Automatic Skin Cancer Detection system

By
Azadeh Noori Hoshyar

Submitted in partial fulfilment of the requirements for the
Master of Engineering

Faculty of Engineering and Information Technology
UNIVERSITY OF TECHNOLOGY, SYDNEY

November, 2014

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

Acknowledgment

Over the last two years I have had the privilege of working with a variety of people who have made my time at University of technology, Sydney an enjoyable and intellectually stimulating experience.

I would like to thank all the people who have helped me along the way and contributed to this dissertation. I am especially grateful to my Supervisor Associate Professor **Dr. Adel Al-Jumaily**. In working with Adel, I have learned how to pursue research problems with intellectual rigor and how to critically evaluate my work.

Finally, I would like to acknowledge the endless love of my husband and family; who have been constant source of support also who have provided guidance, love and encouragement throughout my life.

Table of Contents

CHAPTER I 1

INTRODUCTION 1

1.1. Motivation 1

1.2. Research Objectives 2

1.3. Thesis Structure 3

1.4. Research Contributions 3

1.5. Research Scope 4

1.6. Publications Resulting from the Thesis 4

CHAPTER 2 6

AN OVERVIEW ON HUMAN SKIN, SKIN CANCER AND DIAGNOSIS TECHNIQUES 6

2.1 Introduction 6

2.2 Human Skin 6

2.3 Cancer 7

2.4 Skin Cancer 7

2.4.1 Malignant melanoma 8

2.4.2 Non-melanoma skin cancer 9

2.4.2.1 Basal Cell Carcinoma 9

2.4.2.2 Squamous Cell Carcinoma 10

2.5 Skin Lesion Imaging Methods 11

2.6 Diagnosis System of Skin cancer 12

2.7 Diagnosis Techniques 12

2.8 Brief Summary on Pathology based Research 13

2.8.1 Histo-pathological Images 14

2.9 What are the pathologic features of a melanoma? 16

2.10 Summary 16

CHAPTER 3 18

LITERATURE REVIEW 18

3.1 Introduction 18

3.2 Computer-aided diagnosis system 18

3.2.1 Image acquisition/ methods for screening skin lesions 19

3.2.2 Pre-Processing 19

3.2.2.1 Image Enhancement 19

3.2.2.2 Image Restoration 22
3.2.2.2.1 Restoration from noise 22
3.2.2.2.2 Restoration from blur 24
3.2.2.3 Removing Thick Hairs 24
3.2.3 Segmentation 25
3.2.4 Feature Extraction 27
3.2.5 Feature Selection 29
3.2.6 Classification 31
3.2.7 Performance Indicators 32
3.3 Summary 34
CHAPTER 4 35
METHODOLOGY REVIEW and PROPOSED SYSTEM 35
4.1 Introduction 35
4.2 Pre-Processing 36
4.3 Segmentation 37
4.3.1 The K-means Algorithm 38
4.3.2 The Level Set Framework 39
4.3.3 The Proposed Active Contour Tracking Method 40
4.3.3.1 The Proposed Algorithm 40
4.4 Feature Extraction 42
4.4.1 Texture properties 42
4.4.2 Shape properties 43
4.5 Feature Selection 44
4.5.1 Sequential Feature selection (SFS) 45
4.5.2 Particle Swarm Optimization 45
4.5.3 Support Vector Machine 46
4.5.4 The Proposed Algorithm 49
4.6 Classification 53
4.6.1 Self-advising SVM 53
4.7 Experimental Evaluation and Comparison of Results 54
4.8 Summary 55
CHAPTER 5 56
DISCUSSION OF RESULTS 56
5.1 Introduction 56
5.2 Pre-Processing 56

