

Review

Review on Techniques for Plant Leaf Classification and Recognition

Muhammad Azfar Firdaus Azlah ¹, Lee Suan Chua ^{1,2,*} , Fakhru Razan Rahmad ³, Farah Izana Abdullah ² and Sharifah Rafidah Wan Alwi ⁴

¹ Department of Bioprocess and Polymer Engineering, School of Chemical and Energy Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, UTM Skudai, Johor Bahru 81310, Johor, Malaysia; Azfar_frds@yahoo.com

² Metabolites Profiling Laboratory, Institute of Bioproduct Development, Universiti Teknologi Malaysia, UTM Skudai, Johor Bahru 81310, Johor, Malaysia; izana.fa@gmail.com

³ Robotic and Computer Aided Detection Laboratory, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn, Parit Raja, Batu Pahat 86400, Johor, Malaysia; FakhruRazan94@gmail.com

⁴ Department of Chemical Engineering, School of Chemical and Energy Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, UTM Skudai, Johor Bahru 81310, Johor, Malaysia; syarifah@utm.my

* Correspondence: chualesuan@utm.my; Tel.: +60-19-7214378

Received: 2 September 2019; Accepted: 16 October 2019; Published: 21 October 2019



Abstract: Plant systematics can be classified and recognized based on their reproductive system (flowers) and leaf morphology. Neural networks is one of the most popular machine learning algorithms for plant leaf classification. The commonly used neural networks are artificial neural network (ANN), probabilistic neural network (PNN), convolutional neural network (CNN), k-nearest neighbor (KNN) and support vector machine (SVM), even some studies used combined techniques for accuracy improvement. The utilization of several varying preprocessing techniques, and characteristic parameters in feature extraction appeared to improve the performance of plant leaf classification. The findings of previous studies are critically compared in terms of their accuracy based on the applied neural network techniques. This paper aims to review and analyze the implementation and performance of various methodologies on plant classification. Each technique has its advantages and limitations in leaf pattern recognition. The quality of leaf images plays an important role, and therefore, a reliable source of leaf database must be used to establish the machine learning algorithm prior to leaf recognition and validation.

Keywords: leaf; pattern recognition; artificial neural network; probabilistic neural network; convolutional neural network; k-nearest neighbor; support vector machine

1. Introduction

Plants are essential for mankind. In particular, herbs have been used as folk medicines by indigenous people since ancient times. Herbs are usually identified by practitioners based on years of experiences through personal sensory or olfactory sense [1]. Recent advances in analytical technology have significantly assisted in herbal recognition based on scientific data. This eases many people, especially those that are lacking experience in herbal recognition. Laboratory-based testing requires skills in sample treatment and data interpretation, in addition to time consuming procedures [2]. Therefore, a simple and reliable technique for herbal recognition is needed. Computation combined with statistical analysis is likely to be a powerful tool for herbal recognition. This nondestructive technique shall be the method of choice to rapidly identify herbs, particularly for those who cannot afford to apply expensive analytical instrumentation.

One of the widely used nondestructive techniques to identify herbs is based on their leaf morphological images [3]. Plant leaves are representative enough to differentiate plant species or variety with high accuracy [4,5]. At the present time, plant recognition is still the specialization of plant taxonomists. The advancement of computing technologies would be another alternative of choice for non-specialists. Nowadays, the morphological characteristics of leaves can be extracted by a mathematical model to put into a software program for recognition. This could reduce false positive results due to human error [6]. Computational morphometric methods can quantitatively measure a leaf geometrically and visualize differences in an effective, reproducible, accurate, and statistically powerful way. Factors which are usually considered in leaf morphological study include length, width, area, perimeter, diameter, shape, and color [7]. The basic geometrical information can be used for herbal classification and identification. The possibility to use mathematical model in herbal plant recognition can be seen from the previous studies which reported the application of computational methods to detect plant disease or infection from leaf appearance and morphology successfully [8,9].

In earlier times, leaf pattern classification using an automated system was difficult, mainly because of no standard leaf pattern available as a benchmark in computing. Researchers usually struggle and spend a lot of time to establish the database by gathering many leaf samples as raw dataset. In 2007, a group of researchers published their works on leaf pattern dataset, which is called the Flavia dataset (<http://flavia.sourceforge.net/>) with the aim to share their algorithm and dataset with other researchers. Their initial works classified 32 plant species based on 12 characters in five principal variables. In the subsequent years, there were many studies using the Flavia dataset to develop a model for plant leaf recognition system.

2. Image Processing in Leaf Pattern Recognition

Leaf pattern recognition usually follows the steps as shown in Figure 1. The most challenging part of this study is to extract distinctive features of leaves for plant species recognition. In this case, different classifiers using high performance statistical approaches have been used to perform leaf features extraction and classification. The advancement in computer vision and artificial intelligence have greatly assisted researchers to classify plants through statistical modeling.

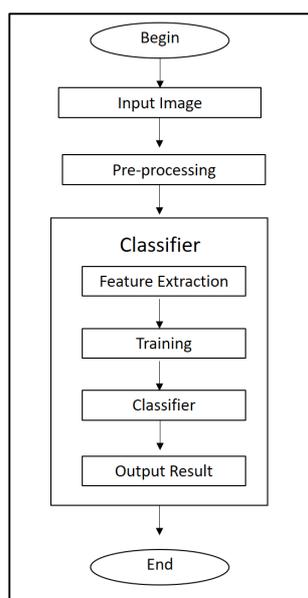


Figure 1. Fundamental of leaf pattern recognition.

The pre-processing step consists of image reorientation, cropping, gray scaling, binary thresholding, noise removal, contrast stretching, threshold inversion, and edge recognition. Image reorientation is

aligning the input image to a standardized position, with the leaf aligned to either the x-axis or y-axis. For leaves that have the greater width: Length ratio, the length is preferably placed in the vertical or upright position [10]. To decrease the amount of computational load that is exerted upon the graphic processing unit, cropping the image is a necessary step to reduce the unnecessary foreground region of the prompt image. Turkoglu and Hanbay [11] suggested that leaf feature extraction could be done by dividing the leaf image into two or four parts, instead of extracting for the whole leaf. The proposed image processing techniques using color, vein, Fourier descriptors (FD), and gray-level co-occurrence matrix (GLCM) methods had proven to achieve 99.1% accuracy using the Flavia leaf dataset.

Gray scale conversion of the image into geometrical data is implemented to optimize the contrast and intensity of images. Later, the thresholding process creates a binary image from the gray scaled image to translate the value of the image to its closest threshold, and therefore having either one of two possible values for each pixel, as presented in Figure 2. Different types of noises, such as grains, and holes, could affect digital images, therefore erosion and dilation are a series of operations implemented in order to remove the background noises. The images are considered homogenous if they do not exhibit substantial differences between one another in terms of contrast stretching. These images, when shown in histogram representation, exhibit very narrow peaks. Inhomogeneity is caused by the lack of uniform lighting upon the image. The image is normalized in order to stretch the narrow range to a more dynamic range. The binary images from the process are inverted during threshold conversion, to convert the background into black. Suzuki algorithm can be utilized to extract the contours of images and further refined by diminishing the contours with small lengths with regards to its largest contour [10]. This process is known as edge recognition.

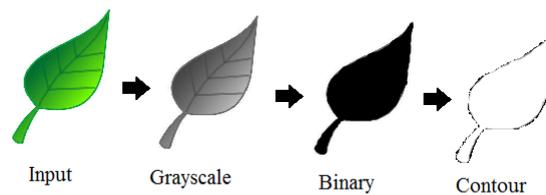


Figure 2. Image pre-processing stage.

Ma et al. [12] suggested that their algorithm was better than the conventional back-propagation algorithm in terms of efficiency and accuracy. They analyzed soybean leaf image to evaluate the nitrogen content by introducing median filter in their preprocessing stage. A hazier image is collected to remove grain noise, which could disrupt image processing due to high frequency properties as a result of grey difference [13]. The subsequent step would be emphasizing the leaf edge to obtain a clear image. In this case, the grey linear transformation technique is used to further add the difference in grey saturation between the leaf image and the background, and thus enhance the image by gaining a comparable threshold value of sample image and background to decrease error rate [14]. Since the image background still has an undesired value, the image is binarized to remove the background value completely before the original image is imposed to the processed image [15]. The output of this preprocessed image had successfully analyzed the nitrogen content of soybean based on colour characteristics [12].

