Mathematics Letters, eISSN: 2948-3735, 2 (1): 34-45, 2023

PLANT LEAF IDENTIFICATION OF KAFFIR LIME LEAVES USING SUPPORT VECTOR MACHINE (SVM)

Nur Afini Ezzulia Ridzuan[1], Nurul Ain Nabilah Jasman[2] and Hanifah Sulaiman[3*]

Pusat Pengajian Sains Matematik, Kolej Pengajian Pengkomputeran, Informatik Dan Media, Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia

[1]afiniy99@gmail.com, [2]aindamel2617@gmail.com, [3*]hanifah@tmsk.uitm.edu.my

(indicates the corresponding author)*

ABSTRACT

Plant identification is very important in day-to-day life especially for biologists, chemists, and scientists to identify every plant species and it is interesting research due to various types of plant species. Numerous studies have been conducted for plant leaf identification especially when the plants have various types of leaf and the plants leaves are almost identical to each other. For this study, two types of different shapes of kaffir lime leaf, one is a figure eight-shape of double leaves and the other one is shaped wide at the top and slightly small at the base has been chosen as the subject of the study. This study focuses on the feature extraction of kaffir lime leaves such as area, parameter, convex hull, major axis, minor axis and the ratio of major and minor axis and determination of the best accuracy between three (3) different kernel functions which are the radial basis function (rbf), linear and polynomial kernel for plant identification of two types of kaffir lime leaves using Support Vector Machine (SVM). A machine learning algorithm, SVM, has been used to create species identification models. The 20 leaves images of figure eight-shape of double leave and 20 leaves images of shape wide at the top and slightly small at the base of the kaffir lime plant were used as the data to determine the features extraction using MATLAB R2021a software. Besides that, the optimization value of the box constraint for three (3) different kernel functions also have been measured automatically by using the function code in the image processing toolbox. The result demonstrated that the most accurate rate of plant identification of two types of kaffir lime leaves is up to 91.67% by using linear kernel function in SVM method.

Keywords: *Two types of Kaffir Lime Leaves, Support Vector Machine, MATLAB R2021, Linear kernel function, Rbf kernel function, Polynomial kernel function.*

Received for review: 13-03-2023; Accepted: 09-04-2023; Published: 14-04-2023

1. Introduction

On earth, there are many types of plant species that are known to science. The plants have their unique characteristics such as shape, texture, and color although they look similar. Identification of plants is an important process to gain required raw materials from the right plants. Plant identification is also important in environmental protection in order to accurately see changes in plant species and population. However, in botany and agriculture, plant identification is frequently exceedingly time-consuming and knowledge-intensive (Hussain et al, 2018). Scientists and biologists have struggled to distinguish the nature of plants so many methods have been developed to solve the problem of identification and recognition of plants.

In this research, the focus is on the identification of Kaffir lime which is commonly known as ‘daun limau purut’. Kaffir lime is a one of the family of Ructatae and a member of the genus Citrus that have fragrant leaves and features aromatic fruits. Commonly, the genus Citrus species stimulates a huge food crop and grows beautifully that consists of lime, lemon and orange (Chase at al., 1999). Kaffir lime is native to SouthEast Asia. The binomen of the kaffir lime is Citrus Hystrix and also called ‘makrut’ in Thailand, ‘jeruk purut’ in Indonesia and ‘limau purut’ in Malaysia. The term ‘kaffir’ is trusted to have been actualized hundreds of years ago and is found in particular languages.

There are two types of Citrus Hystrix leaf that have been found, one is a figure eight-shape of double leaves and the other one is shaped wide at the top and slightly small at the base. The two types of shaped kaffir lime leaves are identical to each other since they came from the same Rutaceae family. The term ‘kaffir’ comes from the Arabic word ‘kafara’ which means ‘infidel’. The Netherlands and Germany commonly used the name Mauritius papeda while the lime name 'combava' is frequently called by France, Italy and Spain. The taste of two kaffir lime leaves of different shapes is the same. The two types of shaped kaffir lime leaves are edible and also can be used in cooking as a kind of aromatic bay leaf. This leaf is said to be unique because both types of kaffir lime leaves grow in pairs and double shape in one leaf. The top of the leaf is deep green and the underside is light green. The difference in shapes and sizes of two types of kaffir lime leaf are used to extract the features using image processing. The leaf area and leaf weight of kaffir lime are the main factors that affect farmers income in agribusiness and both can be measured using destructive and non-destructive procedures (Budiarto et al., 2021).

