

**PLANT LEAF CLASSIFICATION USING SHAPE
FEATURES AND BAYESIAN NETWORK**

**TAN YIN YING
911109-02-5074**

**FACULTY OF COMPUTING AND
INFORMATICS**

2015

**PLANT LEAF CLASSIFICATION USING SHAPE
FEATURES AND BAYESIAN NETWORK**

**TAN YIN YING
911109-02-5074**

**THESIS SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF BACHELOR OF COMPUTER
SCIENCES
(SYSTEM AND NETWORK MANAGEMENT)**

**FACULTY OF COMPUTING AND
INFORMATICS**

2015

DECLARATION

I hereby declare that the material used in this thesis is my own except for quotations, excerpts, equations, summaries and references, which have been duly acknowledged.

8 JULY 2015

TAN YIN YING
BK11160399

CERTIFIED BY

(SUPERVISOR)
DR. MOHD. HANAFI AHMAD HIJAZI

ACKNOWLEDGEMENT

First of all, I would like to thank my supervisor, Dr. Mohd Hanafi Ahmad Hijazi for all his guidance, advices and time for me throughout the research despite of his busy schedule. He has been very patience in guiding me and provided very helpful advice throughout the project.

Secondly, I would like to say thank you to my examiner Dr. Jason Teo for his advice given during presentation, and precious time in checking and marking my project as it is not an easy task.

I would also want to thanks all my lecturers and seniors that taught me during my university studies which giving me invaluable knowledge so that I can applied the knowledge to complete my research.

Last but not least, I would like to thank all my friends and my family members who advised and encouraged me along the way.

ABSTRACT

Environment pollution and overwhelming development by human in recent decades has lead to rise in global warming which cause the plants are at risk of extinction. Therefore, it is very important to have a plant protection plan to manage and archive them from being forgotten by our future generation. With the rapid development of computer technology, plant classification has broken through the traditional methods, rapid identification for plant can be achieved by the image processing and pattern recognition technology with the leaf images provided. There are several features on the leaf that can be used to perform classification. Among of these features, shape is chosen for classification in the research presented in this report because the shape of a leaf is the most distinguishing feature compared to other features. The research presented proposed an approach to classify plant leaf using shape features extracted from the leaf shape images, which will then be forwarded as an input to classifier. A supervised method Bayesian Network that applied search and scoring metrics approach for structure learning is used as classifier. The experiments result is measured in term of average accuracy and standard deviation. The best accuracy obtained from the method proposed is 71.38%. Although the classifier proposed does not perform well, the proposed features are able to produce an accuracy of 83.56% with Neural Network classifier. This result outperformed most of the work in literature reviews that used different features. This shows that the features proposed is capable to produce better accuracy.

ABSTRAK

Pengelasan Tumbuh-tumbuhan Menggunakan Ciri Bentuk dan Rangkaian Bayesian

Pencemaran alam sekitar dan pembangunan bandar yang tidak terhenti-henti oleh manusia dalam beberapa dekad kebelakangan ini telah membawa kepada kenaikan pemanasan global dan seterusnya menyebabkan tumbuh-tumbuhan menghadapi risiko kepupusan. Oleh itu mempunyai pelan perlindungan tumbuhan untuk mengurus dan merekodkan tumbuh-tumbuhan supaya mereka tidak dilupakan oleh generasi yang akan datang adalah sangat penting. Dengan perkembangan yang pesat dalam bidang teknologi komputer, cara klasifikasi tumbuh-tumbuhan telah berevolusi dari kaedah tradisional ke kaedah yang lebih canggih, iaitu tumbuh-tumbuhan boleh dikenal dengan pantas melalui teknik pemprosesan imej dan teknologi pengiktirafan corak dengan imej daun yang disediakan. Terdapat beberapa ciri pada daun boleh digunakan untuk membuat pengelasan. Antara ciri-ciri tersebut, bentuk telah dipilih untuk menjalankan pengelasan dalam penyelidikan yang dibentangkan dalam laporan ini. Ini adalah kerana ciri bentuk daun mempunyai perbezaan paling besar di antara daun-daun dari spesies yang berbeza berbanding dengan ciri-ciri lain. Kajian ini membentang cadangan menggunakan ciri bentuk yang diekstrak daripada imej bentuk daun dan seterusnya digunakan sebagai input kepada pengelas. Kaedah Rangkaian Bayesian dengan pendekatan carian dan pemarkahan metrik telah digunakan sebagai pengelas. Hasil eksperimen akan diukur dari segi ketepatan purata dan sisihan piawai. Keputusan terbaik diperolehi daripada kaedah yang dicadangkan iaitu menggunakan ciri-ciri bentuk diekstrak dan pengelas rangkaian Bayesian, adalah 71.38%. Walaupun ketepatan menggunakan pengelas dicadangkan tidak menunjukkan prestasi yang baik, namun menggunakan ciri-ciri yang dicadangkan mampu mencapai ketepatan 83.56% dengan pengelas Rangkaian Neural. Keputusan ini mengatasi sebahagian besar pekerjaan dalam tinjauan literatur yang menggunakan ciri-ciri berbeza. Ini menunjukkan bahawa ciri-ciri kami mencadangkan lebih mampu dalam menghasilkan ketepatan yang lebih baik

