

**Computer-Aided Plant Identification Through Leaf Recognition
Using Enhanced Image Processing And
Machine Learning Algorithms**

By

N. Valliammal

Supervisor

Dr. S.N.Geethalakshmi

A Thesis Submitted To
Avinashilingam Institute for Home science and Higher Education for Women,
Coimbatore – 641 043

In Partial Fulfillment of the Requirement for the Degree of

Doctor of Philosophy in Computer Science

October, 2013


CERTIFICATE

This is to certify that the thesis entitled "**Computer-Aided Plant Identification Through Leaf Recognition Using Enhanced Image Processing And Machine Learning Algorithms**" submitted to the Avinashilingam University for Women, Coimbatore, for the award of the degree of **Doctor of Philosophy in Computer Science**, is a record of original research work done by **N.Valliammal** during the period of her study in the **Department of Computer Science**, Avinashilingam University for Women, Coimbatore, under my supervision and guidance and the thesis has not formed the basis for the award of any Degree/Diploma/Associateship/ Fellowship or similar title to any candidate of any University.


Signature of the ^{31/10/13}
Head of the Department

Dr. G. PADMAVATHI
M.Sc. M.Phil., Ph.D.
Professor and Head
Department of Computer Science
Avinashilingam Institute for Home Science
and Higher Education for Women
Coimbatore - 641 043


Signature of the Supervisor ^{21/10/2013}



Signature of the Dean

Dr. A. PARVATHI M.Sc., D.Ed., M.Phil., Ph.D.
Dean, Faculty of Science
Principal, School of
Department of Home Science
Avinashilingam Institute for Home Science
and Higher Education for Women
Coimbatore - 641 043

DECLARATION

I hereby declare that the matter embodied in the thesis entitled **“Computer-Aided Plant Identification Through Leaf Recognition Using Enhanced Image Processing And Machine Learning Algorithms”** is the result of investigations carried out by me in the Department of Computer Science, Avinashilingam University for Women, Coimbatore, under the supervision and guidance of **Dr.S.N.Geethalakshmi**, Professor, Department of Computer Science, Avinashilingam University for Women, and that it has not been submitted for the award of any Degree/Diploma/Associateship/Fellowship or similar title of any University or Institute.


Signature of the Guide


Signature of the candidate

ACKNOWLEDGEMENT

*The investigator extends her sincere thanks to Late **Thiru.T.K.SHANMUGANANDAM**, B.A., B.L., Former Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for providing infrastructural facilities and support for the conduct of the study.*

*The investigator records her sincere thanks to **Thiru Dr. T.S.K. MEENAKSHI SUNDARAM**, M.A., M.Phil., Ph.D., Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing the infrastructural facilities for the conduct of the study. His guidance had helped in completing the research work successfully.*

*The Researcher expresses her gratitude and sincere thanks to Hony.Col. **Dr (Tmt). SAROJA PRABHAKARAN**,M.A.,Dip.Ed.,(Madras),Ph.D(Mother Teresa), Former Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women. The researcher places on record her thanks for the support provided and opportunity given to undertake the research work.*

*The Researcher extends her immense gratitude to **Dr.(Mrs) SHEELA RAMACHANDRAN**, M.Sc., P.G. Dip., Ph.D. Vice Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her constant encouragement and support for the completion of the study.*

*The investigator expresses her special thanks to **Dr.(Mrs)GOWRI RAMAKRISHNAN**, M.Sc., M.Phil., Ph.D. Registrar, Avinashilingam Institute for Home Science and Higher Education for Women, for extending her precious help.*

*The investigator expresses her sincere gratitude to **DR.P.SANTHANA KRISHNAN**, Director (Research and Consultancy) for his affiable help and constant encouragement in carrying out the research work.*

*The investigator records her gratefulness to **Dr. (Mrs) R.PARVATHAM**, M.Sc., Dip. Ed.,M.Phil., Ph.D., Former Dean, Faculty of Science, Professor and Head, Department of Bio-Chemistry, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore for her timely help and encouragement in carrying out the research work.*

*The researcher places her thanks to **Dr.(Mrs.) A.PARVATHI**, M.Sc.,Dip.Ed, M.Phil., Ph.D.,Dean,Faculty of Science, Professor and Head, Department of Mathematics, Avinashilingam Institute for Home Science and Higher Education for Women for her motivation and encouragement during the course of the study.*

Heartfelt thanks are expressed to **Dr.(Mrs) G.PADMAVATHI**, M.Sc., M.Phil., Ph.D. Professor and Head, Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her constant encouragement, concern and willing help rendered during the course of the study.