The kaffir lime tree is usually small, having a height of 3 to 5 m and a width of 2.5 to 3 m. The small dark green fruits with small white flowers, dark green fruits about 4 cm wide with warty skin (Bown, 2002). The diameter of kaffir lime leaves are 7.5 to 10 cm long and the color is green. Leaves of kaffir lime are special between the citrus kind and, for some certain applications, are more significant than the fruits. The dark green color and glossy sheen are the features of the kaffir lime leaves. The size of the leaves are quite a bit different and the colors of bigger leaves are commonly darker colors. The fruits are used in cooking and also in the manufacturing of shampoo. This kaffir lime produces an essential oil from its leaves which is very useful for human mouth epidermal carcinoma (KB) and murine leukemia (P388) cell lines which has been shown to have anti-proliferative activity. Citrus Hystrix or kaffir lime leaves are commonly used as a spice in cooking, as herbs and medicines such as abdominal pains, hypertension, and flu, as aromatics and also as a treatment for scurvy (Saptanti et al., 2021). Kaffir lime plants can be found in Southeast Asia and Southeast China. The bioactivities such as antibacterial, antioxidant and antileukemic can be displayed by the essential oil of kaffir lime leaves. Furthermore, the essential oil of kaffir lime also can be used as food ingredients, fragrance, and cosmetic 3 product ingredients (Suwannarach et al., 2022). Thailand has the most kaffir lime leaves products produced followed by Indonesia and Malaysia. The demand for the kaffir lime leaves products keeps increasing and the plantation area for kaffir lime plants also keep increasing because it gives a lot of benefits such as producing a healthier lifestyle.

There are various types of image processing techniques to process a plant image. The image processing refers to the process of applying operations on an image in order to improve it or extract relevant information from it. It is an application of calculating algorithms where the input is an image collected and the output is the features related to that image. Previously, the reason for processing the image was to improve the condition of the images including restoration, encoding and magnification. Since a lot of benefits have been seen while using digital image processing, nowadays, there are many applications and extensive usage of digital image processing such as image sharpening, color processing, encoding, medical field and pattern recognition. Digital computer can execute a mathematical or algorithmic operation on pictures from images with the high-speed general purpose (Haralick et al., 1973). The image processing using a digital computer is usually kept as a two-dimensional array to achieve the aim such as enhancement, classification, identification, and restoration. In recent years, a huge amount of computer processing on images has happened from processing normal images to X rays (Andrews et al., 2009). Image processing is important because it also can be applied in forensic science, medical technologies, and industrial machine control.

One of the types of image processing techniques is Support Vector Machine (SVM). SVM is a universal learning machine that carries out the strategy of keeping the value of the empirical risk fixed, minimizing the confidence interval that was first introduced by Cortes and Vapnik (1995). SVM was proposed by Vapnik and his collaboration in early 1990s (Bafitlhile et al., 2019). Nowadays, SVM are important in numerous real-world situations in order to complete the learning tasks. SVM also has been applied to many fields recently such as biomedical field, protein prediction, power system electroencephalography (EEG) and computer vision with a lot of algorithms applied on it (Wang & Pardalos, 2014).

SVM also used pattern recognition problems such as text classification, image recognition, face detection and faulty card observation. The SVM machine learning forms a model that gives a set of training examples in which every sample belongs to one of the many categories. There are more capability to generalize the problem that is the goal in statistical learning by SVM. The algorithm of SVM is generally used for classification and regression of two types of patterns such as linear and nonlinear. Linear patterns can be easily separated and classified with the help of hyperplanes. When the testing data were added, the hyperplane will determine the class that we assign.