TABLE OF CONTENTS

DECLARATION	ii
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
ABSTRAK	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	xii
LIST OF TABLES	xiii
LIST OF APPENDIX	xv
CHAPTER 1 INTRODUCTION	1
1.1 Overview	1
1.2 Problem Background	2
1.3 Research Questions	3
1.4 Research Objectives	4
1.5 Research Scopes	4
1.5.1 Dataset	4
1.5.2 Features Used in Plant Leaves Classification	5
1.6 Organization of Report	5
CHAPTER 2 LITERATURE REVIEW	7
2.1 Overview	7
2.2 Plant Leaf Recognition	7
2.2.1 Leaf Classification Using Shape Features	8
2.2.2 Leaf Classification Using Texture Features	11
2.2.3 Leaf Classification Using More Than One Features	12
2.3 Feature Selection	19
2.3.1 Information Gain Based Ranking Method	20

2.3.2 Support Vector Machines (SVM) Attribute Evaluator (SVMAttributeEval)	21
2.4 Bayesian Network	25
2.4.1 Search and Scoring Metrics Method	27
2.4.2 Constraints Based Method	29
2.5 Plant Leaf Classification Using Bayesian Network	30
2.6 Platform and Parameters Used	32
2.6.1 Eclipse	32
2.6.2 MATLAB	32
2.6.3 WEKA	34
2.6.4 Search Algorithm	35
2.6.5 Parameter Learning Using Simple Estimator	39
2.6.6 Scoring Function: Bayesian Metric	40
2.7 Summary	41
CHAPTER 3 METHODOLOGY	42
3.1 Overview	42
3.2 Software	42
3.3 Hardware	43
3.4 Digital Image Processing	44
3.5 Features Extraction	44
3.6 Features Selection	45
3.7 Experimental Settings	47
3.7.1 Bayesian Network Classifier in WEKA	48
3.7.2 Other Classifiers Used for Comparison	49
3.7.3 Experiment	52
3.8 Validation of Result	53
3.9 Summary	54
CHAPTER 4 IMPLEMENTATION	55

4.1 Overview	55
4.2 Framework	55
4.3 Image Pre-processing (for certain images only)	56
4.3.1 Binarization	57
4.3.2 Noise Elimination	57
4.4 Feature Extraction	59
4.4.1 Area	61
4.4.2 BoundingBox	61
4.4.3 Centroid	61
4.4.4 ConvexArea	62
4.4.5 Eccentricity	62
4.4.6 EquivDiameter (Equivalent Diameter)	63
4.4.7 EulerNumber	63
4.4.8 Extent	63
4.4.9 Extrema	64
4.4.10 Perimeter	64
4.4.11 Solidity	65
4.5 Feature Selection	65
4.6 Convert File From Excel to ARFF	69
4.7 Experimental Setup in WEKA	71
4.7.1 Full Features (Method Proposed)	72
4.7.2 Reduced Features Set on the Proposed Features	76
4.7.3 Applied Existing Features Available in the Dataset	77
4.7.4 Comparison with Other Classifiers	78
4.8 Test Options	86
4.8.1 Test Options Used	86
4.9 Classification	87

4.10 Summary	88
CHAPTER 5 RESULTS AND ANALYSIS	89
5.1 Overview	89
5.2 Experiment Summary	89
5.3 Experiment 1: Results of full features is applied to Bayes Net classifier using different fine-tuned parameters.	92
5.3.1 Experiment 1a: Results of Tuning Simple Estimator's Alpha Value	93
5.3.2 Experiment 1b: Results of Tuning the Value of Maximum Number of Parent (maxNrOfParents)	94
5.3.3 Experiment 1c: Results of Tuning the Search Algorithm (searchAlgorithm)	95
5.3.4 Overall Analysis	95
5.4 Experiment 2: Results of Reduced Features Applied to Bayes Net classifier.	96
5.4.1 Experiment 2a: Results of Using Reduced Features Produced infoGainAttributeEval Method	97
5.4.2 Experiment 2b: Results of Using Reduced Features Produced SVMAttributeEval Method	99
5.4.3 Overall Analysis	101
5.5 Experiment 3: Results Produced by Applied Features Available in Dataset to Bayes Net classifier using different fine-tuned parameters.	103
5.5.1 Experiment 3a: Results of Tuning Simple Estimator's Alpha Value	103
5.5.2 Experiment 3b: Results of Tuning the Value of Maximum Number of Parent (maxNrOfParents)	104
5.5.3 Experiment 3c: Results of Tuning the Search Algorithm (searchAlgorithm)	105
5.5.4 Overall Analysis	106
5.6 Experiment 4: Comparison with Other Classifiers	107
5.6.1 Experiment 4a: Comparison Among the Classifiers using Full and Reduced Features	107

