Classification of *Artocarpus* species based on leaf recognition using multiclass support vector machine

To cite this article: S Daliman and N Abdul Ghapar 2021 *IOP Conf. Ser.: Earth Environ. Sci.* **842** 012073

IOP Conf. Series: Earth and Environmental Science **842** (2021) 012073 doi:10.1088/1755-1315/842/1/012073

Classification of *Artocarpus* species based on leaf recognition using multiclass support vector machine

S Daliman[1,*], N Abdul Ghapar[1]

[1]Faculty of Earth Science, Universiti Malaysia Kelantan, 17600 Jeli, Malaysia
*Corresponding author: shaparas@umk.edu.my

Abstract. The demand of automated tools has been increasing regarding to the lack of people that expert in taxonomist. The aim of this research is to identify the classification of *Artocarpus* species based on leaf recognition using multiclass Support Vector Machine. This study focusses on identification and classification of selected *Artocarpus* species which are *A. heterophyllus*, *A. altilis*, *A. integer* and *A. odoratissimus* that belong to genus Artocarpus and family Moraceae through their morphological and features extraction by using image processing method. Multiclass Support Vector Machine (SVM) will be used to get the highest accuracy for the classification of *Artocarpus* species. The combination of Prewitt algorithm, Canny algorithm and gray level co-occurrence matrix will be used in SVM. This study capable to provide the results for current accuracy data representation of the selected *Artocarpus* species. The development of Graphical User Interface (GUI) for classification of *Artocarpus* species help user to identify and differentiate the species in faster and easier way especially botanist, taxonomist, and researcher. This system can increase the accuracy and speed of the processing and extraction of features from digital images of leaves samples. A Graphical User Interface utilizes a combination of devices and technologies to give a platform where users can interact with and producing information.

1. Introduction

Artocarpus species is a genus which family of the Moraceae. Moraceae family also known as fig family or mulberry family. Moraceae family consists of approximately 60 shrubs and trees that located in Pacific Origin and Southeast Asian. The plants belong to this family are flowering plants that consist of more over 1000 species and 40 genera. The family of moraceae mostly widespread in subtropical and tropical regions. *Artocarpus* species are shrubs or laticiferous trees that are made of twigs, leaves and stem that are able to produce milky latex. The family of Moraceae type is monoecious which is consist both female and male flowers and will produces unisexual flowers within a similar plant or dioecious trees. The plants of *Artocarpus* species bring out greenish, small, female flowers that grow fleshy spikes and short. The plants from *Artocarpus* species have varies characteristics of leaf such as have opposite, or alternate leaves and they can be unlobed or lobed. The fruit of Moraceae family usually submerged in fleshy receptacle, often connected into a syncarp [1]. Some of *Artocarpus* species have edible fruit and some of it used for timber and furniture. It is also having medicinal importance and for traditional uses. Identification by using manual method requires researcher and botanist to have a depth knowledge in herbaria [2].

Nowadays, identification of the plant species has been complicated and difficult. This is because some of plant leaves have almost similar morphological such as pattern, color and shade of leaves. Knowledgeable people that have worked in plants for a long period of time such as botanist and taxonomist enable to identify and classify plant species. However, the common method uses in

identification of plant species is collect plant specimens into herbarium. This method also very popular among researchers and students. This method consumes more time and labor because they are many procedures during conducting this method such as collect specimen, dried the parts of plant, presses, and mount all the specimen to be stored. This makes general public difficult on gaining species information or acquiring knowledge of the plant species. In this modern world, computers have a lot of importance and ease of work for everyone. Computers are now a much-needed technology. There is an increased demand in automating tool for identification process of the plant species. The advantages and availability of high technologies, such as mobile devices, digital cameras, pattern recognition in image processing, and remote access to databases enable the idea of implementation a program for species identification to come true. This method consumes less time and labor because the image processing of leaf can be done automatically in the software. The identification and classification of plant species can be performed in the system.