Bo et al. [16] used the Hu moment invariant for soybean leaf recognition in image processing to distinguish its leaf feature from weeds. Chronologically, they procured the real time image with a camera as the input data which was then converted to grayscale with 2G-R-B [16]. The data of both leaves from soybean and weed was then further processed by erosion algorithm to remove distortion. The next step involved the Hu moment invariant adopting the 16×16 template to declare the invariability of soybean leaf image including rotation, scale, and translation. Succeeding the preprocessing stage would be the identification process, in which the nearest neighbour classifier was used to compare the relative soybean leaf image via its respective markings. The accuracy of the classification along with the image preprocessing could yield 90.5% recognition rate [16].

Grand-Brochier et al. [17] made a comparative study between different types of image preprocessing techniques to analyze the effectivity of each method, in addition to colour distance map and input stroke. The local color, which is the color of the subject with the whole image was matched. One of the methods is simple linear iterative clustering (SLIC), which utilizes super-pixel to cluster them with a predetermined value through repetitive iteration of the nearest neighbor, so that only the data vector with similar value is included [17]. The results of SLIC technique yielded 87.4% precision [17]. Guided active contour (GAC) is another method. In GAC, the Snake segmentation technique is adopted where the initial stage involves iteration to improve the polygonal framework for the elongated leaf shape [18]. This iteration derives an energy equation to guide the polygon to expand within its closed area [19]. The GAC technique managed to procure precision up to 95.2% [17].

In Kurtz algorithm, a hierarchical approach was used to extract segments of interest from the lowest to the highest resolution data, which was the first cluster image material sharing the common colour properties into a group of coarse image patches [20]. The individual patches are depicted as hierarchical construct with a binary partition tree (BPT). This cluster of images can be assumed to be like a forest of BPTs. When the construction of this forest is completed, each tree will be gradually segmented. Eventually, the progressive segmentation will produce a global segmentation of the image. This method could produce precision up to 85.1%. The data of Kurtz algorithm only utilized colour distance map without input stroke as opposed to the two previously reported methods; SLIC and GAC [21].

A comparative result had been established using power watershed for preprocessing of the image. Power watershed used the concept of magnitude reduction to get more accurate data from the adjacent region, which is identical to the Graphcut method, hence improving segmentation of the image. Power watershed technique included multi labelling, contrast, and ratio invariant stage [22]. The Power watershed method compared to the Graphcut technique yielded 63.5% precision.

The discovery of various preprocessing techniques has proven to vastly facilitate the development of an effective machine learning. Since visual-based machine learning depends on how well the feature of a certain image is extracted from the preprocessing stage, it is crucial to understand the desired outcome in preprocessing stage. This is because different problems require different solutions or approaches. Additionally, different situations will present different images, and the preprocessing stage of the predetermined image would revolve around the enhancement of the subject features.

3. Leaf Feature Extraction

Leaf features such as shape, size, and colour are important characteristics in a computational recognition system. The feature extraction can be performed by contour-based or region-based extraction. The contour-based extraction is expressed in length, width, aspect ratio, and leaf diameter as descriptors. The length descriptor is measured based on the main vein of leaf, as it stretches from the main vein to the end tip. The width descriptor is the span of leaf viewed from one side to the other, from the leftmost point to the rightmost point of leaf. Aspect ratio can be determined by dividing the length of leaf to its respective width. The longest distance between two points inside the covered area of leaf is the leaf diameter, while the maximum distance of the area covered by leaf is the perimeter. Alternatively, leaf perimeter can also be obtained by computing the number of pixels that contains leaf margin.

On the other hand, leaf features extracted from the region-based technique uses the descriptors such as shape, rectangularity, compactness, area, and eccentricity. Shape (Convex Hull) is the coordinates of points where the whole area of leaf can be conveniently determined using the convex hull algorithm. Rectangularity (R) is a rectangle which just fits the image drafted outside the image in Equation (1).

$$R = LpWp/A \quad (1)$$

where Lp , Wp , and A represent the length, width, and area of the leaf, respectively.

Compactness is also known as roundness, which is described as the ratio of surface area of leaf to the square of leaf perimeter, as presented in Equation (2) 'A' is the surface area of leaf and 'P' is the perimeter of leaf.

$$\text{Compactness} = 4\pi A/P^2 \quad (2)$$

Area is defined as the product of thresholding process. Pixel which is represented by the number 1 defines the leaf area (Equation (3)). Eccentricity is a definitive characteristic of any conic section of a leaf. This feature can be obtained from Equation (3), where 'b' is the minimum axial length and 'a' is the minimum axial length of an ellipse.

$$\text{Eccentricity} = \sqrt{1 - [(b/a)]^2} \quad (3)$$

Lu et al. [23] discovered that leaves usually have diseases or holes, which could reduce the total area of leaf, and thus compromise the segmentation result in feature extraction. This is because holes in leaves will be identified as the background instead of leaf area. Therefore, it is suggested to use contour extraction with selective area filling to overcome this problem. This method is initiated by searching the contour point of the foreground vector via pixel scanning in order from left to right, and from bottom to top to determine whether the subjected pixel is categorized as background or foreground. For every foreground encounter, this process is break and the next line is scan. This process is iterated for every individual pixel until their identity is declared upon three aforementioned axis in order of top to bottom, right to left and bottom to top. The declared pixel is not scanned after it is identified. The foreground region can be verified by the four contours stretching along the perimeter of leaf region. This extraction process prepares the pixel for the subsequent stage which is region filling. This feature extraction stage is only viable for binary image and for leaf area extraction. The result of this method which was applied to five leaves had yielded an average absolute error of 3.00 [23].

Shivling et al. [2] applied feature extraction using a method called area labelling. After the image is processed through the image preprocessing stage, the output binary image is subjected to area labelling in order to produce an identified region. This algorithm marked a target matrix which has an integer value of '1' and are interconnected with a positive value to differentiate the foreground and background. The scan kernel used eight-connecting area algorithm which implies the kernel will further search for its eight-connecting area, when the pointer found a pixel with the value '1'. If the area is found, the cluster will be labelled with a new positive integer to match with its centre value. In the situation where no nine-connecting area is found to be connected with the pixel which has a value of '1', a new search is initiated on a different region. Hence, multiple region can be labelled independently. This process is completed when all pixels of the area are marked. This method used regionprops function in MATLAB to count pixels in the labelled region of images. The totality of the pixels marked in the image are counted for image extraction, and thus reflecting the feature properties of the leaf image. Ten leaf images had been used to develop the method and the result was found to achieve high accuracy up to 0.1 mm² [2].

Gopal et al. [24] utilized colour features extraction in an attempt to prepare the presented medicinal images for classification purpose. Preceding the classification stage, feature extraction begins by providing an input image from a digital scanner and then followed by a series of image preprocessing stages. After the image is completely conditioned, it is prompted into the program to procure the colour feature of the medicinal plant with regards to its Fourier descriptor. This Fourier descriptor ensures the invariant of shape in terms of rotation, translation, and scaling [25]. This program was trained using 100 leaf images and tested using 50 leaf images, and ultimately managed to obtain an efficiency of 92% [24].

Leaf feature extraction can also be done using the integral contour-angle (ICA) method as reported by Ni et al. [26]. This method uses a mathematical representation of a shape contour, which is laid out according to the perimeter around the contour and the coordinates of the contour point. The complex value of the vector point is projected on two sides of the image; the right and left neighbourhood

beginning and ending at the vector point along the perimeter to obtain the average vector [27]. This ICA representation is crucial to derive the ICA descriptor to further derive the multiscale descriptor for the image. The ICA descriptor is manipulated by changing the parameter size to suit the desired scale size to obtain the K-dimensional feature vector. The feature extracted from the aforementioned stages were tested and achieved 89.0% precision [26].

Indeed, feature extraction plays the most important role in determining the accuracy and precision of a machine learning mechanism. This is stemmed from the fact that the architecture for machine learning essentially depends on the predetermined feature that is prompted into the network. There is no single feature extraction technique that can be considered as the best method since different problems require different approaches and different mechanisms. If the subject feature is well understood, the feature extraction would be easier.

4. Mathematical Classifiers

Upon the completion of leaf pattern extraction, the information which is also known as feature vectors are used for further inspection, comparison before being grouped into their particular classes. There are many mathematical classifiers being used by researchers. Each classifier may have its advantages and limitations.

4.1. Artificial Neural Network (ANN)

One of the relatively superior and competent classifiers is the ANN, especially in terms of its accuracy. This is because ANN is pertinent to resolve non-linear problems like leaf pattern recognition. However, previous research revealed that the leaf with oblong pattern could increase error rate of recognition, possibly due to the uniform structure [26]. The basic structure of ANN is an interconnected set of nodes [28]. There are multiple layers of nodes connected each other to generate the desired output as illustrated in Figure 3.