5.6.2 Experiment 4b: Compare the Performance of Proposed Approach with Approaches in Literature Reviews	110
5.6.3 Overall Analysis	113
5.7 Summary	115
CHAPTER 6 CONCLUSION AND FUTURE WORKS	116
6.1 Overview	116
6.2 Conclusion	116
6.3 Future Works	118
REFERENCES	120
APPENDIX A	126
APPENDIX B	128
APPENDIX C	129

LIST OF FIGURES

Figure 1.1 Examples of binary leaf shape images from the dataset.	5
Figure 2.1: Maximum margin hyperplane (bold blue line) and margins (blue line) for an SVM trained with samples from two classes. Samples on the margin are called the support vectors.	22
Figure 2.2: SVM-train algorithm.	23
Figure 2.3: SVM feature selection algorithm that applied RFE as described in (Guyon <i>et al.</i> , 2002).	25
Figure 2.4 WEKA GUI Chooser interface in WEKA version 3.6.	34
Figure 2.5 The Explorer user interface in WEKA version 3.6.	35
Figure 2.6: Bayesian Network structure that <code>initAsNaiveBayes</code> option is set to false.	37
Figure 2.7: Bayesian Network structure that <code>initAsNaiveBayes</code> option is set to true.	37
Figure 2.8: How Conditional Probabilities Tables are computed in WEKA using simple estimator	40
Figure 3.1: Distribution probability table for Area attribute in Bayesian Network structure.	48
Figure 3.2: Bayesian Network Structure learned by WEKA using Bayes Net classifier setting.	49
Figure 3.3: The Nearest Neighbors of an instance when the $K = 1, 2$ and 3 .	50
Figure 3.4: Naïve Bayes structure which the nodes (X_n) are independent to each other.	51
Figure 3.5: Neural Network	51
Figure 4.1 Flow Chart of the experiment.	56
Figure 4.2: Flow chart of image pre-processing process.	58
Figure 4.3: Pseudo code for feature extraction based on leaf shape images.	60
Figure 4.4: Area of a region.	61
Figure 4.5: The white pixels are region and the green box is the bounding box, the red dot is the centroid.	62
Figure 4.6: Convex area of a region.	62
Figure 4.7 Extent of a region	63

Figure 4.8: Extrema of two types region. Left: each extrema point is distinct. Right: some extrema points (e.g., top-left and left-top) are identical.	64
Figure 4.9: Perimeter of a region.	64
Figure 4.10: The output of "FeatureExtraction.mat" in ".csv" file format.	65
Figure 4.11: Attribute selection interface in WEKA.	66
Figure 4.12: Pseudo code for converting CSV file to ARFF file.	70
Figure 4.13 The Java Build Path setting in project properties window for adding external JARs.	70
Figure 4.14: The ARFF file for leaf shape features data.	71
Figure 4.15: WEKA.gui.GenericObjectEditor which contain options for Bayes Net.	72
Figure 4.16: WEKA.gui.GenericObjectEditor for Hill Climber search algorithm.	74
Figure 4.17: WEKA random dataset source code in Java for cross-validation.	87
Figure 5.1: Leaves from species named Quercus Canariensis (Left) and Quercus Brantii (Right)	113
Figure 5.2 Leaves from species named Alnus Cordata (Left) and Alnus Maximowiczii (Right)	113

LIST OF TABLES

Table 2.1: Summary of related works in plant leaves classification field.	17
Table 2.2: Cases for Learning Bayesian Networks	26
Table 2.3: Bayesian Network learning algorithms for Search and Scoring Metrics method and constraints based method.	30
Table 3.1: Summary of software requirements.	43
Table 3.2: Summary of hardware requirements.	43
Table 3.3: Features used in each experiment.	52
Table 4.1: Parameters values use in for each experiment for full features with Bayes Net classifier.	76
Table 4.2: Features used in each experiment.	77
Table 5.1 Experiment Summary	90
Table 5.2: Results for experiment 1a.	93
Table 5.3: Results for experiment 1b.	94
Table 5.4: Results for experiment 1c.	95
Table 5.5: Results for experiment using top 25 features.	97
Table 5.6: Results for experiment using top 20 features.	97
Table 5.7: Results for experiment using top 15 features.	98
Table 5.8: Results for experiment using top 25 features.	99
Table 5.9: Results for experiment using top 20 features.	100
Table 5.10: Results for experiment using top 15 features.	100
Table 5.11: Results for experiment 3a.	104
Table 5.12: Results for experiment 3b.	105
Table 5.13: Results for experiment 3c.	106
Table 5.14: Results for experiment 4a using full features.	107
Table 5.15: Results for experiment 4a using top 25 features.	108
Table 5.16: Results for experiment 4a using top 20 features.	108
Table 5.17: Results for experiment 4a using top 15 features.	109
Table 5.18: Comparison of results between best results obtained using the combination of classifier and features proposed, and the method proposed in literature review.	110