Pradhan (2015) stated that the optimal hyperplane can maximize the margin of hyperplane such as the gap from hyperplane to the closest point of each pattern. Larger margin size can produce good classification of the pattern. Nonlinear patterns cannot be easily separated and need to use kernel function to make non-separable data into separable data. The pattern is mapped into high dimensional space by the kernel to be classified. Different kernel tasks give different mapping for nonlinear patterns. SVM can give accurate results and is also helpful in high dimensional spaces.

A matrix laboratory also called MATLAB was built to provide easy access to matrix software. MATLAB is commonly used for data acquisition, algorithm development, data analysis, mathematics and computation and development applications. It can combine programming, computation and visualization in order to solve the problems easily using familiar mathematical notation (Gonzalez et al., 2004). The MATLAB system consists of five parts which are the MATLAB language, the MATLAB working environment, handle graphics, the MATLAB mathematical library function and the MATLAB application program interface.

Several tasks in digital image processing using MATLAB are to enhance the clarity and remove noises, extract the features of images and compress images for communication across the network. There are some functions in the MATLAB for digital image processing such as ‘imread’ to read the images, ‘imshow’ to display the images and ‘imwrite’ to write the images to disk. Furthermore, the features of images such as area, perimeter and convex hull can be calculated using MATLAB using appropriate functions. The images can be read into memory and the data assigned into double precision floating point values.

The leaves are the main essential structures of plants that give a lot of benefits to humans as well as all living things on the earth. There are various types of leaf with differences in shape, texture, and color. An identification of plant leaves is important in order to classify the variety of plants and to observe the changes and classification of a plant species. Over the past decades, research has received considerable attention on the utilization of moments for object characterization in in-variant tasks and non-variant tasks from mechanics and statistics field to pattern recognition and image processing and also the mathematical concept of moments has been around for many years (Singh et al., 2015). Currently, image processing is the top learning in the advanced system. The main focus of this study is to identify the features of extraction of two types of kaffir lime leaf such as area, parameter, convex hull, major axis, minor axis and the ratio of major and minor axis. In order to classify different features of plant leaves, a computer-based plant identification system was used using a SVM. This study used two types of kaffir lime leaf images that were captured by the mobile phone camera with high quality results. The images are converted into digital form to perform operations using image processing.

The first objective of this study is to identify feature extraction of kaffir lime leaves such as area, parameter, convex hull, major axis, minor axis and the ratio of major and minor axis. The second objective is to determine the best accuracy between three different kernel functions which are the radial basis function (rbf), linear and polynomial kernel for plant identification of two types of kaffir lime leaves using SVM. In this study, it focuses on the identification of kaffir lime images. SVM will be used to identify the characteristic of the kaffir lime leaves based on the images.

2. Methodology

The methodology used in this study is shown in Figure 1. There are three major steps involved in the process of plant identification of kaffir lime leaves. First is data acquisition. Second is feature extraction of kaffir lime leaves and lastly performance analysis.

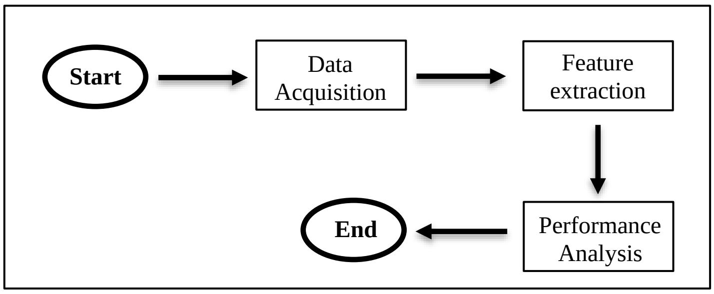

Figure 1. Flowchart of methodology for this study.

