Integrated Deep Hybrid Learning Model Upon Spinach Leaf Classification and Prediction with Pristine Accuracy

Meganathan Elumalai [1*], Terrance Frederick Fernandez [2]
[1, 2] Institute of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India
Email: [1] sakthimega10@gmail.com, [2] frederick@pec.edu
*Corresponding Author

***Abstract*—Over the years, Agriculture has been a mainstay of life for Indians and about half the working population of Tamil Nadu. Spinach is an integral part of everyone's meal and its nutrient content is higher than other veggies. The nutrients are unique for varied varieties so there is a dire need to classify them and thus to predict them. Furthermore, exactitude prediction leads to easy detection of spinach leaves. In this work, we selected 5 varieties of spinach leaves populated under a huge dataset. We implemented the same employing a Deep Hybrid approach which is a fusion of conventional Machine Learning with state-of-the-art Deep Learning using Orange toolkit. Out of the plethora of these AI Domaine approaches, four classifiers, such as Support Vector Machine (SVM), k- Nearest Neighbour(kNN), Random Forest (RF), and Neural Network (NN) were chosen and implemented. Existing methods using these algorithms have achieved promising results, with individual accuracies of 98.80% (RF), 98.20% (KNN), 99.9% (NN), and 99.60% (SVM). However, the IDHLM aims to surpass these individual performances by integrating them into a cohesive framework. This approach leverages each algorithm's complementary strengths to achieve even higher classification accuracy. The abstract concludes by highlighting the potential of the IDHLM for achieving pristine accuracy in spinach leaf classification.**

Keywords*—*Image Classification; Accuracy; Spinach Leaves; Deep Learning; Support Vector Machine; Convolutional Neural Network; Orange Toolkit.

I. INTRODUCTION

Agriculture has always been vital to human life, as seen by the wide variety of crops used to produce the world's food supply. Spinach is particularly noteworthy among these vegetables because of the high concentration of healthful components it contains, including protein, vitamins, and minerals. The health benefits of spinach aren't always worth the difficulty of distinguishing between the various varieties of spinach leaves, especially in densely populated locations. It is super rich in Calcium, Vitamins, Iron, Minerals, and Protein, which supplement the well-being of human skin, and bones, purify the blood, and aid in hair growth. Spinach leaf compliments Vitamin_A, Vitamin_C, Vitamin_E, Vitamin_K, Potassium, Iron, Copper, and Folic Acid, and no wonder the comic character Popeye the Sailor draws super-human Caliber from a spinach leaf can. Distinguishing between various spinach leaf varieties can be challenging, especially for urban residents. Based on this, more than 40 spinach leaf varieties in Tamil Nadu were discovered such as Agathi leaves, Amaranthus, Aritis, Mint, Tropical amaranth, Chinese spinach, Malabar spinach, drumstick leaves, dwarf copper leaf, and many more. The small size of spinach leaves makes it particularly difficult to detect and identify infections in heavily populated metropolitan environments. For the sake of both farmers and consumers, it is essential to find innovative ways to accurately classify spinach leaves. So, to overcome these issues at the beginning we used the Orange data mining tool for classifying the spinach leaves. This research aims to develop a novel method for accurate spinach leaf classification using deep learning and machine learning algorithms within the Orange data mining tool. The primary question is: can a combination of pre-trained deep learning models and machine classification in Orange effectively recognize and classify different spinach varieties? Previous studies have shown that classification tasks, such as plant disease identification, are well-suited to deep learning and machine learning. The difficulties in distinguishing between spinach kinds, especially in city settings, necessitate expanding this success to spinach leaf categorization. Here, we extract features from photographs of spinach leaves and utilize them to classify the leaves using the pre-trained deep learning models included in the Orange data mining tool. Clustering, artificial intelligence, and image processing are also utilized in the study to enhance the accuracy of spinach leaf classification.

To overcome these problems of aptly classifying the type, researchers used Deep_Learning and Machine_Learning algorithms to classify spinach leaves [1]. Additionally, they employed several previously trained models and certain of their own to classify spinach leaves with astounding accuracy. A branch of deep learning (DL) is machine learning. As opposed to that, ML is a component of AI. In deep learning, neural networks mimic or replicate how a human cerebral works. The ML typically carries out feature extraction and transformation [2]. The previous layer's output serves as the next layer's input. DL models on their own can extract the appropriate features. The majority of the time, it works well for large data entry and extraction volumes. Between the outcome and input layers, a Deep Neural Network (DNN), an NN in DL implementation, has several hidden layers [3]. A set of inputs, a complicated operation, and a classification output are the main objectives of NN. To recognize and group the diseases affecting the tomato plants, use CNN-based architecture. For the sake of eradicating the significant features from the input instances and categorizing them, this method [4] utilized three convolutions in addition to max-pooling layers. This method [5] shows improved precision in the classification of tomato diseases, but it has the drawback of over-fitting over a constrained set of classes. The deep residual framework was used instead of the VGG16 [6] model for feature extraction in the researchers' novel Faster-RCNN[7] method. Also, the bounding boxes were grouped using the k-means clustering method.

The method [8] provides better disease classification outcomes for tomato crops but at a higher cost. The author used the Convolutional Neural Network and Haar Cascade algorithm for plant leaf disease detection for good accuracy [9]. This research focuses on utilizing pre-trained deep-learning models within Orange to extract the features from spinach leaf images. A machine learning algorithm will then use these features for classification. The main scope of the work is we extract more features from the input images with the help of a pre-trained network model embedded in orange. We require data mining software in addition to image classification, which is a gadget in Orange-based Python scripting [10]. Orange is a data mining gadget that is helpful for the speculative investigation of data and graphical programming. It has numerous components which are referred to as widgets and supports Windows, Linux, and macOS [11]. The implementation of flora taxonomy originated upon artificial vision is an appetizer to the scientific community. Authors in [12] presented a Zernike polynomial- and artificial neural network-based orange classifier. Principal Component Analysis and Neural Network techniques created an apple classifier [13] based on color and texture. Authors in [14] suggested a methodology to enumerate the fruit volume from visual data. The CNN model, which is capable of categorizing regional **spinach leaf** plants, was used by the authors to create a method for classifying local spinach. Four classifiers showed that image processing of **spinach leaves** with similar appearances could be categorized and result in high output [15]. They trained the dataset utilizing several convolutional neural networks based on deep learning. then used various deep learning models to assess each one's precision. Our datasets all offered high accuracy [16].

Based on these studies, we present a novel deep learning pre-trained model and machine learning classification models in the Orange Data analytics tool to recognize and classify the spinach leaves accurately [17]. It is suggested that researchers use visual programming to combine clustering and categorization with image embedding by trained deep learning models Using a variety of AI techniques, we classified images, and then we compared the results using several parameters, including Area Under the Curve (AUC), Classification Accuracy (CA), F-1 score, precision, and recall. The Orange data mining toolkit is utilized to compare models [18]. Precision in spinach leaf classification might have significant applications in agriculture, particularly in the areas of disease diagnosis and food quality control. More importantly for city inhabitants, it provides the knowledge necessary for consumers to make informed decisions regarding the spinach they consume. Chapter 2 discusses the results of the study, whereas Chapter 1 describes the process used to classify the spinach leaves. Chapter 2 discusses the findings and their significance, while Chapter 3 concludes the study and offers recommendations for further research. The Orange workflow architecture's pre-trained and classification algorithms have identified four varieties of spinach leaves: amaranth, black nightshade, curry, drumstick, and Malabar. The collection has 500 pictures, which are divided evenly across four categories. There are one hundred images in each category. The Kaggle repository was used to obtain the photos. This is the organization of the text: This section will review the procedures followed to arrive at the spinach leaf classification scheme presented in Chapter 2. Chapter 3 will delve into the study's findings. In Chapter 4, we will discuss the findings and how important they are. The concluding chapter will provide our discussion of the study's findings and next steps.

II. METHODOLOGIES

In this section, we discuss the methodology used for **spinach leaf** classification. The workflow diagram for a classification investigation employing the Orange data analysis tool is shown in Fig. 1. The remainder of this section explains each step in the workflow diagram relevant to the proposed work.

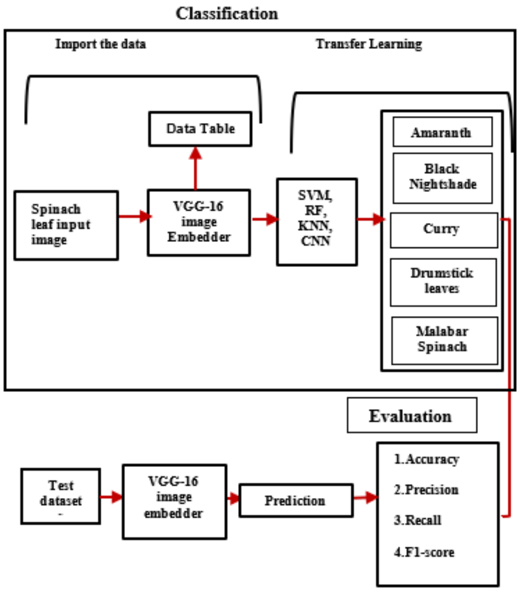


Fig. 1. Architecture diagram for proposed study for image

A. Image Dataset

The comparison is performed using images of Amaranth leaves, black nightshade, curry leaves, drumsticks, and Malabar spinach leaves from the Kaggle repository [19], which is further investigated. The dataset houses two folders, viz. training and testing. The five sub-folders for both training and testing are Amaranth leaves, black nightshade, curry leaves, drumsticks, and Malabar spinach. The training database includes 100 images for every category, and the test dataset contains 30 images. The training and test images, however, were preserved in the same location for image categorization. As a result, 500 images representing each of the five classes of spinach leaves are used in the final comparative analysis.

B. Image Importation

The image importation technique is the first and foremost stage in photo categorization utilizing the 'picture analysis' add-on and the 'Import picture' widget in the Orange data mining toolkit. Now you can access the metadata of the necessary images in the 'Data Table' widget and load them from the dataset location. You must install the 'Image analysis' add-on before you can import the images. Several types of picture analysis will be possible with this. The photos that require classification are subsequently loaded using the 'Import picture' widget. You can see the category, size, and location of the provided dataset, as well as other image details, in the data table widget. To import the photos needed for categorization, the 'Import picture' widget must be loaded before the 'Data Table' widget can access the image details [20]. For further dataset analysis, the 'Image Viewer' widget displays sample input images of curry leaves, drumsticks, curry, Malabar spinach, Amaranth leaves, black nightshade, and curry. The "Import image" widget aids in loading the required images for categorizing images. Image details are viewed in the ‘Data Table’ widget with the image size, category, and location of the input dataset. The sample input images of Amaranth leaves, black nightshade, curry leaves, drumsticks, and Malabar spinach leaves are shown in Fig. 2 with the help of the ‘Image Viewer’ widget from the image analysis add-on for our study purpose.

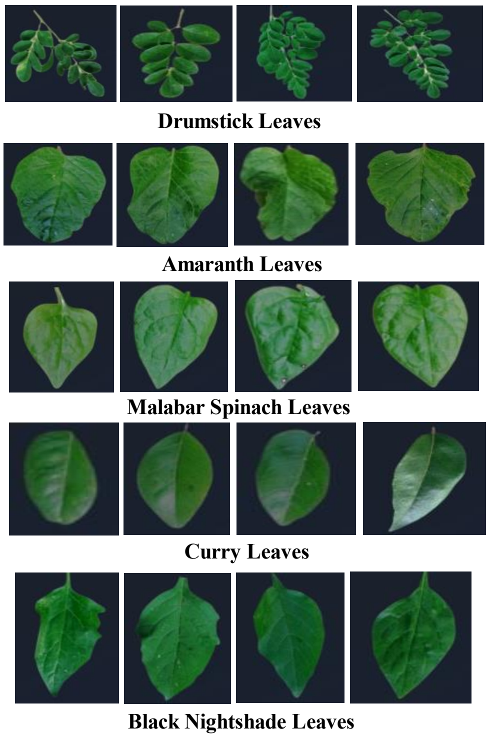


Fig. 2. Sample input spinach images for classification

C. Image Embedding

Make use of the 'Image Embedder' tool found in the Orange toolbox to extract characteristics from embedded photographs. To get the vector representation of the images, feature extraction is done by feeding them into pre-trained deep networks. The feature extraction is crucial, and Orange does it by using a multitude of embedders and uploading the created images to the server. Image feature extraction is a breeze with pre-trained Deep Convolutional Neural Network (DCNN) algorithms like VGG16. The vectors generated by these methods are of great assistance to ML algorithms. Because supervised and unsupervised machine learning classifiers, as well as deep learning classifiers, utilize the extracted features for subsequent processing, this phase is crucial.

