

Computer-aided Interpretable Features for Leaf Image Classification

Abstract

Plant species identification is time consuming, costly, and requires lots of efforts, and expertise knowledge. The studies on medicinal plant identification are often performed based on medicinal plant leaf images. This is because plant leaves contain a large number of diverse set of features such as shape, veins, edge features, apices etc. that are useful in identifying different medicinal plants. In recent, many researchers use deep learning methods to classify plants directly using plant images. While deep learning models have achieved a great success, the lack of interpretability limit their widespread application. To overcome this, we explore the use of interpretable, measurable and computer-aided features extracted from plant leaf images. Image processing is one of the most challenging, and crucial steps in feature-extraction. The purpose of image processing is to improve the leaf image by removing undesired distortion. The main image processing steps of our algorithm involves: i) Convert original image to RGB (Red-Green-Blue) image, ii) Gray scaling, iii) Gaussian smoothing, iv) Binary thresholding, v) Remove stalk, vi) Closing holes, and vii) Resize image. The next step after image processing is to extract features from plant leaf images. We introduced 52 computationally efficient features to classify plant species. These features are mainly classified into four groups as: i) shape-based features, ii) color-based features, iii) texture-based features, and iv) scagnostic features. Length, width, area, texture correlation, monotonicity and scagnostics are to name few of them. We explore the ability of features to discriminate the classes of interest under supervised learning and unsupervised learning settings. For that, supervised dimensionality reduction technique, Linear Discriminant Analysis (LDA), and unsupervised dimensionality reduction technique, Principal Component Analysis (PCA) are used to convert and visualize the images from digital-image space to feature space. All the applications are performed on Flavia and Swedish open source benchmark leaf datasets. The results show that the features are sufficient to discriminate the classes of interest under both supervised and unsupervised learning settings. The results of this study are beneficial for the researchers working in the field of developing automated plant identification and classification systems.

Keywords: Medicinal, Leaf images, Image processing, Feature extraction, LDA, PCA

1 Introduction

Leaf identification is becoming very popular in classifying plant species. Plant leaf contains significant number of features that can help people to identify and classify the plant species. In medical perspective, medicinal plants are usually identified by practitioners based on years of experience through sensory or olfactory senses. The other method of recognizing these plants involves laboratory-based testing, which requires trained skills, data interpretation which is costly and time-intensive. Automatic ways to identify medicinal plants are useful especially those that are lacking experience in medicinal plant recognition. Statistical machine learning techniques play a crucial role in the development of automatic system to identify medicinal plants. In developing such system input features play an important role. The main aim of this paper is to introduce interpretable features that can be computed based on plant leaf images. Image processing and feature extraction play crucial roles in developing a workflow to achieve the aim of this research.

The main aim of image processing is to extract important features by removing undesired noise and distortion (Wäldchen and Mäder 2018). Image processing steps include image segmentation (Anantrasirichai, Hannuna, and Canagarajah 2017), image orientation, cropping, grey scaling, binary thresholding, noise removal, contrast stretching, threshold inversion, image normalization, and edge recognition are some of image processing techniques applied in recent research. These steps can be applied parallel or individually, on several times until the quality of leaf image reaches a specific threshold.

The second step is feature extraction, which identifies and encodes relevant features from leaf images (Wäldchen and Mäder 2018). This is a challenging task due to the structural diversity of the leaf images. Therefore, in recent years, many researchers use deep learning methods to classify plants directly using plant images (Wu et al. 2007a; Azlah et al. 2019; Herdiyeni and Wahyuni 2012). While deep learning models have achieved a great success, the most models remain complex black boxes. The lack of interpretability limit their widespread application.

A digital image is a combination of pixels from three different color planes red, green and blue. Images are stored in computers as three separate matrices corresponding to the red, green and blue color channels of the image. The three separate matrices corresponding to intensities of colors at different positions of the image (Gonzalez and Woods 2006).

The main aim of feature extraction is to reduce dimensionality of this information by obtaining measurable patterns of leaf images. For example, shape, color, and texture are some of the patterns that may be observed. In this paper we introduce a collection of interpretable, measurable and computer-aided features that are useful in image leaf classification. This feature collection includes several pre-established features identified

through a thorough review of literature. Other than existing features, we introduce a number of features computed based on the Cartesian coordinate of the images. Furthermore, we explore the ability of features to discriminate the class of interest under supervised learning setting and unsupervised learning setting.

The paper is organized as follows. Section 2, describes about the steps of image processing. Preprocessing is necessary before extracting features from the images. Section 3, discusses about feature extraction in in-detail because features are highly influenced by the plant species to be classified. Under this section, we discuss about mainly four types of features as shape, color, texture, and scagnostics and how to extract them. Section 4, present empirical application. This section consists of details about the datasets that are used to explain the applications, and visualization of leaf images in the feature space using supervised, and unsupervised dimensionality reduction techniques. In Section 5, consist of summary of the software and packages that use to extract the features. Some discussion about the outputs and concluding remarks are given in last section.

2 Image Processing

Image processing is an essential step to reduce noise and content enhancement while keeping its features intact. (Goyal, Kapil, and Kumar 2018). The workflow we use to process images in this paper is shown in Figure 1. This includes seven main steps. They are: i) converting BGR (Blue-Green-Red) image to RGB (Red-Green-Blue), ii) gray scaling, iii) Gaussian filtering, iv) binary thresholding, v) remove stalk, vi) close holes, and vii) image resizing. Some of these steps are applicable only for specific images. For example, apply remove stalk is applicable only to leaf images which has stalk.

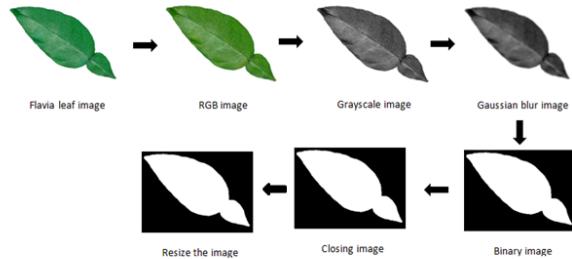


Figure 1: Image processing workflow.

In this study we focus on leaves with simple arrangement as shown in Figure 2c. A single leaf that is never divided into smaller leaflet units is known as a leaf with simple arrangement. This type of leaf is attached to a twig by its stem or the petiole. The margins, or edges, of the leaf can be smooth, lobed, or

toothed (see Figure 2a).

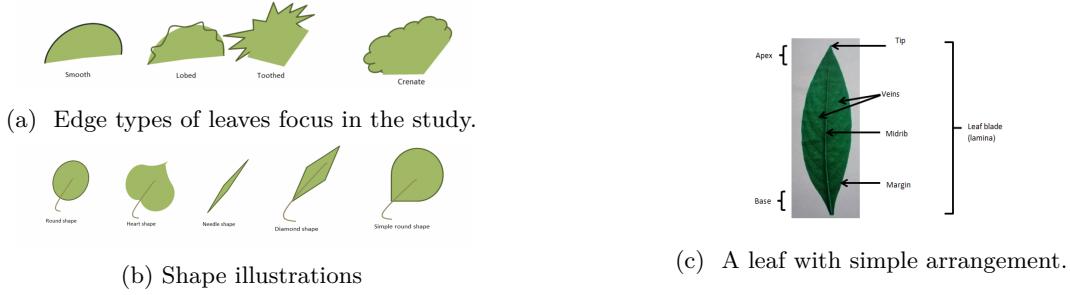


Figure 2

2.1 Step 1: Converting BGR (Blue-Green-Red) Image to RGB (Red-Green-Blue)

BGR and RGB are conventions for the order of the different colour channels. They are not colour spaces. When converting BGR image to RGB, there is no any computations, just switches around the order. Several image processing libraries have different pixel ordering. Therefore to compatible with other libraries we convert the BGR image in to RGB format. For an example when we read an image using OpenCV library in Python by default it interprets BGR format, but when we plot the image it takes the RGB format in matplotlib package in Python.

2.2 Step 2: Grayscaleing

Grayscaleing is the process of converting an image to shades of gray from other colour spaces like RGB. This helps to increase the contrast and intensity of images (Goyal, Kapil, and Kumar 2018). Gray scale images require only one single byte for each pixel where as colour (RGB) image requires 3 bytes for each pixel. Hence, grayscaling reduces the dimension of a image.

In extracting Haralick (Haralick, Shanmugam, and Dinstein 1973) texture features, gray-scale images are used. Another advantage of using gray-scaled image is to reduce model complexity. Consider an example of training neural article on RGB images of $20 \times 20 \times 3$ pixel. The input layer will have 1200 input nodes. Whereas for gray scaled images the same neural network will need only 400 (20×20) input node.

2.3 Step 3: Gaussian Filtering (Gaussian Blurring/Gaussian Smoothing)

Gaussian smoothing is an image smoothing technique. Image smoothing techniques use to remove the noise that can be occurred because of the source (camera sensor). Image smoothing techniques help in smoothing images and to remove low intensity edges.

Gaussian function is used to blur the image. It is a linear filter which is done by using the functions in OpenCV package in Python. By specifying the width and height of the Gaussian kernel that must be positive and odd, and specifying the kernel standard deviation along x and y-axis, Gaussian smoothing is established in OpenCV. When the kernel standard deviation along x-axis is specified, kernel standard deviation along y-axis is taken as equal to the the kernel standard deviation along x-axis. But if both kernel standard deviation are given as zeros, they are calculated by using the kernel size. In our research the width and height of the kernel is defined as 55 and the kernel standard deviation along x-axis is assigned as zero.

An example of applying Gaussian smoothing to an image is shown in Figure 3.