Leaf images are collected from kaffir lime plants. Each type of kaffir lime leaf included 40 samples of leaf images. These leaf images have different shapes and sizes. The data has been collected by capturing 20 images on the figure eight-shape of double leaves and 20 images on the shape wide at the top and slightly small at the base of the kaffir lime leaves. All the leaf images captured will be imported from a document file into the MATLAB as a training data set. The three-dimensional images are transformed into two-dimensional images or grayscale images using coding in MATLAB. The images converted the grayscale images into binary images because computer language can only recognize the binary image for calculation. Then, the features of kaffir lime leaf will be extracted using MATLAB. Some of the features include area, perimeter, convex hull, major axis, minor axis and the ratio of major and minor axis. The features need to be calculated using SVM that will give the numeric value result. After that, we will do a performance analysis of Kaffir lime leaves using the SVM method. Every calculation is completed by using MATLAB.

3. Implementation

For the first step, the primary sources of data collection have been taken directly using digital leaf images. The photos of the 40 Citrus Hystrix leaves are captured using an iPhone 11 pro camera with the white A4 paper used as a background to reduce noise as shown in Figure 2 for one of the examples of each type of leaves. Then, the background has been removed before transferring to a laptop as Figure 3.

Figure 2. The sample images of two types of kaffir lime leaves in three-dimensional.

Figure 3. The sample images of two types of kaffir lime leaves after removing the background.

Next, the leaf images without background are transferred one by one into the laptop and imported into the MATLAB. The MATLAB code can only process one input data image. After collecting all the features information from one data image, repeat the same method for all accessible data images. The 'imread' function in MATLAB is used to import the data images.

Next, the 'rgb2gray' function in MATLAB was used to transform from three-dimensional images into two-dimensional images or grayscale images by eliminating hue and saturation information while retaining the luminance as shown in Figure 4. The function will convert RGB image value to grayscale values by forming a weighted sum of the Red(R), Green(G), and Blue(B) components.

Figure 4. The sample images of two types of kaffir lime leaves in two-dimensional images or grayscale images.

After that, using the function 'im2bw', as Figure 5 below, the transformation of grayscale images into binary images was done. The binary image is also a two-dimensional image. This function converts the grayscale image into a binary image by replacing all pixels in the input image with luminance greater than level with value 1 which represents the white color and replacing other pixels with the value 0 which represents the black color. Therefore, to specify the level in the range of [0,1] depends on the class of the input image.

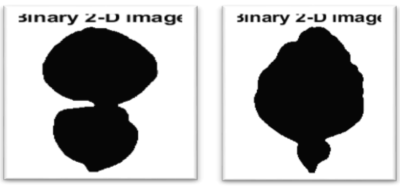

Figure 5. The binary 2-D images of two types of kaffir lime leaves.

All types of image conversions are required to extract the features of the leaves using MATLAB. The reason why it needs to convert the three-dimensional image into a binary image is that computer language can only understand the binary code for any work done in it. This step involves calculating the characteristics of the leaves in the images of kaffir limes leaves, including their area, perimeter, convex hull, major axis, minor axis, and ratio of axis.

i. Area of the leaves

The area of each leaf must first be determined. A basic x-axis and y-axis graph is used in Figure 6 to represent the binary image.

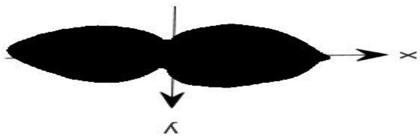

Figure 6. The binary image of eight-shape of double kaffir lime leaf in coordinate axes.

The MATLAB function 'regionprops' is used to determine the leaf's area. This function will determine the precise number of pixels in the leaf's area and return it as a scalar value. Therefore, the area of the leaf will be the returned value.

ii. Parameter of the leaf

In order to calculate the perimeter of a leaf, the function 'bwperim' in MATLAB is used. This function returned a binary image containing only the perimeter pixels of the leaf of the input image. If a pixel has a nonzero value and is connected to at least one pixel with a zero value, it is regarded as being a part of the perimeter. Then, by using the image, another function is used which is the 'nnz' function in the MATLAB. This function measures the perimeter of the leaf for each binary image by counting the nonzero elements.

iii. Convex hull

To calculate the convex hull of a leaf, the set of points on the leaf need to be found. Due to the variations in size, shape, and color of each leaf, each one will have a unique set of points. The 'bwconvhull' function in MATLAB is used to find the set of convex hull points that are present in the leaf. The convex hull of each object in the image will be determined using this function. Then, it will return the binary convex hull image. The input must be in the form of two-dimensional images or grayscale images that were obtained in the previous step. The Figure 11 below shows the points of the convex hull on the vertices of the leave.