Table 6.1: Achievement For Each Objective Defines At The Beginning Of This Research.

117

LIST OF APPENDIX

	Page
Appendix A	126
Appendix B	128
Appendix C	129

CHAPTER 1

INTRODUCTION

1.1 Overview

Image processing is a rapidly growing technologies today, with its applications in various aspects of business. Examples of some of the image processing applications are intelligence transportation systems, remote sensing, biomedical imaging techniques and automatic visual inspection system. Used of image processing in applications in different fields can help to save workloads and reduce cost. Thus, more and more fields start to involve image processing applications in their work place, indirectly it also provides opportunity for Artificial Intelligence (AI) research to make improvement or modification on the current applications. Since image processing has been proved to be useful tool for analysis in many fields, so does the agriculture field use it to analysis parameters like canopy, yield, quality of product, leaves, plants. This research will focus on identify plant leaf types by using image classification approaches.

In this section, a simple description of this research entitled "Plant Leaf Classification using shape features and Bayesian Network" is introduced. Section 1.1 contains introduction about the research and gives a general overview about this chapter. Section 1.2 describes problem background of this research and section 1.3 stated the research questions. Section 1.4 states the objective of this research and research scope is presented in section 1.5. The last section 1.5 describes the organization of this report.

1.2 Problem Background

Plants exist in everywhere we live and they are providing us with oxygen, food and medicine. Other than that, plants also help us to regulate climate change, provide shelters and foods to other living organisms. However, due to the rise in global warming effect plants are at risk of extinction. So it is becoming essential to find a viable alternative method in order to protect the plant types from being forgotten by our future generation. One of the effective way is kept records for each of the plant leaf type. So that in future when the plants is facing critical extinction or no longer exist, there still remain records can be provided to people as references for research or education purpose. Records can be made in either hardcopy or softcopy. Hardcopy records are too risky to keep because they are easy to be destroyed or missing. While softcopy, keep in database are not easy to destroy the data. Although there may have accident happen but compare to hardcopy records, the risk is lower.

Nowadays the imaging and digital data storage technology growth rapidly that we can acquire high definition pictures and store them digitally with less monetary cost. Other methods such as biological method: cell biology and molecule biology, that based on biological characteristics of the plants like cells and tissues structure, or genetics and plants Deoxyribonucleic Acid (DNA) structure of plants required a set of complicated processing steps and expert in that field to work on it all the time. Thus, when compared with biological method, automated or semi-automated leaf images labelling using intelligent systems is the prior choice for users and attract lots of researches in related field. (Agarwal *et al.*, 2006; Zulkifli *et al.*, 2011; Wu *et al.*, 2007).Such systems helps to save man power, botanist expert could reduce their workloads since the system will help them to do the classification.

There are several parts of a plant that can be used to identify plants. Among of these features, leaf is an easier and accurate way to identify plants since leaves are considered as one of the important features to characterize plant species. The classification of plant based especially on leaves is a very challenging task due to the high similarity inter-class and low similarity of intra-class (H. Ling and D.W.

Jacobs. 2009). However, the leaf is an important characteristic of the plant for plant classification which can be collected easily at any season or any stage of plant maturity. Thus, the ability to recognize variations of leaf is still an open challenge for research area to focus on the ability to achieve high accuracy for leaf classification. There are many type of approaches can be applied to classify the extracted plant leaf such as neural network, K-nearest-neighbor, Support Vector Machines and etc. Generally, the classification problems often use empirical method to construct mapping rules. The main problem of applying this method is the lack of sufficient information to construct 100% of correct mapping rules, this lead to a condition where uncertainty occurred. To solve this problem, Bayesian Network will be used to perform the identification for plant leaf based on leaf shapes images. Bayesian Network classifier is based on Bayes' theorem which applied conditional probability in the learning phase to achieve a certain probability sense of correct classification. Conditional probability in Bayes' Theorem helps to solve the uncertainty encounter by classifier which used empirical method.