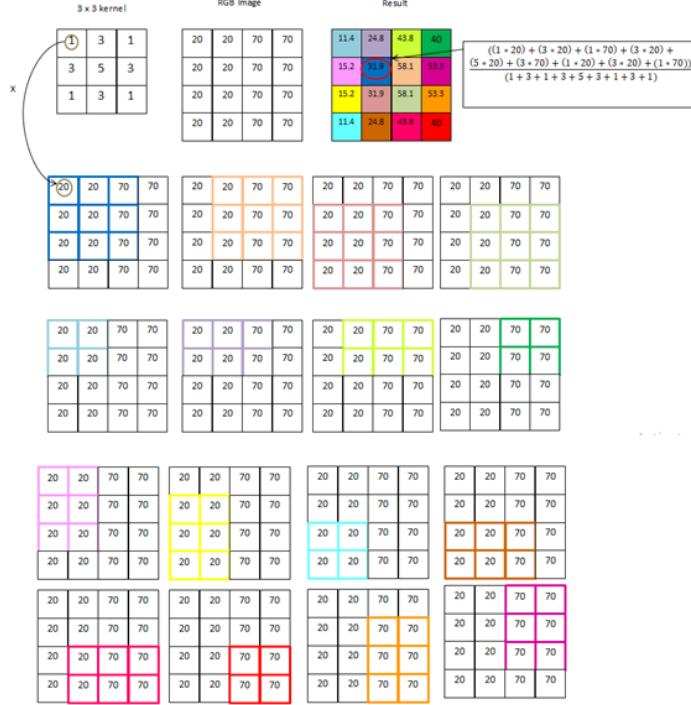


Figure 3: Example of gaussian smoothing

2.4 Step 4: Binary Thresholding

Thresholding is a segmentation technique that is used to separate foreground from its background. Thresholding converts gray-scale image into binary image according the threshold value. If the pixel value is smaller than the threshold value, the pixel value is set as 0, and if not the pixel value is set to a maximum value which is generally 255. Thersholding technique is done on grayscale images in computer vision.

We used Otsu's binarization (Bangare et al. (2015)) which is an adaptive thresholding after Gaussian filtering to convert color images to the binary images. The reason for using Otsu's binarization method is, it automatically determines the optimal threshold value.

Otsu's Thresholding

Nobuyuki Otsu introduced Otsu's method which is defined for a histogram of grayscale values of a histogram ($ghist_I$) of an input image Im . To segment an image Im into two subsets of pixels Otsu's method calculates an optimal threshold τ . The number of pixel locations of the gray scale image is defined as $|\Omega|$. The algorithm maximizes the variance σ^2 between the two subsets (Within-class-variance) to find the threshold τ . The variance σ^2 is defined as

$$\sigma^2 = P_1(\mu_1 - \mu)^2 + P_2(\mu_2 - \mu)^2 = P_1P_2(\mu_1 - \mu_2)^2,$$

where μ is the mean of the histogram, μ_1 and μ_2 are the mean values of first and second subset respectively. The corresponding class probabilities, P_1 and P_2 are defined as follows,

$$P_1 = \frac{\sum_{i=0}^u ghist_I(i)}{|\Omega|}, \text{ and } P_2 = \frac{\sum_{i=u+1}^{255} ghist_I(i)}{|\Omega|},$$

where u is the candidate threshold and the maximum gray level (G_{max}) is assumed as 255. To find optimal threshold τ for segmenting image Im , all candidate thresholds are evaluated this way.

The algorithm of Otsu's method is defined as follows,

For an example, assume that candidate threshold value u is 2. Therefore, the image is separated into two classes, which are class 1 (pixel value ≤ 2) and class 2 (pixel value > 2). Class 1 represents the background and class 2 represents the foreground of the grayscale image. According to Figure 4a, there are 9 pixel locations. The associated measurements are

Create a histogram for the grayscale image
Set the histogram variance $S_{max} = 0$

while $u < G_{max}$ **do**

 Compute $\sigma^2 = P_1 * P_2(\mu_1 - \mu_2)^2$
if $\sigma^2 > S_{max}$ **then**

$S_{max} = \sigma^2$
 $\tau = u$

end if
 Set $u = u + 1$

end while

Table 1: Otsu's method

$$P_1 = \frac{5}{9}, P_2 = \frac{4}{9}, \mu_1 = \frac{(0 * 2) + (1 * 1) + (2 * 2)}{(2 + 1 + 2)} = 1, \mu_2 = \frac{(3 * 3) + (4 * 1)}{(3 + 1)} = \frac{13}{4},$$

$$\sigma^2 = \frac{5}{9} * \frac{4}{9} * (1 - \frac{13}{4})^2 = 1.25.$$

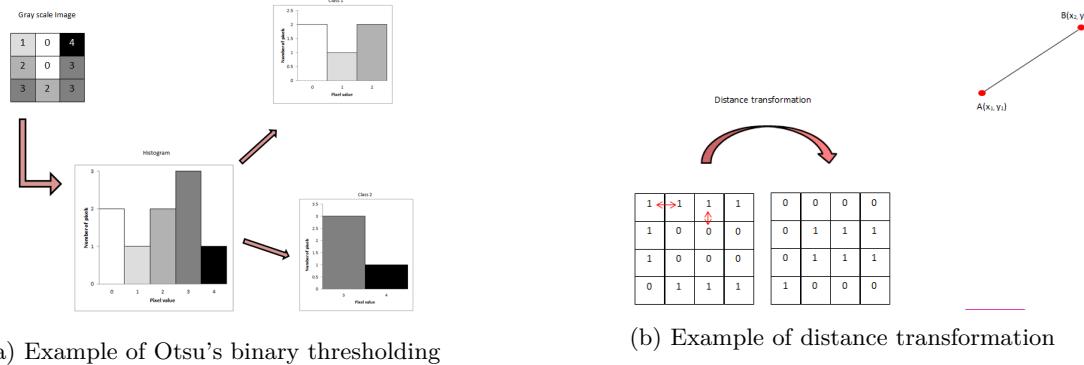


Figure 4

2.5 Step 5: Image Resizing

In this study we use two different benchmark leaf image datasets: i) Flavia: 1907 fully color images of 32 classes of leaves (Wu et al. (2007b)) and ii) Swedish: 1125 images from 15 different plant species (Söderkvist (2001)). Flavia and Swedish images have two different image sizes. To compare the results on different datasets, to improve the memory storage capacity and to reduce computational complexity the leaf images are resized to a fixed resolution. In our study, the leaf images have been resized to [1600 x 1200px] which is the size of Flavia leaf images.

Other than the main image processing techniques discussed above, the following techniques are applied

to some images where necessary after image thresholding.



Figure 5

2.5.1 Remove Stalk

As shown in Figure 5a, this is used to remove the stalk of the image. Remove the petiole (stalk) of leaf image is another version of thresholding process. Thresholding is applied after finding the sure foreground area. To find the sure foreground area, distance transform technique is used. Binary image is used as the input of distance transform technique. In distance transform technique, image is created by assigning a number for each object pixel that corresponds to the distance to the nearest background pixel. The distance is calculated using the Euclidean distance (see Figure 4b). After finding the sure foreground area, Otsu's binarization is applied again as the thresholding technique. The Euclidean distance is defined as

$$\text{Euclidean distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (1)$$

2.5.2 Closing Holes

Holes inside leaf areas occur due to plant disease, light reflection and noise. As shown in 5b closing holes is used to remove small holes inside the foreground objects. This is achieved by pulling background pixels to foreground pixel. The closing holes is also known as dilation and while erosion does the opposite, which is open holes. This step is performed on a binary image.

Erosion

The basic idea of Erosion is that erodes away the boundaries of foreground object. Since the input is binary image, a pixels in the original image is either 1 or 0. If all the pixels under the kernel is 1, a pixel of original image is considered as 1, otherwise made to zero (eroded). Which means that depending upon the

size of the kernel all pixels near boundary will be discarded. Therefore the thickness or size of the foreground object decreases (White region of the image decreases).

Dilation

Opposite of erosion is defined as dilation. If at least one pixel under the kernel is 1, the pixel element is 1 in Dilation. It tends to increase the foreground of the image or the white region of the object.

3 Leaf Image Features

In identification of plant species using leaf images, features of leaves play an important role, because each leaf posses unique feature that it make different from other. In previous studies (Azlah et al. 2019; Mingqiang, Kidiyo, and Joseph 2008; Sun et al. 2017), use deep learning neural networks to classify medicinal plants based on pixel-based images. Given the leaf images deep learning models automatically identify features based on pixel-space of the images. While deep learning models have achieved a great success, the lack of interpretability of features limit their widespread application. To overcome this, we explore the use of interpretable, measurable and computer-aided features extracted from plant leaf images. We identified 52 features. The features are classified into four groups as: i) shape-based features, ii) color-based features, iii) texture-based features, and iv) scagnostic features.

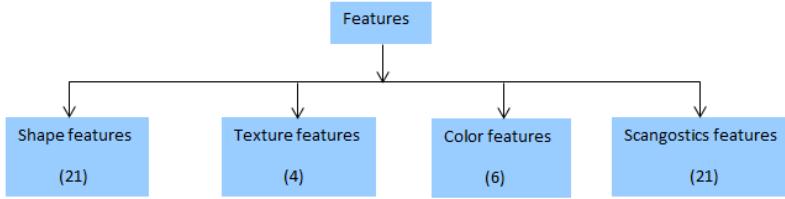


Figure 6: Classification of features and their composition

3.1 Shape Features

When identifying real-world objects, the shape is known as an essential sign for humans. We use the shape descriptors introduced by Wäldchen and Mäder (2018). In addition to that we introduce a number of new shape features such as: number of convex points, x and y coordinates of center, number of maximum and minimum points, correlation of Cartesian contour points, etc. The shape features should be invariant

to certain class of geometric transformation of the object. The main geometric transformation are rotation reflection, scaling and translation (see Figure 7). The shape features we considered in this study are invariant to the rotation and reflection. All the shape features are extracted from the binary images.