2. Materials and Method

2.1. Study area

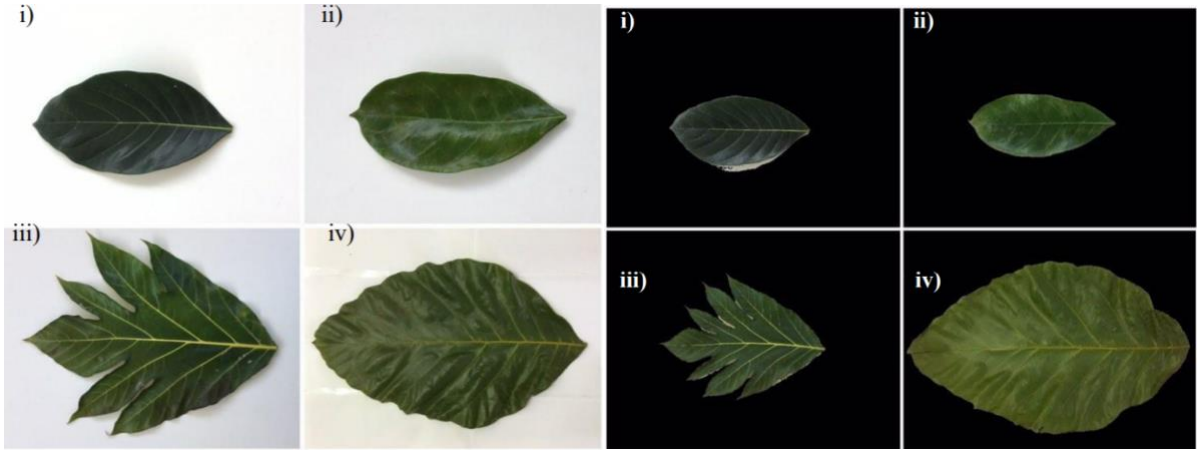
Images of leaves from *Artocarpus* species from four different species are taken in Selangor and Jeli which are *A. integer, A. heterophyllus, A. altilis* are taken in Selangor and *A. odoratissimus* are taken in Jeli, Kelantan. 200 sheets of leaf are collected which consists of 50 samples for each species (A. integer, 50; A. heterophyllus, 50; A. altilis, 50; A. odoratissimus,50;) for this study. All the leaf samples are captured by using iPad and all the photos of leaf samples are changed to the digital images by using a MATLAB software. Compared with other methods, such as cell and molecule biology methods, identification of plants based on leaf image is the most successful and proven method [3].

2.2. Data Collection

200 sheets of leaf are collected which consists of 50 samples for each species. The next step for this study is create leaf database collection. The stem of leaves has been removed because it can affect the feature extraction process. Photos of leaf samples are taken by using iPad with high resolution camera to obtain the better results and efficiency. All images are captured with white background to avoid more unwanted region to be remove and process. All the images are stored in standard jpg format. Then, image processing techniques are applied to these images to extract features from leaf which required for image analysis.

2.3. Data Processing

Image pre-processing of background removal have been applied before leaf recognition been analysed as shown in Figure 1. This is because the edge of the leaf interferes with a shadow from the background of the leaf. Hence, the background of leaf images were removed in MATLAB software. In this process, the original images have been converting to grayscale images. The original images are changed to RGB value and then to HSV value. Next, the segmentation process is carrying out to isolate the desire object to be processed which is leaf only. In this process, Prewitt and Canny algorithms are used as edge detection of the leaf. Imperfect leaves such as small hole and color patch is removed by applying binary gradient mask to improve the appearance of leaf images. In this study, length, width and pattern of the leaf are extracted. Leaf extraction is done by using Gray Level Co-occurrence Matrix (GLCM) analysis in [4] that are performed by using toolbox from MATLAB. After the features has been extracted, all these features are used to identify and classify the leaf by using classifier of multiclass Support Vector Machine (SVM).



(a) Original images (b) Background removal

Figure 1. Image pre-processing of *Artocarpus* species leaf i) *A. heterophyllus*, ii) *A. integer*, iii) *A. altilis*, and iv) *A. odoratissimus*

2.4. Data Analysis

Artocarpus species from four different species which are *A. integer, A. heterophyllus, A. altilis, and A. odoratissimus* are analysed based on their three features which are width, length, and pattern of the leaf by using algorithm. In order to analyze the performance of the feature selection stage, 200 leaf samples are divided into two sets which are training and testing sets to get the highest accuracy of features in classification of four species belong to genus *Artocarpus*. Data samples used for leaf recognition were 80% for training and 20% for testing to get the highest percentage of accuracy class of four selected species under genus *Artocarpus* for multiclass SVM classification.