1.3 Research Questions

There are many aspects need to be concerned in building a Bayesian Network classifier for plant leaf classification. The motivation is to produce a leaf classification approach using leaf shape images and Bayesian Network. In this research, binary leaf shape images are used to extract features. Two research questions derived from the identified motivation:

- I. What are the features that can be extracted from the binary leaf shape images that best represents the leaves?
- II. How Bayesian Network method can best be tuned to classify leaf images using the features extracted in (I)?

1.4 Research Objectives

The identified research questions give raised to three research objectives:

- I. To investigate and identify feature extraction technique applied on the binary leaf shape images that can be used to enrich the shape feature representation of plant leaves.
- II. To fine-tune parameters that maximize the predictive accuracy of Bayesian Network in classifying plant leaves.
- III. To compare the performance of proposed approach with other approaches found in the literature reviews.

1.5 Research Scopes

- I. The research is set to investigate only Bayesian Network to plant leaf classification using shape of leaf and then compare this approach with other approaches.
- II. The dataset used is secondary dataset, which is the data was not collected directly through ourselves experiment but is get it online or from other researcher. Since secondary dataset is used in this research, so there might be other researchers are using the same dataset to conduct research too. The details of the dataset used is described in the following subsection 1.5.1.

1.5.1 Dataset

The plant leaves dataset (binary leaf images) used in this research is secondary type dataset. The original owners of this dataset are Beghin *et al.* (2010). The leaves were collected in the Royal Botanic Gardens, Kew, UK. It consists leaves of 100 plant species and there are 16 instances per species, accumulating to 1600 leaf shape images. Thus, there are total of 1600 samples will be used in this research. All image data in dataset contains the binary form of the leaf samples. The color image is not included in the leaf samples. Besides, all the leaf samples have been labeled with its respective types and organized according to its relative feature vectors. There are three set of extracted features that come along with the dataset. These features are shape, margin and texture.

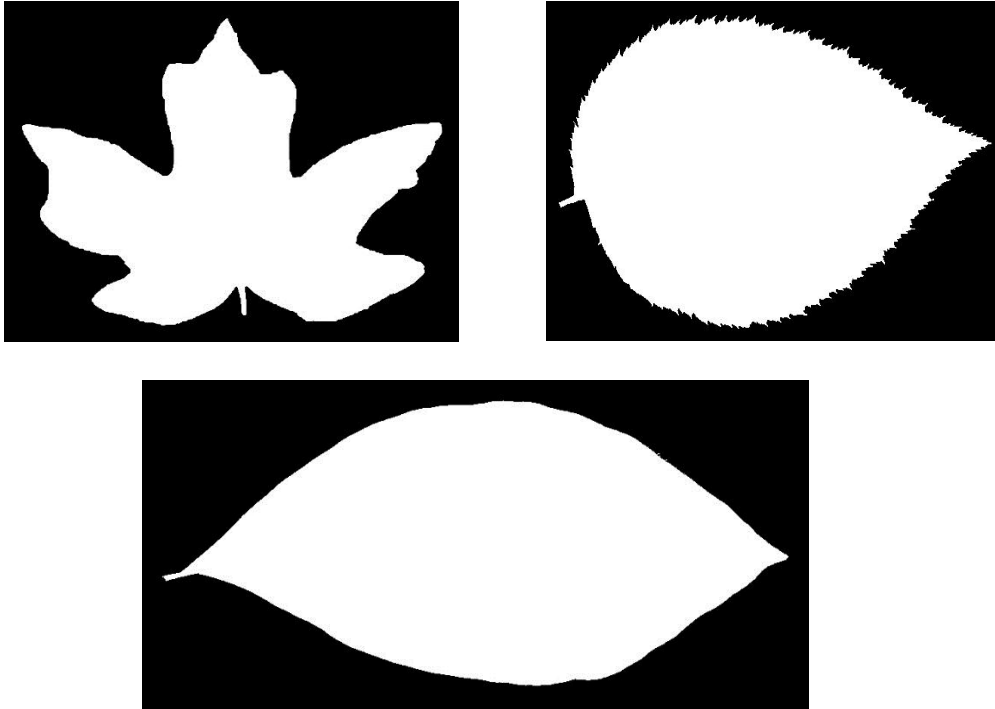


Figure 1.1 Examples of binary leaf shape images from the dataset.

1.5.2 Features Used in Plant Leaves Classification

There are many characteristics on leaves can be used to identify the plant such as shape, texture, color, and margins. Among all of the leaf characteristics, shape analysis will be used as input for Bayesian Network classifier to perform the classification for plant leaf in this research.

1.6 Organization of Report

The report consists of six chapters in total.

- **Chapter 1 Introduction:**

In this chapter, a general overview of the research is provided which include the introduction of the research, problem background, problem statement, objectives, research scope and organization of the report.

- **Chapter 2 Literature review:**

In this chapter, research papers that related to this research is reviewed. A brief description about the techniques and methods used in this research is included in this chapter too. Then a conclusion for all the literature reviews in attached at the end of this chapter.

