UNIVERSITI PUTRA MALAYSIA

AUTOMATED PLANT CLASSIFICATION SYSTEM USING A HYBRID OF SHAPE AND COLOR FEATURES OF THE LEAF

LAITH EMAD HAMID

FK 2016 128

AUTOMATED PLANT CLASSIFICATION SYSTEM USING A HYBRID OF SHAPE AND COLOR FEATURES OF THE LEAF

By

LAITH EMAD HAMID

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

September 2016

DEDICATION

To the ones who loved and supported me endlessly

My Beloved Family

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Master of Science

AUTOMATED PLANT CLASSIFICATION SYSTEM USING A HYBRID OF SHAPE AND COLOR FEATURES OF THE LEAF

By

LAITH EMAD HAMID

September 2016

Chairman : Syed Abdul Rahman Al-Haddad bin Syed Mohamed, PhD
Faculty : Engineering

Automated plant leaf classification is a computerized approach that employs computer vision and machine learning algorithms to identify a plant based on the features of its leaf. The last few decades have witnessed various approaches to implement plant classification systems. Several approaches have been proposed using different features and classifiers. However, the majority of the existing methods either rely on large numbers of training samples or select certain leaves within a dataset to achieve high accuracy rates. The disadvantage of such practices is that the results may not reflect the actual expressiveness of the features to tackle the high interclass similarity among different species. Furthermore, most of the existing systems rely on human intervention to select certain points of the leaf to help the system align the leaf or to select the best result among a few candidates after the classification is done.

An Automated Plant Classification System (APCS) is introduced in this thesis to overcome the aforementioned limitations by proposing an automated alignment algorithm to eliminate the need for human intervention to align the leaf. A new set of Quartile Features (QF) is also proposed to express the partial shape of the leaf. Furthermore, optimizing the performance is also targeted in this research by integrating the proposed Quartile Features with the most discriminant shape and color features in the literature, in order to select the optimal feature vector for the proposed system. The proposed automated alignment algorithm is based on a similarity measure between the vertical and horizontal halves of the leaf. Once the leaf is aligned, the image is sliced into horizontal and vertical quartiles, and the area of each quartile is calculated to extract the proposed Quartile Features.

To optimize the performance and select the final features for the proposed system, Quartile Features and the other categories of shape and color features investigated in this research have been tested and evaluated individually and in

combinations. The most discriminant features in each category are then combined to form the final feature vector input to the classifier. A Nearest Neighbor classifier (1-NN) is used to compute the similarity of a query leaf image with all the samples in the database by calculating the distance between their respective feature vectors.

The experiments in this research have been conducted using two leaf datasets. The first is Flavia dataset which has been used as a benchmark by several researchers in the field of plant recognition. The second dataset is collected by the author, from Putrajaya and Perdana Botanical gardens, containing a total of 396 leaves from 17 species endemic to Malaysia and Tropical Asia. The experimental results and comparisons indicate the efficiency of the proposed automated alignment algorithm and the proposed Quartile Features. The results of using the final selected features have shown an impressive performance, achieving an average accuracy rate of 98.32% for Flavia dataset and 91.29% for Leaves dataset, using k-fold cross-validation.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Master Sains

SISTEM PENGELASAN TUMBUH-TUMBUHAN AUTOMATIK MENGGUNAKAN KOMBINASI CIRI-CIRI BENTUK DAN WARNA DAUN

Oleh

LAITH EMAD HAMID

September 2016

Pengerusi : Syed Abdul Rahman Al-Haddad Bin Syed Mohamed, PhD
Fakulti : Kejuruteraan

Klasifikasi tumbuh-tumbuhan berasaskan daun adalah satu pendekatan berkomputer yang menggunakan penglihatan komputer dan algoritma pembelajaran mesin untuk mengenalpasti tumbuhan berdasarkan ciri-ciri daun tersebut. Beberapa dekad yang lalu telah mencapai pelbagai pendekatan untuk melaksanakan sistem pengelasan tumbuhan. Beberapa pendekatan telah dicadangkan menggunakan ciri-ciri dan pengelasan yang berbeza. Walaubagaimanapun, majoriti kaedah yang sedia ada samada bergantung kepada jumlah yang besar sampel latihan atau memilih daun tertentu dalam set data untuk mencapai kadar ketepatan yang tinggi. Kelemahan cara tersebut adalah keputusan mungkin tidak menggambarkan ekspresi sebenar ciri-ciri untuk menangani persamaan yang tinggi di kalangan spesis yang berbeza. Tambahan pula, kebanyakan sistem yang sedia ada bergantung kepada campur tangan manusia untuk memilih titik-titik tertentu pada daun untuk membantu sistem menyelaraskan daun atau untuk memilih keputusan yang terbaik di antara beberapa keputusan selepas pengelasan dilakukan.