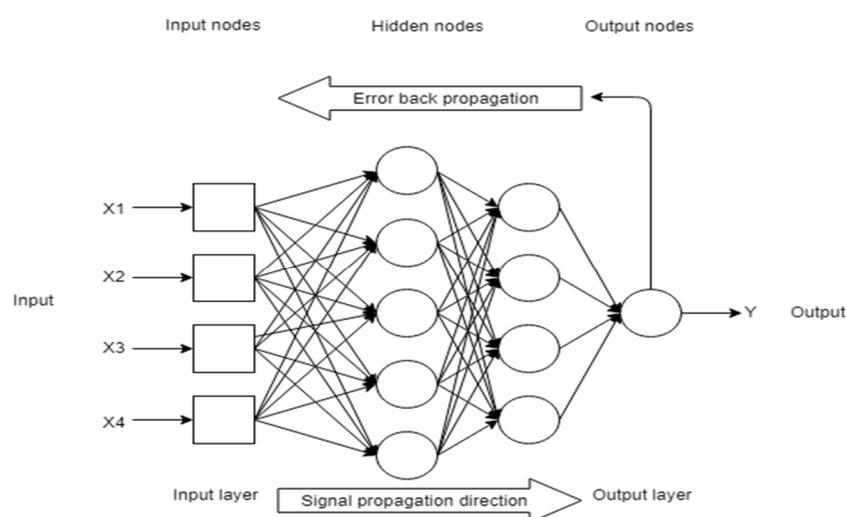


Figure 3. Feed-forward back-propagation mechanism in artificial neural network.

The data processing is directed from the front, when prompts information, to the back. When the system acknowledges the results, further training is performed and this new information is redirected to the front to reset weights to the front side of the neural units in order to increase its accuracy. This process is known as feed-forward back-propagation method, and thus increasing the accuracy of the classifier.

The input nodes of the system will be determined by the extracted features. The number of nodes of the output layers would be determined by the number of plant categories too. The classifiers would

be trained by utilizing back propagation method and the weight of the links will be altered to reduce the error between the expected and actual outputs.

The utilization of ANN as a classifier for leaf pattern recognition is reliable, since it could exhibit the pattern recognition accuracy exceeding 90% [29]. There was a study which could achieve up to 98.6% of recognition accuracy when the size of training set is increased [29].

The success of ANN to classify and identify selected medicinal plant leaves was reported by [30]. The feature extraction was carried out based on shape, colour, and texture of leaf images, and training via the ANN classifier. The image processing and neural network toolboxes were obtained from MATLAB. The implementation of eight input features into the system were found to be optimal in terms of complexity, as it required minimum input, as well as less computational time. The accuracy of the system was 94.4% using 63 leaf images. Fu et al. [31] improved the accuracy of leaf venation extraction for about 10% using the combined thresholding and ANN approach. The accuracy was improved from 84.4% using direct ANN approach to 97.3% using the combined approach. The improvement was also noticed to reduce computing time.

4.2. Convolutional Neural Network (CNN)

CNN implements deep learning to machine image processing in order to classify pictures of leaf samples. Recent development of hardware and information processing technology has made deep learning a self-learning method that utilizes massive sum of data in a more feasible manner as shown in Figure 4.

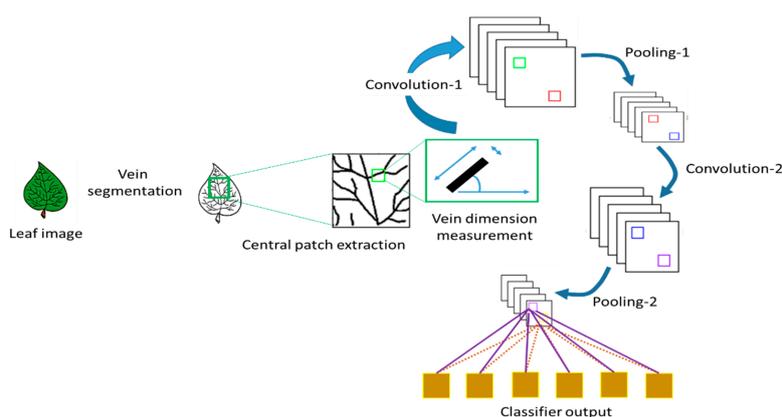


Figure 4. Leaf pattern detection with convolutional neural network.

The CNN initially aims to imitate the visual system of human. The retina identifies edges of an object by strong intensity of light compared to the whole object in the human visual system. This information is sent to lateral geniculate nucleus (LGN). The shape of the information is compressed and delivered to primary visual cortex (V1). The edge and contour of the image is interpreted at V1. Simultaneously, the image information from the retina of the left and the right eyes give depth perception which provides distance information. These data are sent to secondary visual cortex (V2) where recognition of overall shape and color perception of different segments of the image are taking place before being transferred to tertiary visual cortex (V3) to interpret the color of the whole object.

A more advanced CNN uses multiple convolution synchronously in order to procure the required feature vector [32,33]. Factorizing the convolutional layer would be effective to decrease the number of parameters. Instead of using a single convolution filter, a set of smaller layers is recommended. This technique has been proven to significantly minimize the parameters, while successfully extracting key features. Recently, Hang et al. [34] improved the traditional CNN by combining a structure of inception structure and a global layer pooling layer to identify leaf diseases. The combined inception structure reduced the number of model parameters and improved the performance of identifying leaf diseases up to 91.7% accuracy. Similarly, Neuroph was used in the training of a CNN network for

maize leaf disease classification [35]. The method was proven to be effective to recognise three types of diseases, namely northern corn leaf blight, common rust, and gray leaf spot diseases. Deep learning with CNN has been proven to be reliable for plant disease classification by capturing the color and texture of lesions specific diseases, even removing 75% number of parameters that were not attributed to inference, would not affect the classification accuracy [36].

Unlike other classifiers, CNN extracts and identifies the features concurrently; hence it is a faster recognition process. However, this neural network classifier requires users to train numerous sets of data before it is considered to be competent enough for application. Therefore, CNN that incorporates deep learning as its mechanism seems to be remarkably accurate in plant recognition. This is because CNN managed to recognize the leave pattern and reach the accuracy above 94%, even though damaged samples with altered feature vectors could still be processed effectively [37].

Therefore, harnessing the knowledge of deep learning model in CNN could be an effective feature extraction tool to identify vein patterns from the prompted image [32]. Although the network is connected to all output feature vectors of previous convolutional layers, this classifier manages to give huge decrement to its number of trainable parameters. Nevertheless, CNN might lose its ability to generalize feature vector [38].

4.3. Probabilistic Neural Network (PNN)

Similar to CNN, PNN is another branch of ANN which incorporates additional algorithm. PNN utilizes radial basis function (RBF) which measures nonlinear variable in a shape of bell as illustrated in Figure 5.

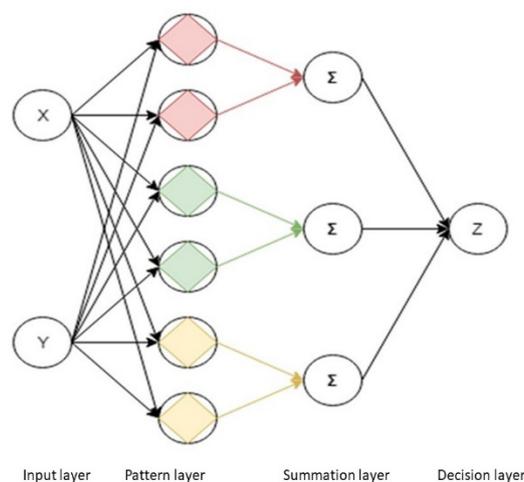


Figure 5. Mechanism of probabilistic neural network.

This PNN classifier trains the loaded feature vector with higher speed rate as compared to that of a backpropagation system. Since the feature characteristic is predetermined, the classification step becomes more straightforward which makes this classifier robust to distortion. This trait also makes this classifier a simple training approach and structure.

Since the weight passes between the layers of nodes are established in the early stage, the pre-existing weights will not be manipulated. Nevertheless, new vectors are placed into weight matrices during the training phase which results in real-time viability. During the process of recognizing the leaf pattern, the feature vector of leaf is classified by the network into a particular class since the assigned class is assumed to have the highest probability to be accurate.

Previous study conducted by Stephen Gang Wu [3] managed to achieve a classification accuracy of 90% using PNN. With the same classifier also, an improved version of PNN produced better result with 93.7% accuracy in the shape feature extraction. When texture feature is incorporated, the accuracy of the classifier could achieve up to 98.3% [39]. In terms of network structure, PNN has the upper hand

compared to ANN since it selects the nodes automatically. This also makes PNN be more flexible to determine the output [40]. Despite being convenient in classifying leaf features, this system possesses limitation where it is susceptible to data overfitting. The development of this neural network is also time consuming because it has intricate network structure [41].