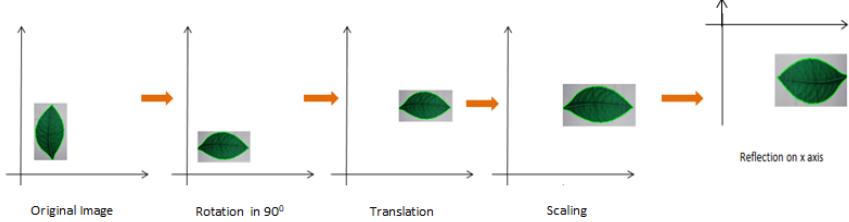


Figure 7: Illustration of geometric transformation

Extraction of image contour plays an important role in measuring the shape of a image. Simply contour (see Figure 8) is a curve joining all the continuous points (along the boundary), having the same color or intensity.

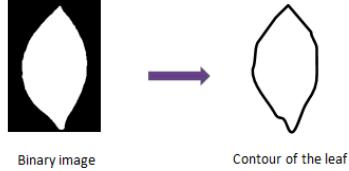


Figure 8: Extract contour of the leaf image

In order to extract the contour of the leaf, the leaf should be placed properly in the center of a white paper. If the image is place as shown in Figure 9, as a result of inappropriate translation (a) and inappropriate scaling (b), problems arises in the calculations of the contour. Furthermore, it is difficult to recognize the contour, when the image is too small.

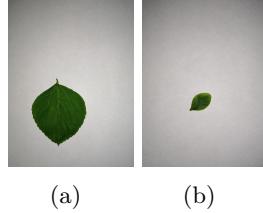


Figure 9: Inappropriate translation (a), Inappropriate scaling (b)

Shape features can be classified into two main categories as contour-based and region-based features (Wäldchen and Mäder 2018). As illustrated in Wäldchen and Mäder (2018), contour-based shape features are computed based on the contour of a shape, whereas region-based shape features are extracted from the whole region of a shape (see Figure 10).

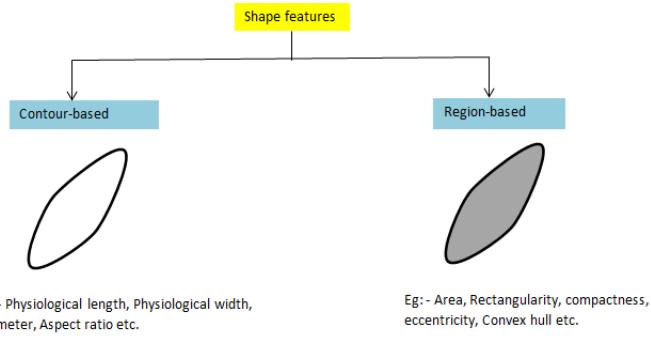


Figure 10: Categorization of shape features

In this study, we use 6 initial shape features that are used to derive 12 shape features. The 6 initial features are: i) diameter, ii) physiological length, iii) physiological width, iv) area, v) perimeter, and vi) eccentricity.

3.1.1 Diameter (F_1)

Diameter is defined as the longest distance between any two points on the margin of the leaf (Wäldchen and Mäder 2018). To calculate the diameter of the leaf image, first we need to find the contour of the leaf image. Then we need to select all pair of contour points and measure the Euclidean distance (equation 2) between the two points separately. Finally, we have to find the maximum distance among the calculated distances (see Figure 11). The distance measurement is

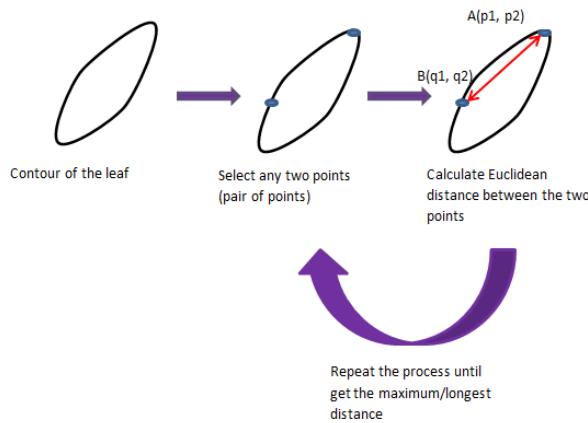


Figure 11: Logic behind calculation of diameter

$$d(A, B) = \sqrt{\sum_{i=1}^2 (q_i - p_i)^2}, \quad (2)$$

$$F_1 = \max(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}); \forall i, j, i \neq j. \quad (3)$$

3.1.2 Physiological length (F_2) and Physiological width (F_3)



Figure 12: Straight (horizontal or vertical) and rotated leaf image in Actual leaf image dataset

According to the authors in Azlah et al. (2019), the physiological length is “measured based on the main vein of leaf, as it stretches from the main vein to the end tip”. We use the definition by Azlah et al. (2019), the physiological width is “the span of leaf viewed from one side to the other, from the leftmost point to the rightmost point of leaf” . There are straight (horizontal or vertical) and angled leaf images in our datasets, Flavia and Swedish (see example Figure 12). There are two types of bounding rectangles.

- i) Straight bounding rectangle: This is a straight rectangle which does not consider the rotation of the object (see Figure 12 (a)).
- ii) Rotated rectangle: This bounding rectangle is drawn with minimum background area. Therefore the rotation of the object is also considered (see Figure 12 (b)).

The straight bounding rectangle is enough to extract physiological length and physiological width of straight (horizontal or vertical) leaf images. However, the straight bounding rectangle is not suitable to compute physiological length and physiological width of angled leaf images. To solve this problem, we considered rotated rectangle rather than bounded rectangle in computing shape features of angled images. As shown in Figure 13, physiological length and width are computed as follows:

F_2 = Length of the rotated rectangle,

(4a)

F_3 = Width of the rotated rectangle.

(4b)

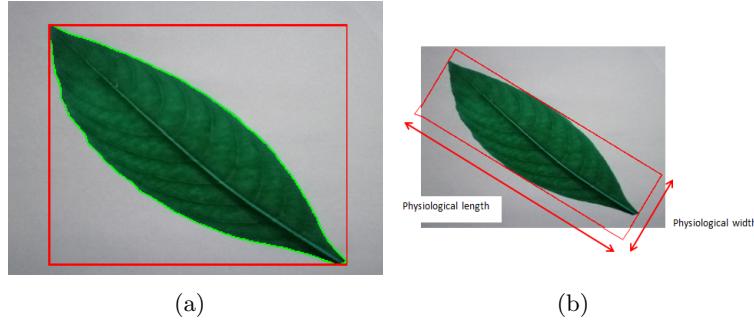


Figure 13: Straight bounded rectangle of a rotated leaf image (a), Rotated rectangle of angled leaf image (b)

3.1.3 Area (F_4)

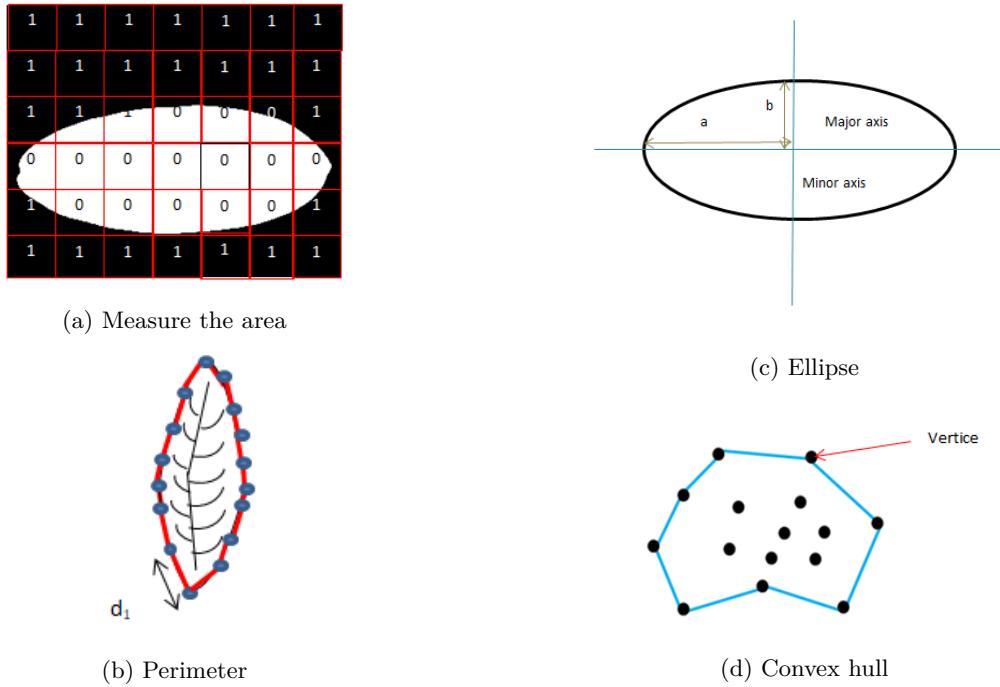


Figure 14

Area is computed after applying the thresholding process. Next, we need to extract the best contour and based on that contour area is measured. Number of 0 pixels covered by the contour is the measure of area of

leaf image (Azlah et al. 2019). As shown in Figure 15a the area measures increase as the area of the leaf increases. The area is defined by

$$F_4 = \text{Number of zero pixels covered by the contour.} \quad (5)$$

3.1.4 Perimeter (F_5)

As shown in Figure 14b perimeter is defined as the summation of euclidean distance of all continuous points in the contour. Let n be the number of distances around the contour, then the perimeter is defined as

$$F_5 = \sum_{i=0}^n d_i. \quad (6)$$

3.1.5 Eccentricity (F_6)

Eccentricity is a characteristic of any conic section of a leaf (Azlah et al. 2019). Eccentricity is defined that how much the ellipse actually varying being circular. Eccentricity is calculated using equation 7 as

$$F_6 = \sqrt{1 - \frac{b^2}{a^2}}. \quad (7)$$

where a is semi-major axis and b is semi-minor axis (see Figure 14c)

Eccentricity of an ellipse is varied between 0 and 1. If eccentricity is 0, then we obtain a circle whereas eccentricity is 1, then we obtain an ellipse (see Figure 15b).