Following feature extraction and image embedding, the hierarchical clustering widget shows attribute-based image clusters, which means class-based clusters are needed for crucial prediction and classification jobs [21]. To further analyze these embedded traits, the next step is to merge traditional machine learning with state-of-the-art deep learning techniques. For better classification results, this fusion combines the best features of k-Nearest Neighbor (kNN), neural networks (NN), support vector machines (SVMs), and Random-Forest (RF)classifiers. This widget contains some pre-trained Deep Convolutional_Neural_Network (DCNN) [22] embedder methods for component mining objectives from the visual data. As Squeeze Net [23], inception V3[24], VGG16 [25], and a few others are available for the feature extraction process. Orange integrates VGG16 as an embedder, which transforms images into suitable machine-learning algorithms. Using the VGG16 embedder, our work transmits the images to a server for feature extraction. utilizing the 4095 properties that the VGG16 network has acquired through its prior learning of object identification from significant photo samples. These features are sent to the ‘Data Table’ widget which is dragged and dropped into the workspace from the ‘Data’ add-on. VGG16 doesn’t classify the images itself. It extracts features from them. These features capture essential characteristics of the single image together with the meta-features of size, width, and height. The extracted features are converted into vectors of numbers. This vector represents the image in a way that machine learning algorithms can understand and use for tasks like classification or clustering.

In our process, depicted in Fig. 1, these image features undergo additional analysis by supervised or unsupervised machine learning classifiers and deep learning classifiers for classification. Images are displayed in their current positions, which are saved in the dataset's location, but after being embedded through an "image embedder" widget with CNN pre-trained models that the developer has selected based on the work, the images' sizes are changed for classification and prediction. Additionally, the hierarchical clustering widget shows that these images form a cluster with the same classes from the embedding.

D. Visual Analytics

The suggested method for mining images employs visual analytics [26], a combination of interactive visualizations and automatic data evaluation, involving machine learning. Orange [27] deals with all of the crucial elements of visual mechanisms for evaluation, including importing data and conversion, visualization of data involving collaboration by users, speculation of data models, and model visualization. Orange regulates data processing and visualization components, in addition to visual programming, it aids data analysts in combining and connecting various data analysis elements to create data evaluation workflows. We produced a set of visual programming tools for image analytics. Using aspects that take up imagery, including them in vector spaces, and analyzing these images to identify image categories or groups, the toolkit helps users build analysis procedures. The toolkit is through data extraction with orange, a flexible Data evaluation methodology that includes elements for grouping, categories, and dynamic data and modeling visualizations. [28][29]. In Orange, data analysis is implemented using workflows. Widget elements that process, model, or visualize data make up a workflow. Data is inputted into widgets, which then show or provide results.

Workflows for data analysis are established by the gadgets chosen and the associations made among them in Orange. In the sequence of actions shown in Fig. 1, for instance, a set of images is loaded from the specified database, they are embedded with vectors of features, the disparities among the vectors and the resulting images are estimated, and the computed distances are then used for aggregating and visual representations of image similarity [30] in the multidimensional escalating plot. Users can examine each intermediate result and keep track of how each step of the Orange workflow is being carried out. For illustration, they can look at imagery that was recently stuffed. see the dendrogram's "hierarchical clustering" widget, and even see the images that have been chosen from a particular dendrogram class before seeing them in the image viewer widget. we show the multi-dimensional scaling plot with are 'Distribution plot' widget, The distribution plot is in Fig. 4, and the Bar plot is shown in Fig. 3.

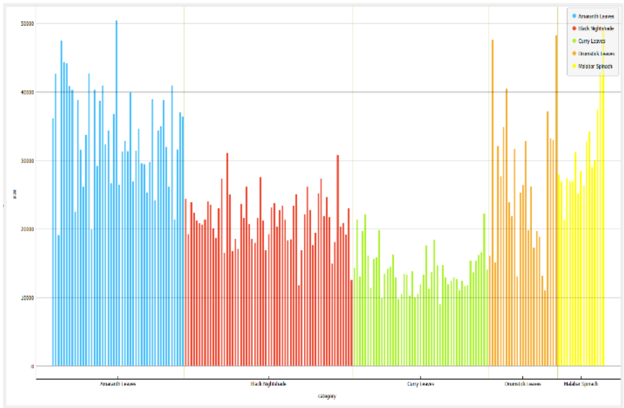


Fig. 3. Bar plot for hierarchical clustering images

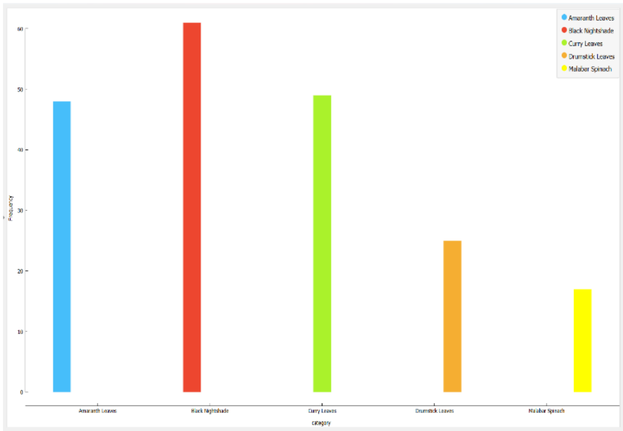


Fig. 4. Distribution chart for selected images in the image viewer

E. Image Classification Process

The approach can classify images by combining the VGG16 model with several classifiers such as SVM, RF, KNN, and NN. All of these classifiers are vital for forecasting and categorizing spinach leaf pictures effectively because they make good use of the unique qualities that the photos acquire. Some of the metrics utilized to evaluate the classifiers' performance were F1 Score, recall, precision, Area Under the Curve (AUC), and Matthews' correlation coefficient (MCC). Training and parameter variability improved performance even further, and merging model outputs improved the accuracy of learning ensembles. So far, we have trained a bunch of photo classifiers using the VGG16 model and features collected from spinach leaves [32]. The classification approach that Orange employs is based on supervised data mining and ML [33]. Learners and classifiers are the two primary types of objects involved in classification. Students want a classifier that can handle class-labeled data. There are a lot of parallels between the classification techniques employed by regression and Orange. Their origins can be traced to supervised data mining, where they were both developed from class-tagged data [34]. When many model outputs are used in a learning ensemble, their combined accuracy is increased. It is possible for different models to make use of different sets of training data samples, or for different sets of learners to be trained on the same datasets. Teachers can attract a more varied student group by adjusting their parameter sets [35]. To be anonymous in Orange, pupils merely wear ensembles [36]. They behave similarly to other students. The information they supply allows their models to make predictions about every given data instance. With 70% of the photos going into training the model and 30% into testing it, we made heavy use of Tenfold cross-validation in our categorization approach. When training and testing, every model makes use of divergent variables. In addition to outlining each model, the paper provides extensive information on the parameters whose values, after extensive experimentation, yielded the best results.

1) Support Vector Machine (SVM) Classifier: SVM is a supervised machine learning technique. Classification and regression are both handled by this supervised ML approach. It finds the best hyperplane and organizes the data into different categories [37]. It is a reliable algorithm for making predictions using a statistical learning system. Koyama et al. [38] suggested classifying spinach using ML. Using a smartphone, pictures of the spinach leaves were captured. The spinach image backgrounds were then removed by the authors. Images with the backgrounds removed are transformed into grayscale, lab, and Colour Saturation Level (CSL) images. Spinach leaves are used to extract the mean, median, and variance. Utilizing the feature detectors Binary Robust Independent Elementary Features (BRIEF) [39] and Features from Accelerated Segment Test (FAST) [40], The regional features are retrieved. The depicts that were chosen are subsequently incorporated into the SVM (Support Vector Machine) machine learning technique for categorizing the spinach leaves. In the case of the two-class dataset, CNN models have an accuracy of 84% [41]. In this work, Iterations are limited to 10, and computational appreciation is set to 0.0010 when using a linear kernel.

$$K = X.Y \quad (1)$$

Based on the linear kernel got 92.9% classification accuracy and a prediction value is 94.7% in all classes' spinach leaves using the orange data mining tool. Both F1-score and precision values are 92.8% while recall is 92.9% for training classification.

2) Random Forest (RF) Classifier: A decision tree is created by the classification algorithm RF [42] using a subset chosen at random from the training dataset. Previously Random Forest model got 85% while using different CNN models for spinach leaves classification [43]. To make predictions, an ensemble learning approach builds many decision trees during training and uses their modes. By determining the behavior of the categories from each particular tree, the result of the class becomes apparent. The number of trees to be produced in the present investigation is set at 5. In our orange model-based classification, we compare the same with other classifiers RF (Random Forest) is given 90.3% accuracy for five different spinach leaves. It had a 90.2% of the F1-score and 90.3% of both precision and recall value to training classification. When test images are embedded with training images for prediction, Random Forest (RF) has attained 86% accuracy.

3) K -Nearest Neighbour (KNN): A simple algorithm that remembers all the samples and sorts new ones using a similarity measure (such as a distance function, for instance). To categorize an image, KNN is a fundamental algorithm for classification [44] that identifies the best 'K' amount of nearby imagery from the training period. An item's closest neighbor is determined utilizing a resemblance metric. In the context of this study, the Euclidean distance is employed to compute the index, and the quantity of acquaintances is assumed to be two (K = 2). The Euclidean distance formula mentions:

$$d = \sqrt{[(X_2 - X_1)^2 + (Y_2 - Y_1)^2]} \quad (2)$$

The coordinates of one point are (X_1, Y_1). The other coordinates are (X_2, Y_2). d represents the separation between (X_1, Y_1) and (X_2, Y_2). In our spinach leaves classification, KNN has achieved 93.4% accuracy in classification at training. It has 93.4% of F1-score and recall with precision value is 93.5% in spinach leaves classification. The prediction value of KNN is 95.3% for all classes in spinach leaves.

4) Neural Network (NN): An algorithm based on the structure of interconnected neurons found in biological brain networks. Neural networks have the ability to learn and execute a variety of tasks, including pattern recognition, classification, and regression. CNN is a particular type of Neural_Network (NN) that enables the input pictures to pull out the basic amenities in a manner that is understandable to humans for better classification [45]. The first benefit of the CNN model is its ability to identify crucial components in images without human assistance. Consider the case where the dataset that the model learns various characteristics for every class category contains both dogs and cats [46]. Numerous pre-trained and CNN models were presented by previous researchers, However, the ability to accurately classify spinach leaves could still use some work [47]. Orange's data mining tool's Neural Network widget provides a full-stack environment for NN model construction, training, and evaluation. Users can design the neural network's architecture by changing the hidden-layer count, neuronal density, activation function, learning rate, and optimization technique. The widget provides backpropagation-based model training on input datasets, cross-validation, and independent validation dataset evaluation. An interactive representation of the neural network's architecture helps users understand the model's structure and optimize its performance via hyperparameter adjustment. Traditional model evaluation metrics like recall, accuracy, precision, F1, score, ROC-AUC curve, and feature importance analysis are available. For tasks that require considerable machine learning, Orange's Neural Network feature is invaluable. It simplifies neural network modeling for classification, regression, pattern recognition, and more. In a manner akin to how brain neuron's function, the Artificial Neural Network (ANN) algorithm for categorization [48] recognizes the hidden relationship between images. A straightforward Neural Network has three layers: an entry layer that receives amenities as input, a covert layer, and a final layer that identifies the group that a given image is owned. The disease-affected and unaffected leaves are divided using ANN (Artificial Neural Network). Additionally, it has provided greater accuracy while requiring less computation time. In this study, 150 hidden layers are assumed, the ReLU function [49] is used as the activation function, Adam [50] is utilized as the optimizer, and 500 iterations are assumed. Using the activation function, it is possible to normalize the output in each layer. We used the Rectified Linear Unit (ReLU) for the activation function. If the input value is less than 0, ReLU will set the value to zero. If not, the raw value of the input is the outcome of the relevant layer. Equation (3) is utilized to Figure out the ReLU activation function.

$$ReLU(P) = \begin{cases} 0 \; if \; P < 0 \\ P \; if \; P \geq 0 \end{cases} \quad (3)$$

In which P displays the source imagery for every layer. Instead of SGD, the Adam optimizer is used to iteratively update network weights based on training data. We achieved a high classification accuracy of 96.6% compared with other models in training for spinach leaf classification. F1-score, precision, and recall all so gave 96.6% in training classification. In prediction, the Neural Network has 96% accuracy than other models.

