Improving Skin Cancer (Melanoma) Detection: New Method

By

Mohamed Khaled Abu Mahmoud

Thesis Submitted as a requirement for the degree of

Doctor of Philosophy

School of Electrical, Mechanical and Mechatronic Systems

Faculty of Engineering and Information Technology

University of Technology, Sydney (UTS)

November, 2014

Abstract

Melanoma, the deadliest form of skin cancer, must be diagnosed early for effective treatment. Rough pigment network and qualities are important signs for melanoma diagnosis using pathologist images. The main focus of this thesis is to improve skin cancer (Melanoma) detection through introducing novel image processing approach for a computer-aided system based on pigment network and elements detection on pathology images. It is important to propose an automated system for differentiating between melanocytic nevi and malignant melanoma. This thesis describes a novel image processing approach for computer-aided pigment network and elements detection on dermoscopy / pathology images. The proposed methods provide meaningful ideas of structures, and extract features for melanoma detection. Additionally, the thesis presents efforts towards prevention of melanoma, by developing a smart system to locate pigment networks.

The thesis aims to cover a complete theoretical model for simulating the processes that takes place when a human interprets an image generated by the eye, through designing a reliable system, that can provide a screening method that "filters" lesions and melanoma in a general practice. The proposed system is to be used with a standard PC with input from a high quality digital camera, dermoscopy / microscopy slides or any other suitable hardware sources. This system analyses the structure of a mole / skin defects, detects cancer, identifies features, makes a decision and provides the result.

The result of the proposed system shows that the Skin Cancer (Melanoma) Detection strategy which uses SVM performs reasonably satisfactorily (accuracy 77.44%, sensitivity 83.60 %, and specify 70.67%). Furthermore, the SVM based wavelet Gabor (SVM-WLG) performs better than the SVM (81.61%, 88.48%, and 74.51 % accuracy, sensitivity, and specify respectively). However, the Swarm-based SVM (SSVM) performs better than the other two algorithms, with average for accuracy, sensitivity, specificity of 87.13%, 94.1% and 80.22%, respectively.

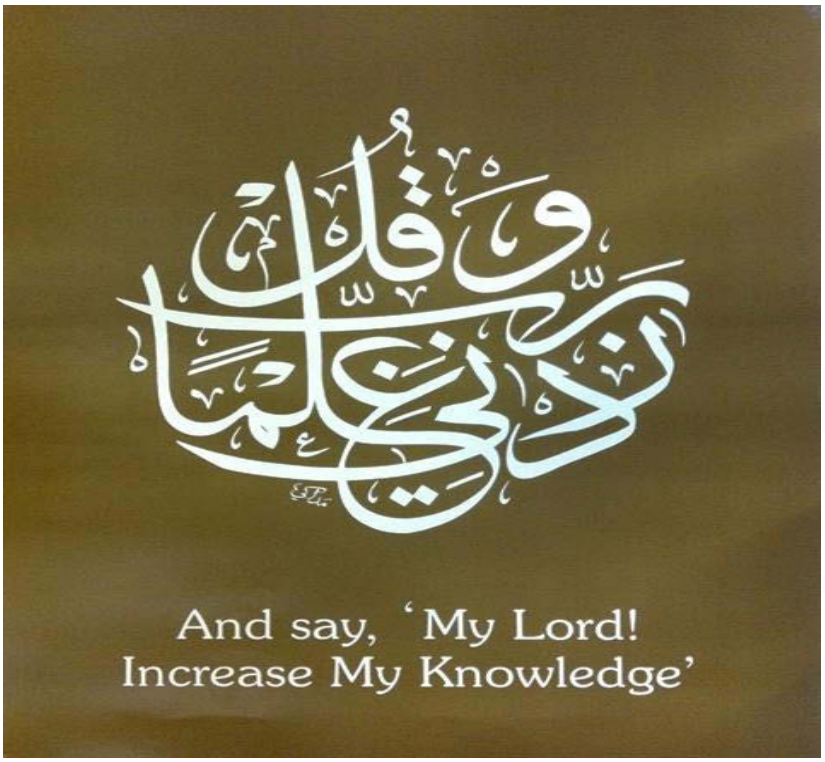
Certificate of Authorship / Originality

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.

Mohamed Khaled Abu Mahmoud

November, 2014



To the soul of my father and my mother

For encouraging me and pounding me with my education

Acknowledgments

Firstly, I would like to express my sincere gratitude to my principal supervisor, Assoc. Prof. Adel Al-Jumaily, who provided such expert guidance and advice throughout this PhD candidature. He possesses an amazing amount of expertise and an ability to critically appraise work, that ensures the end result is worthy and importantly of value to public health.

I would also like to sincerely thank and acknowledge the pathologist Dr. Sarbar Napaki from Southern Pathology laboratory for providing us with pathologist data samples and for updating our medical expertise with the intension of improving our results. Also I would like to extend my thanks to Dr. Ryad Ahmed (for reviewing Chapter 2; Human Skin Biology and Cancer and Related Work in this thesis) and Dr. Rami Khushaba, Dr. Jebrin Sharawneh and Miss. Hayat Al-Dmour who have supported me whenever requested.

I am truly blessed to work in research for an organization dedicated to cancer detection and due to this I would like to express my sincere gratitude to University of Technology for the great experience and knowledge it gave me, my colleagues at the University of Technology, Sydney, Faculty of Engineering and Information Technology, School of Electrical, Mechanical and Mechatronic Systems who always encouraged and supported me, providing that little extra push to enable me to complete this thesis.

I am very lucky to have a wonderful supportive family. My two sons, my three daughters, my three sons in law and my grandkids have always encouraged me in my striving to complete this piece of work. Finally, to my wife, my partner in life for the past 43 years, this would not have been possible without your love and support.

Table of Contents

To the soul of my father and my mother 3

For encouraging me and pounding me with my education ... 3

Acknowledgments 4

Table of Contents 5

Extended Table of Contents 7

List of Figures 12

List of Tables 18

CHAPTER 1 19

Introduction 19

CHAPTER 2 32

Human Skin Biology and Cancer. 32

CHAPTER 3 53

Image Processing and Automatic Skin Cancer Diagnosing 53

CHAPTER 4 92

Pre-processing and Segmentation 92

CHAPTER 5 122

Image Representation and Analysis 122

CHAPTER 6 150

Experimental Results and Discussion 150

CHAPTER 7 185

Conclusion and Future Work 185

Appendices 189

Bibliography 220

Extended Table of Contents

Table of Contents

To the soul of my father and my mother 3

For encouraging me and pounding me with my education ... 3

Acknowledgments 4

Table of Contents 5

Extended Table of Contents 7

List of Figures 12

List of Tables 18

CHAPTER 1 19

Introduction 19

1.1. Research Aims 20
1.2. Problem Description 20
1.3. Objectives 21
1.4. Methodology 22
1.5. Contribution of the Doctoral Research 23
1.6. Structure and Summary of the Dissertation 27
1.7. Publications Presented During the Doctoral Research 29

CHAPTER 2 32

Human Skin Biology and Cancer. 32

2.1. Human Skin Biology 32
2.2. The Cell (Structure and Function) 32
2.2.1. Cell structure 33
2.2.2. Cell Cycle and Replication: 33

2.3. Haematoxylin and Eosin (H & E) 35
2.4. Skin Cell Structure 35
2.4.1. Protection *36*
2.4.2. Sensation *36*
2.4.3. Thermoregulation *36*
2.5. The Layers of the Skin 37
2.5.1. Epidermis *37*
2.5.2. Dermis *37*
2.5.3. Subcutaneous/hypodermis *37*
2.6. Melanocyte 38
2.7. Physical Properties of Human Skin 38
2.8. Human Skin 40
2.8.1. Types of Skin *40*
2.8.2. Layers of Human Skin *40*
2.8.3. Cancer Types *41*
2.8.4. Types of Skin Cancer *41*
2.8.5. Health Behavior of Cancer Survivors *49*

CHAPTER 3 53

Image Processing and Automatic Skin Cancer Diagnosing 53

3.1. Introduction 53
3.2. Vision System & Structure of the Human Eye 53
3.2.1. Human Visual Perception *53*
3.2.2. The Human Eye *54*
3.2.3. Structure of the Eye *54*
3.3. Image Processing & Human skin Imaging 55
3.3.1. Applications of Image Processing *56*
3.3.2. Digital Image Acquisition *58*
3.3.3. Digital Camera *59*
3.3.4. Skin Imaging Techniques *60*
3.3.5. Differential Diagnosis of Pigmented Lesions of the Skin *65*
3.3.6. Automatic Skin Cancer Diagnosing *72*

CHAPTER 4 92

Pre-processing and Segmentation..... 92

4.1. Introduction 92

4.1.1. Digital Image Representation 93

4.1.2. Image Types 93

4.1.3. General approach of developing a Computer Aided Design (CAD) system 94

4.2. Pre-Processing Stage..... 96

4.2.1. Wiener Filter: 98

4.2.2. Gabor Filter 100

4.2.3 Median filter: 102

4.2.4. Adaptive Median Filter (AMF):..... 103

4.3. Image preprocessing: enhancement 104

4.3.1. Image enhancement 105

4.3.2. Intensity transformation 106

4.3.3. Histogram Processing 106

4.3.4. Histogram Equalization 108

4.3.5. Special filtering 111

4.4. Image Segmentation 113

4.4.1. Edge-based Image Segmentation 114

4.4.2. Edge Detection Operations. 114

4.4.3. Thresholding..... 115

4.4.4. Global Thresholding..... 116

4.4.5. Optimal Global Thresholding 118

4.4.6. Pixel Classification through Clustering..... 120

CHAPTER 5 122

Image Representation and Analysis 122

5.1. Introduction 122

5.2. Machine Learning 122

5.3. Main Concepts 123

5.4. Intelligent Image Features Extraction 124

5.4.1. Texture: *125*
5.4.2. Texture features: *126*
5.4.3. Grey-Level Co-occurrence Matrix and Features *127*
5.4.4. Gabor filter feature extraction *128*
5.4.5. Wavelet decomposition of an image *129*
5.4.6. Discrete Wavelet Transform *130*
5.4.7. Classifier distance metric equation *131*
5.5. Feature Extraction and Representation 131
5.5.1. Methodology used in feature extraction *132*
5.6. Classification and Feature selection 133
5.6.1. Feature Selection (FS) *133*
5.6.2. Unbalanced data sets *136*
5.7. SVM for Classification and Regression 137
5.7.1. Linear support vector machine *137*
5.7.2. Nonlinear support vector machine *140*
5.7.3. Support Vector Regression *143*
5.8. Swarm based Support Vector Machine for Melanoma Detection ... 144
5.8.1. Development of a Melanoma Detection based on the Swarm based Support Vector machine *144*
*5.8.2. Optimization of **SVM** parameters using **PSO*** *145*
5.8.3. Fitness function for the optimization *147*

CHAPTER 6 150

Experimental Results and Discussion 150

6.1. Introduction 150
6.1.1. Experiment 1 ..... *152*
6.1.2. Experiment 2 ..... *160*
6.1.3. Experiment 3 ..... *166*
6.1.4. Experiment 4 ..... *171*
6.1.5. Experiment 5 ..... *174*
6.1.6. Experiment 6 ..... *176*

CHAPTER 7 185

Conclusion and Future Work 185

7.1. Conclusion..... 185

7.2. Future Research 187

Appendices 189

A. Appendix A: Glossary of Pathology Terminology 189

B. Appendix B. Glossary of Pathology Terminology 196

C. Appendix C: Glossary of Machine Learning and Computer Vision Terminology 197

D. Appendix D: Margin between two hyper planes..... 203

E. Appendix E: Lagrangian dual optimization..... 204

F. Appendix F: Soft-margin nonlinear support vector machine 206

G. Appendix G: Sequential minimal optimization (SMO) for SVM 208

H. Appendix H: Glossary and Abbreviations..... 211

Bibliography 220

List of Figures

Figure 2.1 shows, Cells structure; source: www.getting-in.com 33

Figure 2.2 The eukaryotic cell cycle, G Phases: growing, S phase: synthesis/ DNA replication, G2 Phase: Growing and preparing for Mitosis, M mitosis. 34

Figure 2.3 The cell cycle; G0 terminally differentiated cell G1 gap phase 1 G2 gap phase 2 M mitotic phase S synthesis phase 35

Source: Wheatear's Functional Histology: A Text and Colour Atlas, 5th Ed 35

Figure 2.4 H&E-stained section of skin. High magnification (40xs) view. Haematoxylin stains the nuclei of cells blue to bluish-purple, and eosin stains other cellular elements in the tissues from pink to red [43]. 35

Figure 2.5 Displayed Cross Section & the main layers of Human Skin 37

Source: nurrashidah2204.blogspot.com 37

Figure 2.6. Skin Color Distribution around the World 38

Source: www.gbhealthwatch.com 38

Figure 2.7. a) Cross section of human skin, b) Structure of Epidermis. 39

Figure 2.8. The Layers of Skin - the Epidermis and Dermis (At the top, the close up shows a squamous cell, basal cell, and melanocyte). 42

Figure2.9- a) Basal Cell Carcinoma (BCC), b) Squamous Cell Carcinoma (SCC) 42

Figure 2.10 – Melanoma 43

Figure 2.11: Malignant Melanoma, World Age-Standardised Incidence Rates, World Regions, 2008 Estimates. 44

Figure 2.12. Lentigomaligna Melanoma. 45

Source: melanomaknowmore.com 45

Figure 2.13. Superficial Spreading Melanoma. 46

Source: melanomaknowmore.com 46

Figure 2.14. Nodular Melanoma. 46

Source: melanomaknowmore.com 46

Figure 2.15.Stages of Cancer Development and Metastasis. 47

Source: http://www.cancervic.org.au/about-cancer/advanced-cancer 47

Figure 2.16: Five-year relative survival from selected cancers by remoteness area, Australia. 2006-2010. 49

Figure 3.1 - Show the function of the eye. 55

Source: http://cdn2.hubspot.net/hub/60407/file-350482438-jpg/images/Master_Eye_logo_short_tag_line_revised_10-13-2013-resized-600.jpg 55

Figure 3.2 - Distance of vision. 55

Source:hyperphysics.phy-astr.gsu.edu/hbase/vision/accom.html.....55

Figure 3.3 - Structure of the Eye.....55

Source:drpion.be/en/bouw-van-het-oog.htm 55

Figure 3.4 Focal length defined [When parallel rays of light strike a lens focused at infinity, they converge to a point called the focal point. The focal length of the lens is then defined as the distance from the middle of the lens to its focal point.] 59

Figure 3.5. Displayed, similarities of overall design in the principal features of an optical microscope, a transmission electron microscope and a scanning electron microscope.....63

Source: www.nslc.wustl.edu.....63

Figure 3.6. Principle of CSLM. Note the same optical path is used for the detector and the source. Optics are used to direct the light towards the detector 64

Figure 3.7. Two-step procedure for the classification of pigmented skin lesions. Adapted from Argenziano.[122] 65

Figure 3.8. Algorithm for the determination of melanocytic versus non melanocytic lesions according to the proposition of the Board of the Consensus Netmeeting. Adapted from Argenziano.[122] 66

Figure 3.9. A, Macroscopic picture of a superficial spreading malignant melanoma (Breslow thickness 0.52 mm; Clark level II). B, Dermoscopy of A shows (atypical) pigment network and branched streaks and can therefore be considered a melanocytic lesion.....66

Figure 3.10. A, Macroscopic picture of a blue nevus. B, Dermoscopy of A shows steel-blue areas (no pigment network, no aggregated globules, and no branched streaks).....67

Figure 3.11. A, Macroscopic picture of a seborrheic keratosis. B, Dermoscopy of A shows comedolike openings (a), multiple milia-like cysts (b), and fissures (c).....67

Figure 3.12. A, Macroscopic picture of a seborrheic keratosis. B, Dermoscopy of A shows comedolike openings and multiple milia-like cysts.....67

Figure 3.13. A, Macroscopic picture of a basal cell carcinoma. B, Dermoscopy of A shows maple lifelike areas, ovoid nests, and arborized telangiectasia. 67

Figure 3.14. A, Macroscopic picture of a basal cell carcinoma. B, Dermoscopy of A shows multiple spoke wheel areas.....67

Figure 3.15. A, Macroscopic picture of an angioma. B, Dermoscopy of A shows red lagoons.....68

Figure 3.16. Asymmetry Border Color Diameter Evolution (ABCD-E) rule for Diagnosis of Melanoma 68

Source: www.webmd.com.....68

Figure 3.17 - Normal moles and Melanomas [Benign and Malignant]. 70

Source: www.skinbychar.com 70

Figure 3.18: Showed: a) Melanoma in skin biopsy with H&E Stain ("This case may represent superficial spreading melanoma."), b) Lymph node with almost complete replacement by metastatic melanoma. The brown pigment is focal deposition of melanin. 71

Figure 3.19. Diagnosis of Melanoma using Three-point Checklist 72

Figure 3.20 - Example of pigmented skin lesion. Left: Traditional imaging technique. Right: Dermoscopy imaging technique. 73

Figure 3.21. (a) Macroscopic Image of a Lesion (b) Dermatoscopic Image of the same Lesion. Source: www.jle.com 74

Figure 3.22- Show the results of skin lesion boundary tracing algorithm, my data experiment shown [a] Image center of mass, [b] Image process 75

Figure 3.23- From left to right: the original image (source image from: http://ijplugins.sourceforge.net/plugins/clustering/), clustered image using 2 clusters (poor), clustered image using 3 clusters (close but one key cluster is missing), clustered image using 4 clusters (optimal), and clustered image using 10 clusters (artifacts are noticeable). 77

Figure 3.24- Structure of fuzzy image processing 78

Figure 3.25- Steps of fuzzy image processing. 78

Figure 3.26. Manual Segmentation by trained pathologists using the 'Aperio Image Scope' software [156]. 82

Figure 3.27 Showed the Radial growth phase melanoma. 83

Figure 3.28 showed the Vertical growth phase melanoma 84

Figure 3.29 Microscopic satellites, (a). Shows neoplastic group is discontinuous (arrow) from the overlying vertical growth phase component. Shows the melanocytes must have a malignant cytomorphology (b). 85

Figure 3.30 Regression: Regression of over 75% of the tumor volume of a melanoma is considered a bad prognostic sign. 86

Figure 3.31: a) An image of a lesion under clinical view (naked eye). b) Shows the same lesion under a dermoscopy with oil immersion. 86

Figure 3.32: a) Colors allow the physician, to some extent, to draw conclusions about the localization of pigmented cells within the skin. Black and brown indicate pigmentation in the epidermis, while grey and blue correspond to pigmented cells within the superficial and deep dermis, respectively [185]. 87

Figure 3.33: Figures (a, b, c, d, and e) show analogue dermoscopy. All of them, except (a), are attachable to digital cameras to function as digital dermoscopy. Figures (f, g, and h) show DinoLite, Handy scope, and Dermoscopy which are modern digital dermoscopy. ... 88

Figure 3.34. The Solar Scan instrument: (a) global appearance, (b) camera, (c) user interface [192]. Source: www.medgadget.com 90

Figure 4.1 Showed: Four stages CAD system for skin lesions. 96

Figure 4.2 true color Pathological Digital Image. 96

Figure 4.4, calculating the median value of a pixel neighborhood. 102

Figure 4.5. The basic structure of the MACWM filter 104

Figure 4.6. Shows the output image of Adaptive median Filtering. 104

Figure 4.7 Showed: (a) grayscale image, (b) Histogram out. 107

Figure 4.8. Showed: intensity image using 109

Figure 4.9. Sobel Edge Detection Image 114

Figure 4.10.Binary Output Image Otsu's Method[..... 114

Figure 4.11, selected thresholds in valleys between peaks 116

Figure 4.12, obtaining the best possible separation measure for two classes 116

Figure 4.13, selecting a threshold by visually analyzing a bimodal histogram. 116

(Principle of histogram peak separation). 116

Figure 5.1 displayed the main categories: texture elements, regularity, randomness, directionality and regularity. 126

Figure 5.2, showed, (a) images (5x 5) matrix with 3 grey levels 0, 1, and 2. (b) The co-occurrence matrix for d= (1, 1). 127

Figure 5.3, displayed the response from convolving image sample with filter. 129

Figure 5.4, (a) Wavelet decomposition of an image (b) Block diagram of the decomposition of an image 130

Figure 5.5 displayed a block diagram of the discrete wavelet transforms (DWT – Gabor approach). 130

Figure 5.6, the three principal approaches of feature Mammographic e selection. The shades show the components used by the three approaches: filters, wrappers and embedded methods. 133

Figure 5.7: Two-out-of-many separating lines; (a) with smaller margin and (b) with larger margin 138

Figure 5.8: Margin m between two supporting hyperplanes. 139

Figure 5.9: Illustration of mapping using a transform $\Psi: \mathbb{R}2 \rightarrow \mathbb{R}3$. 140

Figure 5.10: Introducing slack variable ξ in soft-margin SVM 142

Figure 5.11: Loss Functions 144

Figure 5.12: Melanoma detection using swarm based support vector machine. 145

Figure 5.13: The pseudo of the PSO for the SVM parameter optimization 146

Figure 5.14: The particles of the PSO 147

Figure. 6.2 Skin Cancer Image, (a) Original RGB true colour image 153

Figure 6.3. The basic structure of the MACWM filter 154

Figure 6.4. A snake with traditional potential forces cannot move into the concave boundary region. 156

Figure 6.5 A snake with GVF external forces moves into the concave boundary region. 157

Figure. 6.6. Skin Cancer greyscale image showing: (a) Wiener2 filter has removed the spots effectively (b) noisy image filtered by the median filter 157

(a) Without filtering, (b) with filtering. 158

Figure.6.7 Skin Cancer Image Histogram, ... 158

Figure. 6.8 Skin Cancer greyscale image showing: (a) Sobel segmented image 159

(b) Gradient vector flow (GVF) segmented image 159

Root mean square error 160

Figure 6.9. Result of segmentation by threshold with GHE (bottom right) and LHE (bottom left), Top: shows histogram Local and Global results 162