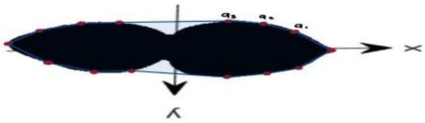

Figure 7. Example of convex hull of a kaffir lime leaf.

The sum function in MATLAB is used to calculate the total convex hull length. Besides, the distance formula is used to determine the length of the set points of the convex hull.

Let $a_1, a_2, a_3 \ldots a_n$ be the sets of convex hull on the vertices of the leaf where $n = 1,2,3, \ldots$

In order to calculate the distance between two pairs of points separately. For example, to find the points at $a_1(x_1, y_1)$ and $a_2(x_2, y_2)$, use equation below.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

where $d_{(1,2)}$ is the distance between points a_1 to a_2.

After calculating all the distances, add up all the distance values.

$$\nabla d = d_{(1,2)} + d_{(2,3)} + d_{(3,4)} \dots + d_{(n,1)} \tag{2}$$

where ∇d is the total distance of the convex hull of a leaf.

iv. Major Axis

The distance from the leaf's base to its tip is known as the major axis. It is necessary to determine the length that is related to the x-coordinates and y-coordinates. Therefore, the distance formula is used to calculate the length from the base of the leaf to the tip of the leaf. The property type "MajorAxisLength" and the function "regionprops" are used in MATLAB. The Figure 8 below shows an illustration of a leaf's major axis.

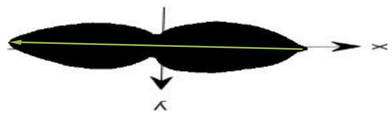

Figure 8. The green line indicated the major axis of an eight-shape of double kaffir lime leaf.

By using equation (1), let the coordinates for the base of the leaf become $(a_{1,} b_1)$, and the coordinates for the tip of the leaf become (a_2, b_2). Therefore, it can be expressed as

$$M_a = \sqrt{(a_2 - a_1)^2 + (b_2 - b_1)^2} \tag{3}$$

Where M_a represents the major axis of the leaf.

v. Minor Axis

The major axis and this feature are relatively similar, however the minor axis is the length times the width of the leaf surface. To determine the length and width, choose an appropriate coordinate that will get the best results for the minor axis. Nevertheless, the minor axis will have a range of values depending on the coordinates. In this study, we seek out the neighborhood that is closest to the selected points. In MATLAB, by using the binary image, the function 'regionprops' is used with the properties of 'MinorAxisLength' to calculate the minor axis. The minor axis of a Kaffir lime leaf is depicted in the Figure 9 below.

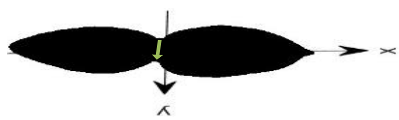

Figure 9. The green line indicated the minor axis of an eight-shape of double kaffir lime leaf.

