tag{8}$$

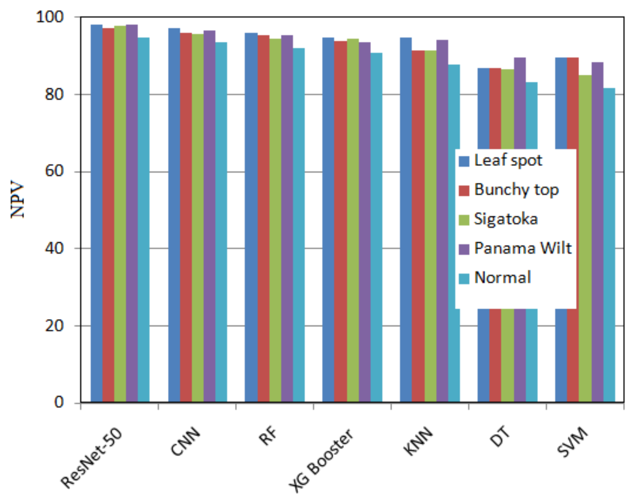

Figure 14: Comparison graph of NPV

The false positive rate is a crucial statistic for evaluating the model's capacity to avoid making false positive predictions. This metric can be especially crucial in applications where false positive can have serious repercussions, like plant leaf disease detection. Equation 8 is used to find FPR of proposed algorithm. "Fig. 15" shows the comparison of FPR between different classifiers used in our study. The proposed ResNet-50 provided sensitivity of 2.7 for leaf spot, 2.7 for bunchy top, 1.9 for sigatoka, 2.5 for panama wilt and 2.4 for normal leafs. While CNN provided 4.6 for leaf spot, 3.8 for bunchy top, 15 for sigatoka, 4.1 for panama wilt and 3.4 for normal leafs. Similarly, RF provided 5.2 for leaf spot, 5.3 for bunchy top, 4.7 for sigatoka, 5.9 for panama wilt and 5.2 for normal leafs. XG booster provided 7 for leaf spot, 6.4 for bunchy top, 5.2 for sigatoka, 7.3 for panama wilt and 6.5 for normal leafs. While KNN provided 9 for leaf spot, 6 for bunchy top, 10.8 for sigatoka, 9 for panama wilt and 7 for normal leafs. DT provided 14 for leaf spot, 11 for bunchy top, 15.1 for sigatoka, 12.9 for panama wilt and 10.9 for normal leafs. SVM provided 15 for leaf spot, 11.9 for bunchy top, 15.6 for sigatoka, 13.4 for panama wilt and 12.1 for normal leafs.