Figure 6.10.Segmentation by Thresholding (ROI) 162

Figure 6.11. Segmentation by SRM 163

Figure 6.12. Display of the image and its transform (wavelet coefficients) 164

Figure. 6.13: A single curvelet with width $2 - 1$ and length $2 - 12$. 164

Figure 6.14. Rectangular frequency (basic digital) tiling of an image with 5 level curvelet. 165

Figure 6.15. Curvelet demising image that contains oriented texture and cartoon edges (a) Original (b) Noisy, and (c) Curvelet. 166

Figure. 6.16. The results showed, original image, image after median filter, gray scale image, ... 172

Figure. 6.17. - A Wavelet Packet decomposition tree 178

Figure. 6.18 shows the melanoma detection employing SVM (SVMR, SVMP, and SVML); the input is pathology images and the output is melanoma (-1) or benign (+1). 181

Figure A1, showed a physician's hands are seen performing a needle biopsy to determine nature of lump either fluid-filled cyst or solid tumour 189

[This image was released by the National Cancer Institute, an agency part of the National Institutes of Health, with the ID 1973 (image)] 189

Figure A2, the diagnostics showed: Micrograph of a needle aspiration biopsy specimen of a salivary gland showing adenoid cystic carcinoma. (Source: Pap stain. MeSH D044963). 189

Figure A3, showed a Normal Epidermis and Dermis with Intradermal Nevus 10x-cropped. (Source, Kilbad). 190

Figure A4, this image shows a crossection of the vascular tissue in a plant stem. Showed as an example for bright field micrograph. (Source: Wikipedia). 190

Figure A5, Principle of confocal microscopy 191

Figure A6 showed the nucleolus is contained within the cell nucleus. (Source: Wikimedia Foundation, Inc.) 192

Figure A7 displayed the negative of the logarithm (base 10) of the optical Density (OD). (Source: semrock@idexcorp.com) 193

Figure A8, showing a pathologist examines a tissue section for evidence of cancerous cells while a surgeon observes. (Source: wikipedia.org) 193

Figure A9, showed: a) A TEM image of the polio virus is 30 nm in size, b) Layout of Optical components in a basic TEM. (Source: Wikimedia)Appendix B: Optical Density of Transmission Microscopy Images 194

Figure B1. Example optical density image 196

Figure C1, showed a graph with three connected components. 197

Figure C2, displayed the machine learning and data mining which used to solve problem in the areas of Classification, Clustering, Regression, Anomaly detection, Association rules, Reinforcement learning, Structured prediction, Feature learning, Online learning, Semi-supervised learning, and Grammar induction. 198

Figure D1: Margin *m* between two supporting hyperplanes 203

List of Tables

Table 2.1: Breslow's depth 48

Table 2.2: Survival figures from British Association of Dermatologist Guidelines 2002 .. 48

Table 2.3: Observed incidence (2009), and morality (2010) of melanoma of the skin and estimated for 2012. 51

Table 2.4: Observed incidence (2009), and morality (2010) of non-melanoma of the skin and estimated for 2012 52

Table 6.1 Sensitivity and Specificity comparison. 152

Table 6.2 Comparison between proposed operators (Sobel, Roberts, Prewitt, Laplacian, Canny and Otsu) and Gradient vector flow (GVF) segmented images. 160

Table 6.3 BNN Classification Test with Different Wavelet. 161

Table 6.4. BNN Classification Test for SRM & Thresholding. 163

Table 6.6. Displaying the comparison between wavelet and curvelet features results**.** 166

Table 6.7. svm classification for all the dataset (including dermoscopy images) with features extracted from wavelets 169

Table 6.8. svm classification excluding dermoscopy images with features extracted from wavelets 170

Table 6.9. svm classification for all the dataset (including dermoscopy images) with features extracted from curvelets 170

Table 6.10. svm classification excluding dermoscopy images with features extracted from curvelets 170

Table 6.11. sensitivity and specificity statistics [310] 171

Table 6.12. Described sensitivity and specificity 174

Table 6.13. show result after training or the (svm) network. using (sfs) technique 174

Table 6.14. the result after training of the (svm) network with using (sfs) technique. 175

Table 6.15. the result after training of the (svm) network without using (sfs) technique. 175

Table 6.17 The Result after Training of the (SVM) Network, with using (SFS)Technique. 180

Table 6.18 The Result after Training of the (SVM) Network. with using (SVM+WLG+PSO) 180

Table 6.19 Described Sensitivity and Specificity [31] 181

CHAPTER 1

Introduction

Australia has one of the highest rates of skin cancer in the world, at nearly four times the rates in Canada, the US and the UK [1]. Two in three Australians will be diagnosed with skin cancer by the time they are 70. It has been estimated 115,000 new cases of cancer per year are diagnosed and more than 43,000 people are expected to die from cancer per year according to the Cancer council of Australian 2014 [2]. The Chair of Public Health Committee pronounced that more than 430,000 cases are treated for non-melanoma, and more than 10,300 people are treated for melanoma, with 1430 people dying each year[2].

Skin cancer is the most expensive cancer [3]. In 2001, it was estimated the treatment of non-melanoma skin cancer cost $264 million and melanoma $30 million. The skin cancer malignant melanoma is the deadliest form and accounts for 75% of all cancer deaths [4] [5]. It can be removed by simple surgery if it has not entered the blood stream. The annual rate of melanoma is increasing at the rate of 6% Worldwide[5]. If the melanoma gets deeper than 3 millimeters[6], the chance of survival is 59%[7]. A patient can recover from melanoma if it is detected and treated in the early stages and this can achieve cure ratios of over 95%. Early diagnosis is obviously dependent upon patient attention and accurate assessment by a medical practitioner. The variations of diagnosis are sufficiency large and there is a lack of detail of the test methods. The diagnostic is mostly based on visual inspection. The traditional imaging is just a recording of what the human eye can see using a digital camera while the Dermoscopy known as Epiluminescence Light Microscopy (ELM) images need a professional with experience to get the required image.

Australia is one of many countries in which skin cancer is widely spread in comparison to other types of cancer[8], and researchers from the School of Medicine, University of Queensland, Australia, found that Australian melanoma rates are the highest globally at almost four times the rates seen in Canada, the United Kingdom and the United States[9] [10] with mortality two-fold that of Southern and central Europe[11].

1.1. Research Aims

The aims of this research are to improve the quality of existing diagnostic systems by proposing 283 new feature extraction, selection and classification methods. These methods are aimed for early detection of pathologic melanoma. This process provides a better tool for screening and detecting lesions that are considered to be suspicious allowing early treatment and improved survival rate. The research focuses on the proposed algorithm that is a combination of segmentation methods for skin cancer image to detect the lesion border (edges), followed by feature extraction, selection and classification. The diagnostic results of the proposed system are compared with other algorithms as well.

1.2. Problem Description

It is difficult to visually differentiate a normal mole from an abnormal one and general practitioners (GP) do not usually have sufficient expertise to diagnose skin cancers. Skin cancer specialists will significantly improve the identification rate by over 75% but are often severely overloaded by referrals from regional practices. For this reason the thesis focuses on the development and implementation of a skin cancer screening system that can be used in a general practice by non-experts to 'filter' the cases, so that in the case of an abnormal mole, a patient can be referred to a specialist.

The automatic diagnostic system for melanoma has been developed and designed for use with a standard PC with input from a quality digital camera, / dermoscopy / microscopy slides or any other suitable hardware sources. It analyses the structure of skin defects, detects cancer identifying- features, makes a decision using a knowledge database and outputs a result. Skin cancer experts create a knowledge database by training the system using a number of case-study images.

One commercial product, Solar Scan by Polar technics, has an accuracy rate of 92%. Solar Scan is a complex system, taking high quality Epiluminescence Light Microscopy (ELM) images and using advanced image analysis techniques to extract a number of features for classification which makes it not suitable for normal person usage. This Dermoscopy equipment (known as Epiluminescence Light Microscopy) needs a professional with experience to get the required image. Researchers in [12] [13] [14] reported that the performance improved by using Oxford Recognition Skin Cancer Screen System

(ORSCSS), which was developed by Oxford Recognition Limited (ORL) in collaboration with Loughborough University. This process investigates the applications of pattern recognition using fractal geometry as a central processing kernel, which allowed the design of a new library of pattern recognition algorithms including the computation of parameters. Also this process of object detection, recognition and classification in a digital image where the classification method is based on the application of a set of features which include fractal parameters where it incorporates the characterisation of an object in terms of its texture. However, reported that this texture based analysis alone is not sufficient in order to design a recognition and classification system [12].

A further development for this process is to consider the effect of replacing the fuzzy logic engine used with an appropriate Artificial Neural Network (ANN) even then it is not clear whether the application of an ANN could provide a more effective system and whether it will provide better flexibility according to the type of images used and the classifications that may be required. However, the result of the proposed system shows that the Skin Cancer (Melanoma) Detection strategy which uses SVM performs reasonably satisfactorily (accuracy 77.44%, sensitivity 83.60 %, and specify 70.67%). Furthermore, the SVM based wavelet Gabor (SVM-WLG) performs better than the SVM (81.61%, 88.48%, and 74.51 % accuracy, sensitivity, and specify respectively). However, the Swarm-based SVM (SSVM) performs better than the other two algorithms, with average for accuracy, sensitivity, specificity of 87.13%, 94.1% and 80.22%, respectively. It was concluded that the proposed system gives fast, accurate classification and better results. This thesis concludes that there are some possible factors to improve the accuracy of detecting malignant melanoma by having a higher number of images for training of the SVM network [15].

1.3. Objectives

This thesis targets the above mentioned problems in an attempt to solve them by developing new algorithms and applying these algorithms into novel applications. This thesis aim to develop hybrid artificial intelligence techniques employing two or more of the following methods: **Wavelet Gabor (WLG)**, fuzzy logic (FL), neural networks (NN), k-means clustering, **hybrid evolutionary optimization algorithm based on combining Particle Swarm Optimization (PSO) and Support Vector Machine (SVM) called PSO-SVM**. The process involves:

1. Using combination of methods to target different parts of the automated recognition system like segmentation, features extraction, feature selection, dimensionality reduction techniques, and pattern recognizers or classifiers.

2. Comparing with the available approaches in literature and demonstrate the successfulness and superiority of the new novel methods.

3. Investigating the utilization of the pattern recognition system into novel applications in healthcare, like identifying the deadliest form of skin cancer problems.

4. Using real patient pathology images sourced from a consultant Pathologist.

This thesis will include a practical element to investigate questions, primarily targeted at skin cancer (melanoma) patients. The thesis is trying to direct the work to meet the Australian community needs in the area of skin cancer (melanoma).

1.4. Methodology

Most general practitioners (GP) have less experience in the full range of melanoma forms, which means many cases, are not diagnosed properly [GPs cannot provide cancer care alone. Nor does the solution lie in the recruitment of more specialists, such as dermatologists [16]. Last decade researchers developed automatic early diagnostic aided systems to advance the diagnostic accuracy from 60% to 92% [17] and they became able to provide recommendations for non-specialized users. As mentioned above early detection improves survival rate. The methodology that has been developed relies on extracting and selecting specific information features that can be used to distinguish malignant and benign lesions by setting an automated cancer diagnosis / image processing system. The common approach to skin lesion images combines four main computational intelligent stages [18] as combined on the following modules:

Pre-processing stage: This stage consists of filtering and contrast enhancing techniques to remove any unwanted structures that might corrupt the image. Also the aim of this stage is to eliminate the background noise and improve the image quality for the purpose of determining the focal areas in the image.

Segmentation using thresholding, region based approach or boundary tracing algorithms to localize the lesion.

Feature extraction / selection. This stage quantifies the distribution of the cells across the tissue. It primarily makes use of the grey-level dependency of the pixels and selects the useful information from these images and passes it to the classification stage for training and testing in order to distinguish between malignant or benign lesions.

The classification of image: This stage used Back-propagation Natural Network or Support Vector Machine [19] to find out which algorithm performed better. This step uses statistical analysis of the features and machine learning algorithms to reach a decision.

The variations of diagnosis are sufficiently large and there is a lack of detail of the test methods. This thesis looks at developing a practical approach to extend the functionality of the existing methods and tools for medical diagnostic purposes and in the long run provide an alternative basis for researchers to experiment with new and existing methodologies for skin cancer detection and diagnosis.

1.5. Contribution of the Doctoral Research

The purpose of this research is to improve the quality of existing diagnostic systems by proposing some new feature extraction, selection and classification methods. These methods are aimed for early detection of pathologic melanoma. This research focuses on the proposed algorithm that is a combination of segmentation methods for skin cancer image to detect the lesion border (edges), followed by feature extraction, selection and classification. The methodology that has been developed distinguishes malignant and benign lesions by setting an automated cancer diagnosis / image processing system.

This thesis, presents a segmentation of skin cancer images based on gradient vector flow (GVF) snake algorithm. This algorithm used winner and median filters respectively to remove the unwanted noise and hairs from the image. Since edge detection is the initial step in object recognition, it is necessary to know the differences between edge detection algorithms. In this thesis we classified the most commonly used algorithms into five categories, then these algorithms have been applied to 8 images. A comparative way published in our paper in [20] that shows, Sobel and Prewitt with better performances for an image contaminated with Gaussian noise and filtered with Winner and Median filters.

These operators also have good contrast, edge map strength and low noise content. Prewitt is more acceptable than Roberts and Laplacian while Sobel are more acceptable than Prewitt. It also shows that Laplacian is very much sensitive to noise. Canny operator shows better detection, good response and improving signal to noise ratio. Filtering the noisy image with a suitable filter is an initial process in the edge detection for noisy images. The subjective evaluation of edge detected images show that proposed operators , have evaluated both the performance and Root mean Square of Signal to Noise ratio between the input image (operators) and the output image (GVF).

Building an image classification system for early skin cancer detection includes different stages. The pre-processing resizes the image which improves the speed performance and removes the extra features. Post-processing enhances the image quality and sharpens the outline of the cancer cell. On other hand, segmentation to find the region of interest by removing the healthy skins from the image and keeping the cancer cell whereas the feature extraction decomposes the useful features without prior clinical knowledge.

The classification results, which based on a database of mixed type of images, showed in our publication [21] are as follows: 58.44 % for back-propagation neural network when using Wavelets and 86.57 % when using Curvelet. This thesis concludes that there are some possible factors for low classification results; one of the main factors that affect reaching high accuracy is the type of database used. A mix of images that were acquired from different sources mainly normal digital images and dermoscopy images has been used. While the proposed technique has tried to generalize the system working area, it has traded off the accuracy. A possible improvement and generalization can be achieved by dealing with the effect of large variation and overcoming it using larger databases.

In this thesis, the Support Vector Machine (SVM) has been implemented for classification of benign from malignant skin tumor. MATLAB package is used to implement the software in the current work; these features were carried out to generate training and testing of the proposed SVM. The present work is a new application based on histo-pathological images of skin lesions that required finding out new features and the correlation of the reduced numbers of features and getting better accuracy. It was required to modify many of the mentioned techniques to make them working for such application. 42 images (28 benign images and 14 melanoma images) were sampled and split into a

training set (mixed 28 images) and test set (mixed 14 images) were sampled from microscopic slides of skin biopsy which have been used in the current work. The performance of the algorithm was evaluated by computing the percentages of Sensitivity (SE) equation (1.1), Specificity (SP) equation (1.2) and Accuracy (AC) equation (1.3): the respective definitions are as follows:

$$SE = \frac{TP}{(TP+FN)} X\ 100 \quad 1.1$$

$$SP = \frac{TN}{(TN+FP)} X\ 100 \quad 1.2$$

$$AC = \frac{(TP+FN)}{(TN+TP+FN+FP)} X\ 100 \quad 1.3$$

Where TP is the number of true positives (expects malignant as malignant), TN is the number of true negatives (expects benign as benign), FN is the number of false negatives (expects malignant as benign), and FP is the number of false positives (predicts benign as malignant). To estimate the performance of classifier based on the classification of benign and malignant skin cell nuclei, sensitivity, specificity and accuracy of prediction have been calculated according to the above equations for all of the testing data.

The higher accuracy of diagnosis proposed work is calculated to display. The results proved to be one of the best results comparing with expertise / physicians results displayed in our publication in [22]. The obtained accuracy of the system is 88.9 % whereas the sensitivity and specificity were found to be equal 87.5 % and 100% respectively. By comparing the results in [22], it was concluded that the proposed system gives fast and accurate classification and better results of skin tumors than physicians do. This thesis concludes that there are some possible factors to improve the accuracy of detecting malignant melanoma by having a higher number of images for training of the SVM network.

A Novel feature extraction methodology based on histopathalogical images and subsequent classification by Support Vector Machine was used. The results show that using the sequential forward selection (SFS) technique produces one of the best results compared with expertise / physicians results displayed in our paper in [23]. The obtained accuracy of the system is 81 % whereas the sensitivity and specificity were found to equal 76 % and 100% respectively. through comparing the results in [23] and from our experimental results CPU time which is faster than the previous experiment it produced more accurate classification results, than physicians do.

The contribution of this thesis proposes an automated non-invasive system for skin cancer (melanoma) detection based on Support Vector Machine classification. The proposed system uses a number of features extracted from the Wavelet or the Curvelet decomposition of the grayscale skin lesion images and colour features obtained from the original colour images. The recognition accuracy obtained by the Support Vector Machine classifier used in this experiment is 87.7 % for the Wavelet based features and 83.6 % for the Curvelet based ones. The proposed system has resulted in a sensitivity of 86.4 % for the case of Wavelet and 76.9% for the case of Curvelet. It has also resulted in a specificity of 88.1% for the case of Wavelet and 85.4% for the case of Curvelet. The obtained sensitivity and specificity results are comparable to those obtained by dermatologists[24].

A hybrid system for skin lesion detection: based on Gabor wavelet and support vector machine (SVM) for classification of benign from malignant skin tumor has been implemented in this thesis. 79 images sampled from microscopic slides of skin biopsy have been used in the current work which uses a new application based on histo-pathological images of skin lesions that required finding out new features getting the reduced numbers of features and getting better accuracy. It was required to modify many of the mentioned techniques to make them work for such an application. The higher accuracy of diagnosis of the proposed work is calculated and displayed. The results show that using the particle Swarm Optimization Support Victor Machine (SSVM) to improve the parameters of support victor machine (SVM) technique produces one of the best results compared with expertise / physicians results displayed in [25]. The obtained accuracy of the system is 87.1% whereas the sensitivity and specificity were found to equal 94.1% and 80.2% respectively. The proposed system produced more accurate classification results than physicians did [25].

This doctoral research has provided contributions to Skin Cancer (Melanoma) Detection models. Novel algorithms and SSVM, a particle swarm optimization (PSO) was used to optimize the SVM parameters that have been introduced for Melanoma detection. Considering the results provided by this doctoral research, several peer reviewed papers have been published, either in journal paper or conference papers. The full–reviewed conference papers are presented in conferences organized by IEEE, Engineering in Medicine & Biology and IEEE Computational intelligence Society's

Through this thesis it was discovered that there are some possible factors to improve the accuracy, feature extraction and sequential feature selection to choose the most relevant features to get higher skin cancer detection accuracy of detecting malignant melanoma by having a higher number of images for training of the SVM network. Future research aims to develop hybrid artificial intelligence techniques employing the following genetic algorithms; hybrid of fuzzy inference system (FIS), Ant Colony Optimization (ACO), PSO-ACO, Ontology, Contourlet Transform and others in multifunction pattern recognition systems

1.6. Structure and Summary of the Dissertation

This dissertation consists of 6 chapters including discussion and conclusions, appendixes and references. The contents are organized sequentially from the first chapter until the last chapter. The contents commence with the introduction of melanoma and finish with discussion and conclusion.

Chapter 2: This chapter Human Skin Biology and Cancer presents and discuss the structure and functions of the human tissues and cells. It takes into account the understanding of the study of histology and how microscopes function. It also includes the various staining procedures used to prepare tissues to make them visible through the microscope. Also discussed is the histology, the study of the microscopic anatomy of cells and tissues of humans, its role as an essential tool of biology and medicine. Histology is commonly performed by examining cells and tissues through sectioning, staining and examination under a light microscope or electron microscope. The study of changes in cell anatomy is known as histopathology. A sound knowledge of the normal structure of skin is essential for an understanding of pathology.

Chapter 3: This chapter presents an Image Processing and Automatic Skin Cancer Diagnosing relevant to melanoma detection. In this chapter, basics of human skin biology are described and different types of common skin lesions are briefly reviewed. The new imaging system which is called digital dermoscopy and different clinical diagnostic method are introduced. Also the two important dermoscopy structures, and pigment network are defined and explained. Lastly, the current computer-aided diagnostic systems have been reviewed and evaluated.

Chapter 4: This chapter presents pre-processing and segmentation to melanoma detection. It has been observed that dermoscopy images often contain objects and unrelated elements, which results in loss of accuracy as well as an increase in computational time. Thus, it requires some pre-processing steps to facilitate the segmentation process by using different types of filters to remove unwanted objects or artefacts. Segmentation is an important step in the medical image analysis and classification for radiological evaluation or computer–added diagnoses. Information is also provided about using the segmentation to partition the input images into a number of meaningful and non-overlapped regions based on some features that help in distinguishing between image regions (object and background).

Chapter 5: Over the past two decades Machine Learning has become one of the backbones of information technology, although usually hidden, part of our life. With the ever increasing amounts of data becoming available there is good reason to believe that smart data analysis will become even more universal as a necessary part for technological progress. The purpose of this chapter is to provide the person who reads with an overview over the massive range of applications which have at their heart a machine learning problem and to bring some degree of instruction to the farm of problems. After that, thesis will discuss some basic tools from statistics and probability theory, since many machine learning problems must be expressed to become open to solving. Finally, thesis will outline a set of equally basic effective algorithms to solve an important problem, namely that of classification. A discussion of more problems and a detailed analysis will follow in later chapters of the thesis.