3.1.6 Number of Convex Points (F_{21}) and Verticies

Wilkinson, Anand, and Grossman (2005) defined a convex hull as a “hull that contains all the straight line segments connecting any pair of points in its interior”. The convex hull bounds a single polygon (see Figure 14d). We introduced two new features computed based on the convex hull: i) Number of convex points (F_{21}) and ii) Number of convex vertices (F_7).

3.1.7 Roundness/ Circularity

Roundness is named as aka form factor, circularity, or isoperimetalical factor. Roundness illustrates the difference between the leaf and a circle. Equation 8 is used to calculate roundness. Figure 15c visual

explanation for roundness measure. The roundness is measured by

$$R = \frac{4\pi F_4}{F_5^2}. \quad (8)$$

3.1.8 Compactness

Compactness is closely related to roundness. Compactness measures that how compatible the leaf fits to a circle area (see Figure 15d). The compactness is measured using

$$C = \frac{F_5^2}{F_4}. \quad (9)$$

3.1.9 Convexity

Convexity measures the curvature of the convex hull.

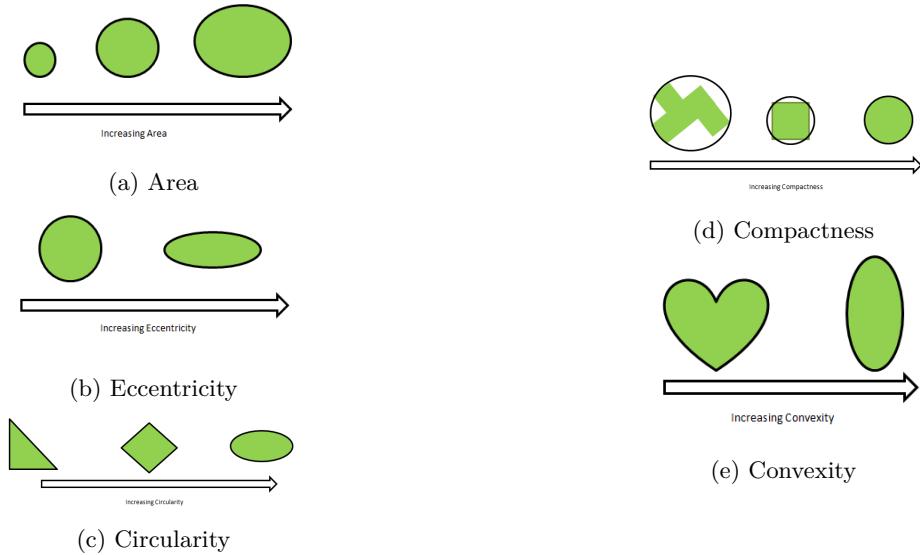


Figure 15

The equations and software packages used to compute shape features are shown in Table 2.

Shape feature	Feature name	Figure	Formula	Range	Software package
F_1	Diameter		eq:3	$[0, \infty]$	combinations, numpy
F_2	Physiological length		eq:4a	$[0, \infty)$	OpenCV
F_3	Physiological width		eq:4b	$[0, \infty)$	OpenCV
F_4	Area		eq:5		OpenCV
F_5	Perimeter		eq:6	$[0, \infty)$	OpenCV
F_6	Eccentricity		eq:7	$[0, 1]$	OpenCV
F_7, F_8	x and y coordinate of center				scipy.ndimage
F_9	Aspect ratio		$F_9 = \frac{F_2}{F_3}$	$[0, \infty)$	-
F_{10}	Roundness/ Circularity		eq:8	$[0, \infty)$	numpy
F_{11}	Compactness		eq:9	$(0, \infty)$	-
F_{12}	Rectangularity		$F_{12} = \frac{F_5^2}{F_4}$	$(0, \infty)$	-
F_{13}	Narrow factor		$F_{13} = \frac{F_1}{F_2}$	$[0, \infty)$	-
F_{14}	Perimeter ratio of diameter		$F_{14} = \frac{F_5}{F_1}$	$[0, \infty)$	-

Table 2 continued from previous page

Shape feature	Feature name	Figure	Formula	Range	Software package
F_{15}	Perimeter ratio of physiological length		$F_{15} = \frac{F_5}{F_2}$	$[0, \infty)$	-
F_{16}	Perimeter ratio of physiological length and width		$F_{16} = \frac{F_5}{F_2 * F_3}$	$[0, \infty)$	-
F_{17}	Perimeter convexity		$F_{17} = \frac{\text{Perimeter of convex hull}}{F_5}$	$[0, \infty)$	OpenCV
F_{18}	Area convexity		$F_{18} = \frac{(\text{Area of convex hull} - F_4)}{F_4}$	$[0, \infty)$	OpenCV
F_{19}	Area ratio of convexity		$F_{19} = \frac{F_4}{\text{Area of convex hull}}$	$[0, \infty)$	OpenCV
F_{20}	Equivalent diameter		$F_{20} = \sqrt{\frac{4 * F_4}{\pi}}$	$[0, \infty)$	numpy
F_{21}	Number of convex points		number of convex points	$[0, \infty)$	OpenCV

Table 2: Summary of shape features.

3.2 Texture Features

Texture features are used to describe the surface or the appearance of the leaf image. Texture can be assessed using a group of pixels. Color is a property of a pixel. Texture is defined as feel of various materials to human touch and texture is quantified based on visual interpretation of this feeling. Leaf surface is a natural texture which has random persistent patterns and do not show detectable quasi-periodic structure (Wäldchen and Mäder 2018). Therefore to describe the natural texture patterns of the leaf fractal theory (Wäldchen and Mäder 2018) is the best approach.

The Haralick texture features (Dean 1999) are functions of the normalized GLCM (Gray Level Co-occurrence Matrix) which is a common method to represent image texture.

$$GLCM = \begin{bmatrix} h(1,1) & h(1,2) & \dots & \dots & h(1,n) \\ h(2,1) & h(2,2) & \dots & \dots & h(2,n) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ h(n,1) & h(n,2) & \dots & \dots & h(n,n) \end{bmatrix}$$

The GLCM is square with dimension n , where n is the number of gray levels in the image. Let $h(a,b)$ is the probability that a pixel with value a will be found adjacent to a pixel of value b . Then $h(a,b)$ is defined as

$$h(a,b) = \frac{\text{Number of times a pixel with value } a \text{ is adjacent to a pixel with value } b}{\text{Total number of such comparisons made}}$$

In order to calculate $h(a,b)$, adjacency can be defined to occur in each of four directions in a 2D (see Figure 18), square pixel image (horizontal, vertical, left and right diagonals - see equation 16).

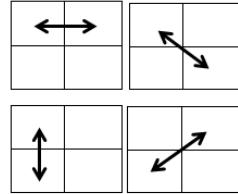


Figure 16: Four directions of adjacency as defined for calculation of the Haralick texture features

The Haralick statistics (Dean 1999) are calculated based on the matrices generated using each of these directions (see Figure 17) of adjacency. Haralick introduced 14 statistics to describe the texture of the image based on the four co-occurrences matrices generated. In this research, we only used the following 4 statistics among 14 of them, because most of the researchers used these 4 statistics as texture features (see Figure 3) of leaf images. All the texture features are extracted from the gray scale image. Texture features are calculated from the mahotas package in Python. Table 3 shows the definitions of texture features.

Let $h(a,b) = \text{Probability density function of gray - level pairs } (a,b)$ and dimension of GLCM is $n \times n$ (Number of columns \times Number of rows). The associated measures are given by

Texture feature	Feature name	Detailed description	Formula	Value range
F_{22}	Contrast	Measures the relation or difference between the highest and lowest gray levels of the GLCM	$\frac{\sum_{a=1}^{columns} \sum_{b=1}^{rows} (a-b)^2 h(a,b)}{\text{Number of gray levels}-1}$	$[0, \infty]$
F_{23}	Entropy	Measures the randomness which means that how uniform the image is Measurement of dependence of gray levels of the GLCM.	$-\sum_{a=1}^{columns} \sum_{b=1}^{rows} h(a,b) \log_2(h(a,b))$	$[-\infty, 0]$
F_{24}	Correlation	It measures that how a particular pixel is correlated to its neighbor pixel over the whole image	$\frac{\sum_{a=1}^{columns} \sum_{b=1}^{rows} (ab)h(a,b) - \mu_x \mu_y}{\sigma_x \sigma_y}$	$[-1, 1]$
F_{25}	Inverse difference moments	It's a measure of homogeneity. Inverse level of contrast that measures how close the values of GLCM to diagonal values in GLCM	$\sum_{a=1}^{columns} \sum_{b=1}^{rows} \frac{h(a,b)}{(a-b)^2}$	$[0, \infty]$

Table 3: Definitions of texture features

$$\mu_x = \sum_{a=1}^{columns} a \sum_{b=1}^{rows} h(a,b), \mu_y = \sum_{b=1}^{rows} b \sum_{a=1}^{columns} h(a,b),$$

$$\sigma_x = \sum_{a=1}^{columns} (a - \mu_x)^2 \sum_{b=1}^{rows} h(a,b), \sigma_y = \sum_{b=1}^{rows} (b - \mu_y)^2 \sum_{a=1}^{columns} h(a,b).$$

3.3 Color Features

Color is an important characteristic of images (Wäldchen and Mäder 2018; Caglayan, Guclu, and Can 2013). Some of leaf images have very similar shape like Hathawariya (Figure 19) and Iramusu (Figure 19). Even though shapes are similar in some leaves, there are some differences in colors of leaf images. Therefore in addition to the shape features, we extract features related to the color. Colors of an image are formed based on Red-Green-Blue (RGB) colour channels of an image. Color properties are defined within a particular color channel (Kodituwakku and Selvarajah 2004; Wäldchen and Mäder 2018). In the field of image recognition, a number of general color descriptors have been introduced. Color moments (Kodituwakku and Selvarajah 2004; Wäldchen and Mäder 2018) are the simple descriptor among them. Mean, standard deviation skewness and kurtosis are the common moments. Color moments are convenient for real-time applications due to its low dimension and low computational complexity.

