Plant Diseases Classification using Machine Learning

Tan Soo Xian², Ruzelita Ngadiran¹²

¹Centre of Excellence Advanced Computing (AdvComp), Universiti Malaysia Perlis, 2Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia. s171020953@studentmail.unimap.edu.my, ruzelita@unimap.edu.my

Abstract. Plant diseases are one of source of obstruction in the quality and productivity of plants which can lead to the shortage of food supply. Therefore, plant disease classification is essential to the agriculture industry. The objective of this research is to classify the plant diseases by assessing the images of the leaves with the application of Extreme Learning Machine (ELM), a Machine Learning classification algorithm with a single layer feed-forward neural network. This work proposed image features as input where the image is pre-processed via HSV colour space and features extraction via Haralick textures. The features are then fitted in the ELM classifier to perform the model training and testing. The accuracy of ELM is then calculated after the testing has been done. The dataset used comprises of tomato plant leaves which is a subset of the Plant-Village dataset. The results produced from the ELM shows a better accuracy that is 84.94% when compared to other models such as the Support Vector Machine and Decision Tree.

1. Introduction
A plant disease prevents a plant from reaching its maximum potential of production. This definition includes non-infectious and infectious diseases [1] that pose threat to the agriculture industry by causing a decline in production and economic as well as reduction in the quality and the quantity of the plant products. A study conducted to report the effect of the plant disease on global production [2]. The research had shown that plant diseases had caused a high yield loss to subtle crops around the world being the wheat (30%), rice (40%), potato (21%), corn (41%) and soybean (30%).

Extreme Learning Machine (ELM) is popular due to its simple design but good generalization. It is formed as single layer feedforward neural network that is used for classification with a minimal amount of modification on the weight of inputs. Thus, the complexity of the application of the model is significantly lower.

In this work, the image features such as Haralick textures, Hue-Saturation-Value Histogram and colour moments are proposed to identify the plant disease from tomato leaves. The performance based on the accuracy rate of ELM classifier is expected to be on par to the with other classification models such as Decision Tree (DT) and Support Vector Machine (SVM). The complexity will be as well reduced.

The remaining paper will highlight the prior research in Section 2. Then, followed by the methods and material in Section 3. The research methods and discussion of the results will be explained in Section 4 and 5. Finally, conclusions are made in Section 6.
2. Literature Review

ELM is proposed by Vijayalakshmi & Murugan [3] to detect plant diseases. Their model follows the basic steps of machine learning model such as feature extraction, training the classifier and classification. In their work, the artificial bee colony clustering (ABC) [4] is used isolate infected areas using the collective behaviour of insects such as bees. Second, the feature extraction is completed using transform encoded local patterns (TELP) to compute the texture analysis [5], gradient features and color histogram techniques from the data. For the training and testing, the author tested both classifiers namely Support Vector Machine (SVM) [6] and ELM, which resulted in ELM leading by 2% at the rate of accuracy at 97%. This work proves that ELM can surpass the conventional model, SVM in terms of classification accuracy.

Another approach on ELM is suggested by Saragih [7] with Simulated Annealing (SA) to classify the Jatropha Curcas Disease. In this work, the dataset of jatropha diseases is used in the process of model training, testing, and optimising the output of the weights of SA [8]. The important role of SA is to tune the best weight value of neuron on ELM classifier. The model is further enhanced using the decision tree classifier to determine the jatropha seed disease. The aim of the decision tree classifier is to create a model capable of making a judgement according to the specialised knowledge. The result of this method gives out a much more desirable rate of accuracy at 94.74% compared to pure ELM at 66.67%.

Maniyath [9] proposed a smartphone application to classify the coffee leaf diseases with ELM. The system recognises diseases such as leaf miners and leaf rust and calculates the rate of its severity. The image processing step involves the identification of a leaf from the image by separating the background and foreground of the image to obtain the Hue Saturation Value (HSV) and YCbCr colour space. This method is done via several segmentation methods that include histogram distribution, edge detection, neighbourhood detection, and clustering. Otsu algorithm [10] and the iterative threshold algorithm are used to perform a comparison to k-means in the process to segment the foliar damages. The results show the practicality of the model to identify feeding damage by obtaining the of the severity rate of 99.095%. Thus, the ability of the ELM model is proved for able to produce a significant result with the given image features of the coffee leaves.