F. Test and Score with Confusion Matrix:

The 'Test and Score' widget takes all the classifiers as input and returns correlated values according to CA, recall, F1-score, AUC, and MCC. The data sequence is passed to a cross-validation device (Test and Score) using vector-based embedding. This device takes an additional input from a machine-learning technique (SVM, KNN, RF, or NN). In the Test alongside Score widget, you can see the AUC, CA, and F1 score of the cross-validated accuracy (the harmonic average of the precision and recall) as well as the evaluation findings for the ambiguity matrix gadget. Misclassification data is presented using the "Confusion Matrix" widget. Make use of Receiver Operating Characteristic Analysis (ROC) to assess how well a categorization model is doing. It shows how changing the discriminating threshold of a binary classifier system affects its diagnostic capacity.

The F1-Score evaluates a model's performance by taking recall and precision into account. Classification difficulties involving unequal classes become much easier to handle with its assistance. Classification accuracy (CA) is the ratio of correctly classified examples to total instances in a classification model [51]. A popular and easy-to-understand measure for gauging the categorization model's overall accuracy. The accuracy with which the model's favorable predictions remain valid is called precision. This metric is derived by dividing the total number of positive and negative cases by the number of genuine positive ones. The model is probably accurate when it accurately predicts a positive outcome. As a measure of how well a model detects positive cases, recall (also called sensitivity) is important. This statistic is determined by dividing the total number of positive results by the total number of negative results [52]. In general, a higher recall indicates that the model correctly identified more instances when tested. As an indicator of a binary classification model's performance, the area under the curve (AUC) can be calculated.

For various classification model thresholds, the area under the Receiver Operating Characteristic (ROC) curve shows the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity). The Matthews Correlation Coefficient (MCC) measures the accuracy of binary classifications by taking into account the total number of correct and incorrect replies. A perfect prediction would have a score of 1, a prediction that is on par with chance would have a score of 0, and a score of -1 would indicate a complete mismatch [53]. The image viewer and bar plot widget can display the classified spinach leaves using the confusion matrix. What follows is an analysis of the results in great detail.

III. Analysis of the Research Results

A. *Result Analysis for Classification*

In this article, a total of 500 images of Amaranth leaves, Malabar spinach, Drumsticks leaves, Curry leaves, and black nightshade are taken for comparative analysis. These input data are divided into Test to training proportion: 70:30. To train the models with four different classifiers for the classification of spinach, we used tenfold cross-validation, which performed the best on our dataset. Utilizing metrics like AUC, CA, F1 score, Precision, Recall, and MCC, the models are contrasted. Orange data analytics tools' 'Test and Score' widget is used to calculate these parameters [54]. The accuracy of the predictions made by the algorithm during training is measured by the Area Under Curve (AUC). The correctly classified training models' total quantity of images is an estimate of classification accuracy [55].

The Orange data mining app has a Test and Score widget for classifier and prediction model evaluation. It gives a complete set of variables for generalizability, robustness, and prediction accuracy for the ML model. Model evaluation, evaluation metrics, cross-validation support, result display, parameter optimization and tuning, and test and score widget explanations are some of the main components. Orange's Test and Score widget is typically used when connecting a prediction model (classifier, regressor, etc.) to its input port. They can then choose assessment metrics and cross-validation parameters to customize the widget. After the widget runs, users can view performance scores and visuals to evaluate their models. Researchers, data scientists, and machine learning practitioners employing Orange data mining need the Test and Score widget. It provides a structured and informative framework for accurate prediction model testing. Recall produces the real positive values that have been properly recognized, while precision makes the percentage of positive values that are truly positive. The concordant sign of recall and precision is provided by the F-1 score. The crucial element in this comparative study is the classifiers' accuracy. In Table I, Neural Network gives the highest accuracy of 96.6% whereas Random Forest has the lower accuracy of 90.3%. SVM (Support Vector Machine) and KNN (K-Nearest Neighbour) classifiers are given 92.9% and 93.4% accuracy in the training of spinach classification. Based on the Test and Score values, spinach leaves are classified using four models and shown in Fig. 5 using the image viewer widget.

TABLE I. Quantities for Each Classifier's Measuring Factors

Model	AUC	CA	F1-Score	Precision	Recall	MCC
RF	98.8	90.3	90.2	90.3	90.3	87.9
KNN	98.2	93.4	93.4	93.5	93.4	91.8
Neural Network	99.9	96.6	96.6	96.6	96.6	95.7
SVM	99.6	92.9	92.8	92.8	92.9	91.1

A 'ROC curve' widget is a graph that displays how well an algorithm for classification performs across all classification boundaries. The TPR and FPR parameters are plotted on this curve. We took a threshold value of 0.5 for ROC analysis based on Test and score values for spinach training classification using four different classifiers which is shown in Fig. 6. So compared to all the classifiers, the Neural Network classifier has high sensitivity by using the below equations,

True_Positive_Rate is specified as:

$$T\_P\_R = \frac{TP}{TP + FN} \tag{4}$$

is also called sensitivity.

False_Positive_Rate is specified as:

$$F\_P\_R = \frac{FP}{FP + TN} \tag{5}$$

is also called specificity.

Whereas TP is True_Positive, TN is True_Negative, FP is False_Positive and FN is False-Negative. Confusion matrix is used to find errors in the spinach classification problem. A confusion matrix contains True_Positive (TP), True_Negative (TN), False_Positive (FP) and False_Negative (FN). The simple symbols TP and TN represent the strategy prediction that all members of the positive category will be optimistic and all members of the unfavorable category will be unfavorable. By the aforementioned performance evaluation, we calculate the spinach classification precision, which is given in equation 6. According to the accuracy, the scenario categorizes all optimistic classes as positive and all negative classes as negative.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Because of its proficiency, the model does not categorize an incorrect value as optimistic. The computation is performed using equation (7).

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

The model categorizes every positive value as positive, according to recall. It is calculated by given equation (8).

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

The accuracy and recall average values are represented by the F1-score. It is calculated by given equation (9).

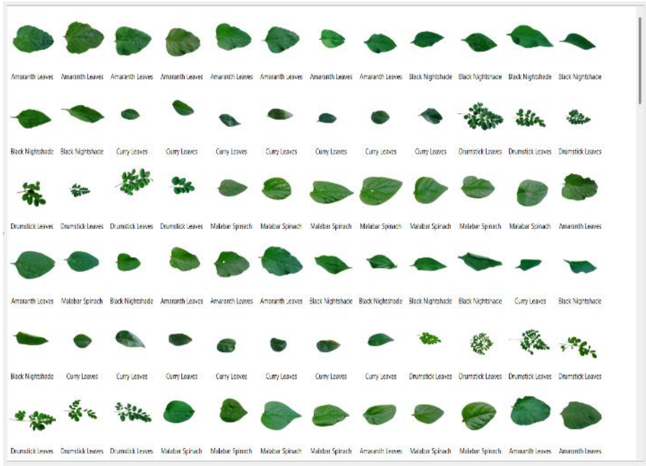
$$F1 - Score = \left(2X \left[\frac{Precision \times Recall}{Precision + Recall}\right]\right) \quad (9)$$

The "Confusion Matrix" gadget of the Orange toolkit creates a matrix of dissonance for every classifier. The model might unintentionally predict from the negative class or make predictions from the positive class that is incorrect. The non-diagonal matrix elements produce the incorrect classification value when classifying the images, while the Diagonal elements produce the correct classification value. Table II depicts the confusion matrix for all of the categorization prototypes utilized to group the spinach leaves. Using the KNN (K-Nearest Neighbour) model in Classification Drumstick Leaves had 100% accuracy while curry leaves got 98.5%. But Amaranth leaves, Black Nightshade, and Malabar Spinach got low accuracy are 89.4%, 88%, and 91.5% shown in Table II.

TABLE II. CONFUSION MATRIX FOR ALL CLASSES USING KNN

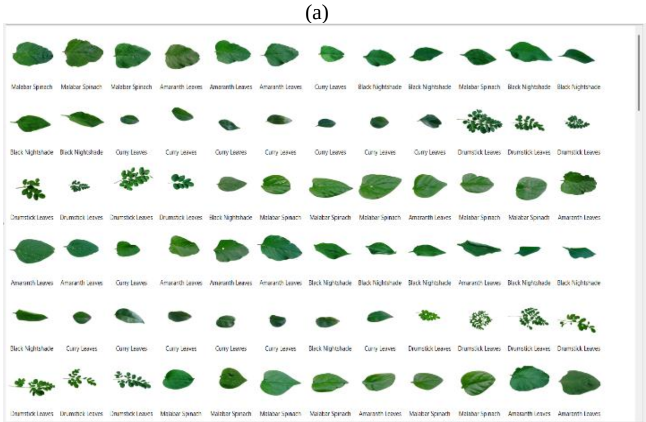
		Predicted					
		AL	**BN**	**CL**	**DL**	**MSL**	∑
Actual	**AL**	89.4%	6.7%	0%	0%	8.5%	70
	BN	4.5%	88%	1.5%	0	0	70
	CL	0	4.0%	98.5	0	0	70
	DL	0	0	0	100%	0	70
	MSL	6.1%	1.3%	0	0	91.5%	70
	∑	66	75	68	70	71	350

(**AL**- Amaranth Leaves, **BN**- Black Nightshade, **CL**- Curry Leaves, **DL**- Drumstick Leaves, **MSL**- Malabar Spinach Leaves).