- **Chapter 3 Methodology:**

This chapter describes image processing, feature extraction techniques used, and the method used to train and test the Bayesian Network classifier. The performance measurements are highlighted in this chapter too.

- **Chapter 4 Implementation:**

The details of this research experiment setup is discussed here. It provides a clear explanation on the feature extraction steps and experiment settings. Besides, this chapter also does contain pseudo code and flow chart of this research.

- **Chapter 5 Results and Analysis:**

Chapter 5 is about the result of the experiments carried out in the research. For each set of experiment, the result will be displayed in table form, then result analysis is done.

- **Chapter 6 Conclusion:**

This chapter summarizes the research and results obtained from the previous chapters and also the future works are discussed.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter reviews relevance work that have been done previously by other researchers which is involving plant leaf classification. This chapter contains five sections. At the beginning of this chapter, section 2.1 provide an overview of the organization of this chapter and section 2.2 will introduce about automated plant leaf recognition. Section 2.2 is divided into three sub-sections for analysis of related work on the approaches that have been used by other researchers in solving plant leaf recognition problems. Section 2.3 introduces the feature selection techniques that will be applied in the research. For this research, Bayesian Network will be used for the recognition of the leaves, so a brief introduction about Bayesian Network methods is discussed in section 2.4. Then follow by section 2.5 which describes the strength of plant leaf classification using Bayesian Network and section 2.6 will discuss the platform and parameters use in the research. Last, the summary of this chapter is included in section 2.7.

2.2 Plant Leaf Recognition

Traditional plant leaves identification can be performed manually by botanist expert but it is time consuming and less effective. Thus, automated plant leaf classification is important towards the people whose working in plant leaves field. Many characteristics on leaves such as shape, texture, and venation can be used as features for plant leaf classification. Many studies on automated plant leaf classification using plant leaf image has been done by researchers. Among these studies, various leaf characteristics is/are used to conduct the computer-aided plant

classification system. There are works which consider only shape (Zulkifli *et al.*, 2011; Wu *et al.*, 2007; Hossain *et al.*, 2010; Uluturk *et al.*, 2012, Du *et al.*, 2005; Florindo *et al.*, 2010; Amina *et al.*, 2013), texture (Cope *et al.*, 2010; Backes *et al.*, 2009) or color information, or take more than one leaf characteristics (Backes *et al.*, 2010; Huang *et al.*, 2008; Zhang *et al.*, 2008; Beghin *et al.*, 2010, Zhang *et al.*, 2012; Tsolakidis *et al.*, 2014; Caglayan *et al.*, 2013; Malleh *et al.*, 2013; Rahmani *et al.*, 2015) for classification.

2.2.1 Leaf Classification Using Shape Features

Among several of the leaf characteristics, shape is the most common features used in leaf image classification. A number of related works have been carried out in analyzing leaf shapes for plant leaf classification. There are a few techniques being introduced in the past few years to improve the performance of leaf recognition using leaf shapes.

Zulkifli, Saad and Mohtar (2011) focused on studies the effectiveness of different moment invariants techniques in extracting region-based shape features from leaf images. These techniques are Zernike Moment Invariant (ZMI), Legendre Moment Invariant (LMI) and Tchebichef Moment Invariant (TMI). After that the feature information from the most effective moment invariants technique are classified using General Regression neural network (GRNN). At the end of the experiment, TMI is found to be the most effective feature extraction technique and the features extracted using TMI were classified using GRNN classifier with 100% correct classification rate.

Wu, Bao, Xu, Wang, Chang and Xiang (2007) use leaf shape features information to classify the leaf too. They employed Probabilistic Neural Network (PNN) for leaf recognition. Five basic geometric features of leaf are extracted, they are diameter, physiological length, physiological width, leaf area and leaf perimeter. Then, 12 digital morphological features (DMFs) are derived from these five basic geometric features. The geometrical features information then are trained and