Sistem Klasifikasi Tumbuhan Automatik (APCS) diperkenalkan dalam tesis ini untuk mengatasi had yang dinyatakan sebelumnya dengan mencadangkan satu algoritma penjajaran automatik untuk menghapuskan peranan campur tangan manusia untuk menyelaraskan daun. Satu set baru ciri-ciri kuartil (QF) juga dicadangkan untuk menyatakan bentuk separa daun. Tambahan pula, mengoptimumkan prestasi juga disasarkan dalam kajian ini dengan mengintegrasikan cadangan ciri-ciri kuartil dengan ciri-ciri bentuk dan warna yang paling diskriminan dalam kesusasteraan untuk memilih vektor ciri-ciri yang optimum bagi sistem yang dicadangkan. Algoritma penjajaran automatik yang dicadangkan adalah berdasarkan ukuran persamaan antara bahagian menegak dan mendatar daun. Sebaik sahaja daun sejajar, imej ini dihiris ke

dalam kuartil secara mendatar dan menegak dan keluasan setiap kuartil dikira untuk mengeluarkan ciri-ciri kuartil yang dicadangkan.
Untuk mengoptimumkan prestasi dan memilih ciri-ciri akhir bagi sistem yang dicadangkan, ciri-ciri kuartil dan kategori lain daripada ciri-ciri bentuk dan warna disiasat dalam kajian ini telah diuji dan dinilai secara berasingan dan secara kombinasi. Ciri yang paling diskriminan dalam setiap kategori kemudiannya digabungkan untuk membentuk vektor ciri-ciri akhir yang merupakan input kepada pengelas. Pengelas Jiran Terdekat (1-NN) digunakan untuk mengira persamaan imej daun yang tidak diketahui dengan semua sampel di dalam pangkalan data dengan mengira jarak diantara vektor ciri-ciri.

Eksperimen dalam kajian ini telah dijalankan dengan menggunakan dua set data daun. Yang pertama adalah set data Flavia yang telah digunakan sebagai penanda aras oleh beberapa penyelidik dalam bidang pengecaman tumbuhan. Set data kedua dikumpulkan oleh penulis dari Taman Botani Putrajaya dan Taman Botani Perdana mengandungi sejumlah 396 daun dari 17 spesis endemik dari Malaysia dan Asia Tropika. Keputusan eksperimen dan perbandingan menunjukkan kecekapan algoritma penjajaran automatik dan ciri-ciri kuartil yang dicadangkan. Keputusan menggunakan ciri-ciri akhir yang dipilih telah menunjukkan prestasi yang mengagumkan mencapai kadar ketepatan purata 98.32% untuk set data Flavia dan 91.29% bagi set data daun menggunakan silang pengesahan k kali ganda.

ACKNOWLEDGEMENTS

Praise be to Allah for giving me the strength and determination to finish this research and reach the moment of writing these lines. My heartfelt love and thanks go to my family, especially my beloved parents for their love and support on every step of the way.

I would also like to express my deepest gratitude and appreciation to my supervisor Assoc. Prof. Dr. Syed Abdul Rahman Al-Haddad for his guidance, patience, and support along the journey of conducting this research. I also extend my gratitude to Dr. Syamsiah bt. Mashohor who inspired the idea of this work, and Prof. Iqbal b. Saripan for their guidance and advice.

Finally, I am thankful to all friends, colleagues, and UPM staff who had helped with any ideas, facilities, or advice to improve this thesis.

I certify that a Thesis Examination Committee has met on 8 September 2016 to conduct the final examination of Laith Emad Hamid on his thesis entitled "Automated Plant Classification System using a Hybrid of Shape and Color Features of the Leaf" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Master of Science.

Members of the Thesis Examination Committee were as follows:

Fakhrul Zaman bin Rokhani, PhD
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Abd. Rahman bin Ramli, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Aini Hussain, PhD
Professor
Universiti Kebangsaan Malaysia
Malaysia
(External Examiner)

NOR AINI AB. SHUKOR, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 22 November 2016

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Syed Abdul Rahman Al-Haddad b. Syed Mohamed, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Syamsiah bt. Mashohor, PhD
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Member)

M. Iqbal b. Saripan, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

BUJANG BIN KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ________________________ Date: __________________

Name and Matric No.: Laith Emad Hamid , GS31508

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ______________________
Name of Chairman of Supervisory Committee: ______________________

Signature: ______________________
Name of Member of Supervisory Committee: ______________________

Signature: ______________________
Name of Member of Supervisory Committee: ______________________

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1-NN	Nearest Neighbor
ANN	Artificial Neural Network
APCS	Automated Plant Classification System
BPNN	Back Propagation Neural Network
CCD	Centroid Contour Distance
CCG	Centroid Contour Gradient
C-R	Centroid Radii
DMF	Digital Morphological Features
DSLR	Digital Single-Lens Reflex
DWT	Discrete Wavelet Transform
EM	Expectation-Maximization
EOH	Edge Orientation Histogram
FFNN	Feed-Forward Neural Network
FFT	Fast Fourier Transform
GBCM	Grid-Based Color Moments
GLCM	Gray Level Co-occurrence Matrix
GRNN	General Regression Neural Network
HF	Horizontal Feature
HLF	Half Leaf Feature
HoCS	Histogram of Curvature over Scale
HQ	Horizontal Quartile
HSI	Hue, Saturation, and Intensity
JPEG	Joint Photographic Experts Group
k-NN	*k*-Nearest Neighbor
L*a*b*	Luminance Color Space