This chapter includes a melanoma detection model using Gabor wavelet to interest point detection. This method displays the characteristics of certain cells in the visual cortex (the outer layer of an internal organ) of some creatures which can be approved by these filters. Discrete Wavelet Transform (DWT) can be used to decomposing the image into non-overlapping sub bands [26], the mean and standard deviation of the decomposed image portions are calculated [27]. Support Vector Machine (SVMs) as classifier directly determine the decision boundaries in the training step and the method can also provide good generalization in high-dimensional input spaces [28]. A hybrid technique of swarm-based support vector machine (SVM) is introduced for melanoma detection using the

pathology images as inputs. In this technique, a particle swarm optimization (PSO) is proposed to optimize the SVM to detect melanoma.

Chapter 6: These chapter used techniques were aiming to be able to provide recommendation for nonspecialized users. However the variations of diagnosis are adequacy large and there is a lack of detail of the test methods. One of the commercial products on the market has an accuracy rate of 92%. These products is a difficult system, taking high quality Epiluminescence Light Microscopy (ELM) images and using advanced image analysis techniques to extract a number of features for classification which it makes it unsuitable for a normal person to use. The traditional imaging is just a recording of what the human eye can see by digital camera while the Dermoscopy known as Epiluminescence Light Microscopy (ELM) images needs an experienced professional to get the required image. This experiment presents a melanoma detection model employing a hybrid of particle swarm optimization (PSO), wavelet Gabor (WLG), and SVM, to improve the performance of the SVM by finding the optimal SVM parameters. This chapter is organized in 6 Experiment Based on Digital and pathological images.

Chapter 7: This chapter presents an early detection include different stages of pre-processing, segmentation, feature extraction and selection, and classification. Since the output of each stage is the input of the next stage, all stages have an important role to escape misdiagnosis. The thesis outcome results are very interesting; it is one of the best results in this type of research as per table 6. 12 described sensitivity and specificity in chapter 6.6.1. results and discussion. This optimal parameter could perform well. Also this chapter include discussion and conclusion.

1.7. Publications Presented During the Doctoral Research

The doctoral research publication presented in journals and conferences are as follows:

Journal paper:

(i) "Wavelet and Curvelet Analysis for Automatic Identification of Melanoma Based on Neural Network Classification", (2013) in International Journal of Computer Information Systems and Industrial Management (IJCISIM), Volume 5 (2013) pp. 606-614 [29].

Conference papers:

(i) "A Hybrid System for Skin Lesion Detection: Based on Gabor Wavelet and Support Vector Machine" The model is submitted to (CISP_BMEI 2014): 7th International Congress on Image and Signal Processing, 7th International Conference on Biomedical Engineering and Informatics, Dalian, China 14-16 October 2014 [15].

(ii) "Automatic Recognition of Melanoma Using Support Vector Machines: A Study Based on Wavelet, Curvelet and Colour Features". The model is presented in (IAICT 2014): International Conference on Industrial Automation, Information and Communications Technology, Bali, Indonesia, Aug 28-30, 2014 [30].

(iii) "Novel feature extraction methodology based on histopathalogical images and subsequent classification by Support Vector Machine". The model is presented by Md Khaled et al. (2014), in (ICCVIA2014) International Conference on Computer Vision & Image Analysis, Ras Al Khaimah, UAE 25-27 March 2014 [31].

(iv) "Classification of Malignant Melanoma and Benign Nevi from Skin Lesions Based on Support Vector Machine" The model is presented by Md Khaled et al. (2013), in Fifth International Conference on Computational Intelligence, Modelling and Simulation, Seoul, South Korea, 24-26 September [32].

(v) "The Automatic Identification of Melanoma by Wavelet and Curvelet Analysis: Study Based on Neural Network Classification". This model is presented by Md Khaled et al. (2011), in 11th International Conference on Hybrid Intelligent Systems (HIS), December 5-8, Malacca, Malaysia [33].

(vi) "Segmentation of skin cancer images based on gradient vector flow (GVF) Snake". The model is presented by Md Khaled et al. (2011), International Conference on Mechatronics and Automation, August 7 - 10, Beijing, China [34].

Chapter Summary

This chapter introduction to skin cancer (melanoma) includes: research aims; problem description; thesis objectives; the methodology used in the thesis; contribution of the doctoral research; structure and summary of the dissertation; followed by seven doctoral research publications one presented in journal and six presented in conferences.

CHAPTER 2

Human Skin Biology and Cancer.

2.1. Human Skin Biology

This chapter will discuss the structure and functions of the human tissues and cells. It takes into account the understanding of the study of histology and how microscopes function. It also includes the various staining procedures used to prepare tissues to make them visible through the microscope[35].

Histology is the study of the microscopic anatomy of cells and tissues of humans, plants and animals. It is commonly performed by examining cells and tissues through sectioning, staining and examination under a light microscope or electron microscope. Histology is an essential tool of biology and medicine [36]. The study of changes in cell anatomy is known as histopathology. A sound knowledge of the normal structure of skin is essential for an understanding of pathology. (more detailed information in appendix A)

2.2. The Cell (Structure and Function)

The cell is the functional unit of all living organisms. The simplest living organisms are unicellular, consisting of a single cell such as bacteria and algae. More complex organisms known as Eukaryotes have membranes surrounding their intracellular organelles. Eukaryotes consist of a defined nucleus and an extracellular matrix. Multicellular organisms, such as humans, display a large diversity in functional and morphological specialization. Specialization of cell structure and function is known as cellular differentiation [37].

Human cells mainly consist of a nucleus and cytoplasm. The cytoplasm contains a number of organelles of which each has a defined function. The nucleus may be considered as being the largest organelle. Figure 2.1 shows cell structure.

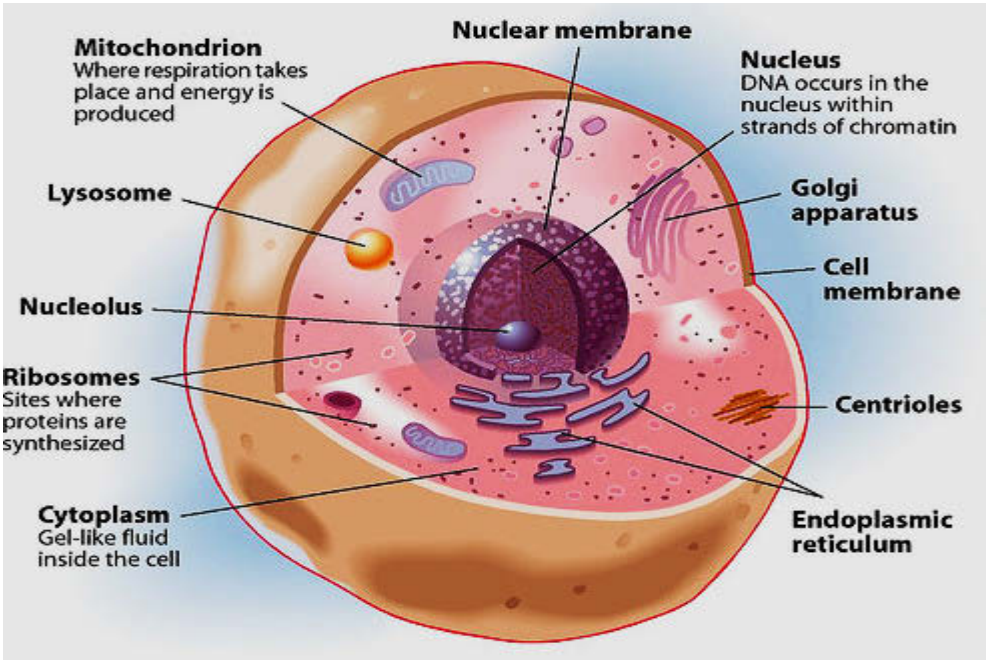


Figure 2.1 shows, Cells structure; source: www.getting-in.com

2.2.1. Cell structure

Nearly all animal and plant cells contain the following parts:

- **The nucleus:** controls the activities that occur in the cell.
- **Cytoplasm:** the substance in which most chemical reactions take place**.**
- **Cell Membrane:** the barrier which holds the cell parts together and controls what enters and leaves the cell.
- **Mitochondria:** where the majority of energy is produced through respiration.
- **Ribosomes:** where protein synthesis takes place.

2.2.2. Cell Cycle and Replication:

Cellular replication is a vital process for growth, repair and development of unicellular organisms which develop into more complex multicellular organisms. In eukaryotes the cell cycle [38], as shown in Figure 2.2, is composed of four separate stages; G1, S, G2 and M. The two (G) stages are the gap stages where the cells conduct a series of checks prior to proceeding to the next stage. The synthesis (S) stage is where DNA replication occurs; these three stages combined form a period known as interphase, where the cell grows and produces proteins necessary for the cell division. The cell subsequently advances to the mitosis (M) stage, where a cell undergoes coordinated nuclear division resulting in the formation of two genetically identical daughter cells [39] [40].

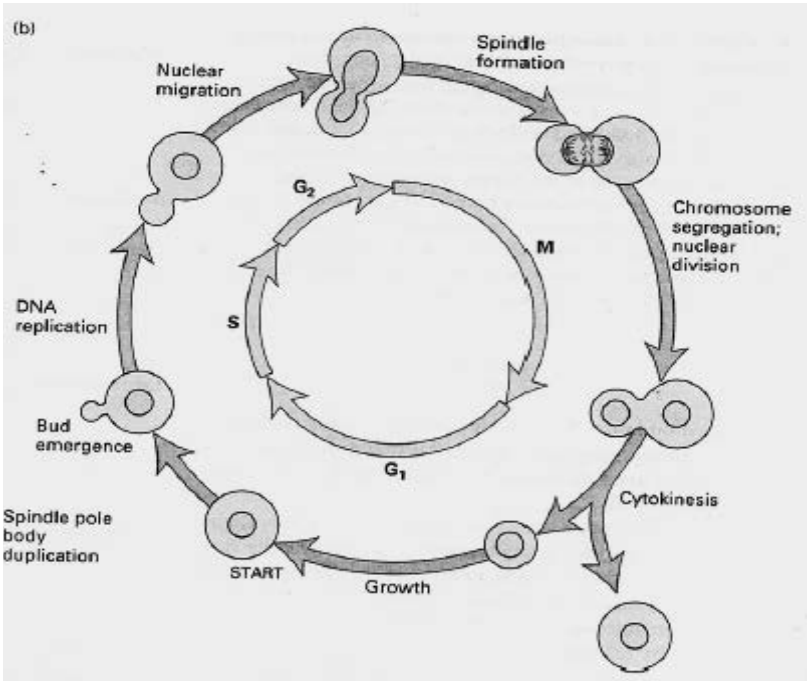


Figure 2.2 The eukaryotic cell cycle, G Phases: growing, S phase: synthesis/ DNA replication, G2 Phase: Growing and preparing for Mitosis, M mitosis.

The cellular progression of the life cycle from one stage to the next is reliant upon passing through a variety of checkpoints. Each checkpoint consists of specialized proteins to determine if the necessary conditions exist within a cell for its progress through the cycle. If the cellular conditions are inadequate when passing through the checkpoints, progression will cease and the cell will remain in a dormant state known as the (G0) stage as seen in Figure 2.3. Checkpoint errors may have significantly destructive consequences on cellular division including cell death or mutations causing unrestrained growth i.e. cancer [41].

The human body is made up of cells, for example, when you have sunburn and your skin peels, you are shedding skin "cells." In the center of each cell is an area called the nucleus, and human chromosomes are located inside the nucleus of the cell. A chromosome is a structure that contains your genes that determine your traits, such as eye color and blood type[42].

The usual number of chromosomes inside every human's cell is 46 total chromosomes, or 23 pairs. One half of the chromosomes is inherited (one member of each pair) from the biological mother and the other half (the matching member of each pair) from the biological father.

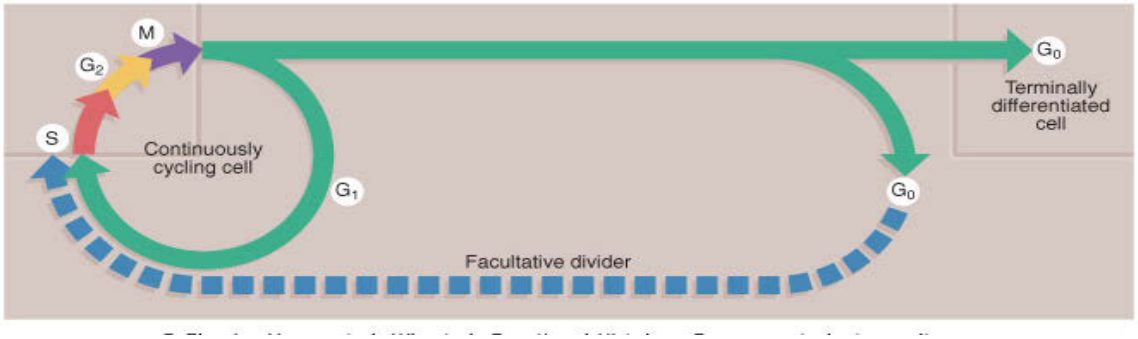


Figure 2.3 The cell cycle; G0 terminally differentiated cell G1 gap phase 1 G2 gap phase 2 M mitotic phase S synthesis phase

Source: Wheatear's Functional Histology: A Text and Colour Atlas, 5th Ed

2.3. Haematoxylin and Eosin (H & E)

Haematoxylin and Eosin is the most popular staining techniques in histology and medical diagnosis laboratories. Haematoxylin is the basic dye component which stains nucleic acids blue/purple, including nuclei, ribosomes and endoplasmic reticulum. Due to their high content of DNA and RNA they have a high affinity for the dye. In contrast, Eosin is

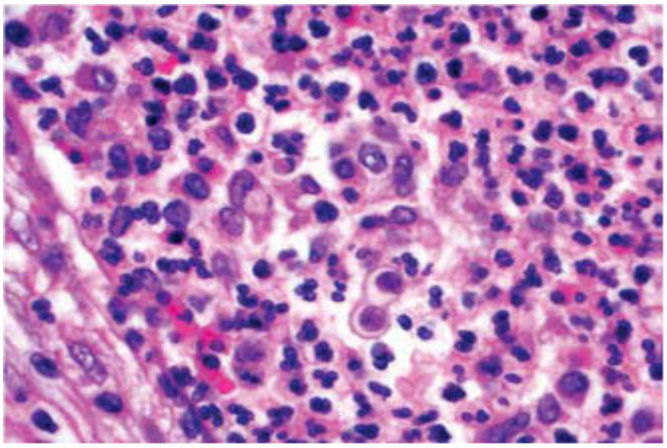


Figure 2.4 H&E-stained section of skin. High magnification (40xs) view. Haematoxylin stains the nuclei of cells blue to bluish-purple, and eosin stains other cellular elements in the tissues from pink to red [43].

the acidic dye component which stains basic structures red/pink, such as basic cytoplasmic proteins within the cytoplasm as shown in Figure 2.4.

2.4. Skin Cell Structure

The skin is the largest organ in the human body and is a highly organized structure consisting of three main layers, called the epidermis, the dermis and the hypodermis as

displayed in figure 2.5. The skin has three main functions represented in the protection, sensation and thermoregulation functions that are explained next.

2.4.1. Protection

One of the most important functions of the skin is its action as a barrier, providing protection against a great variety of external forces. It protects against mechanical impact and pressure e.g. friction, frequently occurring on the soles of feet and the palms of the hands, where skin structure is slightly modified and thickened to resist such friction. It also provides protection and acts as a barrier to water, chemicals and invasion by various microorganisms. An example of this is microorganisms such as bacteria and fungi can live on the skin but unless the skin is damaged, cannot penetrate into underlying tissue.

2.4.2. Sensation

The skin is the largest sensory organ in the body. It consists of a wide range of different receptors and nerve cells to detect and relay changes in the external environment and thus to fulfill different roles such as for touch, temperature regulation, pain and pressure. Interestingly due to regional variation in structure, skin which is most in contact with its environment contains the most receptors. Damage to nerve cells is known as neuropathy causing loss of sensation in affected areas.

2.4.3. Thermoregulation

The skin can act as an organ regulating the effects of several aspects of physiology. Body temperature can be regulated via adipose tissue providing insulation, and sweat glands secreting sweat as illustrated in Figure 2.5. The subsequent evaporation of sweat has a cooling effect. Which are controlled by the hair erector muscles which are attached to each individual hair follicle. Furthermore peripheral circulation responds to changes in flow of heat and responds to certain requirement e.g. to cool the body. Heat can be lost by increasing blood flow through the vascular networks on the surface of the skin onto the skin surface.

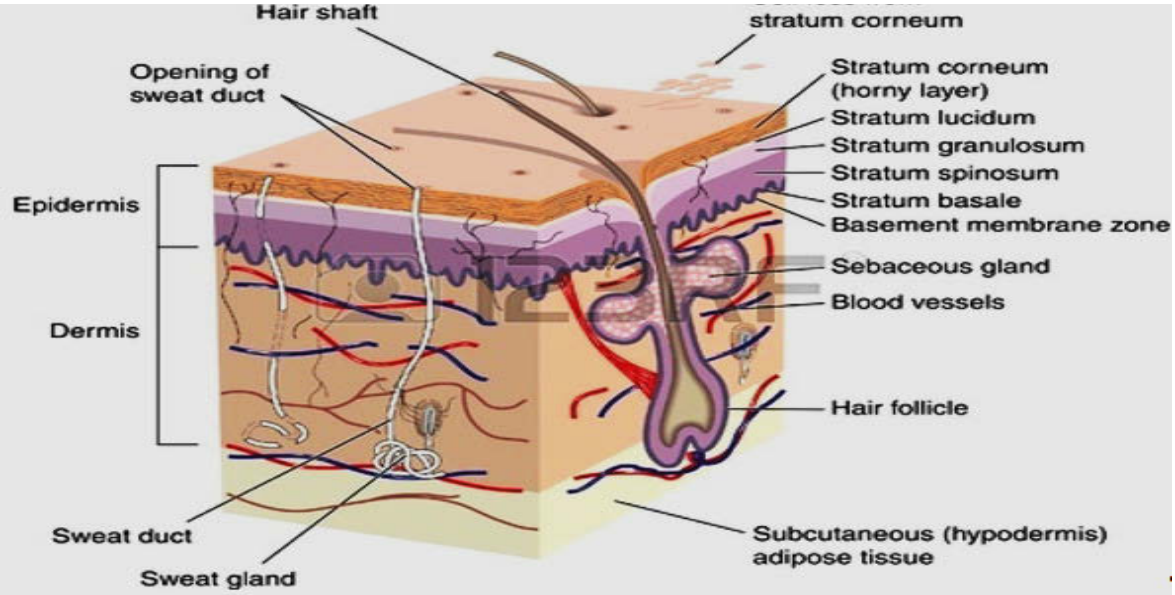


Figure 2.5 Displayed Cross Section & the main layers of Human Skin

Source: nurrashidah2204.blogspot.com

2.5. The Layers of the Skin

The skin cell structure has three main layers as shown in figure 2.5[44]. These layers are Epidermis, Dermis, and Hypodermis and are discussed below.

2.5.1. Epidermis

The external surface of the skin is composed of a renewable stratified squamous epithelium with a surface layer of keratin. Keratin is a protein in direct contact with the external environment.

2.5.2. Dermis

The Dermis is composed of fibro-collagenous and elastic tissue; this layer consists of blood vessels, lymphatic vessels, sensory receptors, nerves, touch and heat mechanoreceptors, sweat glands, hair follicles and apocrine glands.

2.5.3. Subcutaneous/hypodermis

The hypodermis is the deepest layer of skin and is predominantly composed of adipose tissue. This layer contains the larger vessels which supply and drain the dermis.

2.6. Melanocyte

A melanocyte cell produces the protective skin-darkening pigment melanin. Birds and mammals possess these pigment cells, which are found mainly in the epidermis although they also exist elsewhere as for example in the matrix of the hair. Melanocytes could be branched or dendritic and their dendrites are used to transfer pigment granules to adjacent epidermal cells.

2.7. Physical Properties of Human Skin

Figure 2.6 shows how the human skin color is quite variable around the world. It ranges from a very dark brown among some Africans, Australian Aborigines, and Melanesians to a near yellowish pink among some Northern Europeans. There are no people who actually have true black, white, red, or yellow skin. These are commonly used color terms that do not reflect biological reality[45] .