We used mean(M) and standard deviation(SD) of intensity values of red, green and blue channels as color features. Let h be the number of pixels of the image and r is the channel type which can be red, green, or blue. The corresponding colour features are calculated as follows:

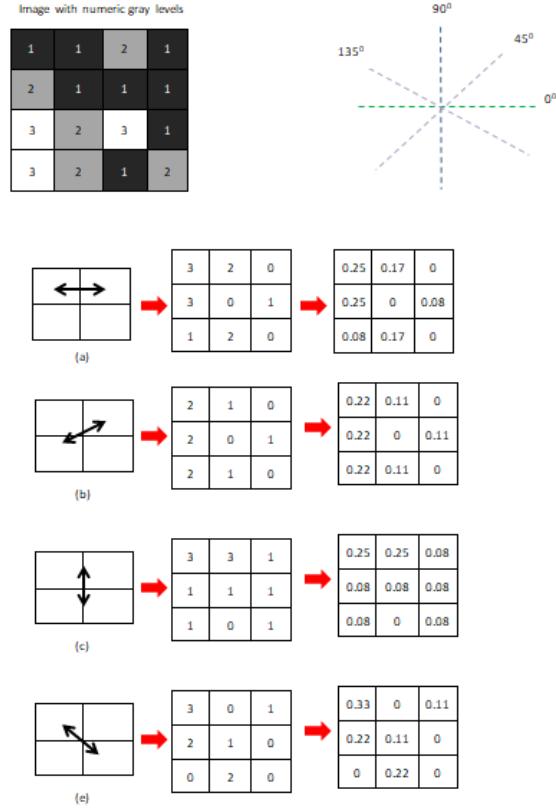


Figure 17: Computing the Haralick texture features from a 4×4 example image step by step

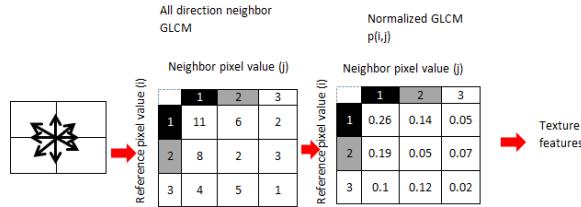


Figure 18: Computing the Haralick texture features from a 4×4 example image with all direction



Figure 19: Images of plant species Hathawariya and Iramusu.

$$M = \frac{\text{Total insensity value of } r^{\text{th}} \text{ channel of the image pixels}}{\text{Total intensity value of the image}}, \quad (10)$$

$$SD = \frac{\sqrt{\sum_{j=0}^h (r^{\text{th}} \text{ channel intensity}_j - r^{\text{th}} \text{ mean value})^2}}{\text{Total intensity value of the image}}. \quad (11)$$

3.4 Scagnostic features

Scagnostic features are used to quantify the characteristics of 2D scatterplot diagrams (Wilkinson and Wills (2008)). Scagnostic measures are calculated based on the appearance of scatterplot. All the scagnostics features are calculated by using the R package called `binostics`. Scagnostic features has the range of [0,1]. There are 9 measures which are classified into three categories as shown in Figure 20. Based on binary images, scagnostic features are extracted. To the best of our knowledge this is the first time, the scagnostic features are used for image recognition.

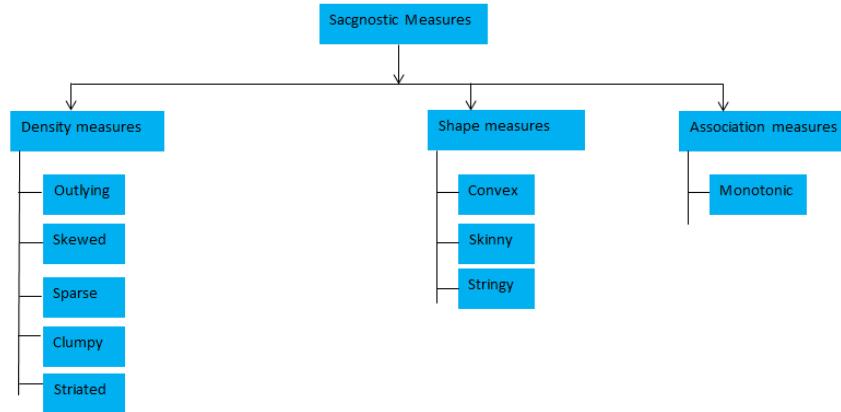


Figure 20: Hierarchy of Scagnostics

We separately measure the scagnostic features based on cartesian and polar coordinate of the contour (see Figure 21a). As the first step, we have to extract the contour of the leaf image (see Figure 21b). Then find the x and y coordinate values of the cartesian and polar separately. The x and y coordinate value is used to calculate the scagnostic features. The following definitions can be useful in understanding scagnostics features.

Geometric Graphs

- i) Graph: A graph $Gr = (Ve, Ed)$ is defined as a set of vertices (Ve) together with a relation on Ve

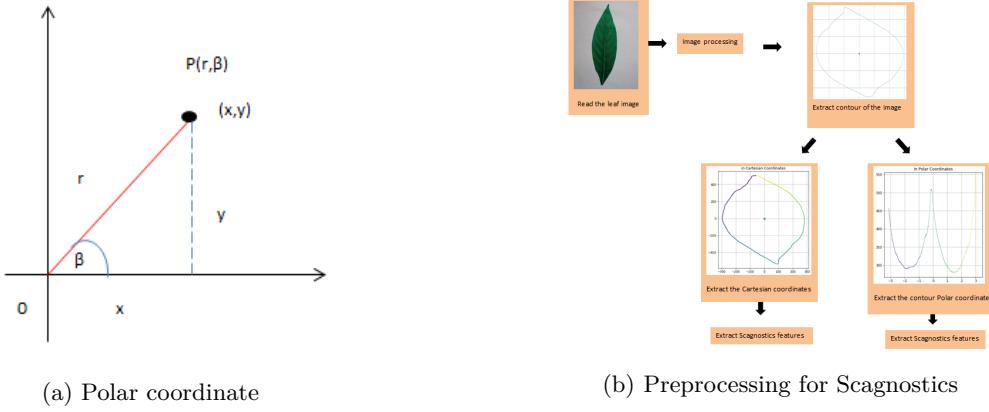


Figure 21

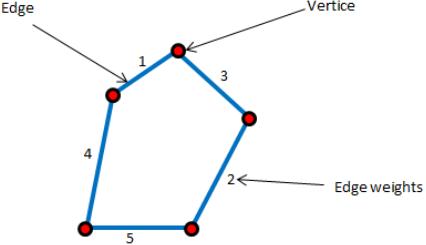
induced by a set of edges (Ed). A pair of vertices is defined as an edge $e(\nu, \omega)$, with $e \in Ed$ and $\nu, \omega \in Ve$.

- ii) Geometric Graph: A geometric graph $G* = [f(Ve), g(Ed), S]$; is a mapping of vertices to points and edges to straight lines to connect points in a metric space S .
- iii) Length of an edge: The Euclidean distance between vertices that connected to edge is defined as the length of an edge, $length(e)$.
- iv) Length of a graph: The sum of the lengths of edges in graph is known as the length of a graph, $length(Gr)$.
- v) Path: A list of successively adjacent, distinct edges are known as a path. If first and last vertices coincide, then the path is closed.
- vi) Polygon: A region bounded by a closed path is known as a polygon (P). A polygon bounded by exactly one closed path with no intersecting edges is known as a simple polygon.
- vii) Perimeter of a simple polygon: The length of boundary of a simple polygon is known as the perimeter of a simple polygon. The area of interior of a simple polygon is known as the area of a simple polygon.

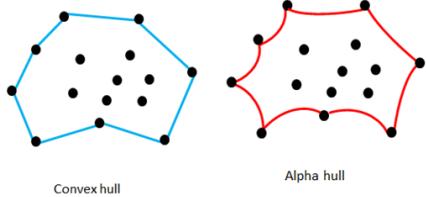
The calculation of scagnostic features require identification of the minimum spanning tree, convex hull and alpha hull.

Minimum Spanning Tree:

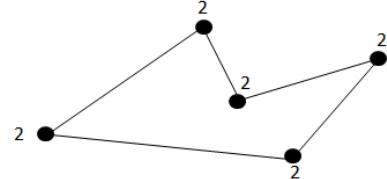
The spanning tree of the graph is defined as $G'(V', E')$, where $V' = Ve$, $E' \subset Ed$ and $E' = |Ve| - 1$. A graph can have more than one spanning tree. Spanning tree should not be disconnected and should not



(a) Graph with 5 vertices and 5 edges



(b) Convex hull and alpha hull



(c) Vertices of degree 2

Figure 22

contain any cycles. A spanning tree whose total length is least of all spanning trees on a given set of points is known as a Minimum Spanning Tree (MST).