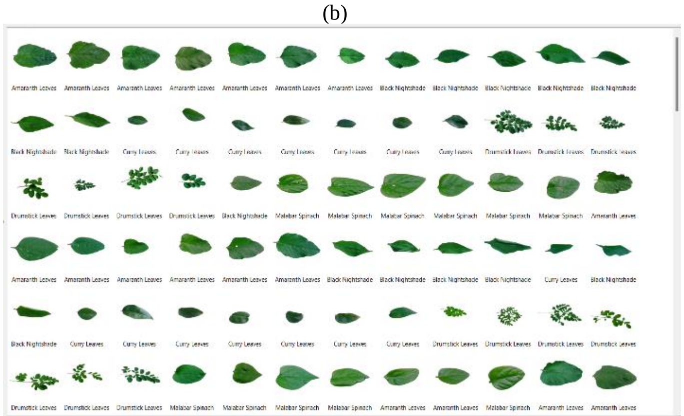


(a)

(b)

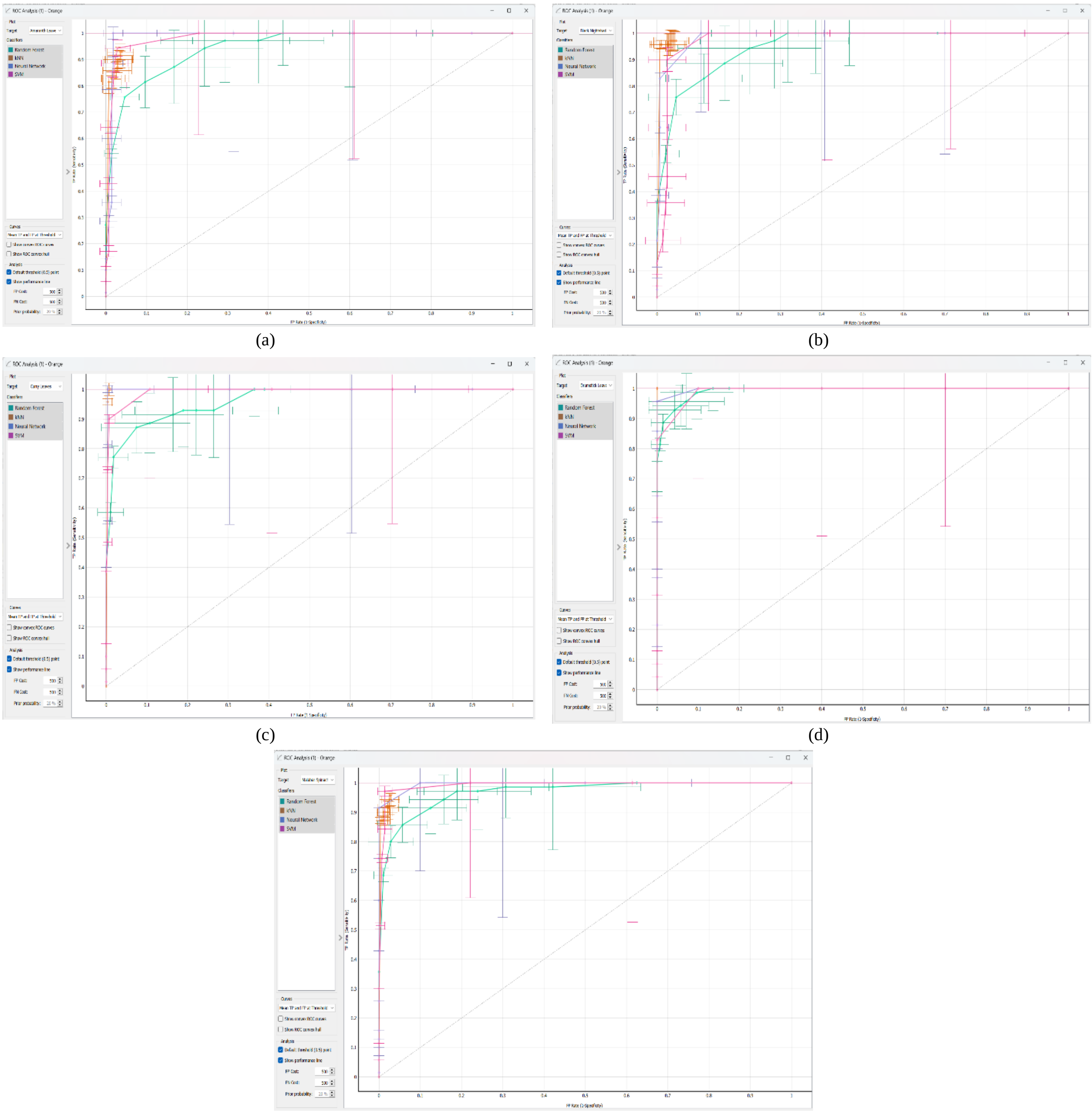


(c)



(d)

Fig. 5. (a), (b), (c) Test and Score classified images using the SVM classifier, (d)Test and Score classified images using Neural Network



(a) (b) (c) (d) (e)

Fig. 6. (a) ROC analysis using Test and Score for Amaranth leaves, (b) ROC analysis using Test and Score for Black Nightshade, (c) ROC analysis using Test and Score for Curry Leaves, (d) ROC analysis using Test and Score for Drumsticks leaves, (e) ROC analysis using Test and Score for Malabar Spinach

In Table III, the Neural Network model is given high accuracy classification to Drumstick leaves at 98.6% and Curry leaves at 97.2% while Amaranth leaves got a low accuracy of 94.2% in classification. Using the RF (Random Forest) classifier Drumstick leaves got a high accuracy of 98.6% and Amaranth leaves attained a low accuracy of 83.3% as shown in Table IV. The SVM (Support Vector Machine) classifier gives high accuracy to Drumstick leaves at 100% and Curry leaves at 95.8% other leaves to have low accuracy are 94.3%, 87.9%, and 86.3% in training classification (Table V).

TABLE III. CONFUSION MATRIX FOR ALL CLASSES USING NN

		Predicted					
		AL	**BN**	**CL**	**DL**	**MSL**	∑
Actual	**AL**	94.2%	2.9%	0%	1.4%	2.8%	70
	BN	2.9%	97.1%	2.8%	0	0	70
	CL	0	0	97.2%	0	1.4%	70
	DL	0	0	0	98.6%	0	70
	MSL	2.9%	0	0	0	95.8%	70
	∑	69	68	71	71	71	350

(**AL**- Amaranth Leaves, **BN**- Black Nightshade, **CL**- Curry Leaves, **DL**- Drumstick Leaves, **MSL**- Malabar Spinach Leaves).

TABLE IV. CONFUSION MATRIX FOR ALL CLASSES USING RF

		Predicted					
		AL	**BN**	**CL**	**DL**	**MSL**	∑
Actual	**AL**	83.3%	6.5%	1.4%	1.4%	11.8%	70
	BN	3.0%	85.7%	2.9%	0	0	70
	CL	1.5%	3.9%	95.7%	0	0	70
	DL	0	1.3%	0	98.6%	0	70
	MSL	12.1%	2.6%	0	0	88.2%	70
	∑	66	77	69	70	68	350

(**AL**- Amaranth Leaves, **BN**- Black Nightshade, **CL**- Curry Leaves, **DL**- Drumstick Leaves, **MSL**- Malabar Spinach Leaves).

TABLE V. CONFUSION MATRIX FOR ALL CLASSES USING SVM

		Predicted					
		AL	**BN**	**CL**	**DL**	**MSL**	∑
Actual	**AL**	87.9%	2.9%	0	0	13.7%	70
	BN	1.5%	94.3%	4.2%	0	0	70
	CL	1.5%	1.4%	95.8%	0	0	70
	DL	0	0	0	100	0	70
	MSL	9.1%	1.4%	0	0	86.3%	70
	∑	66	70	71	70	73	350

(**AL**- Amaranth Leaves, **BN**- Black Nightshade, **CL**- Curry Leaves, **DL**- Drumstick Leaves, **MSL**- Malabar Spinach Leaves).

Based on the confusion matrix for correctly classified and misclassified instances on spinach leaves which is shown on the image viewer widget in Table VI. Different types of spinach leaves were correctly categorized by employing the following algorithms: K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), and Neural Network (NN): At 65 amaranth leaves, 66 black nightshade, 67 curry leaves, 70 drumsticks, and 65 Malabar spinach were correctly identified using the KNN algorithm. One simple and effective method for classifying objects is the K-Nearest Neighbors (KNN) algorithm. When applied to the different Spinach leaves, KNN demonstrated a moderate level of accuracy in this case. On the other hand, the Random Forest (RF) algorithm successfully classified 55 amaranth leaves, 66 Black nightshade, 66 curry leaves, 69 drumsticks, and 60 Malabar spinach. Reduce overfitting and manage data with complex relationships with the help of RF, an ensemble learning technique. When it comes to classifying spinach leaves, the RF model was just as accurate as KNN. Support Vector Machine (SVM) produced somewhat better results than its forerunners in detecting 58 amaranth leaves, 66 black nightshades, 68 curry leaves, 70 drumstick leaves, and 63 Malabar spinach. When it comes to high-dimensional data and identifying the best hyperplane to separate classes, support vector machines (SVMs) shine. The bulk of spinach leaves were better handled by SVM than by KNN and RF, the other two models that were examined. Last but not least, the Neural Network (NN) model completed the most successful leaf classifications of all the algorithms: 68 Malabar spinach, 65 Amaranth leaves, 66 black nightshades, 69 Curry leaves, and 70 Drumstick. One subfield of deep learning, neural networks (NNs) teach themselves complex patterns by sifting through data using vast networks of linked neurons. When it came to accurately classifying the plant species, NN scored best, demonstrating its ability to understand complicated relationships in data as shown in Fig. 7. When it comes to classifying plant species, KNN, RF, and SVM all did decent jobs, but the Neural Network model stood out in every regard. The specific limitations of the classification task, including computational efficiency, dataset complexity, and interpretability, could determine the approach that is chosen.

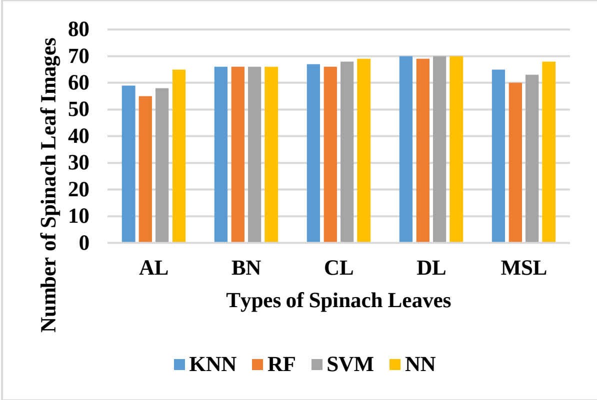


Fig. 7. Comparison of spinach leaf classification

TABLE VI. NUMBER OF CORRECTLY CLASSIFIED IMAGES

		Methods				
		KNN	**RF**	**SVM**	**NN**	**Total**
Types of Spinach Leaves	**AL**	59	55	58	65	237
	BN	66	66	66	66	264
	CL	67	66	68	69	270
	DL	70	69	70	70	279
	MSL	65	60	63	68	256
	∑	327	316	325	338	1306

The K-Nearest Neighbors (KNN) algorithm successfully detected 327 images of spinach leaves. For spinach leaves, the Support Vector Machine (SVM) achieved 325 correct classifications, but Random Forest (RF) achieved 316. The Neural Network (NN) model outperformed all other tested models with 338 correct classifications for images of spinach leaves. The percentage of people who were able to accurately identify plant species also varied. The following leaf types were accurately identified: amaranth (237), black nightshade (264), curry (270), drumstick (279), and Malabar spinach (256). The highest number of plant species successfully identified was Drumstick leaves (279), indicating that their classification performance was rather constant across all methods are shown in Fig. 8. This finding is noteworthy. In conclusion, the Neural Network (NN) outperformed the other approaches in terms of the total number of right classifications for images of spinach leaves. Although the algorithms' exact plant species designations varied, Drumstick leaves consistently ranked first in both accuracy and quantity of correct classifications.

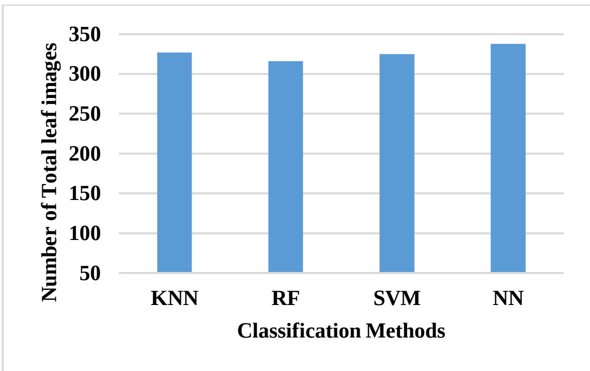


Fig. 8. Classification performance of methodologies

Among the many leaves misclassified by the KNN algorithm were those of 7 amaranth plants, 9 black nightshade plants, 1 curry plant, zero drumstick plants, and 6 Malabar spinach plants. In 9 instances of Malabar spinach, 11 of black nightshade, 3 of curry, one of drumstick, and 11 of amaranth leaves were erroneously identified by the RF algorithm. According to the results of the support vector machine analysis, there were 10 cases of Malabar spinach, 8 instances of amaranth leaves, 4 cases of black nightshade, 3 cases of curry leaves, and zero cases of drumstick leaves that were incorrectly recognized shown in Table VII. Finally, the NN model erred four times when it misidentified amaranth leaves, two times when it misidentified curry, once when it got drumstick wrong, and three times when it got Malabar spinach wrong. Results show that the algorithms' misclassification rates varied between plant kinds. When compared to other models, the NN model reduced the number of misclassifications among all plant species are shown in Fig. 9. Despite often variable results, SVM outperformed KNN and RF in drumstick leaf classification concerning accuracy and error-free classification. On the other hand, they were far more likely to identify most plant species mistakenly.