With glowing sense of gratitude and honesty the researcher places her sincere and deep sense of gratitude and indebtedness to her most honored supervisor a principle centered person **Dr. (Mrs). S.N.GEETHALAKSHMI**, M.C.A.,M.Phil.,Ph.D, Professor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for her dynamic guidance, scholarly advice, constructive criticism, constant encouragement, without which the conduct of the study would have been impossible. As a result, the journey of research life became smooth and rewarding for the investigator.

The investigator expresses her sincere gratitude to her esteemed supervisor's sister **Dr.(Mrs).S.N.PADMADEVI**, M.Sc,M.Phil,B.ed,Ph.D, Associate Professor, Department of Botany PSGR Krishnammal College for Women who started the road for this research through her initial and innovative ideas.

The investigator conveys her special acknowledgement to **TAMILNADU AGRICULTURAL UNIVERSITY**, Coimbatore for their indispensable help and granting her permission for collecting the Real leaf images for her research.

The investigator place her gratitude to **DR.R.GOPALAN**, Ph.D, Scientist(Rtd), Botanical Survey of India ,Professor and Head, Department of Botany, Karpagam University, Coimbatore for his advice and path guidance.

She extends her gratitude to **Dr.R.KARTHIYAYINI**, M.Sc, M.Phil, Ph.D, Assistant Professor, Department of Botany, Avinashilingam Institute for Home Science and Higher Education for Women , Coimbatore for her dedicated and timely help.

The investigator expresses her sincere thanks to her beloved **TEACHERS**, and **FRIENDS** in the Department of Computer Science and Computer Centre, **FACULTY MEMBERS** of other Departments of Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for their help, warmth, responsiveness, valuable suggestions, advice and support in carrying out the study.

The investigator expresses her deep appreciation to her **FRIENDS** **Dr.Mrs.D.MATHIVADHANI, Mrs.G.GEETHA, Dr.Mrs.J.UMAMAHESWARI, Dr.Mrs.S.MALATHI** and other friends for their constant help, warmth, responsiveness, constructive suggestions and consistent encouragement given throughout the study.

My Parents deserve special mention for their inseparable support and prayers. I can never pay the reward to my loving father **Thiru N.NAGARAJAN** who put the fundament in my learning career, showing the joy of intellectual pursuit ever since I was a child, and to my caring mother **Mrs.N.VAIKUNDAPATHY** who is so affectionate and has made my educational foundation through persistent encouragement. I can never pay for their debt of love and affection. She dedicates this thesis to her parents.

The investigator's gratitude goes to her Father-in-law Late **Thiru R.VEERABAHU** and Mother-in-law **Tmt V.ANNALAKSHMI** for their cooperation and valuable assistance for the successful completion of the thesis. The investigator conveys special acknowledgement to all of them, who sincerely raised her, with care and gentle love and remembers her constant support when encountered difficulties.

The investigator thank her husband **Mr.V.PON NARAYAN** for his endless love and encouragement without whose help, co-operation and encouragement, this research would not be successful. He was one of the best counsellors throughout this entire journey and also for the many discussions for providing advises that helped her in staying at the right track in completing her dissertation successfully.

The author is indebted to her most adorable little children **P.RAKSHAA** and **P.SAI KARTHICK** for their unconditional sacrifices, unending love and unconditional support throughout the tenure of the research work.

The chain of gratitude would be definitely incomplete if the investigator would forget to thank her lovable brother **N.NARAYANAN** and family for their inspiring thoughts and guidance in crossing successfully all the tests in the past few years of the doctoral degree.