LDA	Linear Discriminant Analysis
LEOH	Local Edge Orientation Histogram
LMI	Legendre Moment Invariants
LOOCV	Leave-one-out Cross-Validation
LSH	Locality Sensitive Hashing
LWF	Leaf Width Factor
MCH	Moving Center Hyper-sphere
MI	Moment Invariants
PFT	Polar Fourier Transform
PNN	Probabilistic Neural Network
QF	Quartile Feature
RF	Random Forests
RGB	Red, Green, and Blue
ROI	Region of Interest
SIFT	Scale-Invariant Feature Transform
SVM	Support Vector Machine
TMI	Tchebichef Moment Invariants
VF	Vertical Feature
VQ	Vertical Quartile
ZMI	Zernike Moment Invariants

CHAPTER 1

INTRODUCTION

1.1 Background

It is no mystery why plants are important in our lives. They are the primary source of food to almost all animals and humans. They provide Oxygen, without which there would be no life on this planet. Plants also have a great importance in medicine, clothing, fuel and many other industries and fields of life. Therefore, it is essential to understand plants and classify them.

Plant classification is based on studying the morphological features of several parts of the plant such as the leaves, the flowers, the barks, etc. A label is assigned to each plant according to the binomial nomenclature system presented by the Swedish botanist Carl Linnaeus in 1753. Linnaeus introduced a classification system that arranges plants in a hierarchical order according to the similarities amongst them. Since then, plant taxonomy has been conducted based on that system [1]. However, plant classification is almost exclusively performed by professional botanists and taxonomists due to the extensive botanical knowledge required to identify the species of a plant. Considering the massive number of plants all over the globe and the increasing need to classify them, the necessity to implement a computerized plant classification system has emerged.

Computerized plant classification is an interesting field of image processing and computer vision that can be useful in any industrial or commercial application where plants are used. It can save time drastically and identify plants at a glance, especially for non-experts who deal with plants in many scientific or industrial fields. Several approaches have been presented by researchers in the past few years to classify plants using their leaves. Some approaches targeted certain regions such as Leafsnap [2] in The United States, ReVes [3] and Pl@ntNet [4] in Europe, and Clover [5] in Asia. Other approaches focused on classifying medicinal plants [6]–[9], while a few focused on differentiating weed from the crop [10]–[12]. Despite the large number of approaches and the impressive performance of some of them, computerized plant classification is still an area of active research that is fraught with challenges and limitations.

In this thesis, an Automated Plant Classification System (APCS) is proposed for identifying plant species based on the characteristics of their leaves. The proposed system consists of four main stages. The first stage is leaf image acquisition where the leaf samples are digitized. The second stage is preprocessing the leaf to remove the noise, discard the background, and align the leaf. An automated alignment algorithm is proposed in this thesis based on a similarity measure extracted from bisecting the leaf into two halves. The next stage of the system is extracting the features of the leaf. A new set of features extracted from the leaf quartiles is combined with digital morphological features, vein features, and Hu's moment invariants to represent the shape of the leaf. Color information is also taken into consideration, unlike most approaches that usually ignore the color due to the green nature of most leaves. Color features are extracted from the Hue, Saturation, and Intensity components of the (HSI) color space, and combined with the shape features to form the final feature vector that consists of the most discriminative features. The final stage is the classification stage where the features are input into a Nearest Neighbor classifier (1-NN) that predicts the species of a plant by measuring the distance between the feature vector of the input image and the database.

1.2 Motivation

Presenting an automated plant recognition and classification system will not only help people specialized in botany save time and effort to identify a plant but will also open the horizons to non-experts to investigate plants and discover more about them. Researchers in fields like medicine, food industry, and environmental studies can benefit from having a user-friendly tool to recognize the plants without the need for a thorough botanical knowledge. It can also attract nature lovers, enthusiasts, or even tourists who are curious about plant life. Furthermore, in a country like Malaysia with such rich agricultural diversity, implementing an efficient plant classification system can be a helpful tool in all the aforementioned fields.

1.3 Problem Statement

Several plant recognition algorithms have been proposed by researchers, each having its own advantages and disadvantages. Some of those algorithms achieved high accuracy rates and made it out of the prototype phase into the hands of the users as a smartphone application or a web-based system. However, there are several obstacles and limitations facing the implementation of an efficient automated plant classification system.

Due to the irregularity and variation of the leaf shape, one of the challenges against the automation of plant classification systems is aligning the leaf image according to its main vein. Leaf alignment is an important step to extract partial shape features. The main vein of a leaf is represented by the line connecting the base and the apex points. For the leaves whose main vein is the longest line across the leaf blade, i.e., the major axis, the alignment is usually accomplished by calculating the orientation of the major axis of the leaf. The challenging part is aligning the leaves that have a short main vein as the major axis represents the width of the leaf or any line that is longer than the main vein. This normally occurs in lobed leaves or leaves that have a cordate or a reniform shape causing an incorrect feature extraction. To overcome this issue, some approaches rely on the user to detect the main vein by selecting the tip and the base points to align the leaf [13]–[15]. This human interaction, however, is an obstacle against full automation of the classification system. It can also slow down batch processing or even prevent extensive experimentation as each leaf image needs to be aligned manually. Moreover, automating the alignment is not only important for automating the system, it is essential to extract the partial shape features of the leaf as proposed in this research.