5.3 Segmentation 65
5.4 Feature Extraction and Selection 76
5.5 Classification 76
5.6 Results of Feature Selection and Classification 77
5.6.1 Experimental Result by Smart IPSO-SVM Feature Selection Algorithm Followed By SA-SVM Classifier 77
5.6.2 Comparison of the Proposed Algorithm with Different Algorithms 78
5.7 Summary 86
CHAPTER 6 88
SUMMARY AND FUTURE RESEARCH 88
6.1 Introduction 88
6.2 Overall Review on This Thesis 88
6.3 Directions for Future Work 91
6.4 Limitation of the Study 92
REFERENCES 93

List of Figures

Chapter 1

Figure 1.1. Schematic diagram of automatic diagnosis of skin cancer……..…… 2

Chapter 2

Figure 2.1. A squamous cell, basal cell, and melanocyte and epidermis and dermis layers in Human skin ……..... 7

Figure 2.2. Malignant melanoma……... 8

Figure 2.3. Nodulocystic basal cell carcinoma……..… 9

Figure 2.4. Superficial basal cell carcinoma…….. 10

Figure 2.5. Sclerosing (morphoeic) basal cell carcinoma……..… 10

Figure 2.6. Squamous Cell Carcinoma……...… 11

Figure 2.7. Manual Segmentation by Trained Pathologists…… 14

Figure 2.8. Radial and Vertical Growth Phase Melanoma ……..... 15

Chapter 3

Figure 3.1. Tristimulus Curves……... 20

Figure 3.2. Linear Contrast Enhancement Methods…….. 21

Figure 3.3. Non-Linear contrast Enhancement Methods…….….. 21

Figure 3.4. a) Adaptive Histogram Equalization b) Histogram Equalization c) Unsharp Masking…….….. 22

Figure 3.5. a) Image without noise b) Gussian noise c) Poison noise d)Salt and Pepper noise e) Speckle noise……..… 22

Figure 3.6.Common Feature selection process ……..… 30

Chapter 4

Figure 4.1.Skin Cancer detection system……...…. 36

Figure 4.2.Image enhancement……..…. 37

Figure 4.3.Flowchart of proposed SFC- IPSO-SVM……. 52

Chapter 5

Figure 5.1. a) Original image b) Simulated Speckle noise and de-noising by c) Gaussian Filter d) Median Filter e) Mean Filter f) Adaptive Median Filter g) Adaptive Wiener….. 57

Figure 5.2. a) Original image b) Simulated Gaussian noise and de-noising by c) Gaussian Filter d) Median Filter e) Mean Filter f) Adaptive Median Filter g) Adaptive Wiener…… 57

Figure 5.3. a) Original image b) Simulated Salt & Pepper noise and de-noising by c) Gaussian Filter d) Median Filter e) Mean Filter f) Adaptive Median Filter g) Adaptive Wiener……. 57

Figure 5.4. a) Original image b) Simulated Poison noise and de-noising by c) Gaussian Filter d) Median Filter e) Mean Filter f) Adaptive Median Filter g) Adaptive Wiener…… 58

Figure 5.5.Comparison of AHE,HE and UM…… 60

Figure 5.6. Modified Hausdorff Distance…….. 62

Figure 5.7. Euclidean Distance……... 63

Figure 5.8. Correlation……..... 64

Figure 5.9. a) Original Image b) Initial k-means segmentation c) Converting the segmented image to binary d) Segmented result of proposed algorithm after initializing the level set segmentation technique by k-means algorithm……... 65

Figure 5.10. Comparison of ground truth image with the segmented image of proposed algorithm a) Ground truth image b) Segmented result c) Pixels in segmented lesion as well as ground truth (Subscription of „a‟, „b‟ images)…… 66

Figure 5.11. Fuzzy c-means thresholding in two cut-off position a) Original image b) Otsu thresholding c) FCM (sw=0) d) FCM (sw=1)……..... 66

Figure 5.12.Comparison of a) proposed segmentation method b) traditional level set method c) Fuzzy c-means thresholding (sw=0) d) fuzzy c-means thresholding (sw=1)……..… 67