Above all, the investigator raises her humble heart in adoration to **GOD ALMIGHTY**, who has designed and executed the research.

ABSTRACT

Computer aided identification of plants is an area of research that has gained more attention in recent years and is proving to be a very important tool in many areas including agriculture, forestry and pharmacological science. In addition, with the deterioration of environments, more and more rare plant species are at the margin of extinction. Many of the rare plants has become extinct and hence the investigation of plant recognition can contribute to environmental protection.

A general process of a Computer Aided Plant Classification through Leaf Recognition (CAP-LR) contains four steps, namely, leaf image enhancement, leaf segmentation, feature extraction and classification. The preprocessing algorithm proposed in CAP-LR enhances a leaf image by using an approach that simultaneously removes noise, adjusts contrast and enhances boundaries. To extract the leaf image from its background, the CAP-LR system uses an enhanced wavelet based segmentation approach that combine clustering with texture based color features.

From the extracted leaf image, a total of 28 features were extracted which were grouped into five categories, namely, geometric features, color features, texture features, fractal features and leaf specialized features. To enhance the process of leaf recognition for plant identification a fusion of feature categories was proposed. The proposed fusion method combined Genetic Algorithm (GA) and Kernel Principal Component Analysis (KPCA) with shared and merger operations is carried out in the third step. The shared operator selects only those features which were common to GA and KPCA algorithms while the merger operation combined all the features from GA and KPCA algorithm.

The single and fused feature sets are then used by classifier to recognize the leaves and identify the plants. For this purpose, a 2-level classification model was used, where the first level classifier was used to produce an optimized or refined training set, which was used to train the second level classifier.

Two leaf image datasets, namely, standard and real, were used during experiments that evaluated the performance of the proposed algorithms. The experimental results showed that the 2-level classification algorithm improved the

efficiency of recognition and identification in terms of accuracy and speed. The various results showed that the model WNN for the first classifier and SVM for the second classifier that used GA \cap KPCA algorithm for feature extraction and selection produced the best algorithm with the usage of leaf + fractal feature set.

This study proposed techniques to enhance the operation of leaf recognition for plant identification. The positive results obtained from the various experiments prove that the proposed models are effective in discriminating the various leaves and identifying the correct plant corresponding to an input leaf image. Hence, the botanists can safely use it to increase their efforts in saving precious plant leaf and thus improve the quality of human and earth life.

CONTENTS

Chapter	Title	Page No.
	LIST OF TABLES	
	LIST OF FIGURES	
	ACRONYMS	
	ABSTRACT	
1	INTRODUCTION	1-20
1.1	Overview of Plant Kingdom	3
1.2	The Leaf – Important Organ of Plant	5
1.3	Need for Automated Plant Identification	7
1.4	Challenges of Automatic Plant Recognition	8
1.5	General Plant Recognition System	10
1.5.1	Acquisition of Leaf Image	11
1.5.2	Preprocessing	11
1.5.3	Feature Extraction	12
1.5.4	Matching Algorithm	13
1.6	Motivation	16
1.7	Problem Statement and Objectives	17
1.8	Chapter Formulation	18
1.9	Chapter Summary	20
2	REVIEW OF LITERATURE	21-47
2.1	Image Enhancement	21
2.1.1	Noise Removal	21
2.1.2	Contrast and Edge Enhancement	23
2.2	Segmentation	25
2.2.1	General Segmentation Methods	26
2.2.2	Plant and Leaf Related Methods	31
2.3	Feature Extraction	34
2.4	Plant Identification and Leaf Recognition	37
2.4.1	Shape and Contour Feature Based Classification	38
2.4.2	Region Based Classification	44
2.5	Chapter Summary	46