TABLE VII. NUMBER OF WRONGLY MISCLASSIFIED IMAGES

		Methods				
Types of Spinach Leaves		KNN	RF	SVM	NN	Total
	AL	7	11	8	4	30
	BN	9	11	4	2	26
	CL	1	3	3	2	9
	DL	0	1	0	1	2
	MSL	6	8	10	3	27
	∑	23	34	25	12	94

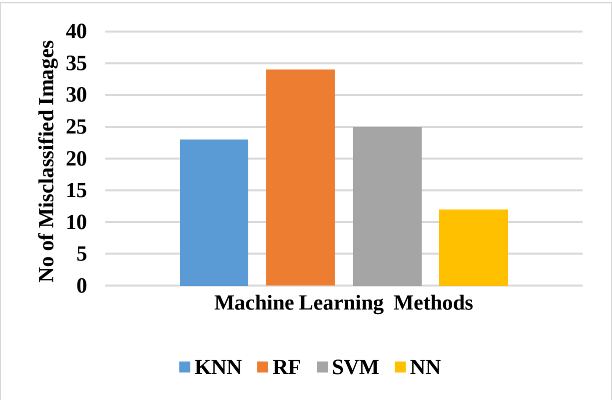


Fig. 9. Comparison of models misclassification spinach leaf images

The KNN algorithm misclassified 23 spinach leaf varieties. SVM and RF algorithms misclassified 25 and 34 spinach leaf pictures, respectively. Only 12 spinach leaf pictures were misclassified by the NN model. Each algorithm misclassified plant species differently. The most misclassified leaves were amaranth (30), black nightshade (26), curry (9), drumstick (2), and Malabar spinach (27). Only two Drumstick leaves were misclassified by all algorithms, showing that this species performed well. Although the Neural Network (NN) model had the fewest false answers for spinach leaves, other techniques generated varied percentages of wrong answers for plant species photos. Only 2 Drumstick leaf occurrences were wrongly recognized across all algorithms, proving its consistent categorization performance. This analysis indicates that misclassification rates and plant species identification accuracy are significant when comparing algorithms. The NN model outperformed the other algorithms in accuracy and misclassifications when detecting spinach leaf photos.

B. *Result Analysis for Prediction*

"Prediction" refers to an outcome of an approach that has been trained on actual data and executed on new data, when predicting the likelihood of a specific outcome [64]. In contemporary society, the development of resources and innovations is speeding up every day. Artificial Neural Networks (ANN) are the foundation of the Bayesian Enhanced Approach (BEA) predictive method for unpredictable enduring combinations of events [56][57]. The illustration, which relates to Artificial Neural Networks, was created employing a Python script that utilizes the employ of the Pytorch library. Using a Random Forest method for disease detection has been proposed. It performs better than DT (Decision Tree) and SVM (Support Vector Machine) when using the same classifier on the same dataset. With an accuracy of 80.56%, DT (Decision Tree) and SVM (Support Vector Machine) methods of classification are outperformed by it [58]. The survey found a research gap in the area of disease detection because there have been many methods proposed, but most of them are less accurate than others [59]. Google-Colab's testing for accuracy, precision, recall, and F1-score to leaf disease detection by using Logistic Regression and XGBoost models were got 94.89% and 79.37% [60] In this research analysis we took 30% of testing data through the 'import image 'widget. These images are passed to the 'Image Embedding' widget so we took the VGG16 embedder with 16-layer image recognition models trained on ImageNet to extract the features of test images. Then we evaluated the prediction values by combining the classified images of the training set using SVM, RF, KNN, and Neural Network classification models and feature-extracted test images through the 'Prediction' widget of the orange data analysis tool [61][62].

The evaluation metrics that consider the total amount to which the models can categorize images of spinach leaves are recall, precision, F1 Score, and area under the curve (AUC). The models' accuracy across different classification thresholds (AUC), precision in detecting positive values (precision), and recall (recall) in locating all relevant occurrences (recall) can be understood by adding together these metrics. By assessing the classification models across several parameters, the different assessment metrics guarantee a thorough comprehension of their performance. Considering the intricacies of recall, precision, and the balance between the two (F1 Score), these accuracy metrics offer a thorough assessment of the models' capacity to correctly categorize images of spinach leaves[63]. Because they show us how effectively the algorithms distinguish between positive and negative instances, these metrics are critical for completely validating the classification models are shown in Table VIII.

TABLE VIII. PREDICTED VALUES WITH ERROR LOSS FOR ALL CLASSES

Spinach leaves	Error rate			
	Random Forest	SVM	Neural Network	KNN
Malabar Spinach	0.081	0.061	0.001	0.277
Drumsticks leaves	0.086	0.019	0.001	0.000
Curry Leaves	0.045	0.031	0.001	0.498
Black Nightshade	0.100	0.007	0.007	1.000
Amarantha leaves	0.005	0.088	0.008	1.000

TABLE IX. PREDICTION ACCURACY OF ALL CLASSES OF SPINACH LEAVES

Model	AUC	CA	F1-Score	Precision	Recall	MCC
Random Forest	97.90	86	86	87.60	86	82.90
SVM	99.80	94.70	94.60	95	94.70	93.40
Neural Network	99.90	96	96	96.20	96	95.10
KNN	97.60	95.30	95.30	95.80	95.30	94.30

The performance score of all Parameters is shown in Table VIII and Table IX, using the ‘Data Table’ widget to display them correctly and wrongly predicted images with labels using different classifiers. In the prediction, using different classifiers to predict the images with labels in the above Table IX. Neural Network classifier had a high accuracy of 96% and Random Forest had a low Prediction accuracy of 86% in spinach leaves prediction. KNN classifier achieved 95.3% and SVM got 94.7% accuracy in prediction. At the same time, the F1 score of Neural Network is 96% and Random Forest is 86% in this spinach prediction study. Precision and Recall values of Neural Networks are 96.2% and 96% which is high compared with other models. We evaluated the prediction result through the confusion matrix with parameters [65] of TP, TN, FP, and FN of predicted values of spinach leaves. In this confusion matrix of predicted values diagonal matrix produces correct predicted values and the non-diagonal matrix produces wrongly predicted values of spinach leaves. It is shown in Table X to Table XIII.

In the confusion matrix (Table X, Table XI, Table XII, Table XIII), curry leaves, Drumstick leaves and Malabar leaves have highly correctly predicted values of 100% compared with other leaves using the KNN model. In comparison, Black Nightshade and amaranth leaves have low correctly predicted values of 87.9% and 90.9%. At the same time, Curry leaves and Malabar spinach had highly wrongly predicted values of 3% and 9.1% while other leaves were 100% correctly predicted. In the confusion matrix, Malabar Spinach, Drumstick leaves and Curry leaves had highly correctly predicted values of 100% compared with other leaves using the Neural Network model. In comparison, Amaranth Leaves and Black nightshade leaves had low correctly predicted values of 90.9% and 90.3% as shown in Table XI. The Random Forest model exhibited an accuracy rate of 63.4% for Amaranth leaves, 93.1% for Black Nightshade, 100% for Curry leaves, and 86.4% for Malabar Spinach leaves. On the other hand, the SVM model got 90.6% accuracy for Amaranth, 88.2% for Black Nightshade, 100% for Curry, and 96.2% for Malabar Spinach. These results demonstrate that the SVM model outperformed the Random Forest model in predicting the types of leaves, except for Black Nightshade is shown in Table XIII. It is worth mentioning that both models accurately predicted Curry and Drumstick leaves, which may suggest that these two types possess distinct and distinctive characteristics.

TABLE X. CONFUSION MATRIX FOR PREDICTED VALUES USING KNN

		Predicted					
		AL	BN	CL	DL	MSL	∑
Actual	**AL**	90.9%	0%	0	0	0	30
	BN	3.0%	87.9%	0	0	0	30
	CL	0	3.0%	100%	0	0	30
	DL	0	0	0	100%	0	30
	MSL	6.1%	9.1%	0	0	100%	30
	∑	33	33	29	30	25	150

TABLE XI. CONFUSION MATRIX FOR PREDICTED VALUES USING NN

		Predicted					
		AL	BN	CL	DL	MSL	∑
Actual	**AL**	90.9%	0%	0	0	0	30
	BN	6.1%	90.3%	0	0	0	30
	CL	0	3.2%	100%	0	0	30
	DL	0	0	0	100%	0	30
	MSL	3.0%	6.5%	0	0	100%	30
	∑	33	31	29	30	27	150

TABLE XII. CONFUSION MATRIX FOR PREDICTED VALUES USING RF

		Predicted					
		AL	BN	CL	DL	MSL	∑
Actual	**AL**	63.4%	3.4%	0	0	13.6%	30
	BN	7.3%	93.1%	0	0	0	30
	CL	4.9%	3.4%	93.1%	0	0	30
	DL	0	0	3.4%	100%	0	30
	MSL	24.4%	0	3.4%	0	86.4%	30
	∑	41	29	29	29	22	150

TABLE XIII. CONFUSION MATRIX FOR PREDICTED VALUES USING SVM

		Predicted					
		AL	BN	CL	DL	MSL	∑
Actual	**AL**	90.6%	0	0	0	3.8%	30
	BN	0	88.2%	0	0	0	30
	CL	0	5.9%	100%	0	0	30
	DL	0	0	0	100%	0	30
	MSL	9.4%	5.9%	0	0	96.2%	30
	∑	32	34	28	30	26	150

At the same time, Black Nightshade had highly wrongly predicted values of 6.1% and 6.5% while Malabar spinach had low wrongly predicted values of 3% are shown in Table X and Table XI. The comparative analysis of all models in each spinach leaf prediction accuracy is shown in Table XIV.

TABLE XIV. COMPARISON OF MODELS' PREDICTION ACCURACY

Spinach leaves types/ Models	KNN	RF	SVM	NN
Amarantha Leaves	90.9%	63.4%	90.6%	90.9%
Black Nightshade	87.9%	93.1%	88.2%	90.3%
Curry Leaves	100%	93.1%	100%	100%
Drumstick Leaves	100%	100%	100%	100%
Malabar Spinach Leaves	100%	86.4%	86.4%	100%

The chart clearly shows in Fig. 10 that the KNN and NN models accurately predicted the emergence of Curry, Drumstick, and Malabar Spinach leaves. Furthermore, KNN achieved an impressive 90.9% accuracy rate when predicting Amaranth leaves. Also, the SVM model did a fantastic job; it predicted Curry and Drumstick leaves with a perfect score of 100% and Malabar Spinach with a score of 96.2%. However, it was 88.2% accurate for black nightshade leaves and just 90.6% accurate for amaranth leaves. Despite achieving 100% accuracy for Drumstick leaf predictions, the RF (Random Forest) model was the least accurate with a total accuracy of 63.4% for Amaranth leaf projections. Generally speaking, the SVM model performed adequately when it came to forecasting Curry, Drumstick, and Malabar Spinach leaves. On the other hand, the KNN and NN models showed consistency in producing the best results for all leaf kinds. When compared to its rivals, the RF model was much behind.