Convex hull and Alpha hull

- i) Convex Hull (see Figure 22b): Given a set of points embedded in 2D Euclidean space, the convex hull of the set is the smallest convex polygon that contains all the points of it.
- ii) Alpha Hull (see Figure 22b): The alpha hull is a generalization of the convex hull. It is defined as the union of all convex hulls of the input within balls of radius alpha. In otherwords, it is a set of piecewise linear simple curves in the Euclidean plane associated a given set of points in the 2D Euclidean space. An open disk ($D(r)$) with radius r is used to define the indicator function to identify the associated points. If a point is on the boundary of D then D *touches* a point and if a point is inside D then D *contains* a point.

3.4.1 Preprocessing steps related to scagnostic features

To improve the performance of the algorithm and robustness of the measures, preprocessing techniques such as binning and deleting outliers are used before computing geometric graphs related features.

- i) Binning

As the first step of binning, the data are normalized to the unit interval. Then use a 40×40 hexagonal grid to aggregate the points in each scatterplot. We reduce the bin size by half and rebind until no more than 250 non-empty cells. If there are more than 250 non-empty cells. When selecting the bins two important points to consider are: i) efficiency (too many bins slow down calculations of the geometric graphs) and ii) sensitivity (fewer bins obscure features in the scatterplots). To improve the performance, hexagon binning is used. To manage the problem of having too many points that start to overlap, hexagon binning is used. The plots of hexagonal binning are density rather than points. There are several reasons for using hexagon binning instead of square binning in a 2D surface. Hexagons are more similar to circle than square. To keep scagnostics orientation-independent this bias reduction is important.

The weight function is defined as

$$\text{weight} = 0.7 + \frac{0.3}{1+t^2}, \quad (12)$$

where $t = \frac{n}{500}$. (n is the number of vertex)

If $n > 2000$ then this function is fairly constant. By using hex binning the shape and the parameters of the function is determined. In computing sparse, skewed and convex scagnostics this weight function is used to adjust for bias.

ii) Deleting Outliers

To improve robustness of the scagnostics, deleting outliers can be used. A vertex whose adjacent edges in the MST all have a weight (length) greater than ω is defined as an outlier in this context. By considering nonparametric criterion for the simplicity and Tukey's idea choose the following weight calculation is

$$\text{weight} = qu_{75} + 1.5(qu_{75} - qu_{25}), \quad (13)$$

where qu_{75} is the 75th percentile of the MST edge lengths and $(qu_{75} - qu_{25})$ is the interquartile range of the edge lengths.

iii) Degree of a Vertex

The degree of a vertex in an undirected graph is the number of edges associated with the vertex. For example, vertices of degree 2 means there are 2 edges associated with each vertex (see Figure 22c).

3.4.2 Definitions of Scagnostic Features

The definitions of scagnostic features are defined as follows. Our notations are as follows: i) Convex hull (CH), ii) Alpha hull (Al) and iii) Minimum spanning tree (MST). In the following section we give a brief description of the calculation of scagnostic measures. For more details on rotation techniques, see Wilkinson and Wills (2008).

3.4.3 Density Measures

- i) Outlying: The outlying measure is calculated before deleting the outliers for the other measures. The outlying measure is

$$F_{sc1} = \frac{\text{Total length of edges adjacent to outlying points}}{\text{Total edge length of the MST}}. \quad (14)$$

- ii) Skewed: The skewed measure is the first measure of relative density which is a relatively robust measure of skewness in the distribution of edge lengths. After adaptive binning skewed tends to decrease with n . The skewed measure is calculated using the equations

$$F_{sc2} = 1 - \text{weight} * (1 - qu_{skew}), \quad (15)$$

where,

$$qu_{skew} = \frac{qu_{90} - qu_{50}}{qu_{90} - qu_{10}}, \quad (16)$$

and the calculation of $weight$ is given in equation 12.

- iii) Sparse: The second relative density measure is sparse measure that measures whether points in a 2D scatterplot are confined to a lattice or a small number of locations on the plane. The sparse is measured by

$$F_{sc3} = \text{weight} * qu_{90}, \quad (17)$$

where the weight function is equation 12 and qu_{90} is the 90th percentile of the distribution of edge lengths in the MST .

The α statistic exceeds unity (e.g., when all points fall on either of the two diagonally opposing vertices of a square), clamp the value to 1 in the extremely rare event (Wilkinson and Wills 2008, @inproceedings44).

- iv) Clumpy: Clumpiness measures the amount of small-scale structures in the 2D-scatter plot (Wilkinson and Wills 2008). In order to calculate this feature another measurement called T is needed which is calculated based on RUNT graph (Hartigan and Mohanty 1992). The clumpy is measured by

$$F_{sc4} = \max_j [1 - \frac{\max_k [\text{length}(e_k)]}{\text{length}(e_j)}]. \quad (18)$$

In the formula below the j value goes over the edges in MST and k runs over all edges in RUNT graph.

- v) Striated: Striated define the coherence in a set of points as the presence of relatively smooth paths in the minimum spanning tree. This measure is based on the number of adjacent edges whose cosine is less than minus 0.75. The stratified is measured by

$$F_{sc5} = \frac{1}{|Ve|} \sum_{\nu \in Ve^{(2)}} I(\cos \theta_{e(\nu,a)e(\nu,b)} < -0.75), \quad (19)$$

where $Ve^{(2)} \subseteq Ve$ and $I()$ be an indicator function.

3.4.4 Scagnostic-based Shape Measures

Both topological and geometric aspects of shape of a set of scattered points is considered. As an example, a set of scattered points on the plane appeared to be connected, convex and so forth, want to know under the shape measures. By definition scattered points are not like this. Therefore to make inferences additional machinery (based on geometric graphs) is needed. By measuring the aspects of the convex hull, the alpha hull, and the minimum spanning tree is determined.

- i) Convex: The ratio of the area of the alpha hall (Al) and the area of the convex hull (CH) is the base of measuring convexity. The convex is measured by

$$F_{sc6} = \text{weight} \times \frac{\text{Area of alpha hull}}{\text{Area of convex hull}}, \quad (20)$$

where the weight function is equation 12.

- ii) Skinny: Skinny is measured by using the corrected and normalized ratio of perimeter to area of a polygon measures. The skinny is measured by

$$F_{sc7} = 1 - \frac{\sqrt{4 \times \pi \times \text{Area of alpha hull}}}{\text{Perimeter of alpha hull}}. \quad (21)$$

Furthermore

$$F_{sc7} = \begin{cases} 0; & \text{if circle} \\ \text{Near 1; if skinny.} & \end{cases}$$

- iii) Stringy: A skinny shape with no branches is known as a stringy shape. By counting the vertices of degree 2 in the minimum spanning tree and comparing them to the overall number of vertices minus the number of single-degree vertices, skinny measure is calculated. To adjust for negative skew in its conditional distribution of n , cube the stringy measure. The stringy is measured by

$$F_{sc8} = \frac{|Ve^{(2)}|}{|Ve| - |Ve^{(1)}|}, \quad (22)$$

where Ve is the number of vertices.

3.4.5 Association Measure

Symmetric and relatively robust measure of association are interested.

- i) Monotonic: To assess the monotonicity in a scatter plot, the squared Spearman correlation coefficient is used. In calculating monotonicity, the squared value of the coefficient is considered to remove the distinction between positive and negative coefficients. The reason is, the researchers are more interested in how strong the relationship is rather than their direction (negative or positive). The measurement is calculated as

$$F_{sc9} = r_{Spearman}^2. \quad (23)$$

3.4.6 Number of Minimum (F_8) and Maximum Points (F_9)

Number of minimum, and maximum points are new measures which are obtained from the polar coordinate of leaf contour. Number of global maximum points are defined as number of maximum points. Number of global minimum points are defined as number of minimum points. The associated features are

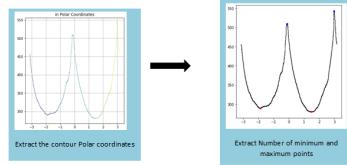


Figure 23: Minimum and Maximum Points

F_8 = Number of global minimum points,

$$(24a) \quad F_9 = \text{Number of global maximum points.} \quad (24b)$$

3.4.7 Correlation of Cartesian Contour (F_{10})

Correlation is another new feature computed based on the cartesian contour. The measure is calculated as

$$F_{10} = \frac{\sum_{i=0}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=0}^m (x_i - \bar{x})^2 (y_i - \bar{y})^2}}, \quad (25)$$

where (x_i, y_i) is the coordinate of cartesian contour and m is the number of points in the cartesian contour

4 Empirical Application

4.1 Data sets

We use two publicly available datasets to demonstrate the applications of features. They are: i) Flavia leaf image dataset and ii) Swedish leaf image dataset

4.1.1 Flavia Leaf Image Dataset

The Flavia dataset contains 1907 leaf images. There are 32 different species and each have 50-77 images. Scanners and digital cameras are used to acquire the leaf images on plain background. The isolated leaf images contain blades only, without petiole. These leaf images are collected from the most common plants in Yangtze, Delta, China (Wäldchen and Mäder 2018). Those leaves were sampled on the campus of the Nanjing University and the Sun Yat-Sen arboretum, Nanking, China (Wäldchen and Mäder 2018) available at <https://sourceforge.net/projects/flavia/files/Leaf%2520Image%2520Dataset/>.