3	METHODOLOGY	48-61
3.1	Research Design	48
3.1.1	Phase I : Preprocessing	49
3.1.2	Phase II: Feature Extraction and Selection	52
3.1.3	Phase III: Classification	54
3.2	Performance Evaluation	54
3.2.1	Phase I : Denoising	55
3.2.2	Phase I : Segmentation	57
3.2.3	Phase II : Feature Extraction and Selection	58
3.2.4	Phase III : Plant Identification through Leaf Recognition	58
3.3.	Datasets	58
3.4	Contributions	60
3.5	Chapter Summary	61
4	DESIGN OF LEAF IMAGE ENHANCEMENT METHOD	62-86
4.1	Basic Concepts	63
4.1.1	Contrast Enhancement	63
4.1.2	Noise Removal	64
4.1.3	Edge Enhancement	65
4.2	Motivation	67
4.3	The WEC Method	67
4.3.1	Step 1: Color Space Transformation	68
4.3.2	Step 2: Wavelet Transformation	70
4.3.3	Step 3: Automatic Contrast Adjustment	70
4.3.4	Step 4: Edge Enhancement	73
4.3.5	Step 5: Noise Removal	79
4.3.6	Step 6: Inverse DWT	80
4.4	Experimental Results	80
4.4.1	Peak Signal to Noise Ratio	81
4.4.2	Figure of Merit(FoM)	81
4.4.3	Mean Structural Similarity Index(MSSI)	83
4.4.4	Speed	83
4.4.5	Visual Results	86
4.5	Chapter Summary	86

5	DESIGN OF LEAF IMAGE SEGMENTATION METHOD	87-112
5.1	Overview of Image Segmentation	88
5.2	Importance of Segmentation in CAP-LR	90
5.3	Proposed Segmentation Method	93
5.3.1	Step 1: Extraction of Texture Features	94
5.3.2	Step 2: Extraction of Color Features	99
5.3.3	Step 3: Feature Space Generation	102
5.3.4	Step 4: Mean Shift Filtering Algorithm	102
5.3.5	Step 5: K Means Algorithm	106
5.4	Experimental Results	108
5.5	Chapter Summary	112
6	FEATURE EXTRACTION AND SELECTION	113-142
6.1	Geometric Features	115
6.2	Color Features	116
6.3	Texture Features	116
6.4	Fractal Features	118
6.5	Leaf Specialized Features	119
6.5.1	Vein Feature	119
6.5.2	Other Leaf Features	126
6.5.3	Morphological Leaf Features	127
6.6	Ripple Feature	128
6.7	Proposed Feature Fusion and Selection Algorithm	129
6.7.1	Feature Fusion and Selection using Genetic Algorithm	131
6.7.2	The Kernel Principal Component Analysis	137
6.8	Experimental Results of Feature Fusion and Selection Algorithms	138
6.9	Chapter Summary	142
7	DESIGN OF LEAF IMAGE CLASSIFICATION MODELS FOR PLANT IDENTIFICATION	143-165
7.1	Machine Learning	144
7.1.1	Types of Generation and Representation	145

7.1.2	Types of Iteration	145
7.1.3	Types of Evaluation	146
7.2	Image Classification	146
7.3	Proposed CL-CL Model	149
7.3.1	Feature Extraction	149
7.3.2	Partitioning Method	149
7.3.3	Design of CL-CL Models	150
7.4	Chapter Summary	165
8	RESULTS AND DISCUSSION	166-195
8.1	BPNN-Based CL-CL Models	168
8.1.1	Experimental Results of BPNN-Based CL-CL Models using Standard Dataset	168
8.1.2	Experimental Results of BPNN-Based CL-CL Models using Real Dataset	173
8.2	SVM-Based CL-CL Models	178
8.2.1	Experimental Results of SVM-Based CL-CL Models using Standard Dataset	178
8.2.2	Experimental Results of SVM-Based CL-CL Models using Real Dataset	182
8.3	WNN-Based CL-CL Models	186
8.3.1	Experimental Results of WNN-Based CL-CL Models using Standard Dataset	186
8.3.2	Experimental Results of WNN-Based CL-CL Models using Real Dataset	189
8.4	Chapter Summary	195
9	SUMMARY AND CONCLUSION	196-199
	FUTURE RESEARCH DIRECTIONS	199
	BIBLIOGRAPHY	200
	LIST OF PUBLICATIONS	218
	Appendix A	220
	Appendix B	225
	Appendix C	233
	Appendix D	241