Maniyath [11] recommends a comparison of various machine learning models such as Support Vector Machine (SVM), K-Nearest Neighbours (KNN), RF and naive bayes algorithm. It uses the public datasets of diseased leaves for training and testing purposes. Second, the histogram of an orientated gradient (HOG) is used to extract features such as hu moments, Haralick textures and colour histogram. The datasets of diseased and healthy leaves are then collectively trained under the aforementioned models. The comparison resulted with random forests having the highest accuracy at 70.14%. The model can be improved with implementing the ELM model as the random tree does not scale well with the amount features compared to ELM.

Panchal, et al. [12] uses Random Forest (RF) as a classifier [13] to detect plant diseases, which share similarities with the previous titles mentioned. Images of leaves with plant diseases such as bacterial spot, early blight and late blight are acquired using a camera or gadgets with similar capabilities to create a dataset. The image processing step comprises of converting the images to grey scale as well as HSV and smoothing of the images using a smoothing filter. Then, the research has utilised the k-means segmentation method to isolate the infected portion of the leaf and Grey Level Co-occurrence Matrix (GLCM) [14] to extract features from leaf. The proposed methodology can recognise and differentiate the plant diseases with the rate of 98% accuracy when the RF classifier is used to determine the plant diseases. Besides, HSV histogram features can be considered the alternative of the GLCM. It is capable of describing an image in a general presentation which is required for a classification model.

Bhatia, et al. [15] suggests, “Application of ELM in Plant Disease Prediction for Highly Imbalanced Dataset”, that implements ELM classifier for plant disease prediction according to an imbalanced dataset consisting of real time images being the Tomato Powdery Mildew Disease (TPMD) dataset. To balance this dataset, several re-sampling techniques such as Random over Sampling (ROS), Synthetic Minority Over-Sampling Technique (SMOTE), Random under Sampling (RUS) and Importance Sampling (IMPS) ahead of performing the classification using ELM. The sample models are further evaluated using the Classification Accuracy (CA) and area under the curve (AUC). This technique resulted to ELM performing better for TPMD dataset after the resampling techniques were applied. The
optimum outputs are found from the IMPS technique that shows 88.57% in CA and 89.19% in AUC. The result can be improved; hence this work will use tomato dataset. According to the literature review, it is observed that image features play a significant role in improving classification result. In addition, ELM, despite using single RBF can give good classification results with suitable features.

3. Methods and Materials
This section presents the material and main classifier models used in this research. Google Colab has been used for the implementation of the current research.

3.1. Dataset
The dataset of tomato leaves is obtained from the Kaggle and it is a subset of the Plant Village dataset [16]. The focused dataset contains 10 classes representing the names of the diseases where each class contains the 1000 images such as shown in Figure 1:

![Figure 1. The distribution of classes of Tomato Leaf Diseases Dataset](image)

3.2. Classifiers

3.2.1. Extreme Learning Machine Model
Extreme Learning Machine (ELM) model was introduced by Huang in 2006 [17]. It is a type of feedforward neural network with the simplest form [18]. The model is consists of multiple hidden nodes to form single hidden layer, where the weights between inputs and hidden nodes are randomly initialised and remain unchanged throughout training. The weights that connect the hidden nodes to the output are trained, but due to the simple structure of an ELM, these weights can be obtained by calculating a system of linear matrix equations [19]. Thus, such neural network requires no iterations which makes it faster with better generalisation performance than a back-propagation neural network. It is considered a supervised learning model because of the requirement of input nodes (features) and output nodes (labels/classes).

ELM can be modelled as shown in Eq. (1):

$$\sum_{j=1}^{I} \beta_j g(w_j x_j + b_j) = y_j, j = 1,...,N$$

(1)

Where $x_j$ is the input, $y_j$ is the output, $N$ is the number of samples, $w_j$ is the input weight, $b_j$ is the bias of the hidden layer and $\beta_j$ is the weight vector connecting the $i_{th}$ hidden node and output nodes.

The above Eq. (1) can be written compactly as

$$T = H\beta$$

(2)
\[ H = \begin{bmatrix} g(w_1 x_1 + b_1) & \cdots & g(w_L x_1 + b_N) \\ \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & \cdots & g(w_L x_N + b_N) \end{bmatrix}_{N \times L} \] (3)

\[ \hat{\beta} = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_L^T \end{bmatrix}_{N \times m} \] (4)

Where \( H \) is the hidden layer output matrix of the neural network; the \( i \)th column of \( H \) is the \( i \)th hidden node output with respect to inputs \( x_1, x_2, x_3, \ldots, x_N \). The \( \hat{\beta} \) parameter is obtained by \( \hat{\beta} = H^\dagger T \). The Moore–Penrose generalized inverse of \( H \) is shown with \( H^\dagger \) [20].