Figure 5.13. Difference of Border error, Similarity, Hammoude distance and Rms error of KLS and TLS, KLS and FCM (sw=0), and KLS and FCM (sw=1)……. 75

List of Tables

Chapter 4

Table 4.1.GLCM features……...… 43

Table 4.2. Volumetric zone length type texture…….. 43

Table 4.3.Extracted Shape properties……..……...…. 44

Chapter 5

Table 5.1.Comparison of PSNR for a skin cancer image after simulating different noises and de-noising by filters in different densities…… 59

Table 5.2. The most effective Filters on different noises with densities between 10% - 80%..... 60

Table 5.3 .Result table of Modified Hausdorff Distance……..…… 61

Table 5.4.Result table of Euclidean Distance……. 62

Table 5.5. Result table of Correlation……. 64

Table 5.6 . Comparison of Proposed method, TLS, FCT (w=0), FCT (w=1) by the Border Error, Similarity, Hammoude Distance and Rms Error …….…. 67

Table 5.7. Anova test on Bordererror, Similarity, Hammoude distance and Rmserror metrics between the “proposed method and TLS”, “KLS and FCT (w=0)” , “KLS and FCT(w=1)”……..…… 68

Table 5.8.Difference of Border error achieved by LS and KLS…… 68

Table 5.9.Difference of Border error by achieved KLS and FCT(w=0)…… 69

Table 5.10. Difference of Border error achieved by KLS and FCT(w=1)…… 69

Table 5.11. Difference of Similarity achieved by LS and KLS…….…… 70

Table 5.12. Difference of Similarity achieved by KLS and FCT(w=0)……..… 70

Table 5.13. Difference of Similarity achieved by KLS and FCT(w=1)…….…… 71

Table 5.14. Difference of Hammoude Distance achieved by LS and KLS…… 71

Table 5.15. Difference of Hammoude Distance achieved by LS and FCT(w=0)…….. 72

Table 5.16. Difference of Hammoude Distance achieved by LS and FCT(w=1)…….. 72

Table 5.17. Difference of Rms Error achieved by LS and KLS…… 73

Table 5.18. Difference of Rms Error achieved by KLS and FCT(w=0)…… 73

Table 5.19. Difference of Rms Error achieved by KLS and FCT(w=1)…… 74

Table 5.20. Difference of Computational cost achieved by LS and KLS ……. 76

Table 5.21. Average performance of each set of features by proposed "Smart IPSO-SVM" algorithm for detecting the melanoma …… 78

Table 5.22. Average performance of "SVM Classification without Feature Selection" and "SVM Classification with Sequential Feature Selection"……..…… 79

Table 5.23. The T-test result between "SVM Classification without FS" and "SVM Classification with SFS"……..… 80

Table 5.24. Average performance of "SVM Classification without Feature Selection" and "SVM Classification with Sequential Feature Selection"……..…… 81

Table 5.25. The T-test result between "SVM Classification with Feature Selection" and "SA-SVM Classification with Sequential Feature Selection"…… 82

Table 5.26. Average performance of "SA-SVM Classification with Sequential Feature Selection" and "SA-SVM Classification with IPSO-SVM Feature Selection"……..… 83

Table 5.27. The T-test result between "SA-SVM Classification with Sequential Feature Selection" and "SA-SVM Classification with IPSO-SVM Feature Selection"……..… 84

Table 5.28. Average performance of "SA-SVM Classification with Sequential Feature Selection" and "SA-SVM Classification with Smart IPSO-SVM Feature Selection"..…… 85

Table 5.29. The T-test result between "SA-SVM Classification with Sequential Feature Selection" and "SA-SVM Classification with Smart IPSO-SVM Feature Selection"…..… 86

List of Abbreviations

Adaptive histogram equalization (AHE)

Artificial Neural Networks (ANN)

Analysis of variance (ANOVA)

Diagnostic odds ratio (DOR)

Epilumence microscopy (ELM)

Fuzzy c-means thresholding (FCM)