There was also a research using mobile application for Indonesian medicinal plants identification. The classifier was PNN combined with a fuzzy local binary pattern and fuzzy colour histogram [42]. The combination via the product decision rule was capable of extracting leaf image texture and colour. The fusion methods of fuzzy local binary pattern and fuzzy colour histogram increased the overall accuracy of plant identification and yielded an accuracy of 74.5%. The accuracy of the system without the fusion features was found to be 59.6% for fuzzy local binary pattern and 50.8% for fuzzy colour histogram, respectively. In another study, leaf species classification was conducted using a botanical shape sub-classifier strategy [43]. The implementation of fusion strategy and its corresponding random-forest-based sub-classifiers as part of a leaf recognition system. The fusion technique provided a significant accuracy improvement, while providing necessary information for educational purposes.

Mahdikhanlou and Ebrahimnezhad [44] reported the improvement of system accuracy using the method of centroid distance and axis of least inertia in PNN. The experimental results showed that the accuracy of the system was 82.1% using the Flavia dataset, and 80.1% using Swedish leaf dataset (<https://www.cvl.isy.liu.se/en/research/datasets/swedish-leaf/>). The accuracy is better than the fuzzy based techniques as reported by [42]. The methods are invariant to translation and rotation for leaf image classification. Canary operator was implemented for binary images converted from RGB (Red, Green, and Blue) to detect and thin out the edges of leaf images before the shape is traced. The centroid distances of the points, as well as the distance of sampling points from the axis of least inertial lines are used for computation.

4.4. Support Vector Machine (SVM)

The idea of SVM is to define decision boundaries of feature vectors on decision plane which separates features unanimously. Since the distinction in feature between the images are evident, the images will be classified into their respective class with little to no complication. New sample is appointed into one of the classes that match their characteristics in this training algorithm. The samples are then charted onto the same region of class where they shall be. A fine separation between classes is naturally attained by the hyperplane that marks the furthest distance to the nearest data point of each group. This is also due to the fact that the larger margin results in small generalization error of the classifier. Applying invariant feature to the SVM will vastly increase the accuracy of the classifier. The application of the invariant features goes with the contour-based descriptor. This classifier is exceptionally accurate because it is capable to recognize images without the problem of image rotation, translation, scale, and inversion or mirroring. The only drawback is that the system is not capable to differentiate leaves that have almost the same shape. In response to this situation, the principal component analysis (PCA) could be applied to enhance the performance of this system. PCA operates by converting data set into a new set of data and the variables are arranged according to their priorities. This causes the components to be uncorrelated and sorted in the order that the components with the largest variation occurs first, followed by the components with the smallest variation which is then terminated. Finally, one of the more superior classifiers would be SVM which integrates the Hu moment invariant algorithm to its model which could achieve up to 94.1% accuracy [45]. Zhang et al. [46] classified plant species using leaf shape and texture. They used SVM classifier for a dataset of 1900 leaves belonging to 32 different species. A new classification method based on the generation of space features by combining local texture features using wavelet decomposition, co-occurrence matrix statistics, and global shape features. The method was able to extract features of the plant leaf images and yielded an accuracy of over 93.8%.

A method of extracting 15 features of leaf images via canny edge detector, and SVM as the classifier was reported by Salman et al. [47]. The extracted features were convex area, filled area, perimeter (P),

eccentricity (E), solidity, perimeter ratio of diameter, orientation, narrow factor, extent, Euler number, diameter, circularity, rectangularity, perimeter ratio of length and width, complexity, and compactness. The study used 2220 images from 22 plant species in the Flavia dataset and achieved an accuracy of 85 to 87% only.

Khmag et al. [4] aimed to create a recognition system for leaf images based on leaf contour and centroid. An image processing algorithm was developed by utilizing the variant to scaling shift, spin technique, scaling approach, and filtering processes, as well as SVM as the classifier for the leaf contour. The proposed method used 70 samples taken from the Flavia dataset and their geometrical and shape features were extracted and yielded an accuracy of 97.7% which is the highest achievement reported in literature.

Different classifiers had been used by Srivastava et al. [48] who did a comparative analysis for leaf classification and recognition. The study extracted 14 leaf features using shape detector and SVM as the classifiers of the system. Sixteen different plant species from the Flavia database comprising of 480 images were used as the training dataset and 14 features extracted for optimal performance. Quadratic SVM yielded the highest accuracy of 90.9% as compared to medium Gaussian SVM and cubic SVM which gave 89.4% and 89.8% test accuracy, respectively.

A multiple classifier system was also used by Araujo et al. [49] based on texture and shape features of leaf images. SVM and neural network classifiers were used to train four different features, namely local binary pattern (LBP), histogram of gradients (HOG), speed of robust features (SURF), and Zernike moments (ZM). Then, a static classifier selection method was used to search for the ensembles that maximize the average classification score. The ImageCLEF 2011 [50] and ImageCLEF 2012 [51] datasets were used for the experiments. The results found that the proposed method was able to improve for 11.7% and 4.2% in the scan category and 4.06% and 5.87% for the scan-like category in the average classification score relative to the best results reported in the literature for ImageCLEF 2011 and ImageCLEF 2012 datasets, respectively. Hence, the multiple classifier system overcame the performance of monolithic methodologies.

A combined classifier consisted of Naïve Bayes and a decision tree along with SVM was discovered to yield maximum accuracy, 93.1% [52] in leaf image classification. The classification was based on visual attributes such as shape, colour, texture, and vein pattern. The accuracy was slightly lower when using the individual classifier, namely 91.8% for SVM, 88.3% for Naïve Bayes, and 85.7% for decision tree. The high performance of Naïve Bayes classifier can also be seen from the findings of Eid et al. [53]. The computational model was proposed using digital plant images based on biometric features such as shape and vein patterns. The model could achieve an average accuracy, 97% of leaf identification based on 10 different biometric information extracted from 1907 leaf images of 32 plant species taken from the Flavia dataset.

The classifier of SVM also proved to be effective to detect plants in their natural habitat, which was purposely designed in an environment with heavy interferences and overlapping issues [54]. The system used three plant species with 300 leaf images for the experiments with 86.7% accuracy in identification. The images were captured and segmented using a marker-controlled watershed segmentation, which was generated automatically using morphological operations. The shapes were then converted into features via Hu moments. The accuracy of the system can be further improved by the addition of more features extracted via the system operation, as well as increasing the dataset used for the experiments.

The significant advantage of SVM is its robustness as a feature vector classifier mechanically, which is based on digital morphological paper [55]. It is a potent classifier with its simplicity in terms of feature extraction [56]. To the best of our knowledge, this classifier could achieve the highest accuracy (85%–98%) in leaf identification. This could be due to the high number of leaf images which usually ranges from hundreds to thousands used in model development. Probably, this also explains that many studies would like to use the SVM classifier in recent years. However, this system needs to be trained for long duration because a new decision boundary is drawn each time when new data is tested.

4.5. K-Nearest Neighbor (KNN)

KNN algorithm classifies images by comparing the input images to the closest training samples from the feature space. This technique is considered as the simplest machine classifier where the training process of this algorithm only involves two fundamental steps of keeping the feature data and information of the training images. Ambiguous points are justified by assigning them to the information of KNN. Generally, the images are grouped with respect to the label of KNN by the majority of votes. Figure 6 describes the sample analysis with KNN.

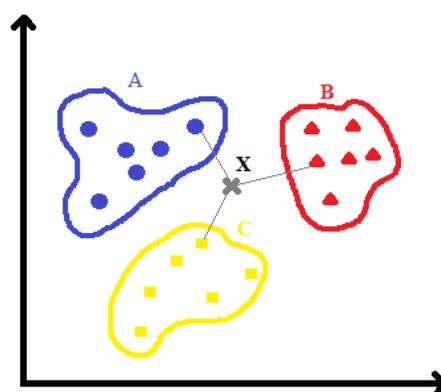


Figure 6. Sample analysis with k-nearest neighbor.

K is always an odd integer since there are only two classes. For example, images that are subjected to test will be directly classified as the class of the image close to it, if $k = 1$. In some cases, when testing for multiclass classification, tied outcome could be encountered when k is an odd integer. This issue could be fixed by measuring the Euclidean distance after acquiring the vector of fixed-length of the image that had been converted to real number.