Another major issue that makes leaf classification a challenging task is the high interclass similarity among different species and the low intraclass similarity within the leaves of some species. To overcome this problem, most researchers use a large number of training samples to increase the accuracy rates [13]–[19]. However, such high accuracy rates cannot be a real indicator of the expressiveness of the extracted features as they are often attributed to a large number of training samples versus a limited number of testing samples.

The green color of the majority of the leaves is another challenge that discouraged researchers from utilizing color information focusing mostly on the shape features. However, a few methods have been proposed to involve color features in the classification process, such as calculating the average of the color proposed by Pornpanomchai et al. [20], dividing the color space into cells proposed by Iwata et al. [21], and color moment extraction proposed by Prasad et al. [22]. The results of those approaches indicated improvements in the accuracy after including color features. Therefore, color features are investigated in this research as well.

Based on the issues and challenges stated earlier, the following research questions are addressed in this thesis:

a) How to automate the leaf alignment and eliminate the human intervention required to align the leaves whose main vein is shorter than the major axis?
b) How to extract features that describe the partial shape of the leaf and the relation among its multiple parts in order to overcome the high interclass global shape similarity?
c) What are the most discriminative global shape and color features that can be integrated to optimize the classification accuracy?

1.4 Research Objectives

The aim of this research is to implement an Automated Plant Classification System (APCS) that combines the most discriminative features of the leaf to optimize the classification accuracy using a Nearest Neighbor Classifier (1-NN). In order to achieve this target, the following three main objectives have to be achieved:

a) To implement an automated alignment algorithm that can correctly align the leaves whose main vein is not represented by the longest line across the leaf blade, i.e., the major axis.
b) To extract a new set of shape features from the leaf quartiles to describe the partial shape of the leaf and the relation among those parts.
c) To optimize the system performance using a Nearest Neighbor classifier (1-NN) by integrating the proposed Quartile Features with the most discriminative shape and color features in the literature.

1.5 Scope of Research

The scope of this thesis covers the classification of plant species using the images of their leaves. The system is designed to classify simple leaves which consist of one entire, undivided blade. The system can also classify lobed leaves as long as the lobation does not reach the main vein. Compound leaves, on the other hand, are beyond the scope of this research. The leaf samples used in this study are photographed against a white sheet of paper. Otsu's global thresholding is used for segmenting the leaf. Therefore, the input leaves must also be captured on a light solid-colored background.

The features extracted in this research are shape and color features. Shape features consist of a set of partial shape descriptors in addition to digital morphological features, moment invariants, and vein features that can be described as global shape descriptors. On the other hand, statistical features are extracted from the components of the HSI color space to describe the color features of the leaf.

The classification algorithm used in this research is k-Nearest Neighbor algorithm using a City Block distance measure and the nearest neighbor, i.e., 1-NN classifier. All the experiments conducted in this study in addition to the optimization of performance are implemented using this classifier.

Two datasets are used to implement and test the proposed algorithms. The first is called Flavia dataset [13] with a total of 1907 leaf images from 32 species, and a second dataset of 396 leaf samples from 17 species endemic to Malaysia and Tropical Asia. The latter was collected by the author from Perdana and Putrajaya Botanical Gardens.

1.6 Contribution

The contribution of this research can be highlighted in the following three main points:

a) Proposing an automated alignment algorithm that can correctly align different leaf shapes including the wide and lobed leaves that usually require human intervention.

b) Proposing a new set of partial shape features extracted from the horizontal and vertical quartiles of the leaf.
c) Optimizing the classification accuracy using a simple Nearest Neighbor classifier (1-NN) and limited numbers of training samples by combining the most discriminative features in each category of the investigated features.

1.7 Thesis Outline

The layout of this thesis is organized as follows:

Chapter 1 is an introduction to plant classification and the principles of computerizing this task, in addition to the motivation and the objectives of this research, and the issues to be addressed.

Chapter 2 starts with an introduction to common leaf shapes and the structure of the leaf followed by a review of the related literature. The literature is categorized and ordered based on each stage of the classification system.

Chapter 3 describes the methodology and the system design adopted to implement the algorithms starting with the preprocessing stage, followed by a detailed description of the features investigated in this thesis and the classification mechanism.

Chapter 4 presents the experiments conducted in this research and evaluates the results of each stage. A comparison of the system performance with the current existing systems is also presented in this chapter.

Chapter 5 concludes this thesis and suggests ideas for future research work.

REFERENCES

[1] D. M. Spooner, W. L. A. Hetterscheid, R. G. Van den Berg, and W. A. Brandenburg, *Plant Nomenclature and Taxonomy: An Horticultural and Agronomic Perspective*, vol. 28. 2003.

[2] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, I. Lopez, and J. V.B.Soares, "Leafsnap: A Computer Vision System for Automatic Plant Species Identification," *Comput. Vis. (ECCV)*, pp. 502–516, 2012.

[3] G. Cerutti, V. Antoine, L. Tougne, J. Mille, D. Coquin, A. Vacavant, G. Cerutti, V. Antoine, L. Tougne, J. Mille, and L. Valet, "ReVes participation-tree species classification using random forests and botanical features," in *Conference and Labs of the Evaluation Forum (CLEF)*, 2012, pp. 1–15.