REFERENCES

- A. Engelbrecht, G.Pampara. 2007. Binary Differential Evolution strategies. Evolutionary Computation, 2007. CEC 2007. IEEE Congress on. 2007 IEEE. Pp.1942-1947.
- Ali Caglayan, Oguzhan Guclu, Ahmet Burak Can. 2013. A Plant Recognition Approach Using Shape and Color Features in Leaf Images. A. Petrosino (Ed.): ICIAP 2013, Part II, LNCS 8157. Springer-Verlag Berlin Heidelberg 2013. Pp.161–170.
- Anang Hudaya Muhamad Amina, Asad I. Khan. 2013. One-Shot Classification of Leaf Shapes using Distributed Hierarchical Graph Neuron (DHGN) Scheme with k-NN Classifier. Procedia Computer Science 24 (2013). 17th Asia Pacific Symposium on Intelligent and Evolutionary Systems, IES2013. 2013 The Authors. Published by Elsevier B.V. Pp. 84 – 96.
- André R. Backes, Odemir M. Bruno. 2009. Plant Leaf Identification Using Multi-scale Fractal Dimension. A. Elmoataz et al. (Eds.): ICISP 2010, LNCS 6134. Springer-Verlag Berlin Heidelberg 2010. Pp.463–470.
- André R. Backes, Odemir M. Bruno. 2010. Plant Leaf Identification Using Color and Multi-scale Fractal Dimension. A. Elmoataz et al. (Eds.): ICISP 2010, LNCS 6134. Springer-Verlag Berlin Heidelberg 2010. Pp.463–470.
- Arief Rakhman, Goeij Yong Sun, Rama Catur APP. 2009. Building Artificial Neural Network Using WEKA Software.
- C.K. Chow, C.N. Liu. 1968. Approximating discrete probability distributions with dependence trees. Information Theory, IEEE Transactions on. IEEE. Pp 462–467.

- C. S. Sumathi, A. V. Senthil Kumar. 2014. Feature Selection for Plant Leaf Classification Based on Information Gain. *International Journal of Advanced Research in Computer Science and Software Engineering*. Volume 4, Issue 10. Pp.856-861.
- Caner Uluturk, Aybars Ugur. 2012. Recognition of Leaves Based on Morphological Features Derived From Two Half-Regions. *Innovations in Intelligent Systems and Applications (INISTA), 2012 International Symposium on*. 2012 IEEE. Pp.1-4.
- Charles D Mallah, James Orwell. 2013. Probabilistic Classification from a K-Nearest-Neighbour Classifier. *Computational Research* 1(1): 1-9, 2013. 2013 Horizon Research. Pp.1-9.
- Choochart Haruechaiyasak. 2008. A Tutorial on Naïve Bayes Classification.
- Christopher J.C. Gurges. 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*. Kluwer Academic Publishers. Pp 121–167.
- Cover, Thomas M., Joy A. Thomas. 2012. *Elements of information theory*. John Wiley & Sons, 2012.
- D. Koller, N. Friedman. 2009. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press.
- Dimitris G. Tsolakidis, Dimitrios I. Kosmopoulos, George Papadourakis. 2014. Plant Leaf Recognition Using Zernike Moments and Histogram of Oriented Gradients. A. Likas, K. Blekas, and D. Kalles (Eds.): *SETN 2014, LNAI 8445*. Springer International Publishing Switzerland 2014. Pp. 406–417.
- Fan, Z.-G., Wang, K.-A., Lu, B.-L. 2004. Feature Selection for Fast Image Classification with Support Vector Machines. In *Neural Information Processing*. Springer Berlin/Heidelberg. Pp.1026-1031.

- Gaurav Agarwal, Peter Belhumeur, Steven Feiner, David Jacobs, W. John Kress, Ravi Ramamoorthi, Norman A. Bourg, Nandan Dixit, Haibin Ling, Dhruv Mahajan, Rusty Russell, Sameer Shirdhonkar, Kalyan Sunkavalli, Sean White. 2006. First Steps toward an Electronic Field Guide for Plants. *Taxon*, 55(3), 597-610.
- H. Ling, D.W. Jacobs. 2009. Shape Matching for Foliage Database Retrieval. *Semantic Mining Technologies for Multimedia Databases*, Chapter 4.
- Hall, M. A., Smith, L. A. 1997. Feature subset selection: a correlation based filter approach. In 1997 International Conference on Neural Information Processing and Intelligent Information Systems. Berlin: Springer. Pp.855-858.
- Hang Zhang, Shangsong Liang, Paul Yanne. 2012. Plant Species Classification Using Leaf Shape and Texture. 2012 International Conference on Industrial Control and Electronics Engineering. 2012 IEEE. Pp.2025-2028.
- Huang Lin, He Peng. 2008. Machine Recognition for Broad-Leaved Trees Based on Synthetic Features of Leaves Using Probabilistic Neural Network. 2008 International Conference on Computer Science and Software Engineering. 2008 IEEE. Pp.871-877.
- Isabelle Guyon , Jason Weston , Stephen Barnhill , Vladimir Vapnik. 2002. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, 46(1-3), Pp.389-422.
- James S. Cope, Paolo Remagnino, Sarah Barman, Paul Wilkin. 2010. Plant Texture Classification Using Gabor Co-occurrences. G. Bebis et al. (Eds.): ISVC 2010, Part II, LNCS 6454. Springer-Verlag Berlin Heidelberg 2010. Pp. 669–677.