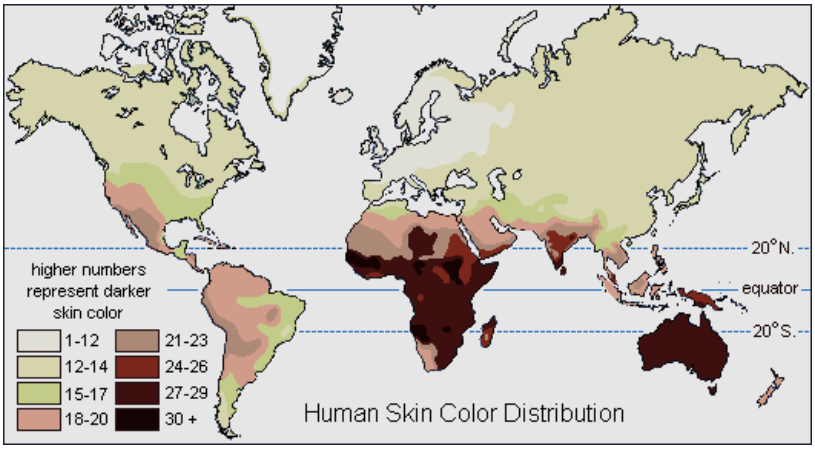
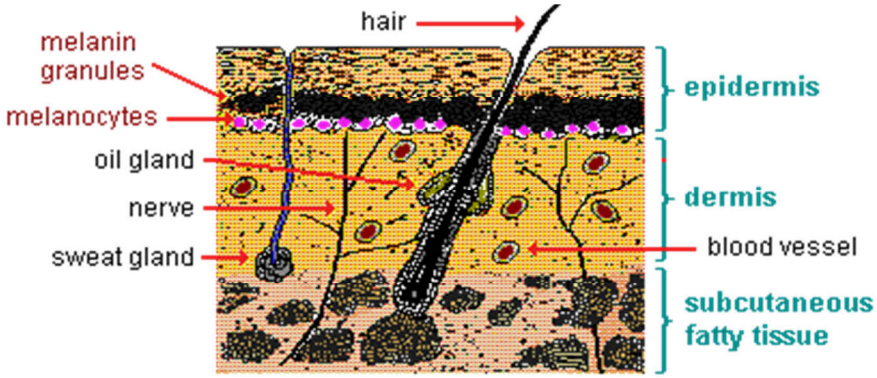


Figure 2.6. Skin Color Distribution around the World

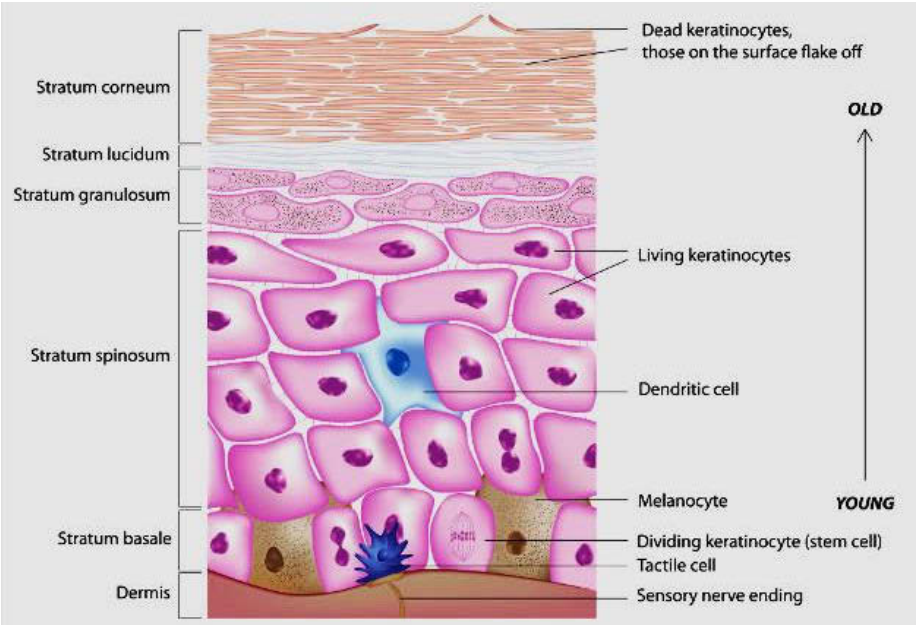
Source: www.gbhealthwatch.com

As displayed in figure 2.6, the skin color is due primarily to the presence of a pigment called melanin, which is controlled by at least 6 genes. The number and size of melanin particles differs in each individual. These two variables are most important in determining skin color in comparison to the percentages of the different kinds of melanin. In lighter skin, color is also affected by red cells in blood flowing close to the skin. The presence of fat under the skin or carotene can also be contributing factors. Also, hair color is another factor controlled by the presence of melanin[45].

As shown in figure 2.7(b), melanin is normally located in the epidermis, or outer skin layer. These melanocyte cells have photosensitive receptors, similar to those in the eye, that detect ultraviolet radiation[46] from the sun and other sources. In response, they produce melanin within a few hours of exposure. Melanin acts as a protective biological shield against ultraviolet radiation which helps prevent sunburn damage which may cause DNA changes and, subsequently, several kinds of malignant skin cancers. Ultraviolet radiation reaching the earth usually increases in summer and decreases in winter. The skin's ability to tan in summertime is adjusted to this seasonal change. Tanning is



a) Cross section of human skin,
Source: anthro.palomar.edu



b) Structure of Epidermis
Source: silver-botanicals.com

Figure 2.7. a) Cross section of human skin, b) Structure of Epidermis.

primarily an increase in the number and size of melanin granules due to the stimulation of ultraviolet radiation.

It would be harmful if melanin acted as a complete shield. A certain amount of shortwave ultraviolet radiation (UVR) [47] must penetrate the outer skin layer in order for the body to produce vitamin D. Approximately 90% of this vitamin in people normally is synthesized in their skin and the kidneys from a cholesterol-like precursor chemical with the help of ultraviolet radiation. The remaining 10% comes from foods such as fatty fish and egg yolks.

2.8. Human Skin

The skin is the largest organ of the body. It provides protection against heat, sunlight, injury, and infection. Skin also helps in controlling body temperature and stores water, fat, and vitamin D.

2.8.1. Types of Skin

There are two main kinds of human skin:

- One is Glabrous skin (non-hairy skin), found on the palms and soles; it is grooved on its surface by continuously alternating ridges and sulci, which has individually unique configurations known as dermatoglyphics. It is characterized by a thick epidermis which is divided into several well-marked layers, including a compact stratum cornea, by the presence of encapsulated sense organs within the dermis, and by a lack of hair follicles and sebaceous glands.
- "The second type is the hair-bearing skin which has both hair follicles and sebaceous glands but lacks encapsulated sense organs. There is also wide variation in skin types between different body sites.

2.8.2. Layers of Human Skin

Skin is mainly composed of three primary layers as displayed on Figure 2.7(a)

- ***The epidermis***, which provides waterproofing and serves as a barrier to infection [48].
- ***The dermis***, which serves as a location for the appendages of skin

- ***The hypodermis*** (subcutaneous adipose layer)

2.8.3. Cancer Types

Cancer is a disease of the body's cells that in the end destroys living organs, which can lead to death. Cancer is the term used to describe a lot of different diseases including malignant tumors, leukemia (a disorder of white blood cells), sarcoma of the bones, Hodgkin's disease and non-Hodgkin's lymphoma (affecting the lymph nodes) in which uncontrolled cell growth threatens the rest of the body. There are more than 100 different types of cancer, but the most common five types are: ***bowel cancer, breast, prostate, melanoma and lung cancer*** (excluding non-melanoma skin cancer) and these form 60% of all cases. In Australia, the most common cancers in men are prostate, colorectal, lung cancers and melanoma, but in women the most common is breast cancer, colorectal cancer, melanoma and lung cancer.

2.8.4. Types of Skin Cancer

Like all body tissues, the skin is made of tiny building blocks called cells. These cells can sometimes become cancerous, from exposure to ultraviolet (UV) radiation[49]. The top layer of skin contains three different types of cells that can be affected: *basal cell cancer, squamous cell cancer and melanoma,* as shown in figure 2.8.

Australia has the highest incidence of skin cancer in the world,[49-51] [52] [53] and non-melanoma skin cancer (NMSC) is the most commonly diagnosed cancer in Australia[54].

Skin cancer in Australia's most expensive cancer, with more than $264 million (9% of the total costs for cancer) spent on diagnosis and treatment in 2000–01[55].

In 2009 the review article[56] concludes that: In the Australian patient population, the whole body skin examination is essential to avoid missing concurrent lesions. Patients with concurrent lesions have a high risk of developing future NMSC

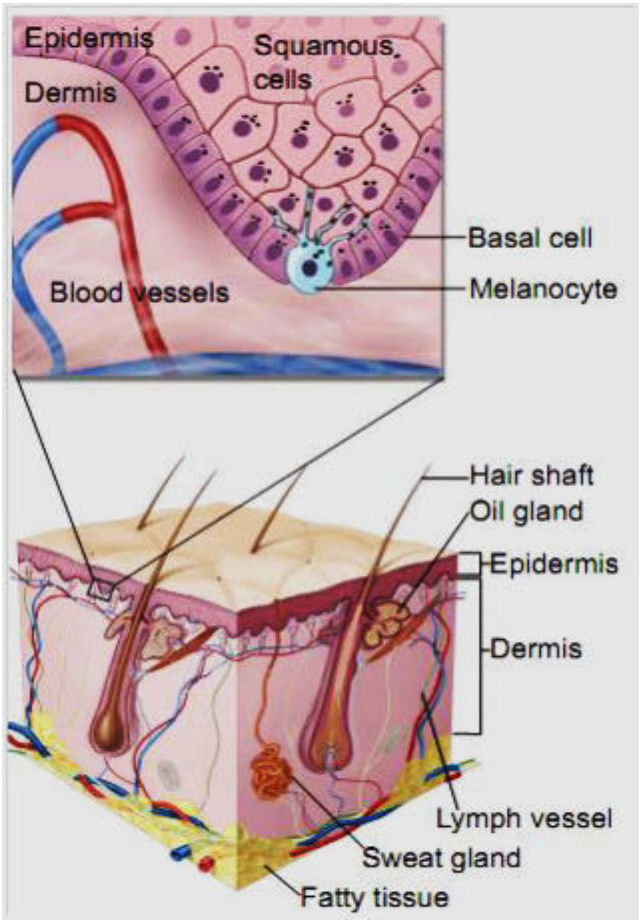


Figure 2.8. The Layers of Skin - the Epidermis and Dermis (At the top, the close up shows a squamous cell, basal cell, and melanocyte). Source: www.uchospitals.edu

.

Figure 2.9(a), shows the most common and least dangerous skin cancer Basal Cell Carcinoma (BCC). It appears as a lump or scaly area and can be red, pale or pearly in color. It grows slowly – usually on the head, neck or upper torso – and can become ulcerated as it grows.

Figure2.9- a) Basal Cell Carcinoma (BCC), b) Squamous Cell Carcinoma (SCC)

Figure 2.9(b), shows the Squamous Cell Carcinoma (SCC) cancer; it grows over a period of weeks or months and may spread to other parts of the body if not treated promptly. It

occurs most often (but not only) on areas exposed to the sun. This can include the head, neck, hands and forearms. This cancer looks like thickened, red, scaly spots.

2.8.4.1. Melanoma

Melanoma is the most dangerous type of skin cancer; it results from sun damage and other causes. A skin biopsy can identify melanoma [57] Melanoma grows quickly and develops over weeks to months. If caught early, it is usually curable, but if it spreads to other parts of the body, it can be very difficult to cure. Melanoma appears as a new spot or as an existing spot, freckle or mole that changes color, size or shape. It usually has an irregular, smudgy outline and is often more than one color.

Melanoma as shown in Figure 2.10 can occur anywhere on the skin, even on the soles of the feet. Melanocytes in the eye, respiratory system, nervous system, even on the cortex of the brain and mucous membranes (e.g. lining of the mouth and nasal passages) can also become cancerous. These types of melanoma are rare; 90% of the cases appear on skin.

Figure 2.10 – Melanoma

Melanoma is malignant tumor of melanocytes [58]. Melanocytes produce the dark pigment, melanin, which is responsible for the color of skin. These cells predominantly occur in skin, but are also found in other parts of the body, including the bowel and the eye. Melanoma can originate in any part of the body that contains melanocytes. It causes the majority (75%) of deaths related to skin cancer[59]. The majority of malignant melanomas are caused by heavy sun exposure in white-skinned populations [60]. Incidence rates are highest by far in Australia/New Zealand (37 per 100,000 in 2008), where it is the third most common cancer in both males and females, accounting for one in nine (around 11% in 2008) of the total cases [61]as displayed in figure 2.11.

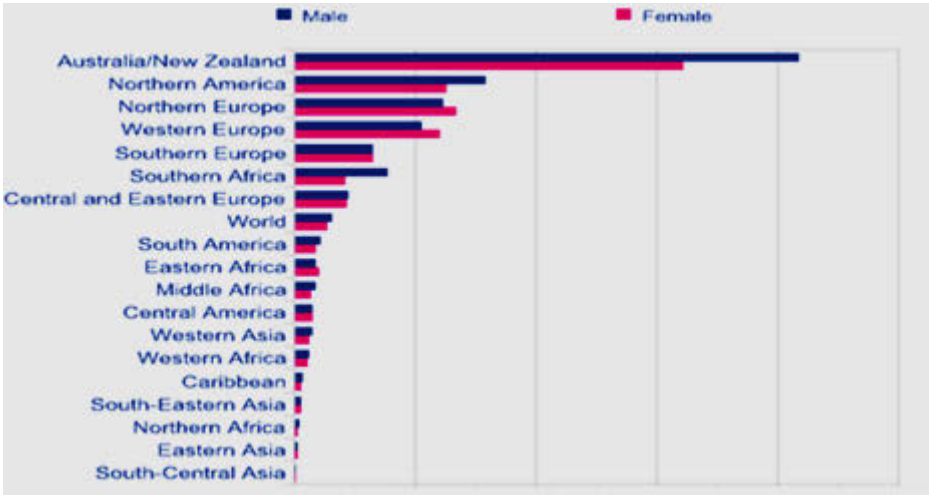


Figure 2.11: Malignant Melanoma, World Age-Standardised Incidence Rates, World Regions, 2008 Estimates.

At later stages, the mole may itch, ulcerate or bleed[56]. Anyone who has more than 100 moles is at greater risk for melanoma. It is so important to get to know your skin very well and to recognize any changes in the moles on your body. Have a check-up with the physician, and then contact the specialist for early diagnosis[56].

All cancers are caused by damage to the DNA inside cells. This damage can be inherited in the form of genetic mutations, but in most cases, it builds up over a person's lifetime and is caused by factors in their environment. DNA damage causes the cell to grow out of control, leading to a tumor. Melanoma is usually caused by damage from UV light from the sun, but UV light from sunbeds can also contribute to the disease[62].

A number of rare mutations, which often run in families, are known to greatly increase one's susceptibility to melanoma. Several different genes have been identified as increasing the risk of developing melanoma. Some rare genes have a relatively high risk of causing melanoma; some more common genes, such as a gene called "MC1R" that causes red hair, have a relatively lower elevated risk. Genetic testing can be used to determine whether a person has one of the currently known mutations.

The earliest stage of melanoma starts when the melanocytes begin to grow out of control. Melanocytes are found between the outer layer of the skin (the epidermis) and the next layer (the dermis). This early stage of the disease is called the radial growth phase, and the tumor is less than 1 mm thick. Because the cancer cells have not yet reached the blood vessels lower down in the skin, it is very unlikely that this early-stage cancer will spread to other parts of the body. If the melanoma is detected at this stage, then it can usually be

completely removed with surgery. When the tumor cells start to move in a different direction — vertically into the epidermis and into the papillary dermis — the behavior of the cells changes dramatically[63].

Visual diagnosis of melanomas is still the most common method employed by health professionals[64]. Moles that are irregular in color or shape are often treated as candidates for melanoma. The diagnosis of melanoma requires experience, as early stages may look identical to harmless moles or not have any color at all. People with a personal or family history of skin cancer or of dysplastic nevus syndrome (multiple atypical moles) should see a dermatologist at least once a year to be sure they are not developing melanoma. There is no blood test for detecting melanomas.

2.8.4.2. Melanoma types:

Melanoma is divided into the following types [65] [66]:

- Lentigomaligna melanoma: Begins as a large freckle in an area of skin that gets a lot of sun exposure, such as the face and upper body as displayed in figure 2.12. It may grow slowly and superficially over many years, later forming lumps as it grows deeper into the skin.

Figure 2.12. Lentigomaligna Melanoma.

Source: melanomaknowmore.com

- Superficial spreading melanoma: Forms 70% of the cases, it grows in the top layer (epidermis). It is dangerous when it invades into the lower layer (dermis) and it is dark brown or black in color as displayed in figure 2.13.

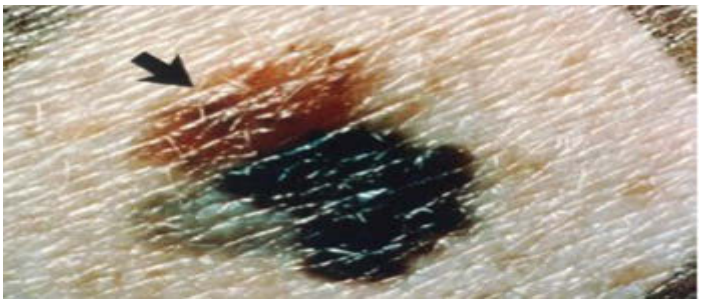


Figure 2.13. Superficial Spreading Melanoma.

Source: melanomaknowmore.com

- Acral lentiginous melanoma: Most commonly found on the palms of the hands and soles of the feet or under the nails. More common in people with darker skin.
- Mucosal melanoma.
- Nodular melanoma: Forms 15% of the cases, it is very dark brownish black or black in color but can be pink or red and forms a lump on the skin surface; also it invades deeper into the skin as displayed in figure 2.14.

Figure 2.14. Nodular Melanoma.

Source: melanomaknowmore.com

- Polypoid melanoma.
- Desmoplastic melanoma.
- Amelanotic melanoma.
- Soft-tissue melanoma.
- Unclassified: up to 5% of cases.

2.8.4.3. Staging

Also of importance are the "Clark level" and "Bristow's depth", which refer to the microscopic depth of tumor invasion[67].

Figure 2.15 represents melanoma stages [68] and the survival rates in 5 year:

Stage 0: Melanoma in situ (Clark Level I), 99.9% survival

Stage I / II: Invasive melanoma, 89–95% survival.

Stage II: High risk melanoma, 45–79% survival

Stage III: Regional metastasis, 24–70% survival

Stage IV: Distant metastasis, 7–19% survival

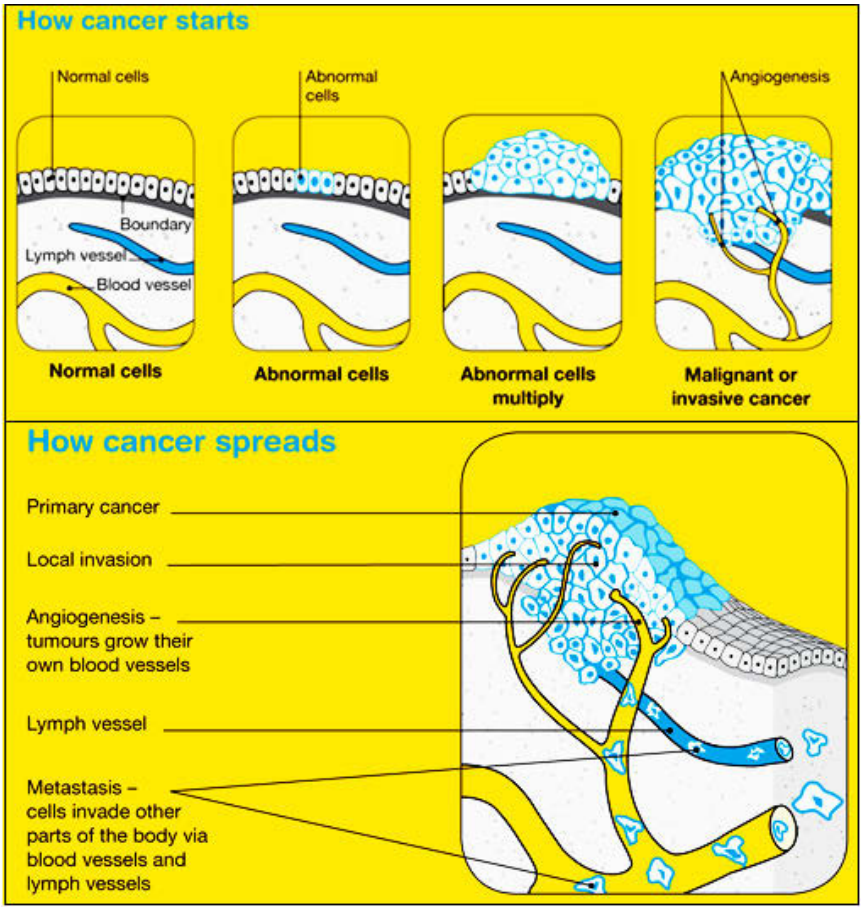


Figure 2.15.Stages of Cancer Development and Metastasis.

Source: http://www.cancervic.org.au/about-cancer/advanced-cancer

2.8.4.4. Clark's level

Clark's level is a related staging system, used in conjunction with Bristow's depth, which describes the level of anatomical invasion of the melanoma in the skin[69].

Five anatomical levels are recognized, and higher levels have worsening prognostic implications. These levels are:

- **Level 1 :** Melanoma confined to the epidermis (melanoma in situ)
- **Level 2 :** Invasion into the papillary dermis
- **Level 3 :** Invasion to the junction of the papillary and reticular dermis
- **Level 4 :** Invasion into the reticular dermis
- **Level 5 :** Invasion into the subcutaneous fat[69].

2.8.4.5. Bristow's depth

In medicine, Breslow's depth was used as a prognostic factor in melanoma of the skin. It is a description of how deeply tumor cells have invaded. Currently, the standard Breslow's depth has been replaced by the AJCC depth. As shown in table 2.1, Breslow's depth was originally divided into 5 stages[70].

Table 2.1: Breslow's depth

Stage	Depth
Stage I	less or equal to 0.75mm
Stage II	0.75 mm - 1.5mm
Stage III	1.51 mm - 2.25mm
Stage IV	2.25 mm - 3.0mm
Stage V	greater than 3.0 mm

The above studies showed that depth was a continuous variable correlating with prognosis. However, for staging purposes table 2.2 shows the most recent AJCC guidelines use cutoffs of 1 mm, 2 mm, and 4 mm to divide patients into stages.

Table 2.2: Survival figures from British Association of Dermatologist Guidelines 2002

Tumor Depth	Approximate 5 year survival
<1 mm	95-100%
1 - 2 mm	80-96%
2.1 - 4 mm	60-75%
>4 mm	50%

2.8.5. Health Behavior of Cancer Survivors

Advances in the early detection and treatment of cancer has resulted in a dramatic increase in the number of cancer survivors,[5] and the numbers are expected to grow considerably [71-75]. More than 60% of cancer patients will survive more than five years after diagnosis. Also the survival rate for many common cancers has increased by more than 30 % in the past two decades [75]. The chart on table 3.3, shows incidence (2009), and morality (2010) of melanoma of the skin of Australia [76]. The chart on table 3.4, shows observed incidence (2009) and morality (2010) of non-melanoma of the skin and estimated for 2012 in Australia.