The image information from the pre-processing stage will determine the features that are going to be compared by rationing the characteristics of leaf to obtain the white area ratio, perimeter to hull, hull is a ration to get the ratio of area of leaf to the area of hull which will be used for vector computation and many others. These ratios are then normalized to have a value between zero and one.

Chronologically, the image that needs to be identified will go through the same process as the images that have been stored in the database up to normalizing the ratios. The tested image is compared to each image from the database afterward. On the other hand, comparison by color histogram could be used to find out the identity of a new leaf if the dataset from previous processes are composed of leaves from different plants.

KNN has advantages of using all input leaf images as test subjects before storing the collected information to the database. Increasing the number of images for testing will definitely increase the accuracy of the classifier. In other words, adding plant species to the database can also reduce the accuracy of the system. Somehow, Kherkhah and Asghari [57] reported that applying Cosine KNN classifier and PCA algorithm in gist feature vector was outperformed compared to the approaches of Patternnet neural network and SVM.

This classifier is sensitive to inadequate lighting which will affect the accuracy of the system. The feature extraction is too simple with basic characteristic requirement. Future work could incorporate the algorithm that would make the system more stable to light disparity. Hence, simple classifier like KNN could obtain 83.5% accuracy [10]. Although the feature extraction process is simple and quick, the low accuracy result is relatively weak to be accepted. This classifier is not stable enough to handle distortion from the extracted samples and would lead to inaccuracy during the classification process. A method incorporated with prescribed color histogram had proven to increase the reliability of this classifier up to 87.3% accuracy [10]. KNN algorithm approaches the subject by directly comparing the feature characteristics of each sample individually and grouped before being represented by their

respective class with the closest similarities [58]. Kumar et al. [59] reported the improvement in leaf classification based on shape and edge features with KNN classifier. The method was tested on 32 plant species in Flavia dataset and yielded an average accuracy of 94.4% in classification. The improvement was attributed to the size and rotation invariant which are being the novelty of the algorithm.

The advantages and disadvantages of the classifiers were summarized in Table 1.

Table 1. Comparison between advantages and disadvantages of classifiers.

Classifiers.	Advantages	Disadvantages
Artificial Neural Network (ANN)	<ol style="list-style-type: none"> 1. Capable in distinguishing complex nonlinear relationship between independent and dependent variables. 2. Simplistic statistical training. 	<ol style="list-style-type: none"> 1. Great tendency of data overfitting. 2. Bigger computational load.
Convolutional Neural Network (CNN)	<ol style="list-style-type: none"> 1. Multiple features can be extracted simultaneously. 2. Robust to noise. 	<ol style="list-style-type: none"> 1. High computation level. 2. No capable for generalization.
Probabilistic Neural Network (PNN)	<ol style="list-style-type: none"> 1. High distortion resistance. 2. Flexible to changing data. 3. Specimen can be classified into multiple output. 	<ol style="list-style-type: none"> 1. Long training duration. 2. Complicated network layout. 3. Have tendency for overfitting with too many traits.
Support Vector Machine (SVM)	<ol style="list-style-type: none"> 1. Great generalization potential. 2. Exceptionally robust. 	<ol style="list-style-type: none"> 1. Speed and size constraint for both training and testing. 2. Complex algorithm structure. 3. Slow training.
K-Nearest Neighbor (KNN)	<ol style="list-style-type: none"> 1. No training needed. 2. Robust in term of research space. 3. Simplest classifier. 	<ol style="list-style-type: none"> 1. Susceptible to noise. 2. Lazy learning. 3. Costly testing for each instance.

ANN is a unique classifier to identify the complex nonlinear relationship between independent and dependent variables. Since leaf pattern recognition is considered as a complex nonlinear problem, ANN would be able to interpret the variables effectively. ANN is also considered to be simple in terms of its statistical training. The iteration of the training mechanism is conducted one at a time, and the result is feed-forwarding into the system as a new information. Consequently, this system may cause data overfitting due to the fact that it iterates a large sum of data. This will end with huge computational load to sustain the feed-forward back-propagated data.

CNN may involve multiple features extraction, and at the same time, providing detail and quick detection. Therefore, this mechanism makes CNN be more resilient towards unnecessary noise. This classifier also utilizes a high computation level to extract multiple features. This system is not appropriate in generalizing features. Similar to CNN, PNN has a high noise resistance too. With its capability to foresee potential information, it is flexible to do data changing. Since the weight of the vector is predetermined, the specimen could be classified into multiple outputs depending on the closeness of the sample to the vector. The drawback for this classifier is the needs for long duration of training when new weight vector is applied. This system has an intricate layout structure. In case excessive traits are involved, it is prone to overfitting also.

SVM is different from other neural networks, it has great generalization potential and is exceptionally robust. The disadvantages of this system are its speed and size which are restricted for both training and testing. This classifier has a complicated algorithm structure to determine the feature vector of each sample individually. Since its generalization potential is great, it takes time for the training process before it could classify a particular set of samples.

The research team of Shah et al. [60] utilized a dual-path deep CNN to learn leaf image features and optimize for classification in order to classify the leaf based on marginalized shape context and shape-texture dual-path deep CNN. The study revealed that the dual-path CNN method outperformed other CNN methods such as uni-patch CNN, texture-patch CNN, marginalized shape context with SVM classifier, multiscale distance matrix with SVM classifier, curvature histogram. The method gave near-perfect-top-1-match result using Flavia dataset and top-three-match results for other datasets.

A more distinct approach brought by KNN classifier is to emphasize the differences between features and compare the close feature vector, and thus make training absolutely unnecessary. This is also robust in terms of research space. The characteristics make KNN the simplest classifier. KNN classifier is very susceptible to noise which may alter the results during the classification process. Lack of training makes this classifier likely to be a lazy learning system. The prediction step in KNN is rather expensive, since every prediction is made, the classifier determines the nearest neighbor by comparing to the entire dataset.

5. Pattern Recognition Method

Various pattern recognition methods have been proposed and new techniques will be discovered continuously to meet the needs of technological advancement. Depending on the situation, different issues require different image classification methods. Various methodologies and approaches describe how flexible leaf recognition system could be, catering to various criteria and prerequisite as shown in Table A1.

In the study of Pankaja et al. [61] who stated that grey level co-occurrence matrix, discrete wavelet transform, and hierarchical centroid could assist in obtaining special information of leaves accurately and improve the accuracy of plant leaf recognition. Three hundred leaf samples of 30 different species sourced from the Flavia dataset were used for the experiment and the model achieved an accuracy of 96.7%. This result was obtained after preprocessing, feature extraction, and classification of leaf samples based on shape and texture, via grey level co-occurrence matrix and hierarchical centroid based technique.

Chaki et al. [62] purposed a methodological recognition of 32 pre-defined classes of plants using a neuro-fuzzy classifier and compared it to other classifiers such as KNN and neural network classifiers. In the study, 640 leaf sample images with various sizes, shapes, and orientations were tested and found to show improvement compared to k-NNC and NNC with an accuracy of 97.5% and 79.7%, respectively. A limitation of the proposed method is that it works reliably only when deformations do not alter the major and minor axis lengths.

In the work of neural network-based leaf recognition, the implementation of pulse coupled neural network in the feature extraction of leaf images was used to obtain the entropy sequence, aspect ratio, Zernike moments, Hu's invariants, form factor, rectangularity, circularity, and area [63]. The accuracy of the system was better than the other methods, probably because the entropy was taken as the key feature in the leaf classification, yielding an accuracy of 92.0%.

The study of Chaki et al. [64] developed a plant leaf recognition using a layered approach. The layered architecture with each layer handles a specific type of visual characteristic using a set of features to form a customized data model. The different layers are then fed into an array of custom classifiers for robust recognition. Colour-based modelling is used for non-green leaves, while shape-based modelling is used for green, simple, and compound leaves, using a layered approach. The database of 600 leaf images from 30 different classes of leaves were used to test the layered model and yielded an accuracy of over 90% for all the different approaches. The utilization of different visuals, segmentation to reduce computational load, different customization in the classifiers, and addition of different layers make the model more robust than all the other models proposed in previous studies.

In the works of Harish et al. [65], the classification of plant leaves using morphological features and Zernike moments was reported to achieve optimal results. They used Flavia dataset of 32 species (1907 leaf images) and six medicinal plants (180 leaf images). The proposed model utilized leaf morphological features and Zernike moments which are independent of leaf growth and image translation, rotation, and scaling. Different classifiers had been tested and found that SVM and PNN showed to have comparably higher accuracy (87–89%) than KNN and Naïve Bayes (65–68%) classifiers due to the latter being lazy learner algorithms.