### 3.2.2. Support Vector Machine Model

A SVM model obtains a boundary or a hyperplane from features extracted to from the data to perform the classification operation. SVM is used to transform features into much higher dimensions [21], which makes it a good comparison to ELM that can handle large data in the same way. The SVM model is created to determine the linear discriminant function with the largest margin that separates each class of the data. The learning data that is the closest to the boundary is the support vector [22]. This classifier provides a reliable solution in many areas such as image and object recognition, voice recognition, fingerprint recognition, and handwriting recognition.

### 3.2.3. Decision Tree Model

A decision tree is a popular technique used to classify a big number of data and to extract the data with similar characteristics. The data are split into smaller subsets addition to that the tree is developed given in a final result consisting of the tree and leaf nodes [23].

### 3.3. Activation Functions

Activation functions (AFs) are mathematical functions commonly used to calculate weights and biases of a neural network, which in this study is ELM. An activation function generates outputs of ELM and patterns of the dataset. In the research, the activation functions introduced are such as:

1. **Linear Function.**
   It is a straight-line function where the activation generates an exact copy of the input as the output of the weight[24], which is the weighted sum from neuron. It is defined in Eq. (5)
   \[ f(x) = x \] (5)

2. **Sigmoid Function**
   It is a non-linear function that exists in between the range of \([0, 1]\) thus is suitable in binary classification problems, which is defined as Eq. (6)
   \[ f(x) = \frac{1}{1 - \exp^{-x}} \] (6)

3. **Hyperbolic Tangent Function (TanH).**
   It exists in between the range of \([-1, 1]\) which is calculated as Eq. (7)
   \[ f(x) = \frac{1}{1 - \exp^{-x}} \] (7)

4. **Rectified Linear Unit (ReLU) Function**
   It is the most common used activation function in convolutional neural networks or deep learning. It exists in between the range of \([0, \infty]\), which is defined as
   \[ f(x) = \begin{cases} -1, & \text{if } x < -1 \\ x, & \text{if } -1 \leq x \leq 1 \\ 1, & \text{if } x > 0 \end{cases} \] (8)
3.4. Performance Metrics

To evaluate the performance of ELM in a plant disease classification, the classification accuracy (CA) [15] and confusion matrix (CM) [25] are used to analyse the performance of the model. Classification accuracy is a method to measure the overall closeness of original labels to detected labels. To obtain a better understanding on the distribution of the accuracy across the labels, and the CM is used. It compares and shows the accuracy of the classification of each label when a prediction is performed.

4. Research Methodology

The proposed methodology contains four steps that are shown in Figure 2:

![Figure 2. Proposed Methodology for Plant Disease Classification](image)

4.1. Image Pre-processing

This section involves image pre-processing methods that will be used to obtain optimum data from images of tomato leave. The methods involve image resizing, colour space conversion and image segmentation. After each image in the dataset are resized to 256*256-resolution in Figure 3, their colour spaces are then converted to HSV for further processing [14]. The HSV colour space represents three elements that describe the colour (Hue), the level of colour dominance (Saturation) and the brightness level (Value). This particular colour space separates the image intensity from the colour information, thus the colour texture, the level of dominance and the brightness of an image can be identified [26]. Figure 3 shows the HSV scatterplot of tomato late blight disease for image segmentation and we observe that grey and black pixels were present in the saturation level of 0 to 60. Thus, by thresholding the saturation ranged from 60 and higher were performed to create a mask that remove the background of the image and retained as much of the information in the image.

Next, the respective images were smoothened using the Gaussian filter to distinguish the edges [27]. A mask is created to represent the shape of the leaf from the threshold, we created. After the closing morphological method was used to remove all the noise around the mask, a new segmented image is created by overlap the mask on the original image, as shown in Figure 3 (c).