(a) Sample of Flavia dataset



(b) Sample of Swedish leaf dataset

Figure 24

4.1.2 Swedish Leaf Image Dataset

The Swedish dataset contains 1125 images. The images of isolated leaf scans on a plain background of 15 Swedish tree species, with 75 leaves per species. This dataset has been captured as part of a joined leaf classification project between the Linkoping University and the Swedish Museum of Natural History (Wäldchen and Mäder 2018) available at <https://www.cvl.isy.liu.se/en/research/datasets/swedish-leaf/>.

4.2 Image processing and Feature Extraction

We applied the aforementioned image processing techniques and computed features from each image. Example of image processing and feature extraction for a leaf in the Flavia data set is shown in Figure 25a and Figure

25b respectively.

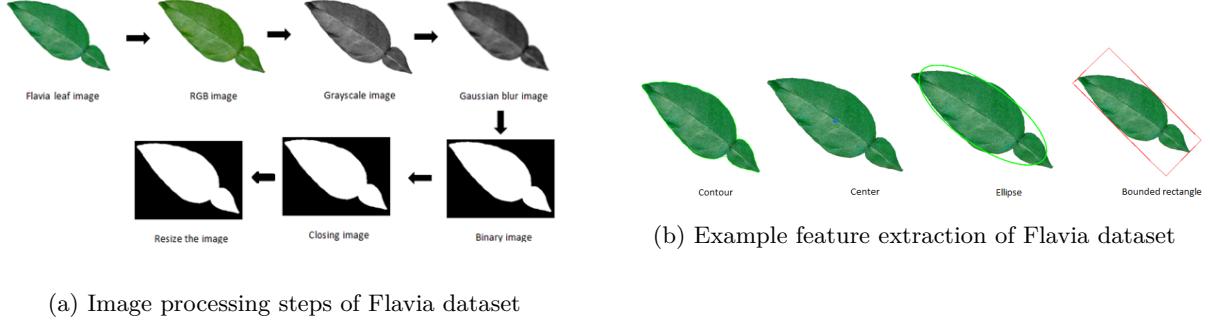


Figure 25

4.3 Visualization of ability of features to distinguish classes of interest

In order to identify ability of features to distinguish classes of interest, we first label features according to their shapes as: i) diamond, ii) heart shape, iii) needle shape, iv) simple round and v) round shape (see Figure 2b). These morphological characteristics are identified by observing images in the medicinal plant repository maintained by Barberyn Ayurveda resort and University of Ruhuna available at <http://www.instituteofayurveda.org/plants/>. Our observed results are converted into a open source R software package called MedLEA: Medicinal LEAf which is available on Comprehensive R Archive Network (Lakshika and Talagala 2021).

We explore the ability of features to classify images under supervised learning setting and unsupervised learning setting. For this purpose we use Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA). LDA is a supervised dimensionality reduction technique, and PCA is unsupervised dimensionality reduction technique. In this section, we visualize, and compare the results obtained using LDA, and PCA on Flavia and Swedish datasets. To compute LDA projection shape label is taken as the response variable. There are 5 main shape categories as diamond, simple round, round, needle, and heart shape.

4.3.1 Swedish Dataset

The visualization of PCA projections of Swedish dataset is shown in Figure 26. The first three principal components (PCs) accounting for approximately 80% of the total variance in the original data, while the first 5 PCs accounts for 90% of the total variance in the original data. Hence, the first five PCs are plotted against each other to visualize data in the PCA space. The LDA projections of Swedish data are shown in Figure

27. According to the LDA results on Swedish dataset, LDA1, LDA2, and LDA3 shows a clear separation of shapes of leaf images.

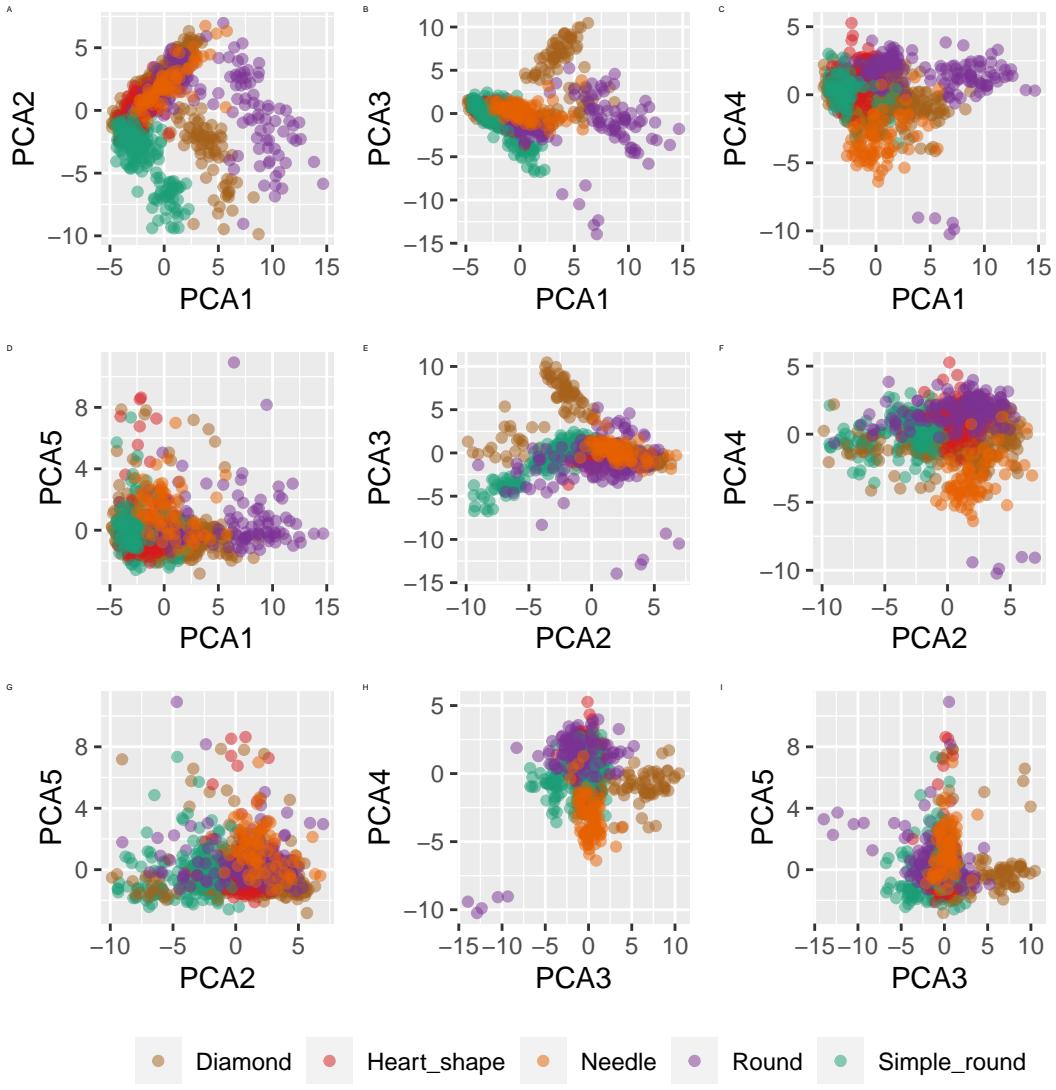


Figure 26: Distribution of Swedish leaf images on the projection space created based on principal component analysis. All projected points are coloured according to their shape labels.

4.3.2 Flavia Dataset

The LDA projections of Flavia data are shown in Figure 28. The visualization of PCA projections of Flavia dataset is shown in Figure 29. The first three principal components (PCs) accounting for approximately 83% of the total variance in the original data, while the first 5 PCs accounts for 90% of the total variance in the original data. Hence, the first five PCs are plotted against each other to visualize data in the PCA projection

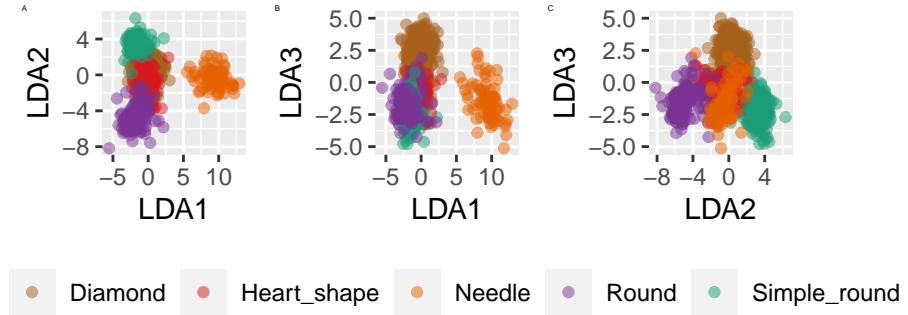


Figure 27: Distribution of Swedish leaf images on the projection space created based on linear discriminant analysis. All projected points are coloured according to their shape labels.

space. According to the LDA results on Flavia dataset, LDA1, LDA2, and LDA3 shows a clear separation of shapes of leaf images. Under both experimental settings class separation is more clearly on the LDA space than the PCA space. The reason could be LDA is a supervised learning algorithm while PCA space is an unsupervised learning algorithm.

According to both PCA and LDA visualizations on Swedish and Flavia data sets we can see a clear separation of classes in their corresponding projection spaces. This reveals our features are capable of distinguishing classes under both supervised learning and unsupervised learning settings.