LIST OF FIGURES

Figure No	Title	Page No
1.1	Parts of Plant body	4
1.2	Features of Leaf	6
1.3	Example of Leaf Variation	9
1.4	General Architecture of Plant Identification System	10
1.5	Features - Distinction between Good and Poor Features	12
1.6	Classifier and Decision Boundaries	14
1.7	Basic Classification model	15
2.1	Nine maple Leaf Shapes of the Study	43
2.2	Four Classes of Chrysanthemum Leaf Images	43
2.3	An example of Self Interaction	44
3.1	Research Methodology	50
3.2	Standard Dataset Test Images	59
3.3	Real Dataset Test Images	59
4.1	Original and Gaussian Noise Degraded Images	65
4.2	Steps in WEC Method	69
4.3	Pseudo Code for AHE	72
4.4	CLAHE Algorithm	74
4.5	Proposed Edge and Contrast Enhancement Algorithm	75
4.6	Rule Based Edge Classification Algorithm	78
4.7	PSNR (dB)	82
4.8	Figure of Merit (FoM)	82
4.9	Mean Structural Similarity Index(MSSI)	83
4.10	Visual Results of Noise Removal(Standard Dataset)	84
4.11	Visual Results of Noise Removal(Real Dataset)	85
5.1	Steps in Image Analysis	89
5.2	Examples of Leaf with Color Variation	92
5.3	Proposed WCF Method	94
5.4	Discrete Wavelet Frame decomposition of the ith level	97
5.5	Discrete Wavelet Frame decomposition	98
5.6	Traditional K Means Algorithm	107
5.7	Speed of Segmentation(Seconds)	109
5.8	Segmentation Results of Standard Dataset	110
5.9	Segmentation Results of Real Dataset	111
6.1	Veins of a Leaf	119

6.2	Appearance of Random Points and Broken Veins	122
6.3	Linking Broken Veins	123
6.4	Elimination of Random or Isolated Points	123
6.5	Leaf Vein Extraction Process	124
6.6	Relationship between Physiological Length and Physiological Width	127
6.7	Proposed Fusion and Selection Algorithm	130
6.8	Feature Selection and Fusion Algorithm	131
6.9	Encoding of Feature Subset in GA – A L-Dimensional Binary Vector	134
6.10	Roulette Wheel Selection	136
6.11	Effect of Leaf Features on Classification	139
6.12	Single Feature and Fusion Feature Sets	140
6.13	Effect of Feature Fusion on Number of Features selected	140
6.14	Effect of Proposed Feature Fusion and Selection Algorithms	141
7.1	Paradigms in Image Classification	147
7.2	Hold-Out Method	150
7.3	Dataset Split into Training and Testing set	150
7.4	Back Propagation Neural Network	152
7.5	BPNN Architecture	155
7.6	Support Vector Machine Hyperplane	157
7.7	Architecture of WNN	160
7.8	Classification Process	163
8.1	Efficiency Gain of WNN-BPNN Model using Standard Dataset over Recognition Rate	170
8.2	Speed of BPNN-Based CL-CL Models (Single Feature Sets)	173
8.3	Speed of BPNN-Based CL-CL Models (Fused Feature Sets)	173
8.4	Efficiency Gain of WNN-BPNN Model using Real Dataset over Recognition Rate	175
8.5	Error Rate of BPNN-based CL-CL Models(Single Features)	176
8.6	Error Rate of BPNN-based CL-CL Models(Fused Features)	176
8.7	Recognition Rate of SVM-Based CL-CL Models(Single Feature Sets-Standard Dataset)	178
8.8	Recognition Rate of SVM-Based CL-CL Models(Fused Feature Sets- Standard Dataset)	179
8.9	Speed of SVM-Based CL-CL Models (Single Feature Sets-Standard Dataset)	181
8.10	Speed of SVM-Based CL-CL Models(Fused Feature Sets-Standard Dataset)	181