- Javed Hossain, M. Ashraful Amin. 2010. Leaf Shape Identification Based Plant Biometrics. Proceedings of 13th International Conference on Computer and Information Technology (ICCIT 2010) 23-25 December, 2010, Dhaka, Bangladesh. 2010 IEEE. Pp.458-463.
- Jixiang Du, Deshuang Huang, Xiaofeng Wang, Xiao Gu. 2005. Shape Recognition Based on Radial Basis Probabilistic Neural Network and Application to Plant Species Identification. J. Wang, X. Liao, and Z. Yi (Eds.): ISNN 2005, LNCS 3497. Springer-Verlag Berlin Heidelberg 2005. Pp.281–285.
- João B. Florindo, André R. Backes, Odemir M. Bruno. 2010. Leaves Shape Classification using Curvature and Fractal Dimension. A. Elmoataz *et al.* (Eds.): ICISP 2010, LNCS 6134. Springer-Verlag Berlin Heidelberg 2010. Pp.456–462.
- Kohavi, R., John, G. 1997. Wrappers for feature subset selection. *Artificial Intelligence*. 97(1), Pp 273-324.
- Lei Yu, Huan Liu. 2003. Feature selection for high-dimensional data: A fast correlation-based filter solution. In Proceedings of the Twentieth International Conference on Machine Learning (ICML 2003). Vol. 3, pp. 856-863
- Lei Zhang, Jun Kong, Xiaoyun Zeng, Jiayue Ren. 2008. Plant Species Identification Based on Neural Network. Fourth International Conference on Natural Computation. 2008 IEEE. Pp.90-94.
- Mingwei Wang, Jingtao Zhou. 2014. A Bayesian Network-based Classifier for Machining Error Prediction. ASME International Conference on Advanced Intelligent Mechatronics (AIM). 2014 IEEE. Pp.841-844.

- Mladenić, D., Brank, J., Grobelnik, M., Milic-Frayling, N. 2004. Feature selection using linear classifier weights: interaction with classification models. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval. ACM. Pp.234-241.
- Mohamed Elhadi Rahmani, Abdelmalek Amine, Mohamed Reda Hamou. 2015. Plant Leaves Classification, The First International Conference on Big Data, Small Data, Linked Data and Open Data. Pp 75-80.
- Nir Friedman, Dan Geiger, Moises Goldszmidt. 1997. Bayesian Network Classifiers. Machine Learning. Kluwer Academic Publishers. Pp 131-163.
- Remeo R. Bouckaert. 2008. Bayesian Network Classifiers in WEKA for Version 3-5-7. University of Waikato, New Zealand.
- Sameer Singh, John Haddon, Markos Markou. 2001. Nearest-neighbour classifiers in natural scene analysis. Elsevier Science Ltd. Pattern Recognition 34(8). Pp 1601–1612.
- Shang, C., Barnes, D. 2010. Combining support vector machines and information gain ranking for classification of mars mcmurdo panorama images. In Image Processing (ICIP), 2010 17th IEEE International Conference on IEEE. Pp.1061-1064.
- Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang Qiao-Liang Xiang. 2007. A leaf recognition algorithm for plant classification using probabilistic neural network. Signal Processing and Information Technology, 2007 IEEE International Symposium on. 2007 IEEE. Pp.11-16.
- Svetlana S. Aksenova. 2004. WEKA Explorer Tutorial for WEKA Version 3.4.3. School of Engineering and Computer Science Department of Computer Science, California State University, Sacramento, California, 95819.

- Thibaut Beghin, James S. Cope, Paolo Remagnino, Sarah Barman. 2010. Shape and texture based plant leaf classification. J. Blanc-Talon et al. (Eds.): ACIVS 2010, Part II, LNCS 6475. International Conference on Advanced Concepts for Intelligent Vision Systems (ACVIS). Springer-Verlag Berlin Heidelberg 2010. Pp.345-353.
- Weidong Cao, Xiangnong Fang. 2010. An Improved Method for Bayesian Network Structure Learning. 2010 Sixth International Conference on Natural Computation (ICNC 2010). 2010 IEEE. Pp.3133-3137.
- Xiaohui Lin, Ping Ma, Xiaolan Li, Jiaxue Jiang, Niyi Xiao, Fufang Yang. 2012. A Learning Method of Bayesian Network Structure. 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012). 2012 IEEE. Pp.666-670.
- Xin Wang, Peng Guo. 2012. A Novel Binary Adaptive Differential Evolution Algorithm for Bayesian Network Learning. 2012 8th International Conference on Natural Computation (ICNC 2012). 2012 IEEE. Pp.608-612.
- Zalikha Zulkifli, Puteh Saad, Itaza Afiani Mohtar. 2011. Plant Leaf Identification Using Moment Invariants & General Regression Neural Network. 2011 11th International Conference on Hybrid Intelligent Systems (HIS). 2011 IEEE. Pp.430-435.