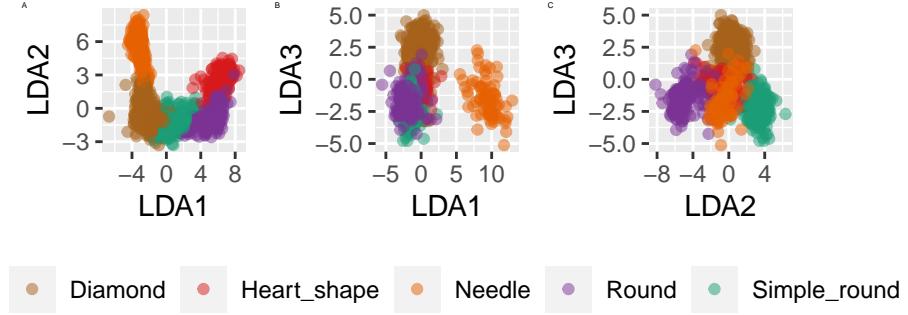


Figure 28: Distribution of Flavia leaf images on the projection space created based on linear discriminant analysis. All projected points are coloured according to their shape labels.

5 Discussion and Conclusions

In this paper we introduce computer-aided, interpretable features for image recognition. There are four main categories of features that are used to classify leaf images. Many research were based on shape,

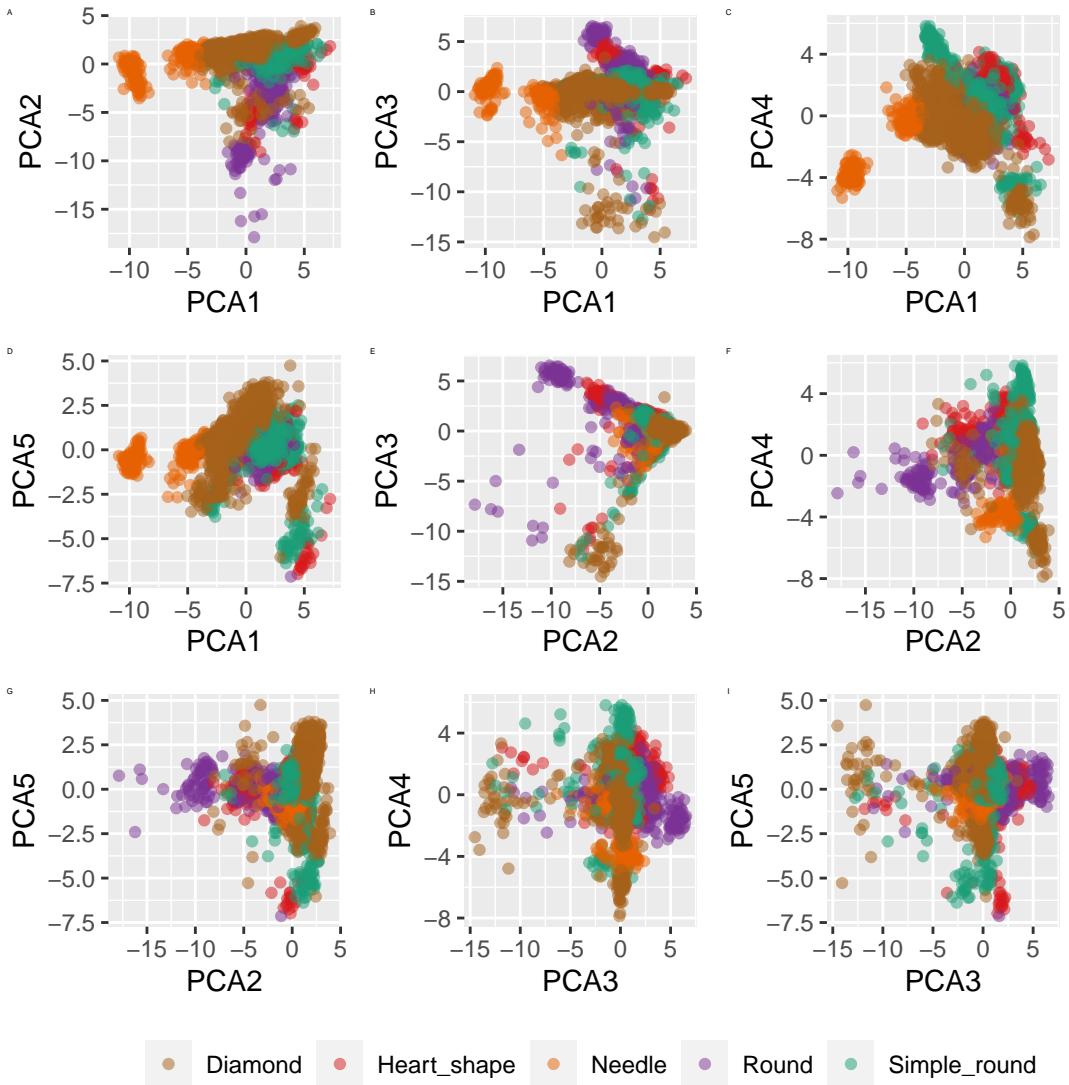


Figure 29: Distribution of Flavia leaf images on the projection space created based on principal component analysis. All projected points are coloured according to their shape labels.

color, and texture features. In this research paper, we introduce new feature category called scagnostics for image classification. Other than that correlation of cartesian coordinate, number of convex points, number of minimum and maximum points are introduced as new shape features. We explore the ability of features to discriminate the classes of interest under supervised learning and unsupervised learning settings using principal component analysis and linear discriminant analysis. Under both experimental settings clear separation of classes are visible in their projection spaces. In our next paper we introduce a meta-learning algorithm to classify plant species using the features introduced in this paper. A more detailed look at the most important features and approach of classification plant species will be presented in our next paper. We hope our analysis opens various research agendas in the field of image recognition. Reproducible research code to reproduce the results of this paper and codes to compute features are available here: <Removed to keep the paper blind.>.

References

- Anantrasirichai, Nantheera, Sion Hannuna, and Nishan Canagarajah. 2017. “Automatic Leaf Extraction from Outdoor Images.” *arXiv Preprint arXiv:1709.06437*.
- Azlah, Muhammad Azfar Firdaus, Lee Suan Chua, Fakhrul Razan Rahmad, Farah Izana Abdullah, and Sharifah Rafidah Wan Alwi. 2019. “Review on Techniques for Plant Leaf Classification and Recognition.” *Computers* 8 (4): 77.
- Bangare, Sunil L, Amruta Dubal, Pallavi S Bangare, and ST Patil. 2015. “Reviewing Otsu’s Method for Image Thresholding.” *International Journal of Applied Engineering Research* 10 (9): 21777–83.
- Caglayan, Ali, Oguzhan Guclu, and Ahmet Burak Can. 2013. “A Plant Recognition Approach Using Shape and Color Features in Leaf Images.” In *International Conference on Image Analysis and Processing*, 161–70. Springer.
- Dean, C. 1999. “Quantitative Description and Automated Classification of Cellular Protein Localization Patterns in Fluorescence Microscope Images of Mammalian Cells.” PhD thesis, PhD thesis, Carnegie Mellon University.
- Gonzalez, Rafael C., and Richard E. Woods. 2006. *Digital Image Processing*. Prentice-Hall, Inc.
- Goyal, N., Kapil, and N. Kumar. 2018. “Plant Species Identification Using Leaf Image Retrieval: A Study.” In *2018 International Conference on Computing, Power and Communication Technologies (Gucon)*, 405–11.

- Haralick, Robert M, Karthikeyan Shanmugam, and Its' Hak Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man, and Cybernetics*, no. 6: 610–21.
- Hartigan, John A, and Surya Mohanty. 1992. "The Runt Test for Multimodality." *Journal of Classification* 9 (1): 63–70.
- Herdiyeni, Yeni, and Ni Kadek Sri Wahyuni. 2012. "Mobile Application for Indonesian Medicinal Plants Identification Using Fuzzy Local Binary Pattern and Fuzzy Color Histogram." In *2012 International Conference on Advanced Computer Science and Information Systems (Icacsis)*, 301–6. IEEE.
- Kodituwakku, Saluka Ranasinghe, and S Selvarajah. 2004. "Comparison of Color Features for Image Retrieval." *Indian Journal of Computer Science and Engineering* 1 (3): 207–11.
- Lakshika, Jayani P. G., and Thiyanga S. Talagala. 2021. "MedLEA: Morphological and Structural Features of Medicinal Leaves." <https://CRAN.R-project.org/package=MedLEA>.
- Mingqiang, Yang, Kpalma Kidiyo, and Ronsin Joseph. 2008. "A Survey of Shape Feature Extraction Techniques." *Pattern Recognition* 15 (7): 43–90.
- Söderkvist, Oskar. 2001. "Computer Vision Classification of Leaves from Swedish Trees."
- Sun, Yu, Yuan Liu, Guan Wang, and Haiyan Zhang. 2017. "Deep Learning for Plant Identification in Natural Environment." *Computational Intelligence and Neuroscience* 2017.
- Wäldchen, Jana, and Patrick Mäder. 2018. "Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review." *Archives of Computational Methods in Engineering* 25 (2): 507–43.
- Wilkinson, Leland, Anushka Anand, and Robert Grossman. 2005. "Graph-Theoretic Scagnostics." In *IEEE Symposium on Information Visualization (Infovis 05)*, 157–58. IEEE Computer Society.
- Wilkinson, Leland, and Graham Wills. 2008. "Scagnostics Distribution." *Journal of Computational and Graphical Statistics* 17 (June): 473–91. <https://doi.org/10.1198/106186008X320465>.
- Wu, Stephen Gang, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang, and Qiao-Liang Xiang. 2007a. "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network." In *2007 Ieee International Symposium on Signal Processing and Information Technology*, 11–16. IEEE.
- . 2007b. "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network." In *2007 IEEE International Symposium on Signal Processing and Information Technology*, 11–16. IEEE.