By using equation (1), the length and width of the leaves, or called it the minor axis can be defined as:

$$M_i = \sqrt{(c_2 - c_1)^2 + (d_2 - d_1)^2} \quad (4)$$

Where M_i represents the minor axis of the leaf.

vi. Ratio of axes

The last feature is the ratio of axes, which is quite simple compared to the others. The major axis is divided by the minor axis to determine the ratio of the axes. For example, let R be the variable in the ratio. Thus, rewrite the general formula for a leaf's axis ratio as follows:

$$R = \frac{M_a}{M_i} \quad (5)$$

$$R = \frac{\sqrt{(a_2-a_1)^2+(b_2-b_1)^2}}{\sqrt{(c_2-c_1)^2+(d_2-d_1)^2}} \quad (6)$$

In the last step, the SVM method with three (3) different kernels is used in this step to distinguish between eight-shape double and shape wide at the top and slightly small at the base of Kaffir Limes Leaves. It is possible to identify these two shapes of plants without using SVM, but it is very time-consuming. Since the values of every feature eight-shape double leave and shape wide at the top and slightly small at the base leave have been calculated earlier, all information will be input into two tables for each of the leaf. Then, the minimum and maximum values for each feature of eight-shape double and shape wide at the top and slightly small at the base will be determined.

In order to analyze the leaves features, SVM has been applied using MATLAB version R2021a. In this method, MATLAB software has been used to execute the SVM with the existing toolbox that has been provided in MATLAB. As mentioned earlier, the dataset has been separated into two parts, 70% for the training set and 30% for the test set where 28 images will be the training set and 12 images will be the test set. Training set will learn from six (6) feature using SVM algorithm whether it is leave 0 or leave 1. Then, by using the model that train set have been learnt, the testing set will predict whether it is leave 0 or 1.

Then, the ‘fitcsvm’ function is used to find the test accuracy that exists for two types of kaffir lime leaves. This function can solve the optimization value of the model. As mention earlier, the properties of box constraint will be resulted in order to know about the cost of the model and the kernels scale also will be resulted to know the scaling of the parameter before input into the kernel function.

Next, the kernel functions which are linear have been used to find the best test accuracy for this SVM method. After that, by using the function ‘OptimizeHyperparameters’ and set it as ‘auto’, it will be attempted to minimize the crossvalidation loss (error) for the ‘fitcsvm’ by varying the parameters. Therefore, it will result in the optimization value for both box constraint and kernel scale. By using the available functions, ‘predict’ and ‘sum’ in MATLAB, they will be resulting in the accuracy value for kernel function in the model.

The same process from importing the training and testing data set was repeated by using another two (2) different kernels function which is (rbf) and polynomial to find out which kernels can give the best prediction and best accuracy between different types of kaffir lime leaves.

4. Results and Discussion

Tables 1 and 2 below show the result of sample attributes of two types of kaffir lime leaves in numeric form. Some of the attributes have been rounded off to 4 decimal points.

Table 1. The features of eight-shape double kaffir lime leaves.

No.	Area	Parameter	Convex hull	Major axis	Minor axis	Ratio of axis
1	179322	5672	665404	1576.3	838.68	1.8795
2	190751	6156	583680	1820.3	700.35	2.5991
3	201375	6065	627200	1810.6	739.21	2.4494
4	207593	6477	709120	1802.4	849.58	2.1215
5	164413	6293	619520	1795.5	746.86	2.404
6	217222	6266	623360	1756.8	702.08	2.5023
7	297877	6806	808960	1653.1	945.75	1.7479
8	194909	6019	592640	1845.1	704.04	2.6208
9	191065	6159	584960	1777.7	704.76	2.5225
10	205560	6594	675840	1810.3	810.86	2.2326
11	50551	3246	175700	1017.8	386.99	2.6301
12	40727	2158	138520	840.7	302.45	2.7796
13	72647	3765	230400	1172.4	435.78	2.6903
14	65984	3783	228000	1173.1	441.35	2.6579
15	99782	4077	277564	1262.1	474.51	2.6598
16	58470	3298	180780	1016.4	397.18	2.5592
17	109558	4141	299280	1173.9	517.61	2.2679
18	142177	1555	158391	504.5	384.71	1.3114
19	51681	3108	166221	944.1	386.44	2.4432
20	70959	3715	227850	1127.7	445.97	2.5288

Table 2. The features of shape wide at the top and slightly small at the base of the kaffir lime leaves.