Gray Level Coocurrence Matrix (GLCM)

Gray Level Non-Uniformity (GLNU)

Gradient vector flow (GVF)

Histogram Equalization (HE)

High Grey-Level Zone Emphasis (HGZE)

Inverse difference normalized (INN)

Inverse difference (INV)

Inertia based Particle Swarm Optimization (IPSO)

K-means-level set (KLS)

K-nearest-neighbourhood (K-NN)

Low Grey-Level Zone Emphasis (LGZE)

Leave One Out (LOO)

Long Zones Emphasis (LZE)

Long Zone High Gray-Level Emphasis (LZHGE)

Long Zone Low Gray-Level Emphasis (LZLGE)

Non-melanoma skin cancer (NMSC),

Optical Coherence Tomography (OCT)

Peak signal-to-noise ratio (PSNR)

Particle Swarm Optimization (PSO)

Reflectance confocal microscopy (RCM)

Self-Advising Support vector machine (SA-SVM)

Symmetry distance (SD)

Sequential Floating Backward Selection (SFBS)

Sequential Floating Forward Selection (SFFS)

Sequential fature selection (SFS)

Short Zone High Gray-Level Emphasis (SZHGE)

Short Zone Low Gray-Level Emphasis (SZLGE)

Support vector machine (SVM)

Short Zones Emphasis (SZE)

Side transillumination (TLM)

Traditional level set (TLS)

Unsharp Masking (UM)

Ultraviolet (UV)

Cross-polarization epiluminescence (XLM)

Zone Length Non-Uniformity (ZLNU)

Zone Percentage (ZP)

ABSTRACT

During recent decades, the incident of malignant melanoma as the lethal form of skin cancer has been raised. The occurrence in Australia is much higher than US, UK, and Canada with the cases more than 10,000 diagnosis and annual mortality of 1250 people. The persistent raise of this cancer in the worldwide, the high medical cost and death rate have prioritized the early diagnosis of this cancer. The anticipation and cure of melanoma is strictly relevant to its thickness, if it can be detected early, the survival rate would be increased. Although lots of effort has been made to advance the detection of skin cancers, the challenging concerns still about it.

The computer-based detection systems can improve the diagnosis rate of melanoma by 5–30% in comparison with the naked-eye. Since the visual perception often involve some faults, the necessity of second opinion with higher accuracy and reliability is highlighted. On the other hand, it reduces the task and responsibilities that are performed by physicians.

Many researches have been developed in automated detection of melanoma. The potential advantages of such studies are significant and incalculable. Moreover, the difficulties entangle are a lot, and the new contributions in the area are highly appreciated. However, it is extensively acknowledged that the more trustful and reliable detection systems require higher accuracy. The purpose of this thesis is to propose an algorithm for skin cancer diagnosis that is able to classify lesions as malignant or benign automatically.

The different components in an automated diagnosis of skin cancer includes: Pre-processing, segmentation, feature extraction and selection, and classification. In this thesis, after selecting the best image enhancement techniques which are achieved by applying and comparing different noise removal and contrast enhancement techniques on images, the segmentation stage is performed. In this stage, a fully automated segmentation algorithm in dermoscopy images based on k-mean and level-set algorithms are proposed and compared with other algorithms mentioned in this thesis using statistical tools. Proposed algorithm shows the improvement in the results.

In the next stage, after extracting the various features of images, a fully automated feature selection algorithm, Smart PSO-SVM, which optimizes the feature selection stage, is proposed. Comparative study of proposed algorithm with other algorithms is performed to analyse the performance of proposed algorithm among others. The results obtained in the best subset of features which feed the classification stage. In classification stage, the use of SA-SVM as a new classifier in the area of skin cancer detection systems is proposed. The average accuracy and F-score are estimated as 87.0611% and 0.9167 respectively. The statistical evaluation using t-test also shows the superiority of proposed algorithm when compares with other algorithms in this thesis.