It was estimated that more than 120,700 Australians were diagnosed with cancer in the year 2012, excluding basal and squamous cell carcinoma of the skin[77]. Males made up more than half (56%) of these cases. In 2012 the most commonly reported cancers were prostate cancer, bowel cancer, breast cancer, melanoma of the skin and lung cancer.

Cancer accounted for about 3 in 10 deaths in Australia in 2010, were the death rate was more than 42,800. For all cancers combined, the age-standardised mortality rate decreased by 17% from 210 per 100,000 in 1991 to 174 per 100,000 in 2010[77]. As indicated in Figure 2.16 five-year relative survivals from selected cancers by remoteness area, Australia. 2006-2010[77].

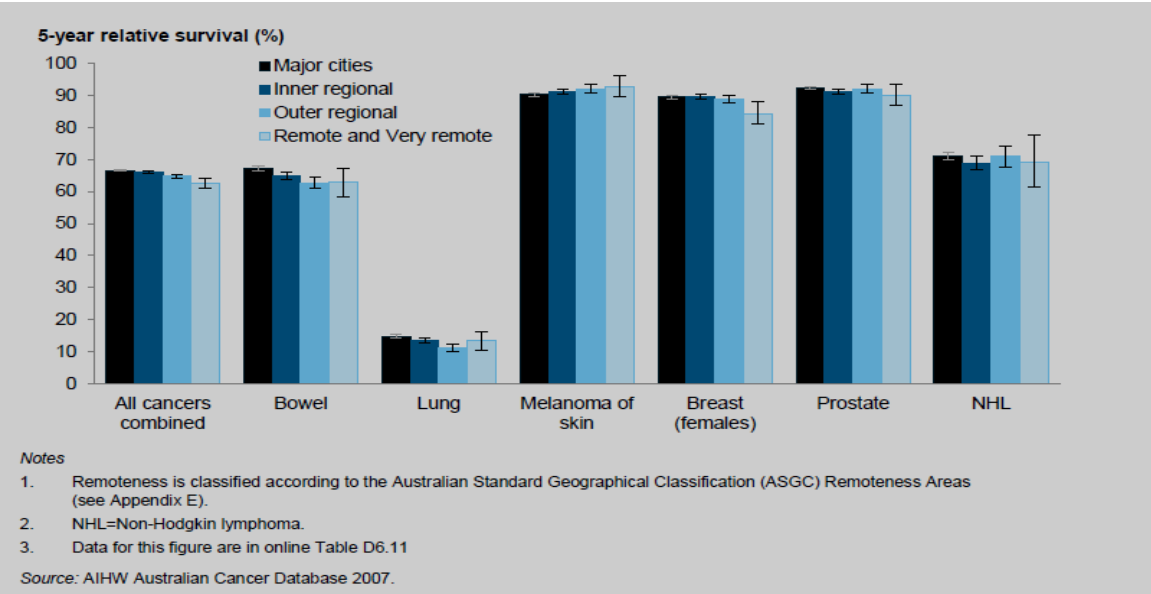


Figure 2.16: Five-year relative survival from selected cancers by remoteness area, Australia. 2006-2010.

It was found that cancer outcomes were different when it came to Aboriginal and Torres Strait Islander status, remoteness area and socioeconomic status. Indigenous Australians experienced higher incidence and mortality rates than non-Indigenous Australians for all cancers combined. Incidence rates and survival were lower for people living in remote areas compared with those in major cities, while mortality rates rose with increasing remoteness. Incidence and mortality rates rose and survival from all cancers fell as a person's socioeconomic status decreased[77] . Table 2.3: displays the incidence (2009) and morality (2010) of melanoma of the skin and estimated for 2012. Also Table 2.4: displays the incidence (2009) and morality (2010) of non-melanoma of the skin and estimated for 2012.

Chapter Summary

This chapter 'Human Skin Biology and Cancer' includes basics of human skin biology; the cell structure and function; hematoxylin and eosin. This chapter also describes staining techniques in histology and medical diagnosis laboratories. In the subsequent sections of this chapter also includes the layers of the skin; melanocyte cell; physical properties of human skin; human skin structure and cancer types like melanoma staging as well as non-melanoma; the chapter also includes study about health behavior of cancer survivors in Australia.

Melanoma of the skin (C43)

Risk factors[a]:

Table B17: Observed incidence (2009) and mortality (2010) of melanoma of the skin, and estimated for 2012

	Incidence			Mortality		
	Males	Females	Persons	Males	Females	Persons
Number						
2009 incidence/2010 mortality[b]	6,764	4,781	11,545	993	459	1,452
2012 (estimated)[c]	7,440	5,070	12,510	1,070	495	1,560
Age-standardised rate[d]						
2009 incidence/2010 mortality	61.7	40.0	49.8	8.9	3.5	5.9
95% CI	*60.3–63.2*	*38.8–41.1*	*48.9–50.8*	*8.4–9.5*	*3.2–3.8*	*5.6–6.3*
2012 (estimated)[c]	62.7	39.9	50.3	9.1	3.6	6.1
Other information for 2009 incidence/2010 mortality						
Per cent of all cancer	10.5	9.6	10.1	4.1	2.5	3.4
Risk to age 75	1 in 22	1 in 33	1 in 27	1 in 181	1 in 410	1 in 253
Risk to age 85	1 in 14	1 in 23	1 in 17	1 in 85	1 in 226	1 in 128
Mean age[e]	62.3	58.8	60.8	69.5	69.0	69.4

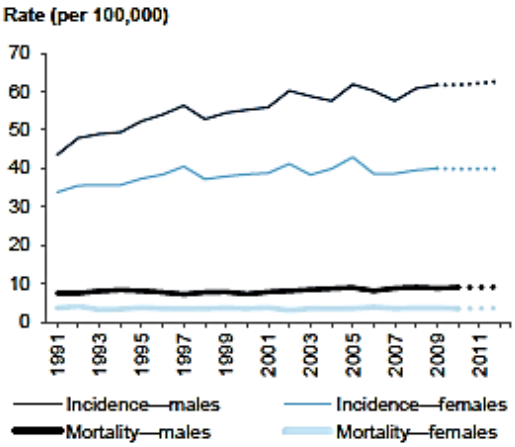


Figure B17.1: Melanoma of the skin incidence and mortality rates[d], 1991–2012

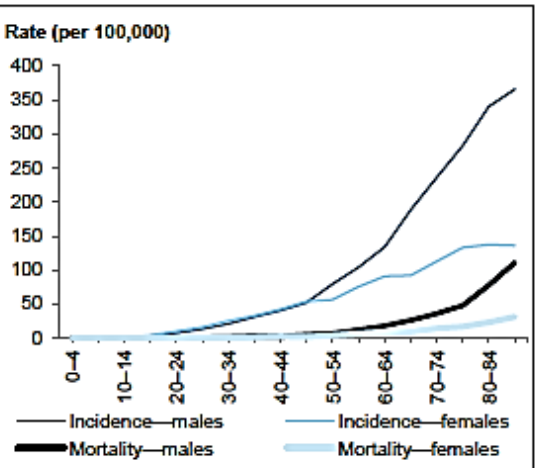


Figure B17.2: Melanoma of the skin incidence (2009) and mortality (2010) rates[f] by age group

(a) Based on the IARC (2008) and WCRF & AICR (2007).

(b) 2009 incidence data include estimates for NSW and the ACT. See Appendix F for more details. Mortality data for 2009 and 2010 are revised and preliminary, respectively, and are subject to further revision.

(c) 2010–2012 estimates for incidence are based on 2000–2009 incidence data. 2011–2012 estimates for mortality are based on 2001–2010 mortality data (see Appendix G). They are rounded to the nearest 10. For estimates less than 1,000 they are rounded to nearest 5. The estimates for males and females may not add to estimates for persons due to rounding. Estimates are displayed on graph as dotted lines.

(d) The rates were age-standardised to the Australian population as at 30 June 2001 and are expressed per 100,000 population.

(e) Mean age for 2009 incidence was calculated excluding NSW and the ACT.

(f) The rates shown are age-specific rates.

Source: AIHW Australian Cancer Database 2009, AIHW National Mortality Database.

Table 2.3: Observed incidence (2009), and morality (2010) of melanoma of the skin and estimated for 2012.

Source: AIHW Australian Cancer Database 2009, AIHW National Mortality Database

Non-melanoma skin cancer (C44)[a]

Risk factors[b]:

Table B24: Observed incidence (2009) and mortality (2010) of non-melanoma skin cancer, and estimated for 2012

	Incidence			Mortality		
	Males	Females	Persons	Males	Females	Persons
Number						
2009 incidence/2010 mortality[c]	469	313	782	304	141	445
2012 (estimated)[d]	490	325	815	350	155	505
Age-standardised rate[e]						
2009 incidence/2010 mortality	4.4	2.4	3.3	2.8	0.9	1.8
95% CI	*4.0–4.9*	*2.1–2.7*	*3.1–3.5*	*2.5–3.2*	*0.8–1.1*	*1.6–1.9*
2012 (estimated)[d]	4.2	2.3	3.1	3.0	1.0	1.9
Other information for 2009 incidence/2010 mortality						
Per cent of all cancer	0.7	0.6	0.7	1.2	0.8	1.0
Risk to age 75	1 in 406	1 in 797	1 in 542	1 in 930	1 in 3,185	1 in 1,449
Risk to age 85	1 in 162	1 in 341	1 in 227	1 in 289	1 in 1,034	1 in 474
Mean age[f]	70.9	70.6	70.8	75.8	81.9	77.8

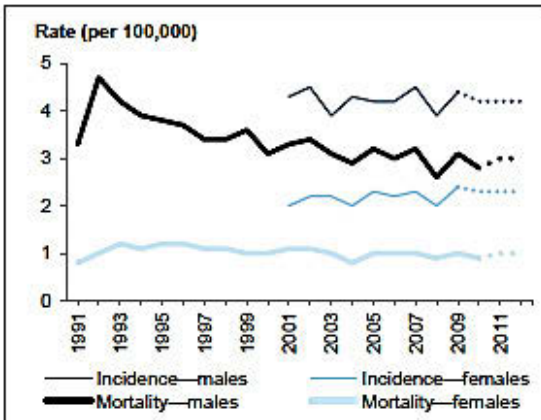


Figure B24.1: Non-melanoma skin cancer incidence and mortality rates[e], 1991–2012

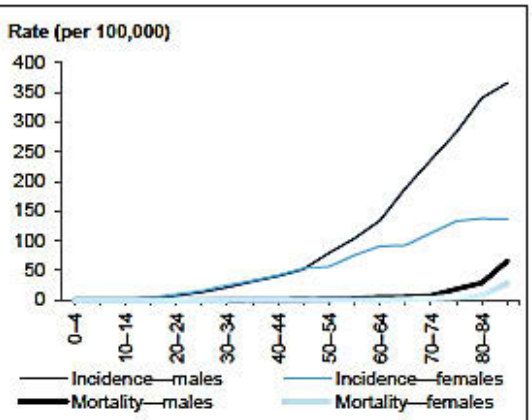


Figure B24.2: Non-melanoma skin cancer incidence (2009) and mortality (2010) rates[g] by age group

(a) For incidence data, those C44 codes that indicate basal or squamous cell carcinoma of the skin are not included.

(b) Based on the IARC (2008) and WCRF & AICR (2007).

(c) 2009 incidence data include estimates for NSW and the ACT. See Appendix F for more details. Mortality data for 2009 and 2010 are revised and preliminary, respectively, and are subject to further revision.

(d) 2010–2012 estimates for incidence are based on 2000–2009 incidence data. 2011–2012 estimates for mortality are based on 2001–2010 mortality data (see Appendix G). They are rounded to the nearest 10. For estimates less than 1,000 they are rounded to nearest 5. The estimates for males and females may not add to estimates for persons due to rounding. Estimates are displayed on graph as dotted lines.

(e) The rates were age-standardised to the Australian population as at 30 June 2001 and are expressed per 100,000 population.

(f) Mean age for 2009 incidence was calculated excluding NSW and the ACT.

(g) The rates shown are age-specific rates.

Source: AIHW Australian Cancer Database 2009, AIHW National Mortality Database.

Table 2.4: Observed incidence (2009), and morality (2010) of non-melanoma of the skin and estimated for 2012

Source: AIHW Australian Cancer Database 2009, AIHW National Mortality Database.

CHAPTER 3

Image Processing and Automatic Skin Cancer Diagnosing

3.1. Introduction

The occurrence of skin cancer (malignant melanoma [MM]) has increased dramatically over the past few decades[41]. Australia is one of many countries in which skin cancer is more widely spread in comparison to other types of cancer[78]. The Health Benefits and Risks of Sun Exposure researchers which reported that solar radiation is the main cause of skin cancers. However, it also is a main source of vitamin D for humans[49]. Researchers found that, primary prevention and early detection continue to be of paramount importance in addressing the public health threat of skin cancer[79] [12].

The diagnosis of a doctor is based on visual inspection and may be reliable, although this procedure takes a lot of time and effort. These routines can be automated rather than the previous procedure saving doctors lots of time and would ensure that the diagnoses are more accurate[80]. There are other good means of storing diagnostic information alongside computer means. These means will help assist further investigations or the creation of new methods of diagnosis based on a vision system.

3.2. Vision System & Structure of the Human Eye

Electromagnetic radiation in the optical band is generated from the visual environment we view which enters the visual system of the eye and is detected in the sensitive cells of the retina. The activities start in the retina, where the signals from neighboring receivers are compared and a coded message dispatched on the optic nerves to the cortex, behind our ears.

3.2.1. Human Visual Perception

An excellent account of human visual perception may be found in[81]. The spatial characteristics of our visual system have been proposed as a nonlinear model in[82] [83]. Although the eyes can detect tranquility and static images, they are essentially motion detectors. The eyes are capable of identification of static objects and can establish spatial relationships among the various objects and regions in a static scene. Their basic

functioning depends on comparison of stimuli from neighboring cells, which results in interpretation of motion. When observing a static scene, the eyes perform small repetitive motions called saccades that move edges past receptors. The perceptual recognition and interpretation aspects of our vision, however, take place in our brain. The objects and different regions in a scene are recognized in our brain from the edges or boundaries that encapsulate the objects or the regions inside the scene.

3.2.2. The Human Eye

The maximum information about the object is embedded along these edges or boundaries. The process of recognition is a result of learning that takes place in our neural organization. The orientation of lines and the directions of movements are also used in the process of object recognition.

Light from an object moves through space and reaches our eyes enabling us to see. Signals are then sent to our brain, and our brain deciphers the information in order to detect the appearance, location and movement of the objects we are seeing. This process would not be possible if the presence of light was not available. Without light, sight would not be possible. The human eye is the organ which allows us to sense sight allowing us to learn more about the surrounding world, more than any of our other five senses. The eyeball is set in a protective cone-shaped cavity in the skull called the orbit or socket and measures approximately one inch in diameter. As Figure 3.1 illustrates, soft, fatty tissues surround the orbit protecting the eye and enabling it to turn easily.

3.2.3. Structure of the Eye

In general the eyes of all human and animals look like simple cameras in that the lens of the eye forms an inverted image of objects in front of it and projects it onto the sensitive retina, which corresponds to the film in a camera. As figure 3.2 shown , the front of the eye mainly contains an elaborate array of structures which are mainly concerned with the refraction (i.e. bending) of light rays and bringing them into focus on the retina. The structure most directly involved is the lens. The eye changes light rays into electrical signals then sends them to the brain, which interprets these electrical signals as visual images via the optic nerve as in figure 3.3.

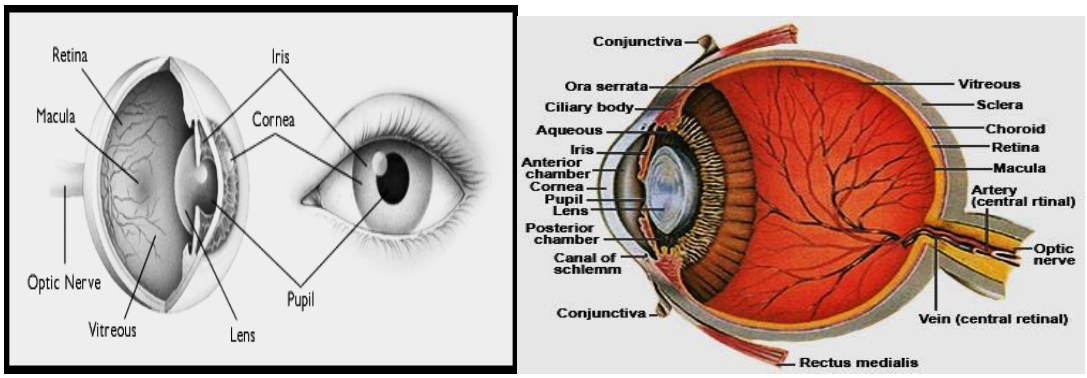


Figure 3.1 - Show the function of the eye.

Source: http://cdn2.hubspot.net/hub/60407/file-350482438-jpg/images/Master_Eye_logo_short_tag_line_revised_10-13-2013-resized-600.jpg

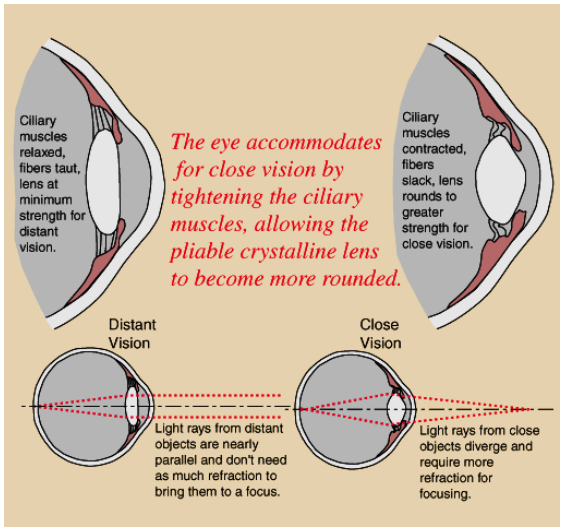


Figure 3.2 - Distance of vision.

Source:hyperphysics.phy-astr.gsu.edu/hbase/vision/accom.html

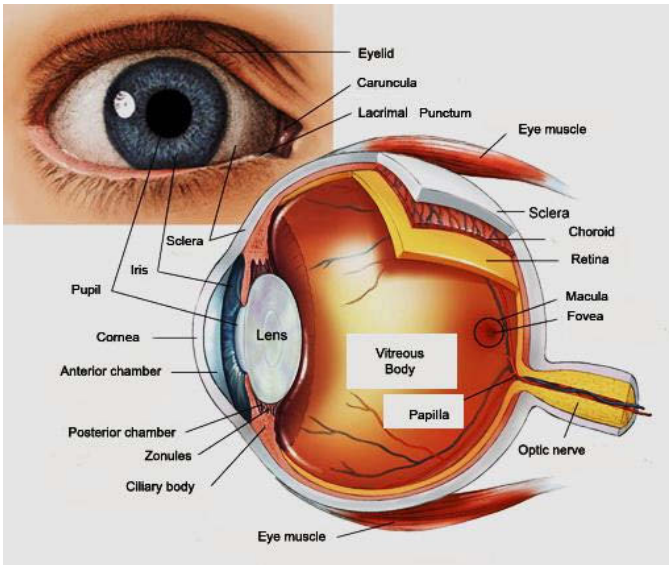


Figure 3.3 - Structure of the Eye.

Source:drpion.be/en/bouw-van-het-oog.htm

3.3. Image Processing & Human skin Imaging

Utilizing the concept of the biological vision system, an image processing system been established. An image processing system contains many components. To start to process digital images, we need to understand how the image processing operates. The brightness variation of an image should be described using its histogram, (making the image brighter or dimmer). We shall also consider thresholding techniques that turn an image from grey level to binary (two values representation). These are called single point operations. As a

result we see how the statistical operations reduce noise in the image, which benefits the feature extraction techniques which need to be considered later. As such, these basic operations are usually for pre-processing to improve display quality.

Digital image processing is based on the conversion of a continuous image field to equivalent digital form. A digital image is a detached function defined over a rectangular grid (lattice) representing the characteristics of the objects being imaged[84]. In 2-D images, each grid element, or "pixel", is defined by a location and a value representing a characteristic of the object at that location. The intensity and brightness are also factors which influence the process. Images, in which each pixel is assigned a single value, as in the above example are called scalar-valued images. Images in which each pixel is represented by a vector are called vector-valued or multispectral images. Typical examples of multispectral images in medicine and biology are red green blue (RGB) color images. As mentioned in[84], in the Matlab Image processing toolbox, there are four types of images: Indexed, Intensity, Binary, and RGB.

- Indexed image: consists of a data matrix and a color map matrix.
- Intensity image: consists of a single data matrix, each element of which denotes the intensity of the corresponding pixel (medical imaging).
- Binary image: is stored as a logical array. (i.e., each pixel is turned on or off), and is represented by a byte.
- RGB image: sometimes referred to as a true- color image, is stored in Matlab as an m-by-n-by-3 data array that defines red, green, and blue color components for each individual pixel.

Image compression techniques help to reduce the redundancies in raw image data in order to reduce the storage and communication bandwidth.

3.3.1. Applications of Image Processing

The Automatic Visual Inspection System improve the productivity and the quality of the product in manufacturing and allied industries[85]. There are many applications of image processing; some are listed below:

- Remotely Sensed Scene Interpretation: Information regarding the natural resources, such as agricultural, hydrological, mineral, forest and geological resources, etc., can be extracted based on remotely sensed image analysis[86] [87].