[4] D. Barthélémy, N. Boujemaa, D. Mathieu, J.-F. Molino, P. Bonnet, R. Enficiaud, E. Mouysset, and P. Couteron, "The Pl@ntnet Project: A computational plant identification and collaborative information system," in *Tech. Rep., XIII World Forestry Congress*, 2009, p. 577.

[5] Y. Nam, E. Hwang, and D. Kim, "Clover: A Mobile Content-Based Leaf Image Retrieval System," *Digit. Libr. Implement. Strateg. Shar. Exp. Lect. Notes Comput. Sci. Springer*, pp. 139–148, 2005.

[6] L. Gao and X. Lin, "A Study on the Automatic Recognition System of Medicinal Plants," in *IEEE 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, 2012, pp. 101–103.

[7] M. S. Sainin and R. Alfred, "Feature Selection for Malaysian Medicinal Plant Leaf Shape Identification and Classification," in *International Conference on Computational Science and Technology (ICCST'14)*, 2014, pp. 1–6.

[8] E. S. Kumar and V. Talasila, "Leaf Features based approach for Automated Identification of Medicinal Plants," in *International Conference on Communication and Signal Processing*, 2014, pp. 210–214.

[9] T. Sathwik, R. Yasaswini, R. Venkatesh, and A. Gopal, "Classification of Selected Medicinal Plant Leaves Using Texture Analysis," in *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, 2013, pp. 1–6.

[10] Z. Bo, W. H. Hua, L. S. Jun, M. W. Hua, and Z. X. Chao, "Research on Weed Recognition Method Based on Invariant Moments," in *Proceeding of the 11th World Congress on Intelligent Control and Automation*, 2014, vol. 11, no. 2012, pp. 2167–2169.

[11] S. Haug, A. Michaels, P. Biber, and J. Ostermann, "Plant Classification System for Crop /Weed Discrimination without Segmentation," in *2014 IEEE Winter Conference on Applications of Computer Vision, WACV*, 2014, pp. 1142–1149.

[12] K. N. Agrawal, K. Singh, G. C. Bora, and D. Lin, "Weed Recognition Using Image-Processing Technique Based on Leaf Parameters," *J. Agric. Sci. Technol.*, vol. 2, pp. 899–908, 2012.

[13] S. G. Wu, F. S. Bao, E. Y. Xu, Y. Wang, Y. Chang, and Q. Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic

Neural Network," in *IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 11–16.
[14] S. Hati and G. Sajeevan, "Plant Recognition from Leaf Image through Artificial Neural Network," *Int. J. Comput. Appl.*, vol. 62, no. 17, pp. 15–18, 2013.
[15] J. Hossain and M. A. Amin, "Leaf Shape Identification Based Plant Biometrics," in *IEEE 13th International Conference on Computer and Information Technology (ICCIT)*, 2010, pp. 458–463.
[16] C. Uluturk and A. Ugur, "Recognition of Leaves Based on Morphological Features Derived From Two Half-Regions," in *International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, 2012, pp. 1–4.
[17] K.-B. Lee and K.-S. Hong, "An Implementation of Leaf Recognition System using Leaf Vein and Shape," *Int. J. Bio-Science Bio-Technology*, vol. 5, no. 2, pp. 57–66, 2013.
[18] C. Arun Priya, T. Balasaravanan, and A. S. Thanamani, "An Efficient Leaf Recognition Algorithm for Plant Classification Using Support Vector Machine," in *Proceedings - International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME)*, 2012, pp. 428–432.
[19] V. Satti, A. Satya, and S. Sharma, "An automatic leaf recognition system for plant identification using machine vision technology," *Int. J. Eng. Sci. Technol.*, vol. 4, no. 5, pp. 874–879, 2013.
[20] C. Pornpanomchai, S. Rimdusit, P. Tanasap, and C. Chaiyod, "Thai Herb Leaf Image Recognition System (THLIRS)," *Kasetsart J. Nat. Sci.*, vol. 45, pp. 551–562, 2011.
[21] T. Iwata and T. Saitoh, "Tree Recognition based on Leaf Images," in *Proceedings of SICE Annual Conference (SICE)*, 2013, pp. 2489–2494.
[22] S. Prasad, S. K. Peddoju, and D. Ghosh, "Mobile Plant Species Classification: A Low Computational Aproach," in *IEEE 2nd International Conference on Image Information Processing (ICIIP)*, 2013, pp. 405–409.
[23] J.-X. Du, D.-S. Huang, X.-F. Wang, and X. Gu, "Computer-aided plant species identification (CAPSI) based on leaf shape matching technique," *Trans. Inst. Meas. Control*, vol. 28, no. 3, pp. 275–285, 2006.
[24] W. P. Armstrong, "Wayne's Word: Botany 115 Terminology, 2001, http://waynesword.palomar.edu/termlf2.htm (Last accessed 7 Nov. 2015)." .
[25] A. Kadir, L. Nugroho, A. Susanto, and P. I. Santosa, "Experiments of Zernike Moments for Leaf Identification," *J. Theor. Appl. Inf. Technol.*, vol. 41, no. 1, pp. 82–93, 2012.
[26] S. Prasad, P. Kumar, and R. C. Tripathi, "Plant Leaf Species Identification using Curvelet Transform," in *IEEE 2nd International Conference on Computer and Communication Technology (ICCCT)*, 2011, pp. 646–652.
[27] O. J. O. Söderkvist, "Computer Vision Classification of Leaves from Swedish Trees," *Master Thesis*, 2001.
[28] H. Goëau, P. Bonnet, A. Joly, I. Yahiaoui, D. Barthélémy, N. Boujemaa, and J.-F. Molino, "The ImageCLEF 2012 Plant Identification Task," in *Conference and Labs of the Evaluation Forum*, 2012, pp. 1–26.
[29] H. Liu, D. Coquin, L. Valet, and G. Cerutti, "Leaf Species Classification