![Figure 3. (a) Tomato Late Blight Image (Resized); (b) Scatter Plot of HSV from Image (a); Mask of Segmented Image (a)](image)
4.2. Feature Extraction

The next section involves feature extraction to reduce the features of the existing ones and then discarding the original features. This new reduced set of features are the features that will summarise most of the information contained in the original image [27]. This process has a significant impact given when dealing with image with as high-dimensional data spaces such as RGB and HSV. For instance, if each pixel of a raw image shown in Figure 3 is to be applied as the features for the machine learning model, it can reach up to 196608 pixels per image.

Haralick Texture: Texture is defined as the frequency of a pattern and colour that are visible in an image or object such diseased spots on the tomato leaves. Thus, Haralick texture is used to represent the texture of an image by using normalized co-occurrence matrices or we call grey level co-occurrence matrices (GLCM) to obtain the distribution of grey values between pixels in a greyscale image to perform texture extraction. It is based on the matrices in the adjacent, which stores the position of the reference pixel and the neighbour pixel values \((i, j)\) [28]. The texture is determined based on the number count of the pixel, \(i\) presence to the position next to the pixel \(j\) such as shown in Figure 4. Among the textures calculated, four of textures are addressed in this work, namely the contrast, correlation, inverse difference moments and entropy.

\[
\text{Contrast} = \Sigma \Sigma (i-j)^2 P(i,j) \tag{9}
\]

\[
\text{Correlation} = \frac{\Sigma \Sigma (i-j)^2 P(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{10}
\]

\[
\text{Inverse Difference Moments} = \Sigma \Sigma \frac{P(i,j)}{1+(i-j)^2} \tag{11}
\]

\[
\text{Entrophy} = \Sigma \Sigma (i-j)^2 P(i,j) \log (P(i,j)) \tag{12}
\]

![Figure 4. GLCM [28]](image)

HSV Histogram: In the context of image processing, a histogram of an image refers to a histogram of the colour pixel intensity and it is applicable throughout different colour spaces such as greyscale, RGB, HSV, CMYK and etc. Each channel presented in an image is 8-bit in size, there are 256 possible intensities available to be displayed in the histogram. By simulating the rest of the channels, a three-dimensional histogram can be constructed by combining the histogram of each channel such as the hue, saturation and values of the colour space [29]. The histogram is presented in Figure 5. The main reason that the histogram of HSV colour space is used in this research because it abstracts the colour (hue) by separating it from saturation and pseudo-illumination (value). The space of the histogram is divided into many ranges, such as arranged as a regular grid, each containing the same colour values.
Colour Moments: Colour moments are the measurements that characterise colour distribution in an image in the form of probability distribution colour channels present within the image. For example, an image with RGB colour space contains three colour channels such as red, green and blue. Colour moments are primarily practised for colour detection purposes as features in image retrieval applications to compare how similar two images are based on colour. In this research, two important components of the colour moments will be utilised namely the mean and standard deviation. The mean signifies an idea where the colour of the pixels is located and allows the summarisation of the colour values of an image. As for the standard deviation, it measures the deviation of a colour value from its mean. When the standard deviation is close to zero, the colour values measured are near the mean but when it is high, the colour values are scattered and distanced away from the mean.

4.3. Dataset Preparation

This section will explain the dataset preparation after image pre-processing and feature extraction is completed. A normalisation process will be performed by subtracting the minimum value in the features and then dividing by the range. The dataset will be split into a ratio of 7:3. The 70% of the dataset, which is 7000 of the tomato leaves images would be used to train the ELM model. The 30% of the dataset, which is 3000 of the images would be used to test the ELM model. The train and test datasets are shuffled to avoid any bias or patterns in the datasets before model training, data frame created were shuffled and the index were recreated so that the it are not be accidentally sorted back to the original arrangement. After the features are processed, the labels are transformed into a numerical value as the machine learning model unable handle categorical variables. This is achieved using One-Hot-Encoding method. Considering that 10 labels (classes) were found in the tomato leaf diseases dataset shown in, each label would be presented in 10 columns, where the values of 0 and 1 are the identifier of the label according to the features.

4.4. Classification

The final process is to classify the disease name of the tomato leaves. To complete this step, the Extreme Learning Machine (ELM) classifier is used. It is a type of feedforward neural network which is considered the simplest form of artificial neural network created. Such neural network does not perform iterations thus making it faster with better generalisation performance than networks trained using a back-propagation method. It contains collection of multiple neurons (nodes) that are arranged in the form of three layers called input layer, single-hidden layer and output layer.