No.	Area	Parameter	Convex hull	Major axis	Minor axis	Ratio of axes
1	293283	6296	715520	1646.1	793.35	2.0749
2	282990	6499	743680	1821	840.21	2.1673
3	268088	6754	619520	1553.1	638.74	2.4316
4	322446	6557	803840	1611.5	912.97	1.7651
5	338013	6302	861440	1857.6	946.01	1.9636

6	292329	6520	706560	1572.8	777.85	2.022
7	303676	6794	727040	1620.3	798.44	2.0294
8	247600	6187	618240	1705.4	695.75	2.4511
9	277891	6023	657920	1712.1	738.47	2.3185
10	258347	5323	620379	1450.7	721.11	2.0118
11	302533	6342	698880	1595.4	776.15	2.0557
12	211533	5424	625050	1688.1	691.70	2.4409
13	278600	5763	739080	1632.8	793.20	2.0584
14	297864	6961	636160	1690.2	691.24	2.4452
15	294220	5467	684091	1499.7	729.70	2.0552
16	269406	5360	608113	1066.3	885.36	1.2044
17	305034	6362	683520	1661.9	767.06	2.1665
18	264842	6301	624640	1590.5	658.64	2.4148
19	206315	6298	634880	1737.2	749.22	2.3187
20	281827	6466	739840	1704.3	860.95	1.9796

Based on Table 2, a few conclusions based on the features of the leaves of kaffir lime can be drawn. First, the smallest area of the leaves is 206315 and the biggest area of the leaves is 338013. Second, the smallest parameter of the leaves is 5360 and the biggest parameter is 6961. For the convex hull, the minimum value for the convex hull is 608113 and the maximum value is 861440. For the major axis, the shortest major axis of the leaves is 1066.3 and the maximum value is 1857.6. For the minor axis, the shortest value is 638.74 and the longest value is 946.01. Lastly, the ratio of the major axis and minor axis. The smallest ratio is 1.2044 and the largest ratio is 2.4511.

Table 3. The feasible point for the box constraint, the suitable value for the kernel scale, the number of misclassification results and the accuracy of the three different kernels are based on the test accuracy generated in MATLAB software.

Types of kernels	**Linear**	**Polynomial**	**Rbf**
Feasible point for box constraint	0.2436	0.0010089	Not found
Suitable value for kernel scale	456.88	999.79	Not found
Number of misclassification results	1	2	6
Accuracy	91.67%	83.33%	50.00%

Based on the Table 3 above, for linear kernel, the best observed feasible point for box constraint is around 0.2436. While, for the kernel scale, the suitable value for the model is 456.88 for the observed points. The MATLAB generated an accuracy test of 91.67% where there is only one (1) misclassification result for this linear kernel.

For polynomial kernel, the best observed feasible point for box constraint is around 0.0010089. While, for the kernel scale, the suitable value for the model is 999.79 for the 25 observed points. The MATLAB generated an accuracy test of 83.33% where there is two (2) misclassification result for this linear kernel.

For rbf kernel, the best observed feasible point for box constraint is not found. While, for the kernel scale, the suitable value for the model is also not found for the observed points. The MATLAB generated an accuracy test of 50.00% where there is six (6) misclassification result for this linear kernel.

So, from here it can be concluded that the best accuracy for plant leaf identification of two types of kaffir lime leaves using SVM is linear kernel because it has the accuracy with the highest value which is 91.67%. With the help of MATLAB, the best accuracy between three different kernel functions which are the radial basis function (rbf), linear and polynomial kernel for plant leaf identification of two types of kaffir lime leaves using SVM was determined.

5. Conclusion and Recommendations

In this study, the feature extraction of kaffir lime leaves such as area, parameter, convex hull, major axis, minor axis and the ratio of major and minor axis were identified. Besides, the best accuracy between three different kernel functions which are the radial basis function (rbf), linear and polynomial kernel for plant leaf identification of two types of kaffir lime leaves using SVM was determined.

Based on the result obtained, using MATLAB with three different kernels is an efficient way to identify between the two types of shape of kaffir lime leaves because it can be used for identification purposes. Despite that, we can save a lot of time and money.