$$False\ positive\ rate\ (FRP) = 1 - SP \tag{9}$$

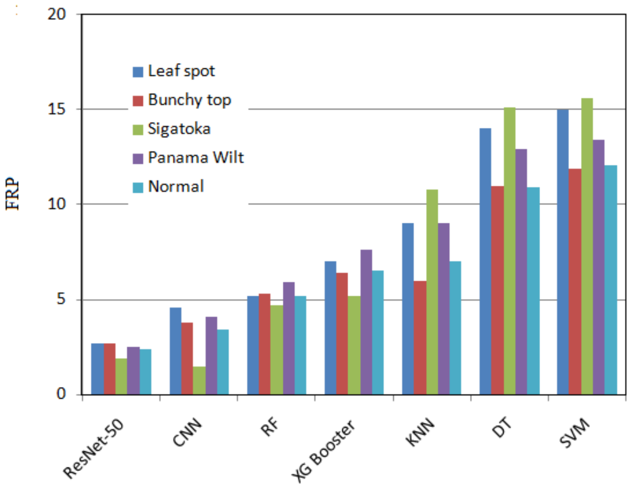

Figure 15: Comparison graph of FRP

Out of all positive predictions produced by the model, FDR reflects the percentage of positive predictions that were erroneous. When dealing with unbalanced datasets or attempting to reduce the number of false positive predictions, FDR is used to manage and comprehend the rate of false positive mistakes in a classification task. Equation 9 is used to find FDR of the proposed model. "Fig. 16" shows the comparison of FDR between different classifiers used in our study. The proposed ResNet-50 provided sensitivity of 11.4 for leaf spot, 10.7 for bunchy top, 13.3 for sigatoka, 10 for panama wilt and 8.8 for normal leafs. While CNN provided 18.7 for leaf spot, 15.3 for bunchy top, 19.4 for sigatoka, 16.4 for panama wilt and 12.9 for normal leafs. Similarly, RF provided 21.6 for leaf spot, 21 for bunchy top, 19.2 for sigatoka, 22.7 for panama wilt and 19.5 for normal leafs. XG booster provided 28.6 for leaf spot, 25.7 for bunchy top, 20.8 for sigatoka, 28.5 for panama wilt and 24.6 for normal leafs. While KNN provided 36.6 for leaf spot, 27.5 for bunchy top, 37.7 for sigatoka, 33.6 for panama wilt and 24.7 for normal leafs. DT provided 55.8 for leaf spot, 51.1 for bunchy top, 53.3 for sigatoka, 50.4 for panama wilt and 42.2 for normal leafs. SVM provided 58.8 for leaf spot, 52.1 for bunchy top, 57.2 for sigatoka, 53.6 for panama wilt and 47.6 for normal leafs.

$$False\ discovery\ rate\ (FDR) = \frac{FP}{TP+FP} \tag{10}$$

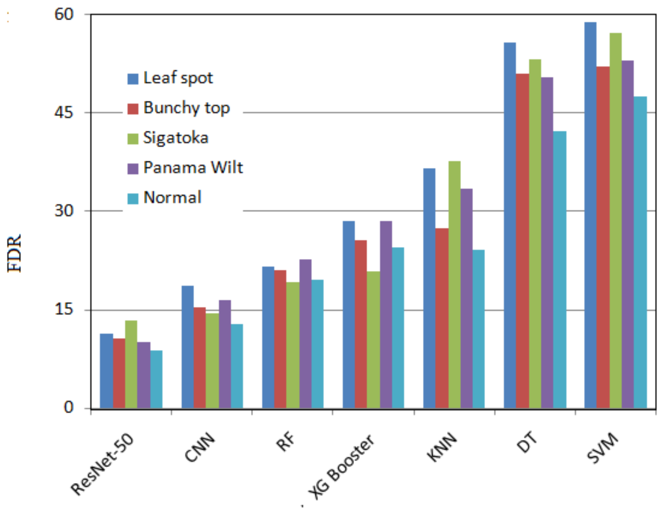

Figure 16: Comparison graph of FDR

4. Conclusion

The images of banana and sunflower were collected by visiting the farms and collecting the images from village dataset for normal and diseased leafs. Pre-processing was performed to reduce the noise in the images and to eliminate the background from the image. RGB image were converted to HSV based on Hue value calculated. Clustering of image was done to obtain the diseased part of the leaf. The segmented part of the leaf was used an input to the proposed model for the identification and classification of the disease. The proposed model was able to identify and suggest the pesticide to control the identified disease. The results obtained from the proposed model were compared with other classifiers to evaluate the performance of the proposed model. When compared with other classifier algorithms provided higher accuracy. The model can be further used to train the diseases in other plants.

Acknowledgement

All Authors express thanks to Sreenivas University, Mangaluru, Bapuji Institute of Engineering and Technology and Jain Institute of Technology, Davangere.

Ethical Clearance

The study was approved the Sreenivas University, Mangaluru.

Conflicts Of Interest

Authors declared that there is no conflict of interest

References:

1. Zhou, C., Zhang, Z., Zhou, S., Xing, J., Wu, Q., & Song, J. (2021). Grape leaf spot identification under limited samples by fine grained-GAN. *Ieee Access*, *9*, 100480-100489.
2. Ramesh, S., & Vydeki, D. (2020). Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. *Information processing in agriculture*, *7*(2), 249-260.
3. Sharma, P., Berwal, Y. P. S., & Ghai, W. (2018, November). KrishiMitr (Farmer's Friend): using machine learning to identify diseases in plants. In *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS)* (pp. 29-34). IEEE.
4. Chakraborty, S., Paul, S., & Rahat-uz-Zaman, M. (2021, January). Prediction of apple leaf diseases using multiclass support vector machine. In *2021 2Nd international conference on robotics, electrical and signal processing techniques (ICREST)* (pp. 147-151). IEEE.
5. Bose, B., Priya, J., Welekar, S., & Gao, Z. (2020, December). Hemp disease detection and classification using machine learning and deep learning. In *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)* (pp. 762-769). IEEE.
6. Shrestha, G., Das, M., & Dey, N. (2020, October). Plant disease detection using CNN. In *2020 IEEE applied signal processing conference (ASPCON)* (pp. 109-113). IEEE.