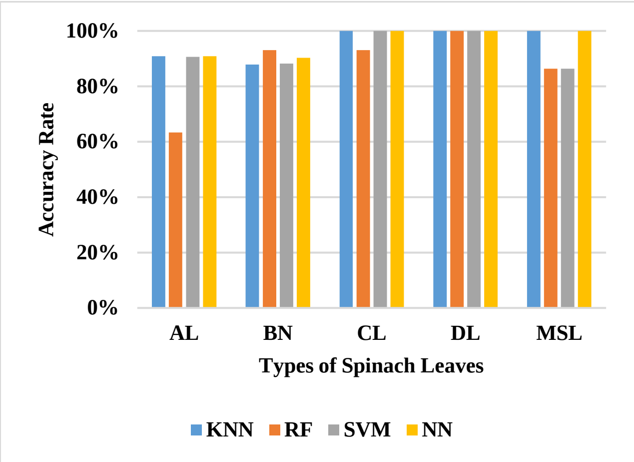


Fig. 10. Comparative analysis of prediction accuracy

When looking at Table XV how well do different machine learning models predict different kinds of leaves, it's clear that they're not all created equal. Findings demonstrated that the K-Nearest Neighbors (KNN) model accurately classified 29 black nightshade plants, 30 drumstick plants, 25 Malabar spinach plants, and 30 amaranth plants. However, out of all the models tested, the Random Forest (RF) model identified 20 cases of Malabar spinach, 29 cases of drumstick, 25 cases of curry, and 28 cases of black nightshades accurately. The SVM model accurately predicted the outcome with 29 amaranth leaves, 30 black nightshades, 28 curry, 30 drumsticks, and 25 Malabar spinach. In the end, the NN model got 31 amaranth leaves, 28 black nightshade leaves, 29 curry leaves, 30 drumstick leaves, and 27 Malabar spinach correctly. When it comes to predicting the different kinds of leaves, every model has its advantages and disadvantages. While the RF model performed better with curry and drumstick leaves, the KNN model flopped with Malabar spinach. The SVM model maintained its high level of performance regardless of the type of Spinach leaf as shown in Fig. 11. While the NN model performed poorly on Malabar spinach and black nightshade, it excelled on drumstick and amaranth leaves.

TABLE XV. PREDICTION OF THE NUMBER OF SPINACH LEAVES

Spinach leaves types/ Models	KNN	RF	SVM	NN	Total
Amaranth Leaves	30	28	29	30	117
Black Nightshade	29	27	30	28	114
Curry Leaves	29	25	28	29	111
Drumstick Leaves	30	29	30	30	119
Malabar Spinach Leaves	25	20	25	27	97
∑	143	129	142	144	558

By correctly predicting 143 images of spinach leaves, the K-Nearest Neighbors (KNN) model showcased remarkable performance in classification. With 142 accurate predictions of spinach leaves, the Support Vector Machine (SVM) model came in second place. The Random Forest (RF) model achieved remarkable performance with 129 spinach leaf photographs accurately predicted, while the Neural Network (NN) model achieved the same feat with 144 images. The amount of correctly anticipated images changes when looking at different types of leaves. The KNN model accurately predicted 176 different kinds of leaves, including 117 Amaranth, 114 Black Nightshade, 111 Curry, 119 Drumstick, and 97 Malabar spinach.

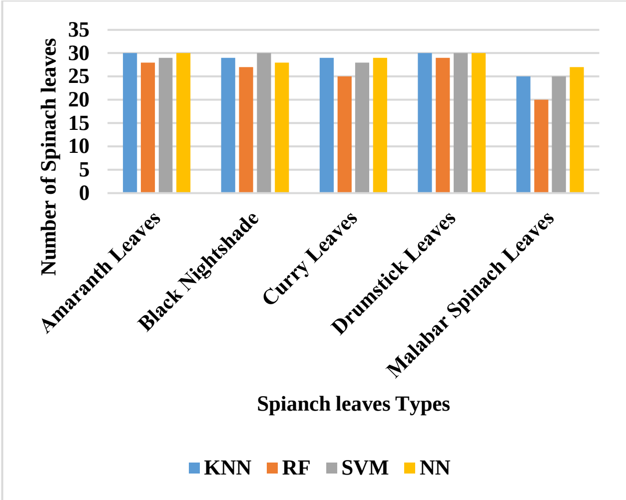


Fig. 11. Comparison of prediction performance of models

The NN model's exceptional performance in accurately predicting 119 drumstick leaves demonstrated its exceptional ability to classify this specific type of leaf. The NN model outperformed all others when it came to predicting images of spinach leaves. It may be difficult to select the optimal model due to the trade-offs between category accuracy and the type of leaf being recognized. It is crucial to find the optimal model for the task, and the outcomes demonstrate that all of the models perform adequately when it comes to forecasting images of spinach leaves.

The accuracy and misclassification rates of the machine learning models differed when tested on various kinds of leaves. Aside from missing Malabar spinach leaves, three kinds of amaranth leaves, four types of black nightshade leaves, curry leaves, and drumstick leaves, the K-Nearest Neighbors (KNN) model got 3 types of amaranth leaves wrong as shown in Table XVI.

TABLE XVI. WRONGLY PREDICTION OF THE NUMBER OF SPINACH LEAVES

		Methods				
Spinach leaves types		KNN	RF	SVM	NN	Total
	Amarantha Leaves	3	11	3	3	20
	Black Nightshade	4	6	4	3	17
	Curry Leaves	0	3	0	0	3
	Drumstick Leaves	0	1	0	0	1
	Malabar Spinach Leaves	0	0	1	0	1
	∑	7	21	8	6	42

On the other hand, the RF model failed miserably at identifying any of the following: (3) curry leaves, (11) amaranth leaves, (6) black nightshade leaves, and (5) drumstick leaves. The SVM model got three amaranth leaves wrong, 4 Black Nightshade right, zero curry wrong, zero drumstick wrong, and one Malabar spinach wrong. The NN model got 3 Amaranth wrong, 3 Black Nightshade wrong, 0 curry wrong, 0 drumstick wrong, and 0 Malabar spinach wrong. The KNN model made a mistake seven times, the RF model twenty-one times, the SVM model eight times, and the NN model six times when it came to spinach leaf shots. By breaking the results down by leaf type, we can see that the model's accuracy in classifying particular leaves varied. Twenty times in one experiment, KNN, eleven times in another, three times in the third, and three times in the second mislabeled amaranth leaves. In addition, NN was shown to be responsible for three misclassifications, KNN for 17, RF for 6, and SVM for 4. A ROC curve [67] performance of the prediction model at all classification thresholds. The prediction performance Curve is shown in Fig. 12(a)-(e) with comparisons of all models.

The Receiver Operating Characteristic (ROC) diagram is a visual depiction of the performance of a binary classification system. The trade-off between sensitivity (the actual positive rate) and specificity (the true negative rate) is shown at different threshold levels. Above, you can see a ROC diagram showing the performance curves of several classifiers. Curves like these encompass Neural Networks, Random Forests, SVM, and KNN. The graphic shows convex ROC curves, which provide insight into the classification performance of each model. Furthermore, by displaying the convex hull of the ROC curves, emphasizes the general limits of categorization and the performance of the model. The underlined ROC curve shows the classification performance at the default threshold value of 0.5. Data on average true positive and false positive rates at various threshold values are also shown, along with the models' classification performance at various decision boundaries. The performance line goes into greater detail about how different classification decisions affect the bottom line. Examining the ROC curves provides a clearer picture of the models' capacity to distinguish between positive and negative instances. To compare the classifiers' efficacy in classifying images of spinach leaves, we may utilize the ROC diagram, which provides a comprehensive visual depiction of the classifiers' performance using the true positive and false positive rates at different decision thresholds.

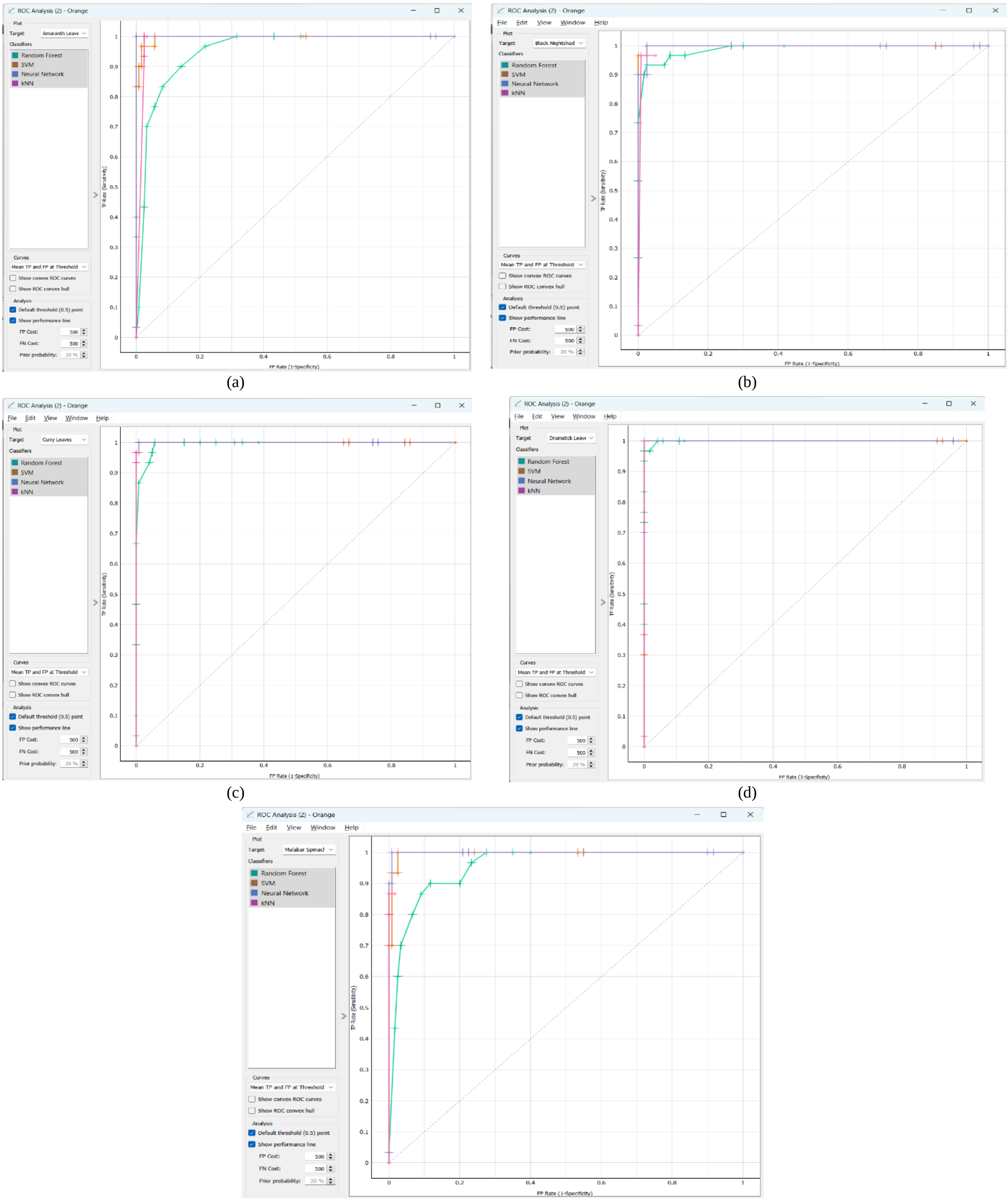
IV. DISCUSSION

The place of work was free and open-source, and we used an image analytics tool. The strategy takes advantage of Orange, a visual programming framework for data mining, and makes use of its ability to build interactive visualizations, workflows, and data models. Computing circumstances, especially those based on Python and enhanced alongside deep learning frameworks like TensorFlow [68], PyTorch [69], and Keras [70], are excellent at supporting contemporary image insights. Although the aforementioned resources ought to be favored by any sophisticated user or statistics scientist, Orange seeks to enhance these toolboxes by offering a user-friendly and collaborative platform that keeps supplying an extensive number of capabilities and perhaps customized to particular requirements using Visualization and the creation of routines tailored to certain issues.