Based on a Botanical Shape Sub-Classifier Strategy," in *IEEE 22nd International Conference on Pattern Recognition (ICPR)*, 2014, pp. 1496–1501.
[30] L. Huang and P. He, "Machine Recognition for Broad-Leaved Trees Based on Synthetic Features of Leaves Using Probabilistic Neural Network," in *Proceedings - International Conference on Computer Science and Software Engineering, CSSE*, 2008, vol. 4, pp. 871–877.
[31] G. Kaur and G. Kaur, "Classification of Biological Species Based on Leaf Architecture," vol. 1, no. 6, pp. 35–42, 2012.
[32] J.-X. Du, X.-F. Wang, and G.-J. Zhang, "Leaf shape based plant species recognition," *Appl. Math. Comput.*, vol. 185, no. 2, pp. 883–893, 2007.
[33] K. Mahdikhanlou and H. Ebrahimnezhad, "Plant Leaf Classification Using Centroid Distance and Axis of Least Inertia Method," in *22nd Iranian Conference on Electrical Engineering (ICEE)*, 2014, pp. 1690–1694.
[34] B. Vijayalakshmi, "A New Shape Feature Extraction Method for Leaf Image Retrieval," in *Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012). Springer India*, 2013, pp. 235–245.
[35] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.
[36] O. Mzoughi, I. Yahiaoui, N. Boujemaa, and E. Zagrouba, "Advanced Tree Species Identification Using Multiple Leaf Parts Image Queries," in *20th IEEE International Conference on Image Processing (ICIP)*, 2013, pp. 3967–3971.
[37] I. Yahiaoui, O. Mzoughi, and N. Boujemaa, "Leaf Shape Descriptor for Tree Species Identification," in *2012 IEEE International Conference on Multimedia and Expo(ICME)*, 2012, pp. 254–259.
[38] A. Akhil, G. Ankit, B. Nitesh, M. Shashwat, and B. Arnab, "A Plant Identification System using Shape and Morphological Features on Segmented Leaflets: Team IITK, CLEF 2012," *CLEF 2012 Eval. Labs Work. Online Work. Notes*, pp. 1–14, 2012.
[39] S. Ridler, T.W. Calvard, "Picture Thresholding Using an Iterative Selection Method," *IEEE Trans. Syst. Man Cybern.*, vol. 8, no. 8, pp. 630–632, 1978.
[40] X.-F. Wang, J.-X. Du, and G.-J. Zhang, "Recognition of Leaf Images Based on Shape Features Using a Hypersphere Classifier," *Adv. Intell. Comput.*, vol. 3644, pp. 87–96, 2005.
[41] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6. pp. 583–598, 1991.
[42] N.-H. Pham, T.-L. Le, P. Grard, and V.-N. Nguyen, "Computer aided plant identification system," in *IEEE International Conference on Computing, Management and Telecommunications (ComManTel)*, 2013, pp. 134–139.
[43] Y. Y. Boykov and M.-P. Jolly, "Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images," *Proceedings. Eighth IEEE Int. Conf. Comput. Vision. ICCV 2001*, vol. 1, pp. 105–112, 2001.