After declaring the number of classes (labels), hidden layers and inputs (features), a new instance of the ELM model will be created to fit to the new training dataset created from feature extraction. The training time, accuracy and loss are recorded. Next, the ELM model created will be validated using the test dataset. The trained model will be saved into the Google Drive for further analysis and testing.
5. Result & Discussion

5.1. Experimental Images
The dataset in this study includes 10 types of diseases found on tomato plants namely the bacterial spot, early blight, late blight, leaf mould, etc. These diseases are commonly observable on the surface of the leaf. 70% (7000) of the images is used as the training dataset, while 30% (3000) of it is used as a testing dataset. The images of the bacterial spot, early blight, late blight, leaf mould are segmented and shown in Figure 6. The background of the images is removed by performing a bitwise operation with the mask created from thresholding the saturation of the image. The characteristics of each leaf are retained with minimal background colour.

Features of the HSV Histogram, Haralick textures and colour moments of RGB colour space are extracted and stored in the storage for further use. A total number of 298 features are extracted from a single image. The composition of the features is shown below:

| Features Extracted | Number of Features |
|--------------------|--------------------|
| HSV Histogram       | 288                |
| Haralick textures   | 4                  |
| Colour Moments      | 6                  |

![Figure 6. Experimental Images](image-url)
5.2. Activation Functions and Neurons
In this section, we perform a parameter tuning with the activation functions, such as linear, ReLU, sigmoid and tanh with a set of selected number of neurons (128, 256, 512, 1024, 2048, 4096).

| Activation Function | Number of Neurons |
|---------------------|-------------------|
|                     | 128               |
| Linear              | 71.40             |
| ReLU                | 69.87             |
| Sigmoid             | 72.30             |
| Tanh                | 70.96             |
|                     | 256               |
|                     | 72.60             |
|                     | 80.43             |
|                     | 82.90             |
|                     | 76.67             |
|                     | 76.33             |
|                     | 81.20             |
|                     | 82.29             |
|                     | 72.60             |
|                     | 72.60             |
|                     | 82.73             |
|                     | 84.43             |
|                     | 72.63             |
|                     | 83.97             |
|                     | 74.30             |
|                     | 71.43             |
|                     | 84.91             |
|                     | 72.63             |
|                     | 84.91             |
|                     | 71.43             |
|                     | 74.30             |

Table 2. Parameter’s Percent Accuracies of ELM

Table 2. shows the performance accuracies of ELM with specified parameters. From results, it is visible that the sigmoid and the tanh activation functions generate better accuracy than the linear and ReLU on activation functions. Moreover, accuracies of the model improve with the increase of neurons in the hidden layer, but slow decline happens from 2048 to 4096 neurons. Although weights and biases in the training and testing are selected at random, the results of the model are still consistent that range in between 69% and 84% of accuracy.

5.3. Performance Analysis
A comparison between ELM and other models such as SVM and decision trees is made in this section. There are no prior changes to features of the dataset, such as the number of features and labels. Table 3 and Table 4 are parameters of aforementioned models used in the comparison.

| Parameters | C | kernel | gamma |
|------------|---|--------|-------|
| Values     | 10| poly   | 1     |

Table 3. Parameters of SVM Model

| Parameters | criterion | max_features | Random_state |
|------------|-----------|--------------|--------------|
| Values     | entropy   | 298          | 1            |

Table 4. Parameters of Decision Tree Model

The heatmap represents the confusion matrix in a two-dimensional form to further analyse the distribution of the accuracy of each model. The values are represented as colours in the graph to provide a coloured visual summary of information. Each class of the disease is shown in numerical form such as ‘0’ for healthy leaves, ‘1’ for bacterial spot in incremental order for other diseases. The classification accuracy of ELM shows 84.94% after the testing is done. This is observed by summing up the CA of each class that is shown in the heat map with a higher temperature (darker).
Figure 7. (a) Confusion Matrix of ELM Model; (b) Confusion Matrix of SVM Model; (c) Confusion Matrix of Decision Tree Model

Table 5. Classification Accuracy of ELM, SVM and DT Model

| Leaf Disease/Healthy | Number of Testing Images | Classification Accuracy % |
|----------------------|--------------------------|---------------------------|
|                      |                          | ELM | SVM | Decision Tree |
| Bacterial spot       | 100                      | 9.20| 9.50| 7.83          |
| Early blight         | 100                      | 7.93| 9.17| 6.77          |
| Late blight          | 100                      | 7.57| 7.80| 6.50          |
| Leaf mould           | 100                      | 7.83| 9.20| 7.60          |
| Septoria leaf spot   | 100                      | 7.97| 8.83| 7.73          |
6. Conclusion