Images may be easily imported into the Orange data mining environment with this comprehensive guide [71]. A detailed explanation of how to load the photos for classification using the 'Image analysis' add-on and the 'Import image' widget should be included in this. For feature extraction from the photos of spinach leaves, the article has to go into detail on the particular pre-trained Deep Convolutional Neural Network (DCNN) embedding methods used, such as Squeeze Net, Inception V3, and VGG16. A thorough rundown of how deep learning classifiers, supervised and unsupervised machine learning classifiers, and features retrieved through embedded images are utilized for categorizing spinach leaves [72]. An explanation of the Orange workflow, including details on the widgets used and why they were connected to load, embed, and classify the images.

Anyway, The visual statistical techniques of Orange are designed to analyze smaller image sets with image embedding as the initial step using a deep network that has already been trained. Data pretreatment is essential to Orange's data mining tool for analysis and model building. One can perform several data preparation tasks using the tool's widgets and actions. Orange handles data loading, cleaning, transformation, feature engineering, imputation, sampling, and machine learning integration. SQL databases, Excel, and CSV files can be imported using Orange's "File" widget. With this widget, the Orange canvas may quickly be loaded with structured data [73]. The "Data Table" widget is ideal for data cleansing. The dataset can be manually cleaned for missing values, outliers, inconsistent data, and extraneous features, including imputation, deleting irrelevant variables, and normalizing. For dataset transformation, Orange offers "Normalize," "Discretize," and "Select Columns" widgets, among others. Normalization, discrete category conversion, and attribute selection widgets simplify statistical analysis [74].

Engineers can utilize widgets like "Feature Constructor" and "Feature Subset" to create new features. Adding attributes, merging features, and selecting features improves the dataset's prediction capabilities. Orange's "Impute" widget offers prediction, mode, median, and mean imputation for missing data. This helps with missing data concerns. Orange's data sampling and splitting widgets enable cross-validation and train-test splits as shown in Table XVII. The "Data Split" and "Data Sampler" widgets manipulate data splitting for model training and evaluation. Orange's seamless integration of data pretreatment and machine learning lets users link preprocessed information to widgets for model creation and evaluation. This allows model creation and data preparation in one spot. So overall Orange workflow on Spinach leaves classification and prediction using SVM, KNN, RF, and NN are performed well. Still, Neural Network had a high classification accuracy value of 96.6% True positive rate of Recall value is 99.4%. The quality of the positive prediction to precision value is 99.3%.



(a) (b) (c) (d) (e)

Fig. 12. (a) ROC analysis for Amaranth leaves using a prediction model, (b) ROC Analysis for Balack Nighshade using a prediction model, (c) ROC Analysis for Curry leaves using a prediction model, (d) ROC Analysis for Drumstick leaves using a prediction model, (e) ROC Analysis for Malabar Spinach leaves using a prediction model

TABLE XVII. A NUMBER OF SPINACH LEAVES DATASETS ARE USED IN THE TRAINING AND TESTING PHASES

Spinach leaves types	Training Dataset	Test Dataset	Total
Malabar Spinach	70	30	100
Drumstick Leaves	70	30	100
Curry Leaves	70	30	100
Black Nightshade	70	30	100
Amaranth Leaves	70	30	100
Total	350	150	500

The combination of precision with recall gives the F1-Score for correct prediction in the overall dataset 99.3% for all Spinach leaves in this study. prediction value of 96.2%. other models attained below 95% Accuracy in the classification of five classes of spinach leaves. Based on the confusion matrix, comparing all models to classification Neural Network was classified 338 images correctly in five groups of spinach leaves [75] and 12 images only wrongly classified but we took a minimum of 350 images of the dataset for this classification with four classifier models. After the classification, these images were passed to the prediction with a new test data set. Compared with all models Neural Network has 96% highest prediction percentage for all spinach leaves. F1-score, Precision, and Recall values of the Neural Network are 96%,96.2%, and 96% and Neural Network and KNN models have 100% Recall values to predict the Amaranth leaves. F1-Score and precision score all so got the same value by using KNN (K-Nearest Neighbour) and NN (Neural Network) in Amaranth leaves. In Black Nightshade, SVM secures a high F1-score value of 93.8% and Random Forest finds a high precision value of 93.1%. On the flip side, SVM (Support Vector Machine) got a high Recall value of 100% compared to other models. NN (Neural Network) and KNN achieve a higher F1-score of 98.3% and apart from RF, other models hit 100% in precision with Recall values of NN and KNN is 96.7% in the Curry Leaves prediction. In Drumstick leaves, 100% of F1-Score and Recall were achieved by SVM, NN, and KNN models with all the models achieving 100% precision value in prediction. Neural Network only has high F1-Score, Precision, and Recall values in Malabar spinach leaves. So, 144 images were correctly predicted and only 6 images were wrongly predicted by Neural Network compared with other classifier models in the spinach leaves in the study.

V. CONCLUSION

We studied the classification of spinach leaf varieties in depth using large machine-learning datasets. With scores above 97%, Neural Networks, 93% for Support Vector Machine (SVM) and k-Nearest Neighbor, and 90% for Random Forest were the four primary algorithmic approaches to prediction. The dataset is too large for Support Vector Machines (SVMs) and k-Nearest Neighbors to effectively forecast the optimal version of spinach leaves, while neural networks' hidden layers perform better. The 96.6% accuracy rate of the Neural Network (NN) in identifying spinach leaves in images is far higher than that of earlier approaches. We also used SVM, RF, and KNN classifiers to assess the model's performance; each had different levels of success. We contrast our models' performance and accuracy rates to show that they can distinguish between various spinach leaf kinds. By analyzing the differences in accuracy across different spinach leaves, we can see how well the models perform in detecting these leaves, and so provide a comprehensive assessment of performance. We evaluated the amount of spinach leaves that were detected properly and incorrectly to assess the effectiveness of the models. Consequently, we have a better grasp of the models' capacity to differentiate between various spinach leaf kinds. By comparing the models on different spinach leaves, we can see their strengths and flaws. Analyzing things in comparison covers all the bases. Further down the lane, this work may be further extended towards disease detection and we already had collected visual Spinach leaf datasets from agricultural sectors as built-in datasets were not available online. The advanced Deep learning and object detection methodologies will be used to detect disease detection is vital as it ensures the edibility of the population that consumes it. Moreover, we will use the Internet-of-Things (IoT) method to take clear images of spinach leaves from agricultural land to perfect disease detection and classification.

REFERENCES

[1] Y. Xu, Y. Zhai, Q. Chen, S. Kong, and Y. Zhou, "Improved Residual Network for Automatic Classification Grading of Lettuce Freshness," in *IEEE Access*, vol. 10, pp. 44315-44325, 2022, doi: 10.1109/ACCESS.2022.3169159.

[2] D. Vaishnav and B. R. Rao, "Comparison of Machine Learning Algorithms and Fruit Classification using Orange Data Mining Tool," *2018 3rd International Conference on Inventive Computation Technologies (ICICT)*, pp. 603-607, 2018, doi: 10.1109/ICICT43934.2018.9034442.

[3] S. A. Hoogenboom, U. Bagci, and M. B. Wallace, "Artificial intelligence in gastroenterology. The current state of play and the potential. How will it affect our practice and when?," *Techniques and Innovations in Gastrointestinal Endoscopy*, vol. 22, no. 2, pp. 42-47, 2020.

[4] M. Agarwal, A. Singh, S. Arjaria, A. Sinha, and S. Gupta, "ToLeD: Tomato leaf disease detection using convolution neural network," *Procedia Computer Science*, vol. 167, pp. 293-301, 2020.

[5] G. Kuricheti and P. Supriya, "Computer Vision Based Turmeric Leaf Disease Detection and Classification: A Step to Smart Agriculture," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 545-549, 2019, doi: 10.1109/ICOEI.2019.8862706.

[6] M. M. Sufian, E. G. Moung, C. J. Hou, and A. Farzamnia, "Deep Learning Feature Extraction for COVID19 Detection Algorithm using Computerized Tomography Scan," *2021 11th International Conference on Computer Engineering and Knowledge (ICCKE)*, pp. 92-97, 2021, doi: 10.1109/ICCKE54056.2021.9721469.

[7] R. Bassiouny, A. Mohamed, K. Umapathy, and N. Khan, "An Interpretable Neonatal Lung Ultrasound Feature Extraction and Lung Sliding Detection System Using Object Detectors," in *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 119-128, 2024, doi: 10.1109/JTEHM.2023.3327424.

[8] Y. M. Bhavsingh, U. Atiya, M. Islam, and S. A. Hossain, "Enhancing Spinach Leaf Classification through Ensemble Learning with Deep Convolutional Neural Networks and Random Forests," *Macaw Int. J. Adv. Res. Comput. Sci. Eng.*, vol. 9, no. 2, pp. 39-45, 2023.

[9] E. Meganathan and T. F. Fernandez, "Enhanced Plant Leaf Disease detection using CNN algorithm in comparison with Haar Cascade Algorithm," *IEEE proceeding of International Conference on Recent Trends in Computer Communication and Business Management (ICRTCCB 2023)*, 2023.

[10] Y. Liu, I. R. Alzahrani, R. A. Jaleel, and S. Al Sulaie, "An efficient smart data mining framework based cloud internet of things for developing artificial intelligence of marketing information analysis," *Information Processing & Management*, vol. 60, no. 1, p. 103121, 2023.

[11] N. Singhal and P. Kumar, "Technique for Recognizing Useful Data from Complex Data Set through Data Miner Tool," *NeuroQuantology*, vol. 20, no. 16, p. 1378, 2022.

[12] M. Recce, J. Taylor, A. Piebe, and G. Tropiano, "High speed vision-based quality grading of oranges," *Proceedings of International Workshop on Neural Networks for Identification, Control, Robotics and Signal/Image Processing*, pp. 136-144, 1996, doi: 10.1109/NICRSP.1996.542754..

[13] D. Unay and B. Gosselin, "Apple defect detection and quality classification with mlp-neural networks," in *Proceedings of the ProRISC Workshop on Circuits, Systems and Signal Processing*, 2002.

[14] K. Fobes. *Volume estimation of fruits from digital profile images*. PhD thesis, Department of Electrical Engineering, University of Cape Town, Cape Town, 2000.

[15] S. Sennan, D. Pandey, Y. Alotaibi, and S. Alghamdi, "A Novel Convolutional Neural Networks Based Spinach Classification and Recognition System," *Computers, Materials & Continua*, vol. 73, no. 1, 2022.

[16] S. Zhu, L. Feng, C. Zhang, Y. Bao, and Y. He, "Identifying freshness of spinach leaves stored at different temperatures using hyperspectral imaging," *Foods*, vol. 8, no. 9, pp. 1–12, 2019.

[17] S. M. Dol and P. M. Jawandhiya, "Use of Data mining Tools in Educational Data Mining," *2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT)*, pp. 380-387, 2022, doi: 10.1109/CCiCT56684.2022.00075.

[18] J. Demšar *et al*., "Orange: data mining toolbox in Python," *the Journal of machine Learning research*, vol. 14, no. 1, pp. 2349-2353, 2013.

[19] Dataset of Amaranth leaves, black nightshade, curry leaves, drumsticks and Malabar spinach images, available https://www.kaggle.com/datasets/ahilaprem/mepco-tropic-leaf. Accessed 24/06/ 2023.

[20] K. Ananth Pai, B. R. Apoorva, D. S. Mendonca, D. S. Hegde, and R. B. Hegde, "Development of an Automated Plant Classification System Using Deep Learning Approach," in *Advances in VLSI, Signal Processing, Power Electronics, IoT, Communication and Embedded Systems: Select Proceedings of VSPICE 2020*, pp. 303-315, 2021.

[21] N. Butale and D. Kodavade, "Survey paper on detection of unhealthy region of plant leaves using image processing and soft computing techniques," *International Journal of Computer Engineering in Research Trends*, vol. 5, no. 12, pp. 232-235, 2018.

[22] P. Khosravi, E. Kazemi, M. Imielinski, O. Elemento, and I. Hajirasouliha, "Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images," *EbioMedicine*, vol. 27, pp. 317–328, 2018.