- Biomedical Imaging Techniques: Various types of imaging devices are used extensively for the purpose of medical diagnosis[88] [89].
- Defense surveillance: Application of image processing techniques in defense surveillance is an important area of study. There is a continuous need for monitoring the land and oceans using aerial surveillance techniques.
- Content-Based Image Retrieval: The features of a digital image (such as shape, texture, color and topology of the objects, etc.) can be used as index keys for search and retrieval of pictorial information from a large image database. Retrieval of images based on such image contents is popularly called the content-based image retrieval[90] [91].
- Moving-Object Tracking: Tracking of moving objects, for measuring motion parameters and obtaining a visual record of the moving object, is an important area of application in image processing[6] [92]. Generally there are three different approaches to object tracking:
    1. Recognition-based tracking.
    2. Motion-based tracking.
    3. Image and Video Compression: This active application area in image processing[91]. Development of compression technologies for image and video continues to play an important role for success of multimedia communication and applications.

There are many algorithms in machine vision which can provide reliable automatic object detection, recognition and classification in an independent environment for understanding two- and three-dimensional objects in a digital image [93],[94],[95],[96]. There are a number of questions concerning vision, such as: ***(i) what are the goals and limitations? (ii) What type of algorithm or set of algorithms is required to affect vision? (iii) What are the Implications for the process, given the types of hardware that might be available? (iv) What are the levels of representation required to achieve vision?*** The levels of representation are dependent on what type of segmentation can and/or should be applied to an image. We are required to consider (a) how to best segment an image and (b) the form this segmentation should take. Segmentation is the first step in automatic image analysis and pattern recognition. In these cases, we need appropriate techniques of refining the images to get better visual quality, which are free from aberrations and noises. Image

enhancement, filtering, and restoration have been some of the important applications of image processing since the early days of the field[97] [98] [84].

Image analysis also includes topics such as feature extraction, shape representation, and object recognition and classification. In this research, we will cover some of the fundamental topics that have found widespread applications in medical image processing.

3.3.2. Digital Image Acquisition

Medical digital images are pictures of distributions of physical attributes captured by an image acquisition system. Medical images may show anatomy including the pathological variation of anatomy if the measured value is related to it or physiology when the distribution of substances is traced. Examples of medical imaging include X-ray imaging, CT, MRI, nuclear imaging, ultrasound imaging, photography, and microscopic images [99].

With respect to digital imaging, four major and several minor imaging techniques meet these requirements. The major techniques are as follows:

- X-ray imaging measures the absorption of short wave electromagnetic waves; this is known to vary between different tissues.
- Magnetic resonance imaging measures the density and molecular binding of selected atoms (most notably hydrogen which is abundant in the human body), which varies with tissue type, molecular composition, and functional status.
- Ultrasound imaging captures reflections at the boundaries between and within tissues with different acoustic impedance.
- Nuclear imaging measures the distribution of radioactive tracer material administered to the subject through the blood flow. It measures function in the human body.

Other imaging techniques include EEG and MEG imaging, microscopy, and photography. All the techniques have in common that an approximate mapping is known between the diagnostic question, which was the reason for making the image and the measurement value that is depicted. This can be very helpful when selecting an analysis technique. If bones need to be detected in an x-ray CT slice, a good first guess would be to select a thresholding technique with a high threshold because it is known that x-ray attenuation in

bone is higher than in soft tissues and fluids[99]. The first major component is a camera that captures the images.

3.3.3. Digital Camera

The sensors which are used in most of the cameras are either Charge Coupled Device (CCD) or CMOS sensors. The CCD camera comprises a very large number of very small photo diodes, called photosets. The electric charges which are accumulated at each cell in the image are transported and are recorded after appropriate Analogue to Digital Conversion (ADC).

In a situation where bright sunlight is present, the aperture which is located behind the camera lens needs to be small as little light is required, whereas on cloudy days more light is needed to create an image so the aperture should be enlarged. This is identical to the functioning of our eyes. The shutter speed gives a measure of the amount of time during which the light passes through the aperture. The opening and closing of the shutter depends on how much light is available. The distance between the focal planes included the focal point and the focal length was measured from the optical 'centre' of the lens and the surface of the sensor array is known as the focal length of a digital camera. Focal length is the critical information in selecting the amount of required magnification which is desired from the camera as shown in figure 3.4.

A digital camera can capture images in various resolutions, (low, medium or high resolution). The cameras that we normally use can produce about 16 million colors, i.e., at each pixel we can have one of 16 million colors. The spatial resolution of an image refers to the image size in pixels, which corresponds to the size of the CCD array in the camera.

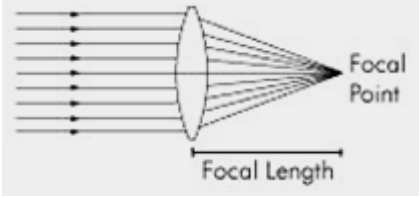


Figure 3.4 Focal length defined [When parallel rays of light strike a lens focused at infinity, they converge to a point called the focal point. The focal length of the lens is then defined as the distance from the middle of the lens to its focal point.]

A digital camera can capture colors in many ways. A practical solution is based on the concept of color interpolation or demo-sailing, which is a more economical way to record the three primary colors of an image. In this method, we permanently place only one type of filter over each individual photo-site. Usually the sensor placements are carried out in accordance to a pattern. The most popular pattern is called the Bayer's pattern[100], where each pixel is indicated by only one color-red, blue, or green pixel. It is possible to make very accurate guesses about the missing color component in each pixel location by a method called color interpolation or demosaicing [101] [102].

In high-quality cameras, however, three different sensors with three filters are used and light is directed to the different sensors by using a beam splitter. Each sensor responds only to a small wavelength band of color. Thus the camera captures each of the three colors at each pixel location.

3.3.4. Skin Imaging Techniques

Dermatology is often described as a visual specialty wherein a majority of diagnoses can be made by visual inspection of the skin. Diagnosis of skin disease in dermatology is largely non-invasive. The physician diagnosis is based on the anatomic distribution, color, configuration, and visible surface changes of a lesion. In some cases, a skin biopsy is performed which again offers the opportunity for a microscopic visual examination of the lesion in question, but there exist limitations in the assessment of depth and size of skin lesions as well as internal features of superficial lesions. Such limitations result in the need for an objective non-invasive means of assessing the skin.

Digital dermatoscopic images firstly have to be parameterized for automatic classification. Extensive study of skin nature has to be done in literature to parameterize it.

3.3.4.1. Skin Imaging

Imaging modalities with major relevance to skin imaging include various modification of electromagnetic wave imaging (Optical, Infrared, Tera Hertz, and Nuclear Magnetic Resonance), acoustical wave imaging and mechanical wave imaging. Usually tomographic images i.e. 2D cross sectional images are acquired with medical imaging systems. Depending on the spatial orientation of the cross section, these images depict information about the tissue over depth or any other direction. Three dimensional tissue volumes are usually imaged by acquiring a stack of consecutive 2D images[103].

The different illumination method called epiluminescence can be used in order to get the image from deeper skin layers. The light is directed straight in to these layers and reflected back giving more information about consistency of light absorbers in these layers. This method of illumination improves diagnosis accuracy up to 10 %.

Another interesting solution for getting more information from skin is by using multi spectral photography, which uses narrow frequency bands of light illumination. Those images give information about consistency and concentration of absorbers and reflectors in the skin. When those photos are made with a range of light waves, the reflectance frequency characteristics of skin can be calculated and compared to normal skin. We can calculate the reflectance frequency characteristics of skin and compare to normal skin. Some of the methods based upon the above classification of skin imaging are described in coming sections[104].

3.3.4.2. Dermoscopy of pigmented skin lesions

In the last two decades an increasing rate of malignant melanoma has been observed [105, 106]. Because of a limited range of effective treatments after the spread of cancer (melanoma), the best treatment currently is still early diagnosis and punctual surgical excision of the primary cancer[105] Dermoscopy (also known as epiluminescence microscopy, dermoscopy, and amplified surface microscopy) is a non-invasive and in live method, that has been reported to be a useful tool for the early recognition of malignant melanoma [107] [108] [109].

The performance of dermoscopy has been investigated by many authors. Its use increases diagnostic accuracy between 5% and 30% over clinical visual inspection depending on the type of skin lesion and experience of the physician [110-113]. This was confirmed by recent evidence-based publications and based on statistics of the literature [107]. This section is a review of the principles of dermoscopy as well as recent technological developments.

3.3.4.3. Microscope

The study of histology is carried out using a variety of microscopes, in order to visualise the structure of body tissues. Pathologists use photomicrographs, captured by the ***light microscope (LM)*** (color images) and the ***electron microscope (EM)*** (black and white images), which differ in optical resolution and available magnification. Resolution

practically refers to the capacity of an optical system to reveal details within a specimen. EM images are therefore said to display cell and tissue ultrastructure. There are two types of electron microscope producing two types of images, either two or three dimensional. The scanning EM which produces a three-dimensional (3D) image is restricted to surface structure and the transmission EM utilizes an electron beam which must pass through the specimen to form an image. (more detailed information in appendix B)

3.3.4.4. Complementarity of Light and electron microscopy

The strengths of LM and EM differ yet complement one another very effectively. Large areas of a specimen can be observed under a light microscope. Identification of cell and tissue features can be accentuated through utilizing a wide range of staining methods. The electron microscope has superior resolution and magnification permitting observation of many features and intracellular structures which simply cannot be seen via the light microscope.

3.3.4.5. Electron microscope scanning EM and transmission EM:

Historically the ***transmission electron microscope (TEM)*** was the first type of Electron Microscope to be developed and is patterned similarly to the light transmission microscope except that a focused beam of electrons is used instead of light.

The advancements in science led to a desire to observe intracellular structures in greater detail and interior structure of organic cells e.g. the nucleus. Light microscopes are limited by the physics of light consequently the electron microscope was developed. This provided required much greater magnification of 10,000 X plus, which was possible using the current optical microscope. Electron microscopes are scientific instruments that examine objects on a very fine scale by using a beam of energetic electrons[114]. The similarities in the overall design and principal features of the optical microscope (OM), the transmission electron microscope (TEM) and the scanning electron microscope (SEM) are shown in Figure 3.5.

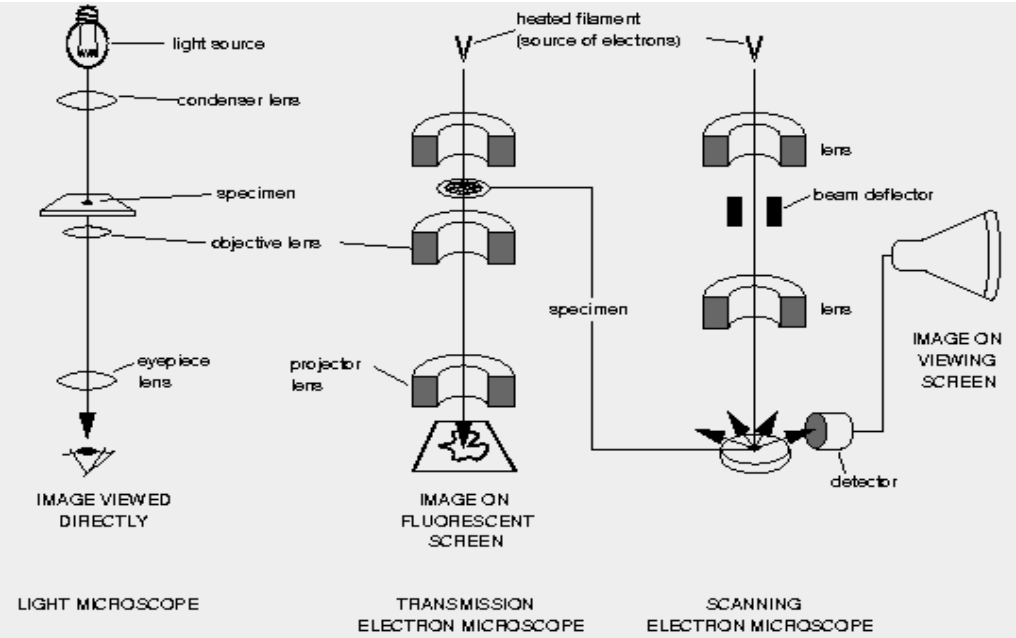


Figure 3.5. Displayed, similarities of overall design in the principal features of an optical microscope, a transmission electron microscope and a scanning electron microscope.

Source: www.nslc.wustl.edu

3.3.4.6. Confocal laser scanning microscopy

Confocal laser scanning microscopy (CLSM) is a new optical microscopic technique, which offers significant advantages over conventional microscopy. CLSM is microscopy of optical sections. Light, which is emitted from regions other than the focal plane, is cut off by introducing a diaphragm in the beam path. The result is an optical "slice", which shows more details because the blurring from out of focus haze disappears. It has been repeatedly used in experimental, but also in diagnostic dermatopathology.

The "in vivo" confocal microscopy, applied directly to the intact skin provides details of living cells in the superficial layers comparable to that of fixed and stained tissue. While the extent of its future applications is hard to predict, its potential for applications in dermatology appears very considerable, particularly for studies of fixed or living tissues, where it is desirable to obtain clear images many micrometers below the surface of the tissue under examination [115].

The confocal microscope is elementally comprised of a coherent scanning laser light source, condenser lens, objective lens, and detector. A laser light point source illuminates a small region within the skin and is scanned across the skin. Light reflected from this focal point propagates back to the detector through a pinhole aperture after a single collision

within the tissue, allowing the observation of reflections from single scattering events. Light from the in focus plane is emanating from above and below the focus plane minimally passes through the pinhole to the detector, resulting in images of high resolution and contrast. The term “confocal” stems from this design in which the pinhole in front of the light source and the pinhole in front of the detector are in optically conjugate focal planes. Figure 3.6 illustrates the principle of confocal microscopy.

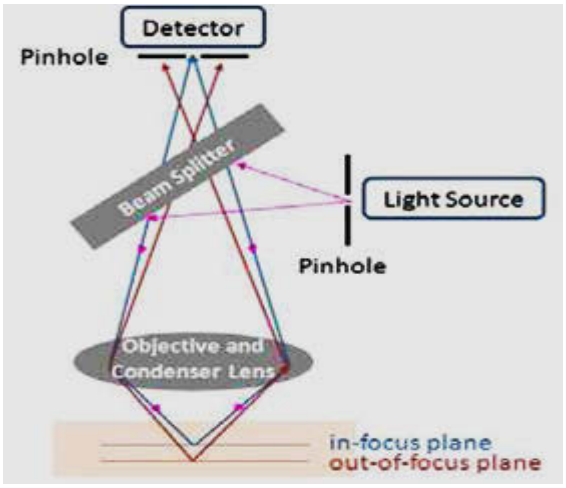


Figure 3.6. Principle of CSLM. Note the same optical path is used for the detector and the source. Optics are used to direct the light towards the detector

Confocal laser scanning microscopy is a novel technique that enables us to non-invasively investigate skin at the cellular level and observe dynamic changes that may occur over a period of time. It was shown that the incident light wavelength in CSLM systems affects signal intensity and structural details at the cellular level. Systems equipped with shorter wavelengths may provide improved resolution in images, while those with longer wavelengths can image deeper into the skin. When comparing systems of the same design and only differing in incident wavelength, the 785 nm system produces images with greater signal intensity and contrast in structural detail compared to the 830 nm system. It is peculiar that imaging human skin with 785 and 830 nm wavelengths produces intensity profiles that appear biphasic, while imaging at 405 nm result in intensity profiles that follow a simple exponential decay[116].

Capturing cellular changes over time with CSLM is of great significance as it enables the study of the physiology of living skin tissue. Skin is the body’s largest functioning organ, consisting of millions of cells that are in constant activity whether it is proliferation, pigmentation, repair, etc. Capturing the same cells over time and observing their activity

provides powerful diagnostic information. By studying healthy skin tissue and understanding how it reacts and functions under average conditions, gives a significant insight into the cells of healthy tissue. Skin tissues can be better characterized and the differentiation between healthy and unhealthy tissue made. Capturing physiological and morphological changes in skin due to aging, treatment and wounding is only the beginning of the fundamental understanding of dynamic cellular changes using CSLM. Having a strong foundation in the knowledge of normal function and activity of healthy tissue is essential to the study and understanding of diseased tissue.

3.3.5. Differential Diagnosis of Pigmented Lesions of the Skin

There are many publications on the subject of the differential diagnosis of pigmented lesions of the skin. The 5 algorithms most commonly used are pattern analysis[117] [113] [118]; the ABCD rule of dermoscopy [41] [119] [119]; the 7-point checklist [105] [117] [120] ; the Menses method[121], [122], [123]; and the revised pattern analysis [124]. The common two-step procedure for the classification of pigmented lesions of the skin is illustrated in Figure 3.7. The first step is the differentiation between a melanocytic and a non melanocytic lesion. For this decision, the algorithm in Figure 3.8 is used. The outcomes image results are displayed on Figures from 3.9 to 3.15.

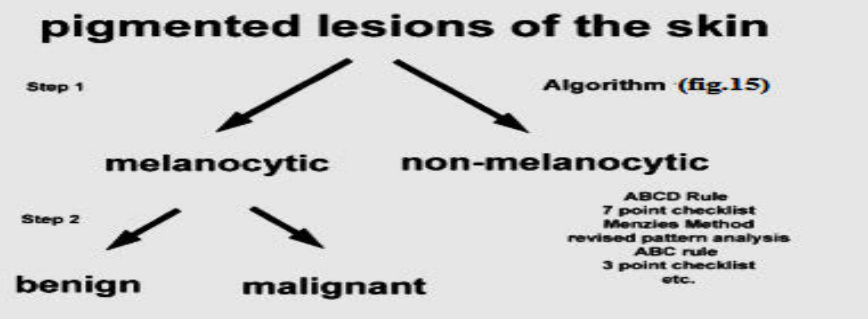


Figure 3.7. Two-step procedure for the classification of pigmented skin lesions. Adapted from Argenziano.[122]

To detect melanomas (and increase survival rates), it is recommended to learn what they look like ("ABCDE" as displayed below), to be aware of moles and check for changes (shape, size, color, itching or bleeding) and to show any suspicious moles to a doctor with an interest and skills in skin malignancy [125].

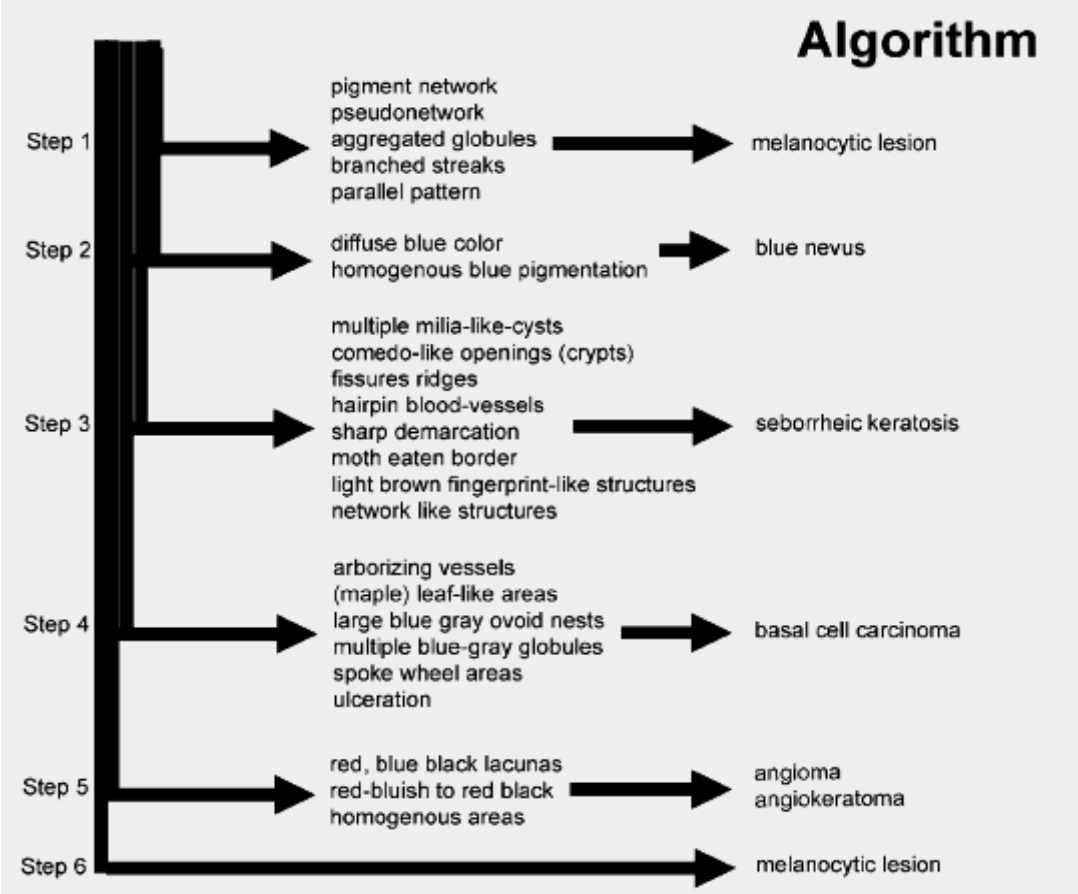


Figure 3.8. Algorithm for the determination of melanocytic versus non melanocytic lesions according to the proposition of the Board of the Consensus Netmeeting. Adapted from Argenziano.[122]

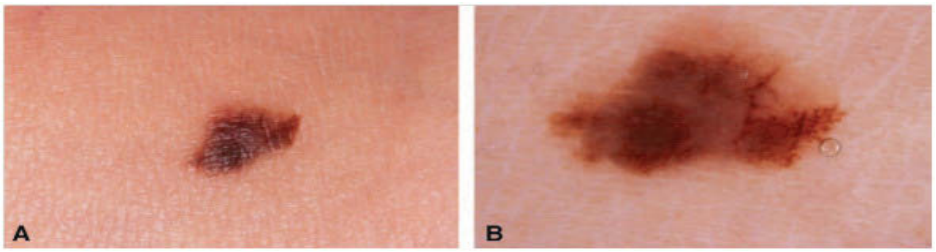


Figure 3.9. A, Macroscopic picture of a superficial spreading malignant melanoma (Breslow thickness 0.52 mm; Clark level II). B, Dermoscopy of A shows (atypical) pigment network and branched streaks and can therefore be considered a melanocytic lesion.