8.11	Speed of SVM-Based CL-CL Models (Single Feature Sets-Real Dataset)	185
8.12	Speed of SVM-Based CL-CL Models (Fused Feature Sets-Real Dataset)	185
8.13	Recognition Rate of SVM-Based CL-CL Models(Single Feature Sets-Standard Dataset)	186
8.14	Recognition Rate of SVM-Based CL-CL Models(Fused Feature Sets-Standard Dataset)	187
8.15	Speed of WNN-Based CL-CL Models(Single Feature Sets-Real Dataset)	192
8.16	Speed of WNN-Based CL-CL Models(Fused Feature Sets-Real Dataset)	193

LIST OF TABLES

Table No.	Title	Page No.
3.1	Features Extracted from Leaf Image	53
6.1	Features Extracted	114
6.2	Color Features	116
6.3	Morphological Operations	121
7.1	Proposed CL-CL Two Level Models	164
8.1	Techniques Used in CAP-LR	167
8.2	Recognition Rate (%) of BPNN-Based CL-CL Models (Standard)	169
8.3	Error Rate(%) of BPNN-Based CL-CL Models (STANDARD)	171
8.4	Recognition Rate of BPNN-Based CL-CL Models(Real)	174
8.5	Efficiency Gain(%) of WNN-BPNN Model for Leaf and Leaf+Fractal Feature Sets (Real Dataset)	177
8.6	Speed of BPNN-Based CL-CL Models (Real)	177
8.7	Efficiency Gain(%) of Recognition Rate for Standard Leaf Dataset	180
8.8	Error Rate of SVM-Based CL-CL Models(Standard)	180
8.9	Recognition Rate of SVM-Based CL-CL Models(Real)	183
8.10	Efficiency Gain(%) of Recognition Rate for Standard Leaf Dataset	183
8.11	Error Rate of SVM-Based CL-CL Models(Real)	184
8.12	Error Rate of WNN-Based CL-CL Models(Standard)	188
8.13	Speed of WNN-Based CL-CL Models(Standard)	189
8.14	Recognition Rate of WNN-Based CL-CL Models(Real)	190
8.15	Error Rate of WNN-Based CL-CL Models(Real)	191
8.16	Details of Winning Models	193
8.17	Performance Ranking Order of the Feature Sets	194
8.18	Performance Ranking Order Of the Proposed CL-CL Models	194

ACRONYMS

Abbreviation	Description
AHE	Adaptive Histogram Equalization
ASM	Angular Second Moment
BPNN	Back Propagation Neural Network
CAP-LR	Computer Aided Plant Identification Through Leaf Recognition
CCD	Centroid Contour Distance
CL	Clip Limit
CLAHE	Contrast Limited Adaptive Histogram Equalization
CL-CL	Classification-Classification
CLFS	Color+Leaf Feature Set
DWF	Discrete Wavelet Frame
DWT	Discrete Wavelet Transformation
FLFS	Fractals+Leaf Feature Set
FOM	Figure Of Merit
GA	Genetic Algorithm
GLCM	Gray Level Co-Occurrence Matrix
GLCM	Gray Level Co-Occurance Matrix
GLFS	Geometric+Leaf Feature Set
GRNN	General Regression Neural Network
HH	High High
HL	High Low
HVS	Human Visual System
KPCA	Kernel Principal Component Analysis

Abbreviation	Description
LH	Low High
LL	Low Low
LMI	Legendre Moment Invariant
MAP	Maximum Posteriori Position
MSSI	Mean Structural Similarity Index
NIR	Near Infrared
NN	Nearest Neighbour
PCA	Principal Component Analysis
PSNR	Peak Signal To Noise Ratio
RGB	Red Green Blue
ROI	Region Of Interest
SVM	Support Vector Machine
TLFS	Texture+Leaf Feature Set
TMI	Tchebichef Moment Invariant
WCF	Enhanced Wavelet Based Segmentation Using Clustering And Texture Based Color Features
WEC	Enhanced Wavelet Based Denoising With Inbuilt Edge Enhancement And Automatic Contrast Adjustment
WNN	Wavelet Neural Network