The advancement of vision technology paralleled with deep learning has brought upon a remarkable benchmark in artificial intelligence, primarily in the field of visual classification. In spite of all the advantages that visual classification has granted to various fields, improvement can be seen in order to have a robust system. For example, a lot of classification systems are still having problems with overfitting which greatly affects how the system views the subject. The machine learning system should provide short processing time in this computational age.

6. Conclusions and Recommendation

Several reliable automated procedures are used for leaf pattern recognition. This paper mainly reviews the advantages of each classifier and compares their compatibility with different leaf features recognition process. A computer vision approach which can completely neglect the background of the image is speeding up the recognition process and it is suitable for highly complex plant leaf samples. A system that neglects distortion tremendously enhances the recognition technology and even makes the recognition of aquatic fauna more feasible since aquatic plants or algae may not have a definitive shape. The current image processing technique should be robust under diverse intensity of lighting. This new algorithm can be developed by tweaking the detection technique which may lead to detection of specific diseases. The advantage can also be applied for herbal plants recognition to prevent adulteration for better quality control, especially for product efficacy and safety.

Author Contributions: Conceptualization, M.A.F.A. and L.S.C.; Methodology M.A.F.A. and F.I.A.; Software, M.A.F.A.; Writing—original draft preparation, M.A.F.A. and F.I.A.; Writing—review and editing, L.S.C., F.R.R. and S.R.W.A.; Visualization, F.I.A.; Supervision, L.S.C.; Project administration, F.I.A.; Funding acquisition, L.S.C.

Funding: This research was supported by the Ministry of Higher Education, Malaysia under the research grant 4J263.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Previous pattern recognition methods.

No.	Goals	Method	Results	References
Artificial Neural Network (ANN)				
1	Identification Of Selected Medicinal Plant Leaves Using Image Features And ANN	Method utilizes feature extraction based on shape, color, and texture of leaf images and training via ANN classifier for leaf classification of the system. The image processing and neural network toolboxes were obtained from MATLAB.	Implementation of 8 input features into the system were found to be optimal in terms of the complexity as it demands minimum input as well as less computational time required. The accuracy of the system was 94.4% against 63 leaf images in the dataset.	[30]
2	Neural Network Based Leaf Recognition	Implementation of Pulse Coupled Neural Network into the feature extraction of leaf images to obtain the entropy sequence, aspect ratio, Zernike moments, Hu's invariants, form factor, rectangularity, circularity, and area. The classifier used is Artificial Neural Network.	The accuracy of the system proposed was better compared to the other methods in the study, where the entropy was taken as the key feature in the leaf classification, yielding an accuracy of 91.98%.	[63]
3	Combined Thresholding and Neural Network Approach for Vein Pattern Extraction From Leaf Images	A proposed model that extracts leaf vein patterns via combining thresholding methods and an ANN classifier.	The experimental results of the method proposed have shown to possess the capability of extracting a more accurate venation pattern of the leaf samples for pattern recognition with 97.3% accuracy while reducing computing time as well compared to direct neural network approach which achieved an accuracy of 84.4%.	[31]
4	Classification of Plant Leaves using Morphological Features and Zernike Moments	The proposed model utilizes leaf morphological features and Zernike moments which are independent of leaf growth and image translation, rotation and scaling and then classified with several different classifiers to achieve optimal result.	The dataset tested were Flavia dataset of 32 species (1907 images, 50-60 leaves each species) and a medicinal plants dataset of 6 species (30 images each). SVM and PNN have comparably higher accuracy than k-NN and naïve Bayes classifiers due to the latter being lazy learners algorithms.	[65]
Convolutional Neural Network (CNN)				
5	Leaf Classification based on Marginalized Shape Context and Shape-Texture Dual-Path Deep Convolutional Neural Network	Utilizing a dual-path deep CNN to learn leaf image features and optimize them for classification compared to vanilla convolutional network systems.	Dual-path CNN method outperforms various other CNN methods, being: uni-patch CNN, texture-patch CNN, marginalized shape context with SVM classifier, Multiscale Distance Matrix with SVM classifier, curvature histogram, giving near-perfect-top-1-match results on Flavia dataset and top-3-match results on all other datasets.	[60]

Table A1. Cont.

No.	Goals	Method	Results	References
Probabilistic Neural Network (PNN)				
6	Mobile Application for Indonesian Medicinal Plants Identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram	Utilizing the combined methods of Fuzzy Local Binary Pattern (FLBP) and the Fuzzy Color Histogram (FCH) in order to identify the medicinal plants found in Indonesia. The combination was done via Product Decision Rule (PDR) method, and is capable of extracting leaf image texture and color. The classifier utilized was PNN classifier.	Fusion of methods FLBP and FCH has shown to be capable of increasing the overall accuracy of plant identification methods and yielded an accuracy of 74.51%. The accuracy of the system without the fusion features were 59.61% for FLBP features and 50.78% for FCH feature respectively.	[42]
7	Plant Leaf Classification Using Centroid Distance and Axis of Least Inertia Method	The utilization of centroid distance and axis of least inertia for plant classification. Canary operator was implemented for the binary images converted from RGB (Red, Green, Blue) to detect and thin out the edges of the leaf images before the shape is traced. The centroid distances of these points as well as the distance of sampling points from the axis of least inertial lines were computed. The classifier used was the PNN classifier for two public leaf databases, Flavia and the Swedish leaf dataset.	Combination of both these methods, which are invariant to translation, rotation and scale into the extraction of feature vectors for leaf image classification has been proven to be able to improve the accuracy. Experimentation results have shown that the accuracy of the system from the Flavia dataset was 82.05% and Swedish leaf dataset was 80.10%, an improvement from previous methods proposed.	[44]
Support Vector Machine (SVM)				
8	Combined Classifier for Plant Classification and Identification from Leaf Image based on Visual Attribute	The identification of the plant images based on varying features such as shape, color, texture, and vein pattern with combined classifier using majority voting technique has been proposed for recognizing leaf's category along with SVM, Naïve Bay's and Decision tree as the combined classifiers	Utilization of the combined classifiers with several features extraction of the leaf images have been shown to yield maximum accuracy output as compared to them individually and all existing methodologies with accuracy of 93.11% while with separately, SVM is 91.8576, Naïve Bayes is 88.2736%, and Decision tree is 85.6678% respectively.	[52]
9	Comparative Analysis of Leaf Classification and Recognition by different SVM Classifiers	A new approach proposed by extracting 14 features of leaf by utilizing shape detector and SVM as the classifier of the system	16 different plants species from the Flavia database comprising of 480 images were used as the training dataset with 14 features extracted for optimal performance. Quadratic SVM yielded the highest accuracy of 90.9% as compared to Medium Gaussian SVM and Cubic SVM gave 89.4% and 89.8% test accuracy separately.	[48]

Table A1. Cont.

No.	Goals	Method	Results	References
10	Multiple Classifier System for Plant Leaf Recognition	The methodology presents a multiple classifier system (MCS) of plant identification based on texture and shape features of the leaf images. SVM and Neural Network classifiers are trained on four different feature sets, namely, Local Binary Pattern (LBP), Histogram of Gradients (HOG), Speed of Robust Features (SURF) and Zernike Moments (ZM). Then, a static classifiers selection method is used to search for the ensembles that maximize the average classification score.	ImageCLEF 2011 and ImageCLEF 2012 datasets were used for the experiments with the proposed MCS methods showing an improvement of 11.76% and 4.23% in the scan category and 4.06% and 5.87% for the scan-like category in the average classification score relative to the best results reported in the literature for ImageCLEF 2011 and ImageCLEF 2012 datasets respectively. MCS approach also overcomes the performances of monolithic methodologies compared in the study.	[49]
11	Identifying leaf in a natural image using morphological characters	Proposed methodology of plant detection in their natural habitat, designed in an environment with heavy interference and overlapping. The images are captured and segmented using marker-controlled watershed segmentation which is generated automatically using morphological operations and the shapes obtained are converted into features via Hu moments. The classifier utilized is SVM classifier.	3 species with 300 different leaf image samples were used for the experimentation of the system and captured in real time with 86.7% identification accuracy. The accuracy of the system can be further improved by the addition of more features extracted via the system's operation as well as increasing the dataset used for the experiment.	[54]
k-Nearest Neighbor (KNN)				
12	Recognition of Whole and Deformed Plant Leaves using Statistical Shape Features and Neuro-Fuzzy Classifier	A method-logical recognition of 32 pre-defined classes of plants using a Neuro-fuzzy classifier and comparing it to other classifiers such as k-Nearest Neighbor and Neural Network classifiers.	640 leaf sample images of varying sizes, shapes, and orientation were tested in the experiment and have been proven to show improvements compared to k-NNC and NNC with an accuracy of 97.5% and 79.7%. A limitation of the proposed method is that it works reliably only when deformations do not alter the major and minor axis lengths.	[62]
13	Leaf Classification based on Shape and Edge feature with k-NN Classifier	A proposed approach via the utilization of edge based and shape based features extraction for the classification of leaf images.	The method was tested on Flavia dataset of 32 plant species and yielded an average classification accuracy rate of 94.4%, an improvement compared to existing methods, with size and rotation invariant being the novelty of the algorithm.	[59]

Table A1. Cont.