[23] H. O. B. Diezma and A. Herrero-Langreo, "A Hybrid Approach to Spinach Leaf Classification: Integrating Convolutional Neural Networks with Genetic Algorithms for Feature Selection," *Frontiers in Collaborative Res.*, vol. 2, no. 1, pp. 21-28, 2024.

[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[25] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[26] D. A. Keim, F. Mansmann, and J. Thomas, "Visual analytics: how much visualization and how much analytics?," *Acm Sigkdd Explorations Newsletter*, vol. 11, no. 2, pp. 5-8, 2010.

[27] D. Sacha, M. Sedlmair, L. Zhang, and J. A. Lee, "What you see is what you can change: human-centered machine learning by interactive visualization," *Neurocomputing*, vol. 268, pp. 164–175.2017.

[28] J. Rockstroma, J. Barron, and A. Lavanya, "Aquatic-Based Optimization Techniques for Sustainable Agricultural Development," *Front. Collab. Res*, vol. 1, no. 1, pp. 12–21, 2023.

[29] J. Demšar, B. Zupan, M. Zitnik, M. Toplak, A. Staric, and L. Umek, "Orange: data mining toolbox in Python," *J. Mach. Learn. Res*., vol. 14, pp. 2349–2353, 2013.

[30] J. Demsar and B. Zupan. *Orange: From Experimental Machine Learning to Interactive Data Mining*. White Paper, Faculty of Computer and Information Science, University of Ljubljana, 2004.

[31] M. Bhavsingh, Y. Alotaibi, and S. Alghamdi, "Fusion of Convolutional Neural Networks and Gradient Boosting Machines for Spinach Leaf Classification and Prediction," *Int. J. Comput. Eng. Res. Trends*, vol. 11, no. 3, pp. 11-18, 2024.

[32] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[33] D. Vaishnav and B. R. Rao, "Comparison of machine learning algorithms and fruit classification using orange data mining tool," in *2018 3rd International Conference on inventive computation technologies (ICICT)*, pp. 603–607, 2018.

[34] S. U. Habiba, M. K. Islam, and S. M. M. Ahsan, "Bangladeshi Plant Recognition using Deep Learning based Leaf Classification," *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*, pp. 1-4, 2019, doi: 10.1109/IC4ME247184.2019.9036515.

[35] L. Guo, J. Li, Y. Zhu, and Z. Tang, "A novel Features from Accelerated Segment Test algorithm based on LBP on image matching," *2011 IEEE 3rd International Conference on Communication Software and Networks*, pp. 355-358, 2011, doi: 10.1109/ICCSN.2011.6013732.

[36] A. Abidin, B. Deng, A. M. Dsouza, M. B. Nagarajan, P. Coan, and A. Wismuller, "Deep transfer learning for characterizing chondrocyte patterns in phase contrast X-Ray computed tomography images of the human patellar cartilage," *Comput. Biol. Med*., vol. 95, pp. 24–33, 2018.

[37] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Stat. Comput*., vol. 14, no. 3, pp. 199– 222, 2004.

[38] A. Sharma and P. Malhotra, "LDA Based Tea Leaf Classification on the Basis of Shape, Color and Texture," *Int. J. Comput. Eng. Res. Trends*, vol. 4, no. 12, pp. 543–546, Dec. 2017.

[39] F. Zhang, F. Ye, and Z. Su, "A modified feature point descriptor based on binary robust independent elementary features," *2014 7th International Congress on Image and Signal Processing*, pp. 258-263, 2014, doi: 10.1109/CISP.2014.7003788.

[40] S. Sivan and G. Darsan, "Computer vision based assistive technology for blind and visually impaired people," in *Proceedings of the 7th International Conference on Computing Communication and Networking Technologies*, pp. 1-8, 2016.

[41] D. Bisen, "Deep convolutional neural network based plant species recognition through features of leaf," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 6443-6456, 2021.

[42] S. R. dos Santos, M. K. Kondo, and M. S. Kiran, "Multimodal Fusion for Robust Banana Disease Classification and Prediction: Integrating Image Data with Sensor Networks," *Frontiers in Collaborative Research*, vol. 1, no. 2, pp. 22-31, 2023.

[43] M. Islam, N. J. Ria, J. F. Ani, A. K. M. Masum, S. Abujar, and S. A. Hossain, "Deep learning based classification system for recognizing local spinach," in *Advances in Deep Learning, Artificial Intelligence and Robotics: Proceedings of the 2nd International Conference on Deep Learning, Artificial Intelligence and Robotics,(ICDLAIR) 2020*, pp. 1-14, 2022.

[44] G. H. Chen and D. Shah, "Explaining the success of nearest neighbor methods in prediction," *Trends® in Mach. Learn*., vol. 10, no. 5–6, pp. 337–588, 2018.

[45] J. Lu, L. Tan, and H. Jiang, "Review on convolutional neural network (CNN) applied to plant leaf disease classification," *Agriculture*, vol. 11, no. 8, pp. 1–18, 2021.

[46] I. K. Raji and K. K. Thyagharajan, "An analysis of segmentation techniques to identify herbal leaves from complex background," *Procedia Computer Science*, vol. 48, pp. 589-599, 2015.

[47] M. O. Ramkumar, S. S. Catharin, V. Ramachandran, and A. Sakthikumar, "Cercospora identification in spinach leaves through resnet-50 based image processing," in *journal of physics: conference series*, vol. 1717, no. 1, p. 012046, 2021.

[48] O. Tsepa and M. M. Pedram, "ShiftSense: A Unified Framework for Comprehensive Detection of Gradual and Abrupt Concept Shifts in Streaming Data," *Frontiers in Collaborative Research*, vol. 1, no. 4, pp. 1-9, 2023.

[49] P. S. Kanda, K. Xia, and O. H. Sanusi, "A deep learning-based recognition technique for plant leaf classification," *IEEE Access*, vol. 9, no. 1, pp. 162590–162613, 2021.

[50] Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," *2018 IEEE/ACM 26th International Symposium on Quality*

of Service (IWQoS), pp. 1-2, 2018, doi: 10.1109/IWQoS.2018.8624183.

[51] A. Wajid, N. K. Singh, P. Junjun, and M. A. Mughal, "Recognition of ripe, unripe and scaled condition of orange citrus based on decision tree classification," *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1-4, 2018, doi: 10.1109/ICOMET.2018.8346354.

[52] Q. Yang and X. Wu, "10 challenging problems in data mining research," *International Journal of Information Technology & Decision Making*, vol. 5, no. 4, pp. 597-604, 2006.

[53] M. Muchová, J. Paralič, and M. Jančuš, "An approach to support education of data mining algorithms," *2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, pp. 000093-000098, 2017, doi: 10.1109/SAMI.2017.7880282.

[54] P. Godec *et al*., "Democratized image analytics by visual programming through integration of deep models and small-scale machine learning," *Nature communications*, vol. 10, no. 1, p. 4551, 2019.

[55] S. K. Dash and M. Panda," Image classification using data mining techniques," *Adv. Comput. Sci. Inform. Technol. (ACSIT)*, vol. 3, no. 3, pp. 157–162, 2016.

[56] T. Saba, A.M. Khattak, A. Khan, R. Ullah, and Z. Ullah, "A comparative study of machine learning algorithms for early detection of tomato leaf diseases," *Computers and Electronics in Agriculture*, vol. 171, p. 105345, 2020.

[57] C. Priya, A. Balasaravanan, and S. Thanamani, "An efficient leaf recognition algorithm for plant classification using support vector machine," *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)*, pp. 428-432, 2012.

[58] S. Shukla, S. Tripathi, and A. Shukla, "Early prediction of plant leaf diseases using XG_Boost algorithm," *International Conference on Sustainable Computing in Science, Technology and Management*, 2020.

[59] W. Ocimati, S. Elayabalan, and N. Safari, "Leveraging Deep Learning for Early and Accurate Prediction of Banana Crop Diseases: A Classification and Risk Assessment Framework," *Int. J. Comput. Eng. Res. Trends*, vol. 11, no. 4, pp. 46–57, 2024.

[60] M. E and T. F. Fernandez, "Quantitative Analysis Upon Plant Leaf Diseases Employing Regressive Machine Learning Algorithms," *2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN)*, pp. 1-5, 2023, doi: 10.1109/ViTECoN58111.2023.10157445.

[61] A. Ishak, K. Siregar, R. Ginting, and M. Afif, "Orange software usage in data mining classification method on the dataset lenses," in *IOP Conference Series: Materials Science and Engineering*, vol. 1003, no. 1, p. 012113, 2020.

[62] S. Weng, S. Yu, R. Dong, F. Pan, and D. Liang, "Nondestructive detection of storage time of strawberries using visible/near-infrared hyperspectral imaging," *International Journal of Food Properties*, vol. 23, no. 1, pp. 269-281, 2020.

[63] S. G. Wu, F. S. Bao, E. Y. Xu, Y. Wang, Y. Chang, and Q. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network," *IEEE International Symposium on Signal Processing and Information Technology*, pp. 11-16, 2007.

[64] N. M. Butale and D. V. Kodavade, "Detection of plant leaf diseases using image processing and soft-computing techniques," *International Research Journal of Engineering and Technology*, vol. 6, no. 6, pp. 3288-3261, 2019.

[65] S. Akter *et al*., "Comprehensive Performance Assessment of Deep Learning Models in Early Prediction and Risk Identification of Chronic Kidney Disease," in *IEEE Access*, vol. 9, pp. 165184-165206, 2021, doi: 10.1109/ACCESS.2021.3129491.

[66] T. H. F. Harumy, M. Zarlis, S. Effendi, and M. S. Lidya, "Prediction Using A Neural Network Algorithm Approach (A Review)," *2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM)*, pp. 325-330, 2021, doi: 10.1109/ICSECS52883.2021.00066.

[67] R. Jerčić, I. Pavić, and I. Damjanović, "New algorithm for identifying network topology based on artificial neural networks," *2019 2nd International Colloquium on Smart Grid Metrology (SMAGRIMET)*, pp. 1-5, 2019, doi: 10.23919/SMAGRIMET.2019.8720364.

[68] Y. Jingyi, S. Rui, and W. Tianqi, "Classification of images by using TensorFlow," *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, pp. 622-626, 2021, doi: 10.1109/ICSP51882.2021.9408796.

[69] R. Lohith, K. E. Cholachgudda, and R. C. Biradar, "PyTorch Implementation and Assessment of Pre-Trained Convolutional Neural Networks for Tomato Leaf Disease Classification," *2022 IEEE Region 10 Symposium (TENSYMP)*, pp. 1-6, 2022, doi: 10.1109/TENSYMP54529.2022.9864390.

[70] C. Shorten and T. M. Khoshgoftaar, "KerasBERT: Modeling the Keras Language," *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 219-226, 2021, doi: 10.1109/ICMLA52953.2021.00041.

[71] P. Radchenko, J. Bilan, and V. Kachurka, "Application of data mining methods for prediction of possible future income values," *2019 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, 2019, doi: 10.1109/DSMP.2019.8787476.

[72] S. Chandra *et al*., "Automated detection and classification of spinach leaf diseases using convolutional neural networks," *Computers and Electronics in Agriculture*, 2019, doi: 10.1016/j.compag.2018.12.023.

[73] F. Marozzo, D. Talia, and P. Trunfio, "A Workflow Management System for Scalable Data Mining on Clouds," in *IEEE Transactions on Services Computing*, vol. 11, no. 3, pp. 480-492, 2018, doi: 10.1109/TSC.2016.2589243.

[74] S. Slater, S. Joksimović, V. Kovanovic, R. S. Baker, and D. Gasevic, "Tools for educational data mining: A review," *Journal of Educational and Behavioral Statistics*, vol. 42, no. 1, pp. 85-106, 2017.

[75] S. Mehta, V. Kukreja, and S. Vats, "Improving Crop Health Management: Federated Learning CNN for Spinach Leaf Disease Detection," *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pp. 1-6, 2023, doi: 10.1109/CONIT59222.2023.10205629.