In this work, Extreme Learning Machine (ELM) classifier is proposed to identify leaf diseases of tomato plants based on image features. First, the HSV colour segmentation is applied to the input image to segment the leaf. Then, the features are extracted using HSV Histogram, Haralick and colour moments of each RGB colour space. These features are then input to the ELM classifier for training and testing purposes to identify the disease found on the tomato leaf. Overall, the result shows that ELM has a good result compared with decision tree classifier using proposed image features. For future work, it would be interesting to consider the disease classification for other plants such as rice, corn and wheat rather than specific plants.

References

[1] University of Nebraska - Lincoln, “Plant Disease: Pathogens and Cycles | CropWatch,” University of Nebraska, 2020. https://cropwatch.unl.edu/soybean-management/plant-disease (accessed Nov. 12, 2020).

[2] S. Savary, L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson, “The global burden of pathogens and pests on major food crops,” Nat. Ecol. Evol., vol. 3, no. 3, pp. 430–439, 2019, doi: 10.1038/s41559-018-0793-y.

[3] S. Vijayalakshmi and D. Murugan, “An Effective Approach for Diagnosis of Plant Disease using ELM,” vol. 4, no. 12, pp. 327–335, 2017.

[4] A. Kumar, D. Kumar, and S. K. Jarial, “A review on artificial bee colony algorithms and their applications to data clustering,” Cybern. Inf. Technol., vol. 17, no. 3, pp. 3–28, 2017, doi: 10.1515/cait-2017-0027.

[5] Z. Pan, Z. Li, H. Fan, and X. Wu, Feature based local binary pattern for rotation invariant texture classification, vol. 88. Elsevier Ltd, 2017.

[6] B. Kalantar, B. Pradhan, S. Amir Naghibi, A. Motevalli, and S. Mansor, “Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN),” Geomatics, Nat. Hazards Risk, vol. 9, no. 1, pp. 49–69, 2018, doi: 10.1080/19475705.2017.1407368.

[7] T. H. Saragih, W. F. Mahmudy, A. L. Abadi, and Y. P. Anggodo, “Application of extreme learning machine and modified simulated annealing for jatropha curcas disease identification,” Int. J. Adv. Soft Comput. its Appl., vol. 10, no. 2, pp. 108–119, 2018.

[8] L. M. R. Rere, M. I. Fanany, and A. M. Arymurthy, “Simulated Annealing Algorithm for Deep Learning,” Procedia Comput. Sci., vol. 72, pp. 137–144, 2015, doi: 10.1016/j.procs.2015.12.114.

[9] G. L. Manso, H. Knidel, R. A. Krohling, and J. A. Ventura, “A smartphone application to detection and classification of coffee leaf miner and coffee leaf rust,” arXiv, 2019.

[10] C. Huang, X. Li, and Y. Wen, “AN OTSU image segmentation based on fruitfly optimization algorithm,” Alexandria Eng. J., vol. 60, no. 1, pp. 183–188, 2020, doi: 10.1016/j.aej.2020.06.054.

[11] S. R. Maniyath et al., “Plant disease detection using machine learning,” Proc. - 2018 Int. Conf. Des. Innov. 3Cs Comput. Commun. Control. ICDI3C 2018, no. April, pp. 41–45, 2018, doi: 10.1109/ICDI3C.2018.00017.

[12] P. Panchal, V. C. Raman, and S. Mantri, “Plant Diseases Detection and Classification using Machine Learning Models,” CSITSS 2019 - 2019 4th Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. Proc., vol. 4, pp. 1–6, 2019, doi: 10.1109/CSITSS47250.2019.9031029.
[13] M. Belgiu and L. Drăgu, “Random forest in remote sensing: A review of applications and future directions,” ISPRS J. Photogramm. Remote Sens., vol. 114, pp. 24–31, 2016, doi: 10.1016/j.isprsjprs.2016.01.011.

[14] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, “Tomatoes classification using K-NN based on GLCM and HSV color space,” Proc. - 2017 Int. Conf. Innov. Creat. Inf. Technol. Comput. Intell. IoT, ICITech 2017, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/INNOCIT.2017.8319133.