[44] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut - Interactive Foreground Extraction using Iterated Graph Cuts," *ACM Trans. Graph.*, vol. 23, no. 3, p. 309, 2004.
[45] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*, 2nd ed. Gatesmark Publishing, Knoxville, TN, 2009.
[46] "Matlab R2013b Image Processing Toolbox documentation. Mathworks Inc. [Online]. Available: http://www.mathworks.com/help/images/index.html." .
[47] J. Chaki, R. Parekh, and S. Bhattacharya, "Plant leaf recognition using texture and shape features with neural classifier," *Pattern Recognit. Lett.*, vol. 58, pp. 61–68, 2015.
[48] A. Aakif and M. F. Khan, "Automatic classification of plants based on their leaves," *Biosyst. Eng.*, vol. 139, pp. 66–75, 2015.
[49] O. Mzoughi, I. Yahiaoui, and N. Boujemaa, "Extraction of Leaf Parts by Image Analysis," in *International Conference Image Analysis and Recognition*, 2012, pp. 348–358.
[50] L. Gao, X. Lin, M. Zhong, and J. Zeng, "A Neural Network Classifier Based on Prior Evolution and Iterative Approximation Used for Leaf Recognition," in *Proceedings of 6th International Conference on Natural Computation (ICNC)*, 2010, vol. 2, pp. 1038–1043.
[51] M.-K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Inf. Theory*, vol. 8, pp. 179–187, 1962.
[52] J. Chaki and R. Parekh, "Plant Leaf Recognition using Shape based Features and Neural Network classifiers," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 10, pp. 41–47, 2011.
[53] K. Tan, B. C. Ooi, and L. F. Thiang, "Retrieving Similar Shapes Effectively and Efficiently," *Multimed. Tools Appl.*, vol. 19, no. 2, pp. 111–134, 2003.
[54] F. Zernike, "Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode," *Physica*, vol. 1, no. 7, pp. 689–704, 1934.
[55] B. S. Harish, A. Hedge, O. Venkatesh, D. G. Spoorthy, and D. Sushma, "Classification of plant leaves using morphological features and Zernike moments," in *IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2013, pp. 1827–1831.
[56] Z. Zulkifli, P. Saad, and I. A. Mohtar, "Plant Leaf Identification using Moment Invariants & General Regression Neural Network," in *IEEE 11th International Conference on Hybrid Intelligent Systems (HIS)*, 2011, pp. 430–435.
[57] H. Goëau, A. Joly, S. Selmi, P. Bonnet, E. Mouysset, L. Joyeux, J.-F. Molino, P. Birnbaum, D. Barthelemy, and N. Boujemaa, "Visual-based plant species identification from crowdsourced data," in *Proceedings of the 19th ACM International Conference on Multimedia*, 2011, pp. 813–814.
[58] K. Lee, K. Chung, and K. Hong, "An Implementation of Leaf Recognition System Based on Leaf Contour and Centroid for Plant Classification," *Ubiquitous Inf. Technol. Appl. Springer Netherlands*, vol. 214, pp. 109–116, 2013.
[59] M. Bong, G. Bin Sulong, and M. Rahim, "Recognition of Leaf Based on Its

Tip and Base using Centroid Contour Gradient," *Int. J. Comput. Sci. Issues*, vol. 10, no. 2, pp. 477–482, 2013.
[60] C. Y. Gwo, C. H. Wei, and Y. Li, "Rotary matching of edge features for leaf recognition," *Comput. Electron. Agric.*, vol. 91, pp. 124–134, 2013.
[61] H. Hajjdiab and I. Al Maskari, "Plant Species Recognition Using Leaf Contours," *IEEE Int. Conf. Imaging Syst. Tech.*, pp. 306–309, 2011.
[62] C. Im, H. Nishida, and T. L. Kunii, "Recognizing Plant Species by Leaf Shapes - A Case Study of the Acer family," in *Proceedings of IEEE Fourteenth International Conference on Pattern Recognition*, 1998, vol. 2, p. 1171--1173.
[63] S. Mouine, I. Yahiaoui, and A. Verroust-Blondet, "Plant species recognition using spatial correlation between the leaf margin and the leaf salient points," in *IEEE 20th International Conference on Image Processing (ICIP)*, 2013, pp. 1466–1470.
[64] A. Joly and O. Buisson, "A Posteriori Multi-Probe Locality Sensitive Hashing," in *Proceedings of the 16th ACM international conference on Multimedia*, 2008, pp. 209–218.
[65] B. Wang, D. Brown, Y. Gao, and J. La Salle, "Mobile plant leaf identification using smart-phones," in *20th IEEE International Conference on Image Processing (ICIP)*, 2013, vol. 1, pp. 4417–4421.
[66] P. Bhandarkar, H. Doshi, M. Kadakia, and R. Ahmed, "Leaf Identification using Morphology and Structural Decomposition," in *International Conference on Signal Processing and Integrated Networks (SPIN)*, 2014, pp. 125–129.
[67] Q. Wu, C. Zhou, and C. Wang, "Feature Extraction and Automatic Recognition of Plant Leaf Using Artificial Neural Network," *Adv. Artif. Intell.*, vol. 3, pp. 5–12, 2006.
[68] N. A. C. Hussin, N. Jamil, S. Nordin, and K. Awang, "Plant Species Identification by using Scale Invariant Feature Transform (SIFT) and Grid Based Colour Moment (GBCM)," in *IEEE Conference on Open Systems (ICOS)*, 2013, pp. 226–230.
[69] H.B. Kekre, D. Mishra, S. Narula, and V. Shah, "Color feature extraction for CBIR," *Int. J. Eng. Sci. Technol.*, vol. 3, no. 12, pp. 8357–8365, 2011.
[70] B. B. Chaudhuri and N. Sarkar, "Texture Segmentation Using Fractal Dimension," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 1, pp. 72–77, 1995.
[71] A. R. Backes and O. M. Bruno, "Plant Leaf Identification Using Multi-scale Fractal Dimension," in *International Conference On Image Analysis And Processing (ICIAP)*, 2009, pp. 143–150.
[72] A. R. Backes and O. M. Bruno, "Plant Leaf Identification Using Color and Multi-scale Fractal Dimension," in *International Conference on Image and Signal Processing*, 2010, pp. 463–470.
[73] L. Zhang, J. K. J. Kong, X. Zeng, and J. Ren, "Plant Species Identification Based on Neural Network," in *IEEE Fourth International Conference on Natural Computation (ICNC)*, 2008, vol. 5, pp. 90–94.
[74] X. Gu, J.-X. Du, and X.-F. Wang, "Leaf Recognition Based on the Combination of Wavelet Transform and Gaussian Interpolation," in