2.5. Development of Graphical User Interface (GUI)

In this study, a user-friendly Graphical User Interface is designed for classification of selected *Artocarpus* species. The first interface is the introduction of the GUI. The title of the GUI is provided for the first interface. The second interface was designed for removal of background image, edge detection of the image by using GLCM, Canny algorithm and Prewitt algorithm. The final result of classification of selected *Artocarpus* species will be shown in the second interface.

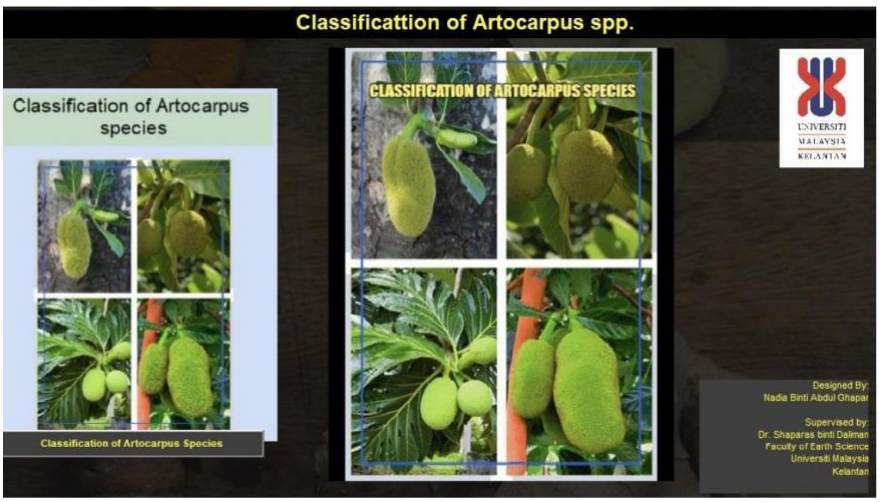


Figure 2. Introduction interface of GUI for classification of *Artocarpus* species leaf

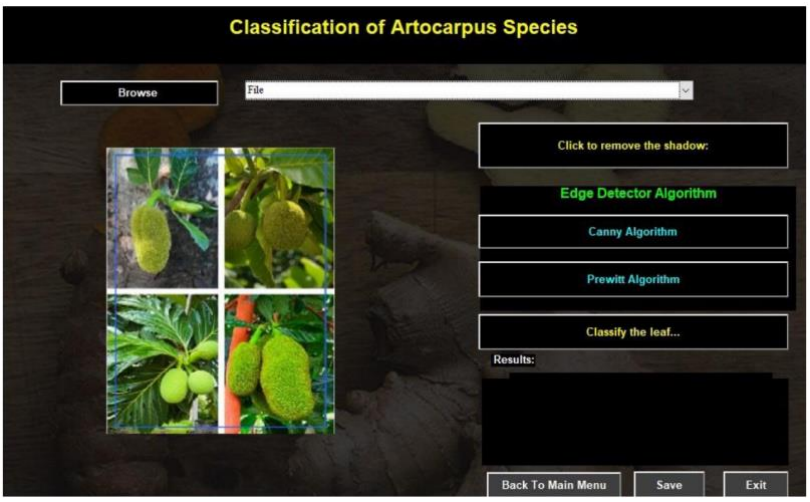


Figure 3. The analysis interface of GUI for classification of *Artocarpus* species leaf

3. Results and Discussion

A multiclass SVM was used in testing and training leaves sample for classification of selected *Artocarpus* species. The features based on GLCM, Canny algorithm, Prewitt algorithm and the combination of two algorithms and combination of all features, are being analysed.