This method also can be used to identify other plant species such Spinacia oleracea and Brassica juncea that are identical to each other. In addition, our study using SVM to identify between different leaf species will contribute to plant taxonomy fields because it can be used to identify various kinds of leaves. Other algorithms that can be used to identify plant leaves instead of SVM are KNN, Naive Bayes or CNN.

Acknowledgement

The authors thanks to Pusat Pengajian Sains Matematik, Kolej Pengajian Pengkomputeran, Informatik Dan Media, Universiti Teknologi MARA (UiTM) for supporting this research.

References

Andrews, H. C., Tescher, A.G., & Kruger, R.P. (2009). Image processing by digital computer. IEEE Spectrum, 9(7), 20-32. https://doi.org/10.1109/mspec. 1972.5218964

Bafitlhile, T. M., & Li, Z. (2019). Applicability of ε-Support Vector Machine and artificial neural network for flood forecasting in humid, semi-humid and semi-arid basins in China. Water (Switzerland), 11(1). https://doi.org/10.3390/w11010085

Bown, D. (2002). The Royal Horticultural Society New Encyclopedia of Herbs & Their Uses, Great Britain, Dorling Kindersley, London.

Budiarto, R., Poerwanto, R., Santosa, E., Efendi, D., & Agusta, A. (2021). Short communication: Allometric model to estimate bifoliate leaf area and weight of kaffir lime (Citrus hystrix). Biodiversitas, 22(5), 2815–2820. https://doi.org/10.13057/biodiv/d220545

Chase, M. W., Morton, C. M., & Kallunki, J. A. (1999). Phylogenetic relationships of rutaceae: A cladistic analysis of the subfamilies using evidence from rbcL and atpB sequence variation. American Journal of Botany, 86(8), 1191–1199. https://doi.org/10.2307/2656983

Cortes, C., Vapnik, V. (1995). Support-vector networks. Mach Learn 20, 273–297. https://doi.org/10.1007/BF00994018

Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2004). Digital image processing using MATLAB (Vol. 624). Upper Saddle River, New Jersey: Pearson Prentice-Hall

Haralick, R. M., Dinstein, I., & Shanmugam, K. (1973). Textural Features for Image Classification. IEEE Transactions on Systems, Man and Cybernetics, SMC-3(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Hussain, A., Li, X., Weng, Y., Liu, Z., Ashraf, M. F., Noman, A., Yang, S., Ifnan, M., Qiu, S., Yang, Y., Guan, D., & He, S. (2018). CaWRKY22 acts as a positive regulator in pepper response to ralstonia solanacearum by constituting networks with caWRKY6, caWRKY27, caWRKY40, and caWRKY58. International Journal of Molecular Sciences, 19(5). https://doi.org/10.3390/ijms19051426

Pradhan, A. (2015). SUPPORT VECTOR MACHINE-A Survey. September 2012.

Saptanti, K., Nisa, F., Pamudji, S., & Wonorahardjo, S. (2021). The effect of kaffir lime (Citrus hystrix) leaves extract in chloroform as ladybug (Epilachna sparsa) antifeedant, a preliminary study. AIP Conference Proceedings, 2353(May). https://doi.org/10.1063/5.0052801

Singh, S., & Bhamrah, M. S. (2015). Leaf Identification Using Feature Extraction and Neural Network. IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), 10(5), 134–140. https://doi.org/10.9790/2834-1051134140

Suwannarach, N., Khuna, S., Kumla, J., Cheewangkoon, R., Suttiprapan, P., & Lumyong, S. (2022). Morphology Characterization, Molecular Identification, and Pathogenicity of Fungal Pathogen Causing Kaffir Lime Leaf Blight in Northern Thailand. Plants, 11(3). https://doi.org/10.3390/plants11030273

Wang, X., & Pardalos, P. M. (2014). A Survey of Support Vector Machines with Uncertainties. Annals of Data Science, 1(3–4), 293–309. https://doi.org/10.1007/s40745-014-0022-8