International Conference on Intelligent Computing, 2005, pp. 253–262.
[75] R. Haralick, K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, 1973.
[76] A. Bhardwaj, M. Kaur, and A. Kumar, "Recognition of plants by Leaf Image using Moment Invariant and Texture Analysis," *Int. J. Innov. Appl. Stud.*, vol. 3, no. 1, pp. 237–248, 2013.
[77] M. Shabanzade, M. Zahedi, and S. A. Aghvami, "Combination of local descriptors and global features for leaf recognition," *Signal Image Process. An Int. J.*, vol. 2, no. 3, pp. 23–31, 2011.
[78] B. . S. Anami, J. D. Pujari, and Rajesh. Yakkundimath, "Identification and Classification of Normal and Affected Agriculture/horticulture Produce Based on Combined Color and Texture Feature Extraction," *Int. J. Comput. Appl. Eng. Sci.*, vol. 1, no. 3, pp. 356–360, 2011.
[79] J. Chaki and R. Parekh, "Plant Leaf Recognition using Gabor Filter," *Int. J. Comput. Appl.*, vol. 56, no. 10, pp. 26–29, 2012.
[80] R. J. Qian, P. J. L. Van Beek, M. I. Sezan, and N. W. P. R. Blvd, "Image Retrieval Using Blob Histograms," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2000, vol. 1, pp. 125–128.
[81] M. Ferecatu, "Image retrieval with active relevance feedback using both visual and keyword-based descriptors," *Unpubl. Dr. Diss. Univ. Versailles, Amsterdam, Holl.*, 2005.
[82] W. H. Rankothge, D. M. S. B. Dissanayake, U. V. K. T. Gunathilaka, S. A. C. M. Gunarathna, C. . Mudalige, and R. . Thilakumara, "Plant Recognition System Based on Neural Networks," in *IEEE International Conference on Advances in Technology and Engineering (ICATE)*, 2013, no. 77, pp. 1–4.
[83] T. Beghin, J. S. Cope, P. Remagnino, and S. Barman, "Shape and Texture Based Plant Leaf Classification," in *International Conference on Advanced Concepts for Intelligent Vision Systems*, 2010, pp. 345–353.
[84] A. Kadir, L. E. Nugroho, A. Susanto, and P. I. Santosa, "Performance Improvement of Leaf Identification System Using Principal Component Analysis," *Int. J. Adv. Sci. Technol.*, vol. 44, pp. 113–124, 2012.
[85] V. Cheung and K. Cannons, "An Introduction to Probabilistic Neural Networks," *http://www.psi.toronto.edu/~vincent/research/presentations/PNN.pdf (Last Accessed in July 2015)*. 2003.
[86] C. S. Sumathi and A. V. S. Kumar, "Plant Leaf Classification Using Soft Computing Techniques," *Int. J. Futur. Comput. Commun.*, vol. 2, no. 3, pp. 2–6, 2013.
[87] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[88] J. Liu, S. Zhang, and S. Deng, "A Method of Plant Classification Based on Wavelet Transforms and Support Vector Machines," *Int. Conf. Intell. Comput.*, pp. 253–260, 2009.
[89] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.

[90] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imaging*, vol. 13, no. 1, pp. 146–165, 2004.
[91] Q. Ye, "A fast algorithm for convex hull extraction in 2D images," *Pattern Recognit. Lett.*, vol. 16, no. 5, pp. 531–537, 1995.
[92] J. C. Russ and R. P. Woods, *The Image Processing Handbook*. CRC Press, 2011.
[93] M. S. Nixon and A. S. Aguado, *Feature Extraction and Image Processing*. Academic Press, 2008.

BIODATA OF STUDENT

Laith Emad Hamid was born in Iraq in 1986. He received his BSc. degree in Electrical Engineering from the University of Babylon in 2008. After graduation, he started working for the International Relief and Development organization (IRD). After that, he was hired by the Foundation of Technical Education in Iraq for three years. He had enrolled the Master of Science program in the field of Computer and Embedded Systems Engineering in Universiti Putra Malaysia. He is also currently a member of the Institute of Electrical and Electronics Engineers (IEEE).

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : ____________________

TITLE OF THESIS / PROJECT REPORT :

AUTOMATED PLANT CLASSIFICATION SYSTEM USING A HYBRID OF SHAPE AND COLOR FEATURES OF THE LEAF

NAME OF STUDENT: LAITH EMAD HAMID

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√)

☐	**CONFIDENTIAL**	(Contain confidential information under Official Secret Act 1972).
☐	**RESTRICTED**	(Contains restricted information as specified by the organization/institution where research was done).
☐	**OPEN ACCESS**	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐	**PATENT**	Embargo from _____________ (date) until _____________ (date)

Approved by:

_____________________	_____________________
(Signature of Student)	(Signature of Chairman of Supervisory Committee)
New IC No/ Passport No.:	Name:
Date :	Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]