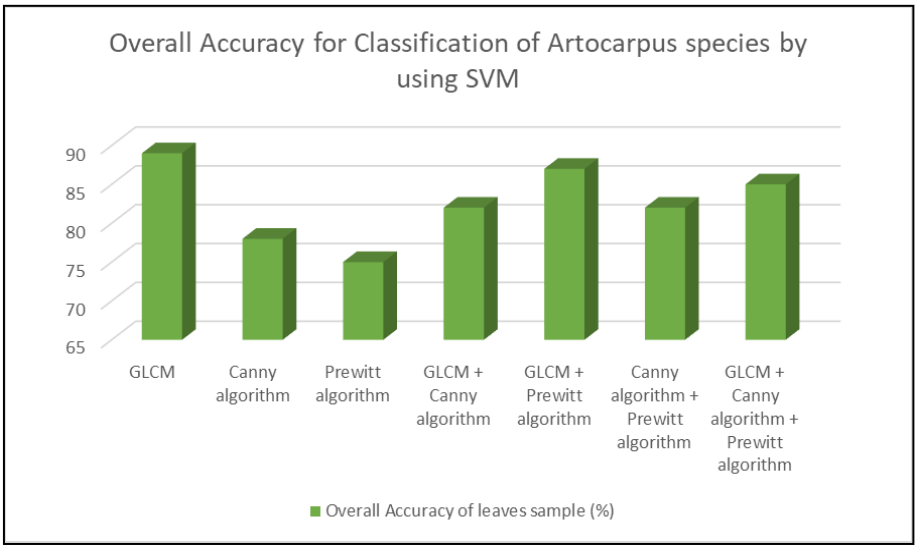


Figure 4. Overall accuracy for classification of *Artocarpus* species by using multiclass SVM

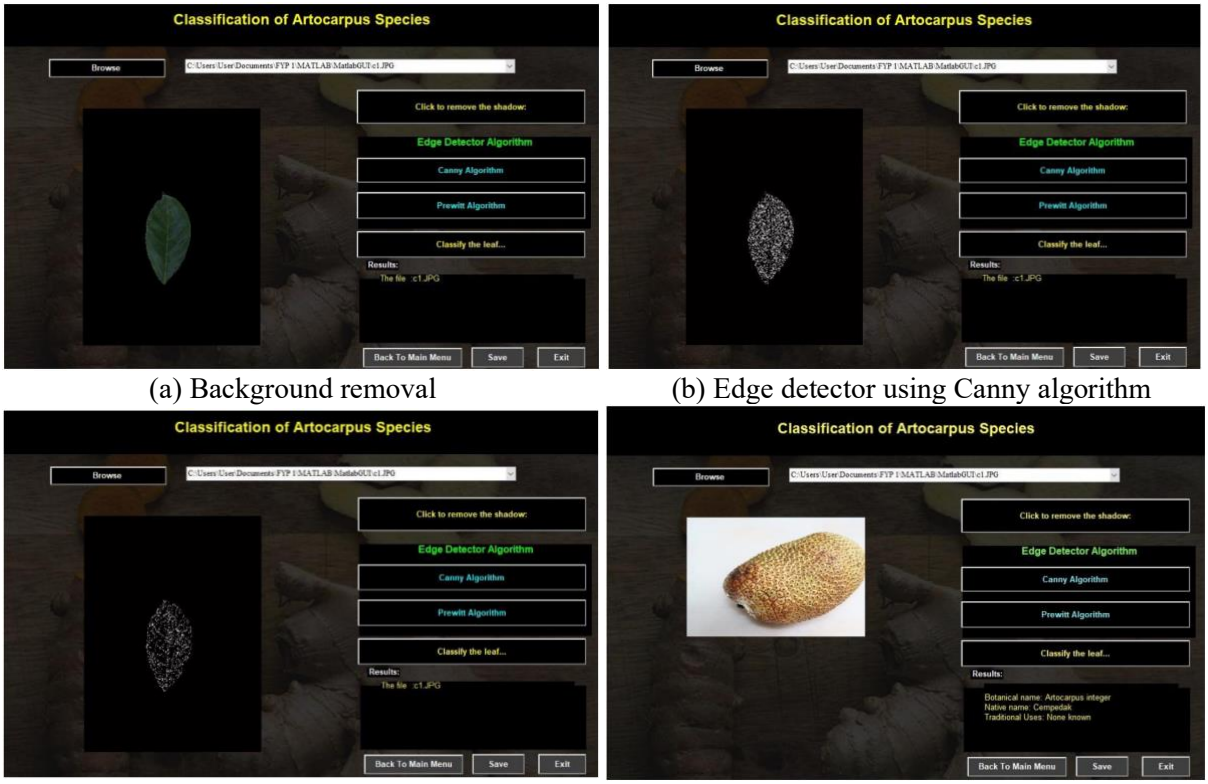
Figure 4 illustrate the percentage of overall accuracy for classification of selected *Artocarpus* species by using multiclass SVM. The results above was compared based on different feature extraction from GLCM, Canny algorithm, Prewitt algorithm, combination of GLCM and Canny algorithm, combination of GLCM and Prewitt, Canny and Prewitt algorithm, and combination of all features.