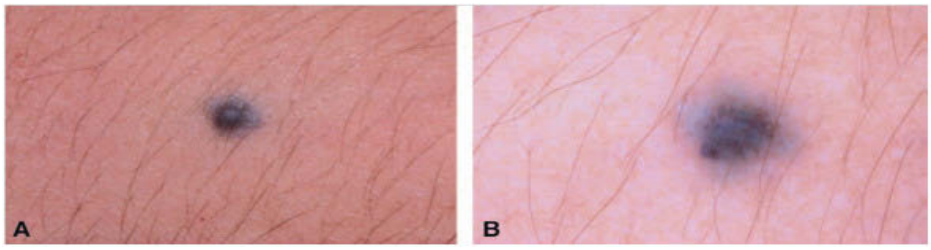


Figure 3.10. A, Macroscopic picture of a blue nevus. B, Dermoscopy of A shows steel-blue areas (no pigment network, no aggregated globules, and no branched streaks).

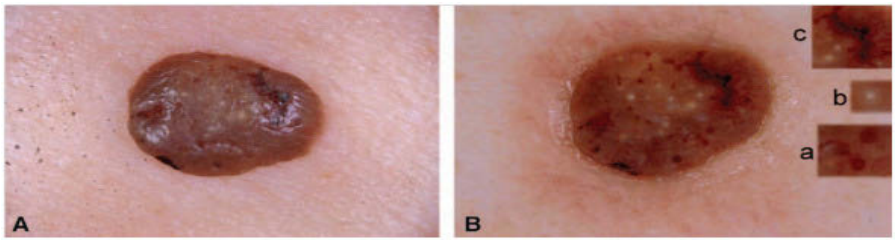


Figure 3.11. A, Macroscopic picture of a seborrheic keratosis. B, Dermoscopy of A shows comedolike openings (a), multiple milia-like cysts (b), and fissures (c).

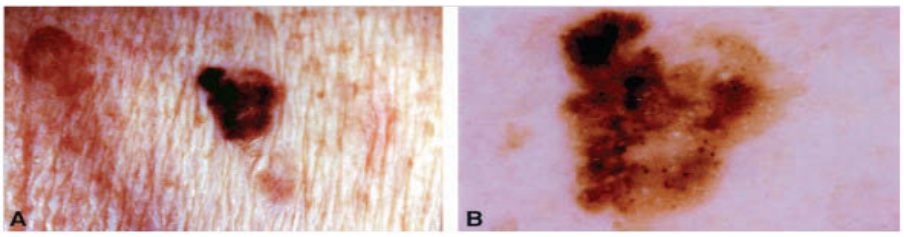


Figure 3.12. A, Macroscopic picture of a seborrheic keratosis. B, Dermoscopy of A shows comedolike openings and multiple milia-like cysts.

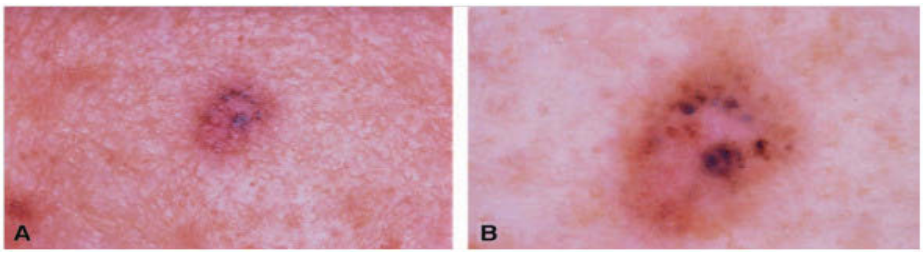


Figure 3.13. A, Macroscopic picture of a basal cell carcinoma. B, Dermoscopy of A shows maple lifelike areas, ovoid nests, and arborized telangiectasia.

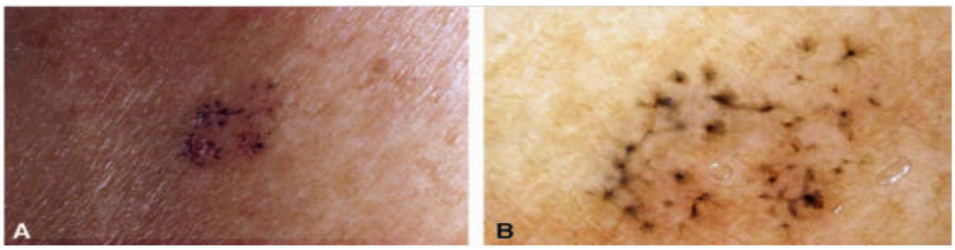


Figure 3.14. A, Macroscopic picture of a basal cell carcinoma. B, Dermoscopy of A shows multiple spoke wheel areas.

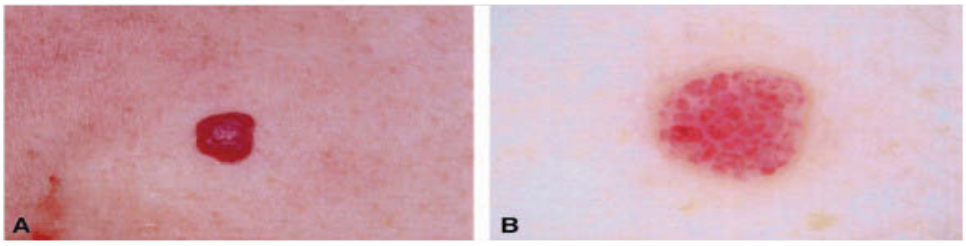


Figure 3.15. A, Macroscopic picture of an angioma. B, Dermoscopy of A shows red lagoons.

3.3.5.1. Basic Principles of the ABCDE (Warning Signs)

Anyone who has more than 100 moles is at greater risk for melanoma. It is so important to know the skin very well and to recognize any changes in the moles on the body. The ABCDE rules of melanoma can help in such case, if a person see one or more of this signs; they should check-up with the physician, and then contact the specialist for early diagnosis.[56]. The Asymmetry Border Color Diameter Evolution (ABCDE) popular method for remembering the signs and symptoms of melanoma is as displayed in figure 3.16. It is based on check the skin lesion as below:

- Asymmetrical skin lesion: shows the two halves will not match.
- Border: shows the border edges may be scalloped or notched.
- Color: A number of different shades of brown, tan or black could appear. A melanoma may also become red, blue or some other color.
- Diameter: shows that melanomas are larger in diameter than the size of the eraser on your pencil (1/4 inch or 6 mm).

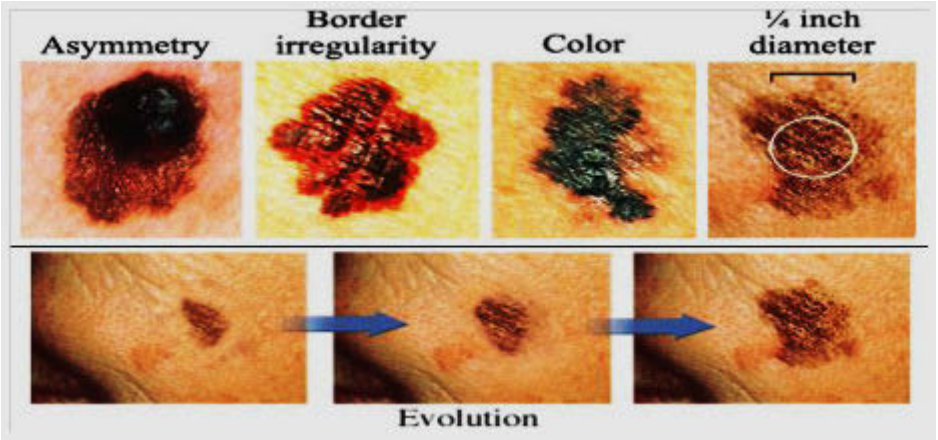


Figure 3.16. Asymmetry Border Color Diameter Evolution (ABCD-E) rule for Diagnosis of Melanoma

Source: www.webmd.com

Evolving: shows the change in size, shape, color, evasion, and bleeding, itching or crusting these points to danger

The pictures on Figure 3.17, shows the typical normal moles and melanomas these are useful in showing people the differences as a start point.[126]

The ABCD rule is introduced here and can be easily learned and rapidly calculated and has been shown to be a reliable method providing a more objective and reproducible diagnosis of melanoma. The semi-quantitative ABCD rule represents the second step of a two-step procedure that was originally proposed by [127] and who later modified the rule to be applied, at least following Stoltz' instructions [128].

To improve diagnostic accuracy the ABCD rule of lesion screening is widely used based on asymmetry (A), border (B), color (C), and differential structure (D). Measuring criteria have to be assessed semi-quantitatively. Then, each of the criteria has to be multiplied by a given weight factor yielding a total dermoscopy value (TDV) which results from the calculation as per rule (1).

TDV = (A score x 1.3) + (B score x 0.1) + (C score x 0.5) + (D score x 0.5) …. (1)

This score contributes to the differentiation between benign and malignant lesions:

1.00 – 4.75 – Benign skin lesion

4.75 – 5.45 – Suspicious lesion

More than 5.45 – Highly suspicious for melanoma.

Nodular melanomas do not fulfill these criteria, having their own mnemonic, "EFG":

- Elevated: the lesion is raised above the surrounding skin.
- Firm: the nodule is solid to the touch.
- Growing: the nodule is increasing in size.

If the mole is malignant, the mole and an area around it need excision. Elliptical excisional biopsies may remove the tumor (figure 3.18(a)); this is followed by histological analysis and Bristow scoring refers to figure 3.18(b). Punch biopsies are contraindicated in suspected melanomas, for fear of seeding tumor cells and hastening the spread of the malignant cells.

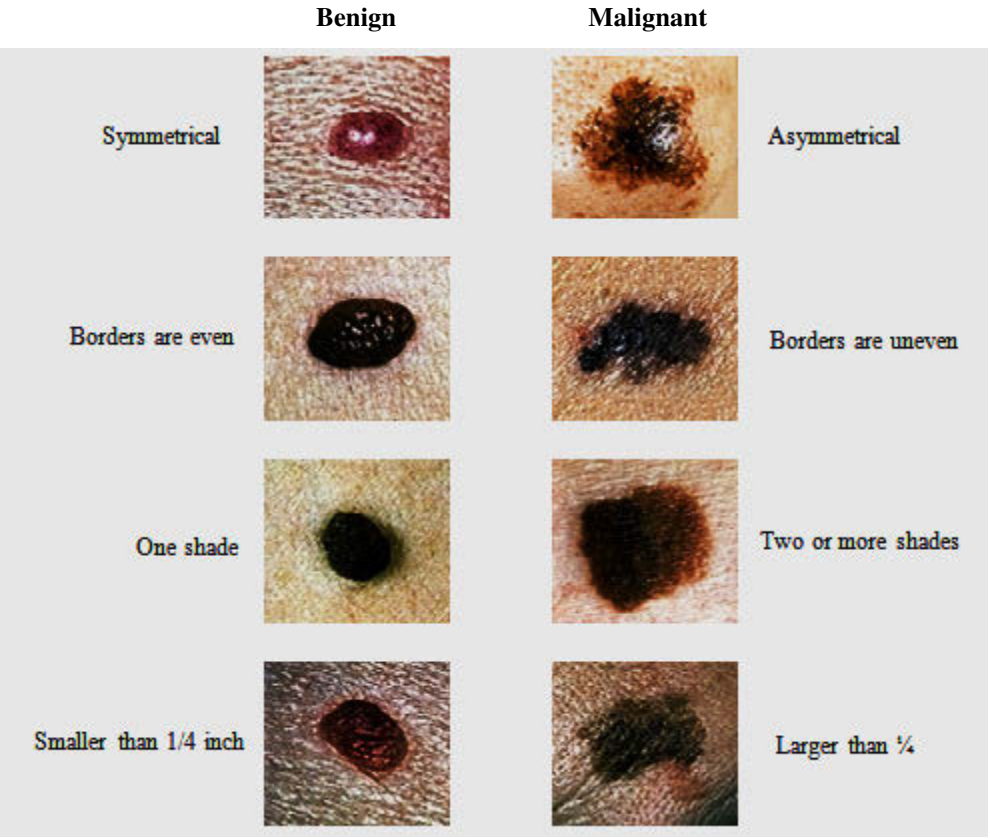


Figure 3.17 - Normal moles and Melanomas [Benign and Malignant].

Source: www.skinbychar.com

Total body photography, which involves photographic documentation of as much body surface as possible, is often used during follow-up of high-risk patients. The technique has been reported to enable early detection and provides a cost-effective approach (being possible with the use of any digital camera), but its efficacy has been questioned due to its inability to detect macroscopic changes[64]. The diagnosis method should be used in conjunction with (and not as a replacement for) dermoscopy imaging, with a combination of both methods appearing to give extremely high rates of detection.

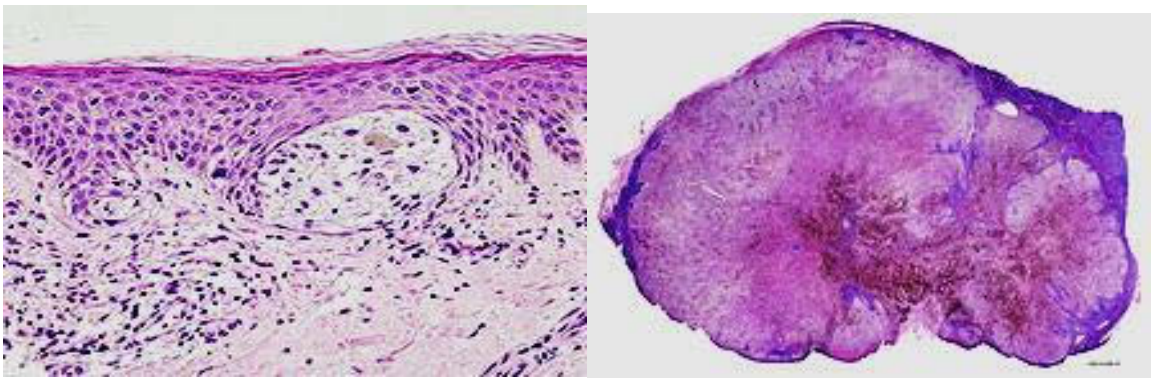


Figure 3.18: Showed: a) Melanoma in skin biopsy with H&E Stain ("This case may represent superficial spreading melanoma."), b) Lymph node with almost complete replacement by metastatic melanoma. The brown pigment is focal deposition of melanin.

3.3.5.2. 7-Point Checklist

In 1998 Argenziano [4] and his colleagues described a 7- point checklist based on the analysis of 342 pigmented skin lesions. In[45],[129],[130] they distinguish 3 major criteria and 4 minor criteria. Each major criterion has a score of 2 points while each minor criterion has a score of 1 point. A minimum total Score of 3 is required for the diagnosis of malignant melanoma.

3.3.5.3. Menses method

According to Menses[131] for diagnosing melanoma, both of the following negative features must not be found: a single color (tan, dark brown, gray, black, blue, and red, but white is not considered) and ‘‘point and axial symmetry of pigmentation’’ (refers to pattern symmetry around any axis through the center of the lesion). This does not require the lesion to have symmetry of shape. Additionally, at least one positive feature must be found.

3.3.5.4. The Three-Point Checklist

In figure 3.19, the three-point checklist was developed for inexperienced dermoscopists. This is meant to encourage them to use dermoscopy by teaching them a simplified algorithm for the evaluation of pigmented skin lesions. General practice physicians can easily be taught dermoscopy and the three-point checklist, which can be utilized as a screening tool to determine whether a given pigmented lesion should be further evaluated by a dermatologist. The three-point checklist [132] features three criteria that are found to be particularly important in the differentiation between malignant and benign pigmented skin lesions by participants in the Consensus Net Meeting on Dermoscopy [122]. In the

three-point checklist the presence of any two of the three criteria: asymmetry; atypical pigment network; and blue-white structures indicates that the lesion under investigation may be a melanoma.

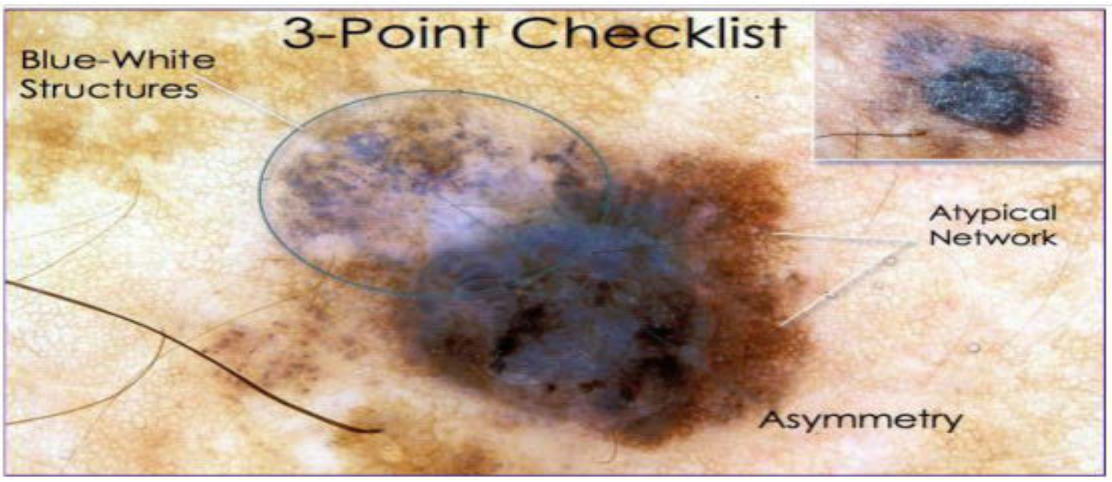


Figure 3.19. Diagnosis of Melanoma using Three-point Checklist

Source: www.dermoscopy.org

3.3.6. Automatic Skin Cancer Diagnosing

There is need for appropriate techniques of refining the images to get better visual quality, free from aberrations and noise. Some methods that need to be considered for describing skin lesion properties for diagnosing processing are:

- Asymmetry of the lesion border.

Transition of the pigmentation from the skin lesion to the surrounding skin. (Edge Gradient)

- Colour distribution of the skin lesion including the blue-white veil.

3.3.6.1. Skin Lesion imaging Methods

We found that there are a number of various imaging methods of skin lesions.

- ***Traditional Photography***

The simplest visualization method is photography. This method gives only the top layer skin image.

- ***Dermoscopy Imaging Techniques***

In order to get deeper layer image oil immersion is used. It reduces reflections of surface and brightens the image off the epidermis – the second skin layer. Better results are reached when photos are made with polarized light source from the top layer of skin as shown in Figure **3.20**. Then it is easier to estimate the lesion structures like dots, globules, nets that are the major indicators to melanoma diagnosis.

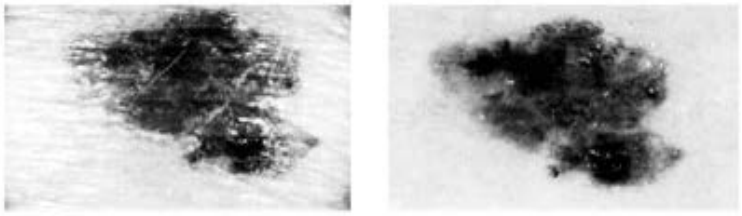


Figure 3.20 - Example of pigmented skin lesion. Left: Traditional imaging technique. Right: Dermoscopy imaging technique.

- ***Multi Spectral Photography***

Another interesting solution of getting more information from skin is by using multi-spectral photography. It uses a narrow frequency band of light illumination. Those images give information about consistency and concentration of absorbers and reflectors in the skin. The idea is that a different pigment of skin absorbs different light wave frequencies, determining the color of our skin. When those photos are made with a range of light waves, we can calculate the reflectance frequency characteristics of skin. These photos can be used to compare with normal skin characteristics and diagnostic decisions can be made about skin pigment consistency.

- ***Ultrasound Visualization***

Ultrasound visualization is usually used to measure depth of melanoma and if there is no melanoma practically there are no differences. When a doctor diagnoses melanoma, then he uses high frequency ultrasound (over 30 MHz) to measure penetration depth in order to make the correct cut during surgery.

- ***Dermoscopy***

It is a non-invasive diagnostic technique used in dermatology for the in vivo observation of pigmented skin lesions. Dermoscopy is an imaging technique that provides a more direct link between biology and distinct visual characteristics. It is known as epiluminescence light microscopy (ELM).

This diagnostic tool allows for a better visualization of surface and subsurface structures and thus permits the recognition of morphologic structures which are not visible to the naked eye. This technique has opened a new dimension for analysing the clinical morphologic features of pigmented skin lesions.

Dermoscopy images are much more detailed than conventional macroscopic (clinical) images. Macroscopic images are merely images of skin lesions seen under a magnifying glass. Figure 3.21 shows a comparison of a macroscopic and a dermatoscopic image of the same skin lesion.