7. Mukti, I. Z., & Biswas, D. (2019, December). Transfer learning based plant diseases detection using ResNet50. In *2019 4th International conference on electrical information and communication technology (EICT)* (pp. 1-6). IEEE.
8. Paymode, A. S., & Malode, V. B. (2022). Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. *Artificial Intelligence in Agriculture*, *6*, 23-33.
9. Lahza, H., Naveen Kumar, K. R., Sreenivasa, B. R., Shawly, T., Alsheikhy, A. A., Hiremath, A. K., & Lahza, H. F. M. (2023). Optimization of Crop Recommendations Using Novel Machine Learning Techniques. *Sustainability*, *15*(11), 8836.1
10. Kumar, K. N., Lahza, H., Sreenivasa, B. R., Shawly, T., Alsheikhy, A. A., Arunkumar, H., & Nirmala, C. R. (2023). A Novel Cluster Analysis-Based Crop Dataset Recommendation Method in Precision Farming. *Comput. Syst. Sci. Eng.*, *46*(3), 3239-3260.
11. Nandish, B. M., & Pushparajesh, V. (2021, December). Simulation of Household Appliances with Energy Disaggrigation using Deep Learning Technique. In *2021 International Conference on Computational Performance Evaluation (ComPE)* (pp. 173-178). IEEE.
12. Ramesh S, Hebbar R, Niveditha M, Pooja R, Shashank N, & Vinod P V, (2018) Plant disease detection using machine learning. In *2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C)*, IEEE, 41-45. https://doi.org/10.1109/ICDI3C.2018.00017
13. Govardhan M, &Veena M B, (2019) Diagnosis of Tomato Plant Diseases using Random Forest. In *2019 Global Conference for Advancement in Technology (GCAT)*, IEEE, 1-5. https://doi.org/10.1109/GCAT47503.2019.8978431
14. Ahmed K, Shahidi T R, Alam S M I, &Momen S, (2019) Rice leaf disease detection using machine learning techniques. *International Conference on Sustainable Technologies for Industry 4.0 (STI)*, IEEE,1-5. https://doi.org/10.1109/STI47673.2019.9068096
15. Ramesh S, &Vydeki D, (2018) Rice blast disease detection and classification using a machine learning algorithm. *2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)* , IEEE, 255-259. https://doi.org/10.1109/ICMETE.2018.00063
16. Sarangdhar A A, &Pawar V R, (2017) Machine learning regression technique for cotton leaf disease detection and controlling using IoT. *International conference of Electronics, Communication and Aerospace Technology (ICECA)* , IEEE, 449-454. https://doi.org/10.1109/ICECA.2017.8212855
17. Pooja V, Das R, &Kanchana V, (2017) Identification of plant leaf diseases using image processing techniques. *IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, IEEE,130-133. https://doi.org/10.1109/TIAR.2017.8273700
18. Prajapati B S, Dabhi V K, &Prajapati H B, (2016) A survey on detection and classification of cotton leaf diseases. *International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)* , IEEE, 2499-2506.DOI: https://doi.org/10.1109/ICEEOT.2016.755143
19. Joshi A A, &Jadhav B D, (2016,) Monitoring and controlling rice diseases using Image processing techniques. *International Conference on Computing, Analytics and Security Trends (CAST)*, IEEE, 471-476. https://doi.org/10.1109/CAST.2016.7915015
20. Waghmare H, Kokare R, &Dandawate Y, (2016) Detection and classification of diseases of the grape plant using opposite colour local binary pattern feature and machine learning for the automated decision support system. *3rd international conference on signal processing and integrated networks (SPIN)*, IEEE, 513-518. https://doi.org/10.1109/SPIN.2016.7566749.