Based on the bar chart, the GLCM recorded the most accurate result with overall accuracy 89% which proved that image processing was successful method for classification of plant based on the feature extraction of the leaf. However, Prewitt algorithm recorded the lowest accuracy which is 75%. However, *A.integer* and *A.heterophyllus* obtained low accuracy for classification using SVM model due to the same shape of both species. The presence of noise also could be a factor of low accuracy. In this study, 35 out of 50 leaves sample for each species were used for training process and the remaining 15

leaves sample are used for testing process in SVM model. The result could be more accurate if the training data increased.

Artocarpus have many traditional medicinal properties that can benefit to people especially to indigenous people who need to use traditional medicine to cure certain diseases. This study capable to provide the results for current accuracy data representation of the selected *Artocarpus* species. It allows botanist and researcher to recognize this species in easier way and explore new drugs for medicine from this species. Since leaves are the organ of plants and their shapes shift between various species, the leaf shape gives significant data to plant identification.

The Graphical User Interface (GUI) allows users to identify and classify the *Artocarpus* species in effective way. Automated tools are significant in pharmacological science, forestry, and agriculture fields [5]. The GUI result of classification of *A.integer* is shown in Figure 5.



(a) Background removal (b) Edge detector using Canny algorithm

(c) Edge detector using Prewitt algorithm (d) Result classification of *A. integer* in GUI

Figure 5. GUI for classification of Artocarpus species by using multiclass SVM

4. Conclusion

By using image processing method of leaf recognition, the efficiency of a classification technique can be obtained and help researcher to differentiate and classify the *Artocarpus* species in a fast way compared to manual process. Plant identification through leaf characterization is the field of research which is achieving maximum significance recently [6]. This study capable to provide the results for current accuracy data representation of the selected *Artocarpus* species. The development of Graphical User Interface (GUI) for classification of *Artocarpus* species help user such as botanist, taxonomist, and researcher to identify and differentiate the species easier. Moreover, they can identify the uses of *Artocarpus* species easily once they identify the species such as for traditional use and medicinal significance. This study also gives an opportunity for general public to contribute the shared knowledge of plant species.

Acknowledgement
Authors would also like to express appreciation to the members of Faculty of Earth Science for their support and valuable comments and ideas. Authors also acknowledge the Universiti Malaysia Kelantan for providing facilities and financial support to proceed this study.

References
[1] Somashekhar M, Nayeem N and Sonnad B 2013 A review on family Moraceae (Mulberry) with a focus on Artocarpus species *World J Pharm Pharm Sci* **2** 2614-21.
[2] Belhumeur P N, Chen D, Feiner S, Jacobs D W, Kress W J, Ling H and L Zhang 2008 Searching the world's herbaria: A system for visual identification of plant species *European Conference on Computer Vision* Springer Berlin Heidelberg 116-129.
[3] Du, J X, Wang X F and Zhang G J 2007 Leaf shape-based plant species recognition *Applied mathematics and computation* 185(**2**) 883-893.
[4] Ramos E and Fernández D S 2009 Classification of leaf epidermis microphotographs using texture features *Ecological Informatics* 4(**3**) 177-181.
[5] Nagamalai D, Renault E and Dhanuskodi M 2011 Trends in Computer Science Engineering and Information Technology Springer.
[6] Waldchen J, Rzanny M, Seeland M and Mäder P 2018 Automated plant species identification—Trends and future directions *PLoS computational biology* 14(**4**) e1005993.