[15] A. Bhatia, A. Chug, and A. Prakash Singh, “Application of extreme learning machine in plant disease prediction for highly imbalanced dataset,” J. Stat. Manag. Syst., vol. 23, no. 6, pp. 1059–1068, 2020, doi: 10.1080/09720510.2020.1799504.

[16] A. Ali, “PlantVillage Dataset | Kaggle,” 2019. https://www.kaggle.com/abdallahalidev/plantvillage-dataset (accessed Jan. 02, 2021).

[17] G. Bin Huang, Q. Y. Zhu, and C. K. Siew, “Extreme learning machine: Theory and applications,” Neurocomputing, vol. 70, no. 1–3, pp. 489–501, 2006, doi: 10.1016/j.neucom.2005.12.126.

[18] K. Erden, “Introduction to Extreme Learning Machines | by Kemal Erden (burnpiro) | Towards Data Science,” May 30, 2020. https://towardsdatascience.com/introduction-to-extreme-learning-machines-c020020f8f28 (accessed Jan. 05, 2021).

[19] M. Jain, W. Andrecopoulos, and M. Stamp, “Convolutional neural networks and extreme learning machines for malware classification,” J. Comput. Virol. Hacking Tech., vol. 16, no. 3, pp. 229–244, 2020, doi: 10.1007/s11416-020-00354-y.

[20] M. A. Rakha, “On the Moore-Penrose generalized inverse matrix,” Appl. Math. Comput., vol. 158, no. 1, pp. 185–200, 2004, doi: 10.1016/j.amc.2003.09.004.

[21] K. Elangovan, “Plant Disease Classification Using Image Segmentation and SVM Techniques,” Int. J. Comput. Intell. Res., vol. 13, no. 7, pp. 1821–1828, 2017, [Online]. Available: http://cs231n.stanford.edu/reports/2017/pdfs/325.pdf.

[22] A. Dixit and S. Nema, “Wheat Leaf Disease Detection Using Machine Learning Method-A Review,” Int. J. Comput. Sci. Mob. Comput., vol. 7, no. 5, pp. 124–129, 2018, [Online]. Available: www.ijcsmc.com.

[23] P. Hamsagayathri and P. Sampath, “Performance Analysis of Breast Cancer Classification Using Decision Tree Classifiers,” Int. J. Curr. Pharm. Res., vol. 9, no. 2, p. 19, 2017, doi: 10.22159/ijcpr.2017v9i2.17383.

[24] S. Sharma, “Activation Functions in Neural Networks | by SAGAR SHARMA | Towards Data Science,” Sep. 06, 2017. https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6 (accessed Feb. 25, 2021).

[25] L. Cai, W. Shi, Z. Miao, and M. Hao, “Accuracy assessment measures for object extraction from remote sensing images,” Remote Sens., vol. 10, no. 2, 2018, doi: 10.3390/rs10020303.

[26] D. S. Y. Kartika, D. Herumurti, and A. Yuniarti, “Butterfly Image Classification Using Color Quantization Method on HSV Color Space and Local Binary Pattern,” IPTEK J. Proc. Ser., vol. 4, no. 1, p. 78, 2018, doi: 10.12962/j23546026.y2018i1.3512.

[27] N. Mohd Ali, N. K. A. Md Rashid, and Y. M. Mustafah, “Performance comparison between RGB and HSV color segmentations for road signs detection,” Appl. Mech. Mater., vol. 393, no. September 2013, pp. 550–555, 2013, doi: 10.4028/www.scientific.net/AMM.393.550.

[28] T. Löfstedt, P. Brynolfsson, T. Asklund, T. Nyholm, and A. Garpebring, “Gray-level invariant Haralick texture features,” PLoS One, vol. 14, no. 2, pp. 1–18, 2019, doi: 10.1371/journal.pone.0212110.

[29] S. Sergyán, “Color histogram features based image classification in content-based image retrieval systems,” SAMI 2008 6th Int. Symp. Appl. Mach. Intell. Informatics - Proc., no. February, pp. 221–224, 2008, doi: 10.1109/SAMI.2008.4469170.

[30] D. JadHAV, G. PhadK, A. KhandagALe, and M. MulIK, “Image Retrieval Based On Higher Order Colour Moments and Texture Features .”

[31] K. Potdar, T. S., and C. D., “A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers,” Int. J. Comput. Appl., vol. 175, no. 4, pp. 7–9, 2017, doi: 10.5120/ijca2017915495.