No.	Goals	Method	Results	References
No Classifier				
14	Leaf Recognition and Classification using GLCM and Hierarchical Centroid Based Technique	Utilization of various techniques for preprocessing, feature extraction, and classification of leaf based on shape and texture, via Gray Level Co-occurrence Matrix and hierarchical centroid based technique.	Gray Level Co-occurrence Matrix, Discrete Wavelet Transform, and hierarchical centroid have been proven to be able to obtain the special information of leaves accurately and improve on the accuracy. 300 leaf samples of 30 different species from Flavia dataset were used for the experiment and the model achieved an accuracy of 96.66%.	[61]
15	Plant Leaf Recognition Using a Layered Approach	A proposed methodology using layered architecture with each layer handling a specific type of visual characteristics using a set of features to form a customized data model. The different layers are then fed into an array of custom classifiers for robust recognition. Color based modeling is used for non-green leaves while shape based modeling is used for green simple and compound leaves, using a layered approach.	The database of 600 leaf images from 30 different classes of leaves were used to test the layered model and yielded an accuracy of over 90% for all the different approaches. The utilization of different visuals, segmentation to reduce computational load, different customization in the classifiers, and addition of different layers make the model more robust than all the other models proposed in previous studies.	[64]
16	Leaf Plant Identification System based on Hidden naïve Bays Classifier	A proposed computational model in plant identification system of digital images of plants utilizing biometric features such as shape and vein patterns via hidden naïve Bays classifier.	1907 leaf image samples from 32 different species of plants taken from Flavia dataset were used in the study in determining the accuracy of the system and the model has proven to have an average identification accuracy of 97% based on 10 different biometric information extracted.	[53]
17	Leaf species classification based on a botanical shape sub-classifier strategy	The implementation of fusion strategy and its corresponding random-forest-based sub-classifiers as part of a leaf recognition system.	The fusion technique provides a significant accuracy enhancement compared to other proposed methods, while providing necessary information for educational purposes.	[43]

References

1. Hoffman, H.J.; Cruickshanks, K.J.; Davis, B. Perspectives on population-based epidemiological studies of olfactory and taste impairment. *Ann. N. Y. Acad. Sci.* **2009**, *1170*, 514. [[CrossRef](#)] [[PubMed](#)]
2. Shivling, V.D.; Singla, A.; Ghanshyam, C.; Kapur, P.; Gupta, S. Plant leaf imaging technique for agronomy. In Proceedings of the 2011 International Conference on Image Information Processing, Shimla, India, 3–5 November 2011; pp. 1–5.
3. Hossain, J.; Amin, M.A. Leaf shape identification based plant biometrics. In Proceedings of the 2010 13th International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 23–25 December 2010; pp. 458–463.
4. Khmag, A.; Al-Haddad, S.A.R.; Kamarudin, N. Recognition system for leaf images based on its leaf contour and centroid. In Proceedings of the 2017 IEEE 15th Student Conference on Research and Development (SCoReD), Putrajaya, Malaysia, 13–14 December 2017; pp. 467–472.
5. Wäldchen, J.; Rzanny, M.; Seeland, M.; Mäder, P. Automated plant species identification-trends and future directions. *PLoS Comput. Biol.* **2018**, *14*, e1005993.
6. Sabu, A.; Sreekumar, K. Literature review of image features and classifiers used in leaf based plant recognition through image analysis approach. In Proceedings of the 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 10–11 March 2017; pp. 145–149.
7. Wu, S.G.; Bao, F.S.; Xu, E.Y.; Wang, Y.; Chang, Y.; Xiang, Q. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. In Proceedings of the 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, Egypt, 15–18 December 2007; pp. 11–16.
8. Singh, V.; Misra, A.K. Detection of plant leaf diseases using image segmentation and soft computing techniques. *Inf. Process. Agric.* **2017**, *4*, 41–49. [[CrossRef](#)]
9. Chaki, J.; Parekh, R.; Bhattacharya, S. Plant leaf classification using multiple descriptors: A hierarchical approach. *J. King Saud Univ. Comp. Inf. Sci.* **2018**, in press. [[CrossRef](#)]
10. Munisami, T.; Ramsurn, M.; Kishnah, S.; Pudaruth, S. Plant Leaf Recognition Using Shape Features and Colour Histogram with K-nearest Neighbour Classifiers. *Procedia Comput. Sci.* **2015**, *58*, 740–747. [[CrossRef](#)]
11. Turkoglu, M.; Hanbay, D. Recognition of plant leaves: An approach with hybrid features produced by dividing leaf images into two and four parts. *Appl. Math. Comput.* **2019**, *352*, 1–14. [[CrossRef](#)]
12. Ma, L.; Fang, J.; Chen, Y.; Gong, S. Color analysis of leaf images of deficiencies and excess nitrogen content in soybean leaves. In Proceedings of the 2010 International Conference on E-Product E-Service and E-Entertainment, Henan, China, 7–9 November 2010; Volume 11541023, pp. 1–3.
13. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 3rd ed.; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2002.
14. Kittler, J.; Illingworth, J. Minimum error thresholding thresholding minimum error decision rule Classification error Dynamic clustering. *Pattern Recognit.* **1986**, *19*, 41–47. [[CrossRef](#)]
15. ASM International. *Practical Guide to Image Analysis*; ASM International: Novelty, OH, USA, 2000.
16. Bo, Z.; Hua, W.H.; Jun, L.S.; Hua, M.W.; Chao, Z.X. Research on weed recognition method based on invariant moments. In Proceeding of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 2167–2169.
17. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2281. [[CrossRef](#)]
18. Cerutti, G.; Tougne, L.; Mille, J.; Vacavant, A.; Coquin, D. Understanding leaves in natural images—A model-based approach for tree species identification. *Comput. Vis. Image Underst.* **2013**, *117*, 1482–1501. [[CrossRef](#)]
19. Chan, T.F.; Vese, L.A. Donald Middleton, L.D.S. *Br. Dent. J.* **1977**, *10*, 266–277.
20. Kurtz, C.; Passat, N.; Gançarski, P.; Puissant, A. Extraction of complex patterns from multiresolution remote sensing images: A hierarchical top-down methodology. *Pattern Recognit.* **2012**, *45*, 685–706. [[CrossRef](#)]
21. Grand-Brochier, M.; Vacavant, A.; Cerutti, G.; Kurtz, C.; Weber, J.; Tougne, L. Tree leaves extraction in natural images: Comparative study of preprocessing tools and segmentation methods. *IEEE Trans. Image Process.* **2015**, *24*, 1549–1560. [[CrossRef](#)] [[PubMed](#)]

22. Couprie, C.; Grady, L.J.; Najman, L.; Talbot, H. Power watersheds: A new image segmentation framework extending graph cuts, random walker and optimal spanning forest. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; Volume 9, pp. 731–738.
23. Lü, C.; Ren, H.; Zhang, Y.; Shen, Y. Leaf area measurement based on image processing. In Proceedings of the 2010 International Conference on Measuring Technology and Mechatronics Automation, Changsha, China, 13–14 March 2010; Volume 2, pp. 580–582.
24. Gopal, A.; Prudhveeswar Reddy, S.; Gayatri, V. Classification of selected medicinal plants leaf using image processing. In Proceedings of the 2012 International Conference on Machine Vision and Image Processing (MVIP), Taipei, Taiwan, 14–15 December 2012; pp. 5–8.
25. Bama, B.S.; Valli, S.M.; Raju, S.; Kumar, V.A. Content Based Image Retrieval Using Advanced Color and Texture Features. *Int. J. Comput. Appl.* **2011**, *2*, 10–14.
26. Ni, F.; Wang, B. Integral contour angle: An invariant shape descriptor for classification and retrieval of leaf images. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1223–1227.
27. Kpalma, K.; Yang, M.; Ronsin, J. Planar shapes descriptors based on the turning angle scalogram. *Int. Conf. Image Anal. Recognit.* **2008**, *5112*, 547–556.
28. How Artificial Intelligence Has Switched-Up the SEO Game—More than you Imagined. 2016. Available online: <http://backlinksforseo.com/artificial-intelligence-switched-up-seo/> (accessed on 31 January 2018).
29. Wu, Q.; Zhou, C.; Wang, C. Feature Extraction and Automatic Recognition of Plant Leaf Using Artificial Neural Network. *Adv. Artif. Intell.* **2006**, *3*, 5–12.
30. Janani, R.; Gopal, A. Identification of selected medicinal plant leaves using image features and ANN. In Proceedings of the 2013 International Conference on Advanced Electronic Systems (ICAES), Pilani, India, 21–23 September 2013; pp. 238–242.
31. Fu, H.; Chi, Z. Combined thresholding and neural network approach for vein pattern extraction from leaf images. *IEE Proc. Vis. Image Signal Process.* **2006**, *153*, 881–892. [[CrossRef](#)]
32. Grinblat, G.L.; Uzal, L.C.; Larese, M.G.; Granitto, P.M. Deep learning for plant identification using vein morphological patterns. *Comput. Electron. Agric.* **2016**, *127*, 418–424. [[CrossRef](#)]
33. Vaidya, A.; Pujari, D.; Desai, H.; Borse, K. Leaf Recognition—A Technical Review. *Int. J. Res. Emerg. Sci. Technol.* **2015**, *2*, 46–51.
34. Hang, J.; Zhang, D.; Chen, P.; Zhang, J.; Wang, B. Classification of Plant Leaf Diseases Based on Improved Convolutional Neural Network. *Sensors* **2019**, *19*, 4161. [[CrossRef](#)]
35. Sibiyi, M.; Sumbwanyambe, M. A Computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks. *AgriEngineering* **2019**, *1*, 119–131. [[CrossRef](#)]
36. Toda, Y.; Okura, F. How convolutional neural networks diagnose plant disease. *Plant Phenomics* **2019**, *2019*, 9237136. [[CrossRef](#)]
37. Jeon, W.S.; Rhee, S.Y. Plant Leaf Recognition Using a Convolution Neural Network. *Int. J. Fuzzy Log. Intell. Syst.* **2017**, *17*, 26–34. [[CrossRef](#)]
38. Zeiler, M.D.; Fergus, R. Stochastic Pooling for Regularization of Deep Convolutional Neural Networks. *arXiv* **2013**, arXiv:1301.3557.
39. Lin, H.; Peng, H. Machine recognition for broad-leaved trees based on synthetic features of leaves using probabilistic neural network. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Hubei, China, 12–14 Decemner 2008; Volume 4, pp. 871–877.
40. McKenzie, J.S.; Jurado, J.M.; de Pablos, F. Characterisation of tea leaves according to their total mineral content by means of probabilistic neural networks. *Food Chem.* **2010**, *123*, 859–864. [[CrossRef](#)]
41. Sawant, S.S.; Topannavar, P.S. Introduction to Probabilistic Neural Network—Used for Image Classifications. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2015**, *5*, 279–283.
42. Herdiyeni, Y.; Kadek, N.; Wahyuni, S. Mobile Application for Indonesian Medicinal Plants Identification using Fuzzy Local Binary Pattern and Fuzzy Color Histogram. In Proceedings of the 2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 1–2 December 2012; pp. 978–979.

43. Liu, H.; Coquin, D.; Valet, L.; Cerutti, G. Leaf species classification based on a botanical shape sub-classifier strategy. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; Volume 5, pp. 1496–1501.
44. Mahdikhanlou, K.; Ebrahimnezhad, H. Plant leaf classification using centroid distance and axis of least inertia method. In Proceedings of the 2014 22nd Iranian Conference on Electrical Engineering (ICEE), Tehran, Iran, 20–22 May 2014; pp. 1690–1694.
45. Lukic, M.; Tuba, E.; Tuba, M. Leaf recognition algorithm using support vector machine with Hu moments and local binary patterns. In Proceedings of the 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics, Herl'any, Slovakia, 26–28 January 2017; pp. 000485–000490.
46. Zhang, H.; Yanne, P.; Liang, S. Plant species classification using leaf shape and texture. In Proceedings of the 2012 International Conference on Industrial Control and Electronics Engineering, Xi'an, China, 23–25 August 2012; pp. 2025–2028.
47. Salman, A.; Semwal, A.; Bhatt, U.; Thakkar, V.M. Leaf classification and identification using Canny Edge Detector and SVM classifier. In Proceedings of the 2017 International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 19–20 January 2017; pp. 1–4.
48. Srivastava, V.; Khunteta, A. Comparative Analysis of Leaf Classification and Recognition by Different SVM Classifiers. In Proceedings of the 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 11–12 July 2018; pp. 626–631.
49. Araújo, V.; Britto, A.S.; Brun, A.L.; Koerich, A.L.; Falate, R. Multiple classifier system for plant leaf recognition. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 1880–1885.
50. Goeau, H.; Bonnet, P.; Joly, A.; Nozha, B.; Daniel, B.; Jean-Francois, M.; Philippe, B.; Elise, M.; Marie, P. The Image CLEF 2011 plant images classification task. In Proceedings of the CLEF'2011: Conference and Labs of the Evaluation Forum, Amsterdam, The Netherlands, 19–22 September 2011; pp. 1–23.
51. Goeau, H.; Bonnet, P.; Joly, A.; Yahiaoui, I.; Barthelemy, D.; Boujemaa, N.; Molino, J.F. The Image CLEF 2012 plant identification task. In Proceedings of the CLEF'2012: Conference and Labs of the Evaluation Forum, Rome, Italy, 17–20 September 2012.
52. Mittal, P.; Kansal, M.; Jhaji, H.K. Combined classifier for plant classification and identification from leaf image based on visual attributes. In Proceedings of the 2018 International Conference on Intelligent Circuits and Systems (ICICS), Phagwara, India, 19–20 April 2018; pp. 188–194.
53. Eid, H.F.; Hassani, A.E.; Kim, T.H. Leaf Plant Identification System Based on Hidden Naïve Bays Classifier. In Proceedings of the 2015 4th International Conference on Advanced Information Technology and Sensor Application (AITS), Harbin, China, 21–23 August 2015; pp. 76–79.
54. James Nesaratnam, R.; Balamurugan, C. Identifying leaf in a natural image using morphological characters. In Proceedings of the 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 19–20 March 2015.
55. Ehsanirad, A. Plant Classification Based on Leaf Recognition. *Int. J. Comput. Sci. Inf. Secur.* **2010**, *8*, 78–81.
56. Zhang, S.; Chau, K. Dimension Reduction Using Semi-Supervised Locally Linear Embedding for Plant Leaf Classification. *Int. Conf. Intell. Comput.* **2009**, *5754*, 948–955.
57. Kherkhah, F.M.; Asghari, H. Plant leaf classification using GIST texture features. *IET Comput. Vis.* **2019**, *13*, 369. [[CrossRef](#)]
58. Rahmani, M.E.; Amine, A.; Hamou, M.R. Plant Leaves Classification. *ALLDATA* **2015**, *82*, 75–80.
59. Kumar, P.S.; Rao, K.N.V.; Raju, A.S.N.; Kumar, D.J.N. Leaf classification based on shape and edge feature with k-NN classifier. In Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, India, 14–17 December 2016; pp. 548–552.
60. Shah, M.P.; Singha, S.; Awate, S.P. Leaf classification using marginalized shape context and shape+ texture dual-path deep convolutional neural network. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 860–864.
61. Pankaja, K.; Suma, V. Leaf Recognition and Classification Using GLCM and Hierarchical Centroid Based Technique. In Proceedings of the 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 11–12 July 2018; pp. 1190–1194.
62. Chaki, J.; Parekh, R.; Bhattacharya, S. Plant leaf recognition using texture and shape features with neural classifiers. *Pattern Recognit. Lett.* **2015**, *58*, 61–68. [[CrossRef](#)]

63. Wable, P.B.; Chilveri, P.G. Neural network based leaf recognition. In Proceedings of the 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune, India, 9–10 September 2016; pp. 645–648.
64. Chaki, J.; Parekh, R.; Bhattacharya, S. Plant leaf recognition using a layered approach. In Proceedings of the 2016 International Conference on Microelectronics, Computing and Communications (MicroCom), Durgapur, India, 23–25 January 2016; pp. 6–11.
65. Harish, B.S.; Hedge, A.; Venkatesh, O.; Spoorthy, D.G.; Sushma, D. Classification of plant leaves using Morphological features and Zernike moments. In Proceedings of the 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, India, 22–25 August 2013; pp. 1827–1831.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).