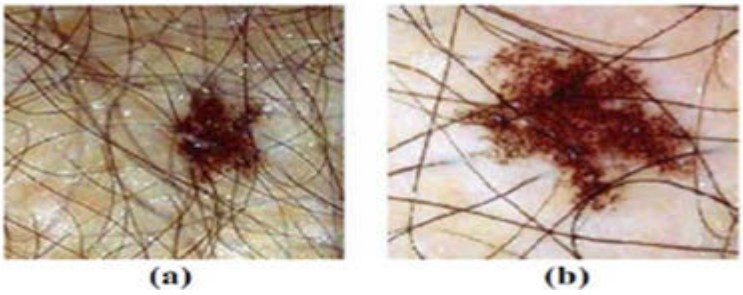
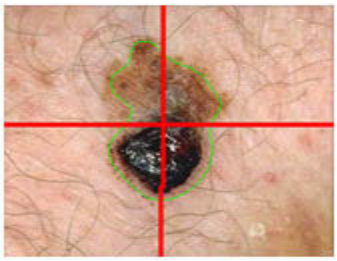


Figure 3.21. (a) Macroscopic Image of a Lesion (b) Dermatoscopic Image of the same Lesion. Source: www.jle.com

Studies have shown that dermatoscopic diagnosis has an accuracy of 75% to 97% while for macroscopic diagnosis this range is 65% to 80% [36] [38]. However, it has also been noticed that dermoscopy may actually lower the diagnostic accuracy in the hands of inexperienced dermatologists [38]. Therefore, due to the lack of reproducibility and subjectivity of human interpretation, the development of computerized techniques is of utmost importance [133].

- ***Skin Lesion Boundary Tracing Algorithm***

Work reported in [134], where it found Matlab very handy tool which allows for automatic detection of skin irregularities and which was used to calculate lesion properties like asymmetry of shape, or border irregularities, which can help in detecting melanoma. There are a lot of investigations in this work; finally, the author provides his own source code to let researchers trial they own data. Figure 3.22. Showed my data results using the trial code.



[A] Image Center of mass.

Original image	Filtered Image
	
Black and White	Traced
	

[B] Image process.

Figure 3.22- Show the results of skin lesion boundary tracing algorithm, my data experiment shown [a] Image center of mass, [b] Image process.

3.3.6.2. Skin Cancer Detection Methods: New and Existing Techniques

Automated border detection is one of the most important steps in this procedure as the accuracy of the subsequent steps essentially depends on the accuracy of this step. Unsupervised approach to border detection in dermoscopy skin lesion images is based on a modified version of the Japan Society of Engineering geology (JSEG) algorithm. This method presented achievements of both fast and accurate border detection in dermoscopy images[135] [136]. Some of the digital processing techniques are contour detection and image segmentation.

Some studies have combined image processing with knowledge-based approaches in order to achieve better results in interpreting images by presenting a general methodology developed for computer interpretation and integration of medical images. Plain anatomical

models, as well as other domain knowledge, are used to facilitate the feature extraction and fusion of images from different modalities.

- ***Image Segmentation***

Segmentation is to break an image up into meaningful segments. Segmented images improve image analysis especially for image querying and retrieval [137]. There are many different approaches to segment images; this section will describe some techniques commonly used with segmentation.

- ***Data Clustering Based Segmentation***

Image segmentation via the use of data clustering is a very effective technique[137]. When it is used with the appropriate distance and probability distribution functions, segments/clusters of an image can be properly distinguished. Often, the most difficult task in data clustering in general is determining the optimal number of clusters[137]. Specifying too few clusters will generate incomplete image segments, and specifying too many will generate redundant segments. Figure 3.23 shows an example of the cluster threshold problem.

Generally, grey scale images are easier to segment compared to color images. This is due to the fact that grey scale images consist of only one color component which is the grey level, ranging from 0 to 255. Color images are generally stored in the RGB color space. The RGB format is not a strong model for calculating distances between color pixels since it is not a uniform color space, and as such, it will not perform well for color segmentation[137]. Calculating distance measures for color pixels requires the use of a uniform color space. Usually, the color spaces are preferred for color segmentation distance calculations.

- ***Texture Segmentation***

Texture segmentation is classified into two main categories, supervised and unsupervised. In cases where the different types of textures are small, supervised segmentation is used where the a-priori knowledge of the textures is known [138]. In cases where the types of textures are large or the assumptions of the texture cannot be made, unsupervised texture segmentation must be used.

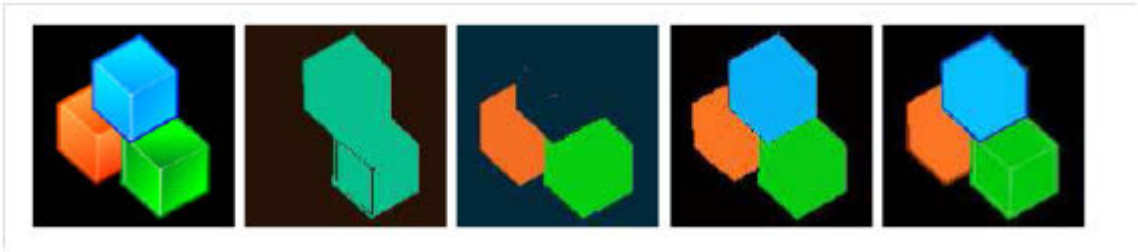


Figure 3.23- From left to right: the original image (source image from: http://ijplugins.sourceforge.net/plugins/clustering/), clustered image using 2 clusters (poor), clustered image using 3 clusters (close but one key cluster is missing), clustered image using 4 clusters (optimal), and clustered image using 10 clusters (artifacts are noticeable).

Texture segmentation involves applying the texture filters on the image and then using segmentation methods such as the clustering technique to segment an image into textured as such segments.

Many texture filters have been implemented in the past. Some of the commonly used filters are the Gabor filters[139] and the wavelet-based decomposition/ transformations [140].

- ***K-Means Clustering (Euclidean Distance)***

K-means is one of the common approaches for data clustering. It attempts to subdivide a region of pixels into smaller regions based on the region's pixel mean value, and by doing so, pixels that are closer to the mean value will be grouped together[141]. The process continues until it reaches a stage where nothing has changed. The naive approach to finding the distance between pixel values in the feature space is to use the Euclidean Distance. Although it is a faster way to determine distances between pixel values, it assumes that clusters are shaped in hyper-spheres only, which is quite restrictive.

The Euclidean distance in equation 3.1:

$$d(X,Y) = \sqrt{(X-Y)^T + (X-Y)} \qquad 3.1$$

Where (X,Y) is a matrix of the difference between the two data vectors X and Y [142].

- ***K-Means Clustering (Mahalanobis Distance)***

The Mahalanobis Distance method models the distance between color pixels more accurately compared to the Euclidean distance model [141]. It calculates the distances using the variance values of a cluster's data set. That is, the distance between color pixels is calculated using the variance values between each color component.

The Mahalanobis distance in equation 3.2:

$$d_{M_j} = \sqrt{(X-Y)^T \sum_j^{-1} (X-Y)} \qquad 3.2$$

Where $(X-Y)$ is a matrix of the difference between the two data vectors X and Y with respect to cluster j, and $\sum_j^{-1}$ the covariance matrix for cluster j [142]

- ***Fuzzy Logic Concepts and Application***

Fuzzy set theory application in image processing is used in areas like enhancement, segmentation, filtering, edge detection, content-based image retrieval, pattern recognition and clustering [143].

Fuzzy logic concepts and application in digital image processing have become apparent. Fuzzy set theory is thus useful in handling computer vision and image-processing applications. Fuzzy logic has three main stages, namely, image fuzzification, modification of membership values, and, if necessary, image defuzzification as shown in figure 3.24. The main part of fuzzy image processing in the membership plane is shown in steps of fuzzy image process in figure 3.25.

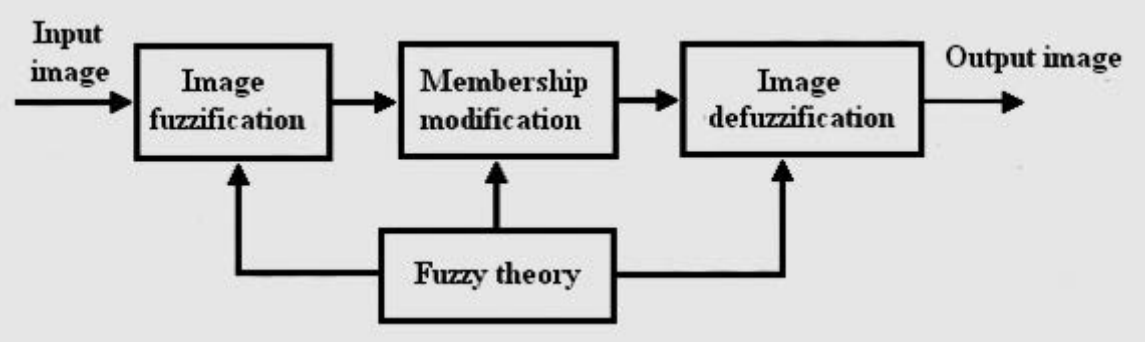


Figure 3.24- Structure of fuzzy image processing.

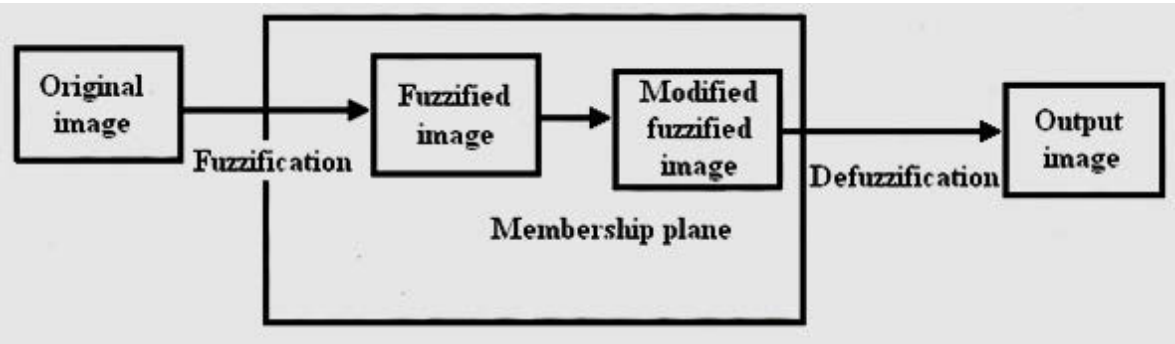


Figure 3.25- Steps of fuzzy image processing.

- ***Fuzzy C-Means Clustering***

Another modification of the K-means clustering is the Fuzzy C-means (FCM) clustering algorithm. Data points in the non-fuzzy clustering methods can only belong to a single cluster [144]. In the FCM approach, data points can belong to more than one cluster based on some predefined membership rules that determine the degree of membership within each cluster for each data point. In the basic form of K-means clustering, each data points' distance from its cluster center contributes equally to its centroid location. This makes the cluster centroid values weak to outliers. The FCM method reduces the force of outliers by reducing the influence of each outlier for those clusters that are least likely to have generated them [145].

- ***Clustering (Expectation Maximization - Gaussian Mixture Model)***

Clustering can be improved by using the Expectation Maximization (EM) algorithm[141]. It computes the probabilities of cluster memberships by maximizing the log likelihood of the data that were generated by the Gaussian mixture model. Other possible probability distributions are the Normal, log-normal, and the Poisson distributions, all of which provide a statistical measure of the image information. Another important note to mention is the EM method can be applied to continuous or categorized data. Categorization of data will allow for weighting of data; thus areas of an image that are considered less important can be successfully separated from the key areas.

- ***Gradient Vector Snakes***

The gradient vector flow (GVF) technique utilizes the snake algorithm to detect the borders of the lesion areas of an image [146]. It uses luminance blurring (for correction of color saturation) to remove noise from a dermatology image during its initialization process. The snake border is then drawn based on the outline of the blurred lesion. It then successively applies the blurring technique to deform the snake border, which would eventually conform to the shape of the lesion. However, this technique is not accurate as weaker outlines/edges in the original image would fade away after the initial blurring step, thus the snake border will not conform to the original shape of the lesion.

Use of a double image buffer to rectify this problem has improved the overall accuracy of the snake border, but the error rates are still not acceptable.

- ***Variable Weighting Factor Markov Random Field Model***

Markov Random Field (MRF) models have been applied to the field of digital image restoration of degraded images [147]. The image restoration process is closely related to image segmentation and boundary detection. MRF models are a powerful technique and consist of two layers, a region labeling component and a feature modeling component, with a constant weighting factor to combine the two. Many researchers have mentioned that the improved MRF model is able to produce accurate results in image segmentation.[148]

3.3.6.3. The Importance of the Pathology report for primary melanoma in the modern era

- ***3.3.6.3.1. Introduction***

Since the 1960s, the clinical characteristics of melanoma, its histopathology and its biological basis have been the subject of intense study at pigmented lesion clinics in North America, Europe, and Australia [149]. The purpose of this future work is to bring to the attention of anatomic pathologists the essential characteristics of the pathology report for primary cutaneous melanoma in the modern era.

The researchers in [150] [151] recognized the importance of anatomic compartment of attack to the melanoma. In 1970 the researcher in [70] described the use of an optical micrometer to specifically measure primary tumor thickness. These two parameters became the recognized prognostic variables available to histo-pathologists in predicting the biological behavior of melanoma [152]. Throughout the 1980s, advances in biology and pathology of melanoma made it apparent that tumors undergo specific growth phases. In particular, the radial growth phase was shown to be unable to generate metastatic generating metastatic events unless a worsening was seen at the primary site. Further work carried out by Massachusetts General Hospital and the University of Pennsylvania established that the existence of microscopic satellites and brisk mitotic activity could modify the behavior of a tumor.[149]

- ***3.3.6.3.2. Dermoscopy Diagnosis of Skin Lesions***

Successful treatment of melanoma, by surgical removal depends upon the early detection of the lesion. The challenge is to diagnose and remove all malignant lesions at an early stage and at the same time minimize the unnecessary removal of benign lesions. However, the differentiation of early melanoma from other non-malignant pigmented skin lesions is

not easy even for the experienced dermatologists. In several cases primary care physicians underestimated melanoma in its early stage [153].

Visual inspection with the naked eye has a relatively low sensitivity and reliability in detecting early melanoma. In this context, several non-invasive diagnostic modalities, such as dermoscopy, total body photography, and reflectance confocal microscopy (RCM), have emerged in recent years that are aimed at increasing diagnostic accuracy. The main developments in this field are the integration of dermoscopy and digital photography into clinical practice.

Dermoscopy (also known as dermoscopy or epiluminescence microscopy) helps in the visualization of anatomical structures within the epidermis and papillary dermis that cannot be visualised by the naked eye. This is achieved with the use of a hand held dermatoscope to magnify the skin surface and reduce the refraction of light by the corneal layer. It has been reported that dermoscopy training for primary care physicians can improve their ability to correctly refer individuals with suspicious lesions and decrease the rate of excision or referral in benign skin lesions.

Pathologists are using the microscopic to perform histo-pathological examination and provide diagnostic information based on their observations. Figure 3.26 shows manual segmentation performed by pathologists using the ‘Aperio Image Scope’ software [154], but there are many problems associated with manual segmentation, being time-consuming and depends on the skill and training of the pathologist doing the analysis. Therefore automatic segmentation needed to bring some kind of consistent accuracy within the process. On the technical report in [155], the automated cancer diagnosis consists of three main computational steps: pre-processing, feature extraction, and diagnosis. The aim of the diagnosis step is (i) to distinguish benignity and malignancy or (ii) to classify different malignancy levels by making use of extracted features. This step uses statistical analysis of the features and machine learning algorithms to reach a decision.

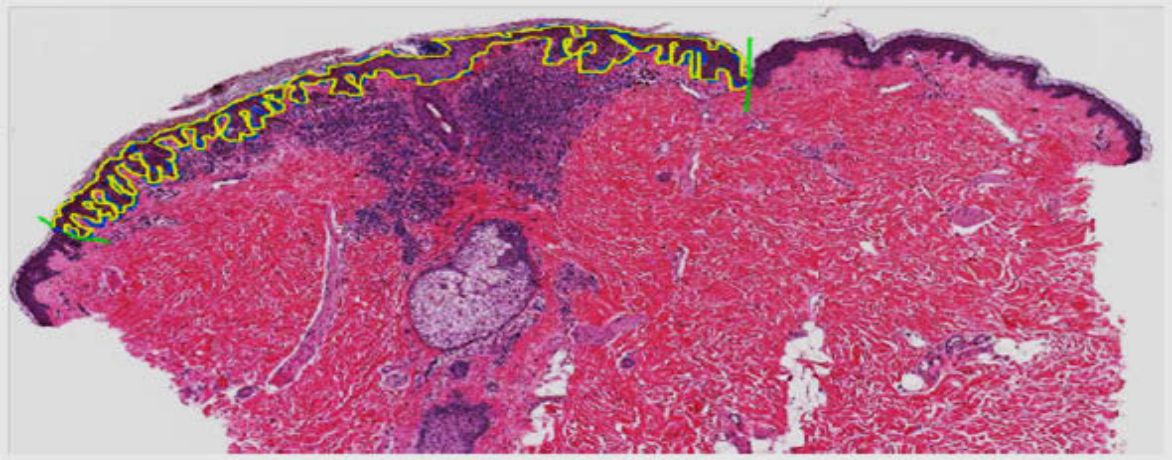


Figure 3.26. Manual Segmentation by trained pathologists using the 'Aperio Image Scope' software [156].

➢ *3.3.6.3.3. Histo-pathological Images*

Malignant melanomas are the most serious form of skin cancer accounting for the majority of skin cancer related deaths. Histo-pathological images of skin tissues are analysed for detecting various types of melanomas. The automatic analysis of these images can greatly facilitate the diagnosis task for dermato-pathologists. The first and primary step in automatic histo-pathological image analysis is to accurately segment the images into dermal and epidermal layers along with segmenting other tissue structures such as nests and melanocytic cells which indicate the existence of cancer.

An efficient automatic segmentation procedure has been proposed for histo-pathological images of skin tissue. The proposed technique is based on the Orientation Sensitive Fuzzy C-Means (OS-FCM) algorithm along with other refinement techniques [154].

➢ *3.3.6.3.4. Radial Growth Phase Melanoma*

The radial growth phase growing in a horizontal array as single cells and small nests is predominantly in an intra-epidermal location.[149] The radial growth phase is incapable biologically of generating metastatic events, save for specific and uncommon circumstances. In this example, there is pagetoid spread of fully transformed malignant epithelioid melanocytes within the epidermis, see (Figure 3.27) the basement membrane zone in concert with a level II component in the papillary dermis (i.e. Clark level II). Note that the dermal component comprises small nests of neoplastic melanocytes (red arrow figure.3.27) with a similar cytology to those seen above the basement membrane zone which separates them from the overlying epidermal component (black arrow figure 3.28).

The dermal nests of radial growth phase melanoma are smaller than those seen in the overlying epidermal component.

- ***3.3.6.3.5. Vertical Growth Phase***

Clark in [157] described the early vertical growth phase and distinguished it from the radial growth phase by criteria that included, most importantly, the existence of a dominant nest within the papillary dermis. This expansile nest is larger than any nest within the epidermis or the surrounding dermis (Figure 3.28). Clark also suggested that the nodule should have 25–50 cells to aid in its identification, and that the cells comprising this nodule are cytological separate from those of the intraepidermal component by virtue of their different characteristics, such as shape, size, nuclear or cytoplasmic features, or the presence or absence of pigment.

The vertical growth phase means the point at which the melanoma becomes biologically capable of producing metastatic events. Note that there is a main able to expand dermal nest of neoplastic melanocytes (arrow) which is larger than any of the overlying the quiz at the dermal–epidermal junction or within the epidermis.

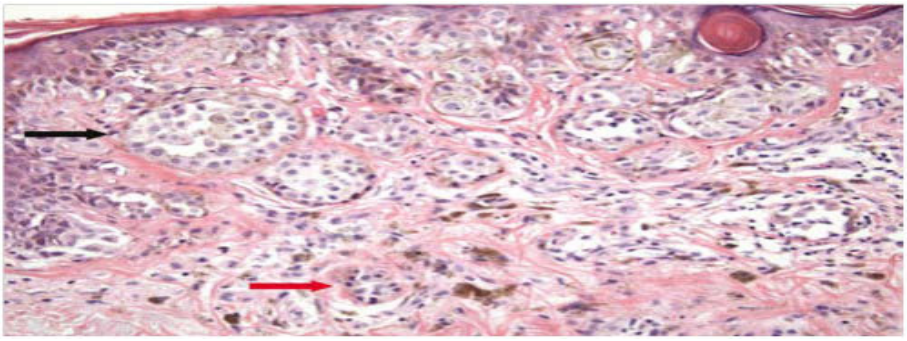


Figure 3.27 Showed the Radial growth phase melanoma.

- ***3.3.6.3.6. Histologic type (The microscopic structure) of melanoma***

Since the 1970s melanomas have been histologically sub-classified as belonging to four major groups: superficial spreading, lentigo malign, acral lentiginous, and nodular melanoma.[151], [158], [159] Some observers suggest that the relatively greater survivorship seen in lentigo maligna melanoma vs. the other histologic subtypes is reflective of smaller thickness at time of diagnosis.[157], [160], [161]

The lifetime risk for developing invasive disease in a lesion of lentigo malign has been estimated to be only in the 5% range. [162] The early work of Fiddler[163] suggests that

those are the blood vessels capable of accommodating the large tumor bolus conducive to the establishment of the metastatic seed and so suggest that this is a critical factor in the mortality of any given acral melanoma.

- ***3.3.6.3.7. Thickness (Breslow)***

In the accumulated experience of several investigators over a long period of time, tumor thickness as measured by ocular micrometer has emerged as the most powerful predictor of outcome in primary cutaneous melanoma [70], [164], [157], [8], [165], [166], [167], [168, 169], [170], [171], [7], [172]. The Breslow measurement is taken from the epidermal surface or, in the event that the surface is ulcerated, from the base of the ulcer, and is made with a calibrated ocular micrometer. [70] Long clinical follow-up indicates that melanoma thickness is a continuous variable associated with dropout of survivors beyond 10 years. The Breslow measurement is the most important means of prognosticating mucosal melanomas that lack the anatomic compartmentalization seen in the skin. [173] In most prognostic studies, the measured depth emerges as the most powerful independent factor for prediction of lymph node metastasis and survival.[174]

- ***3.3.6.3.8. Microscopic Satellites***

A review of the microscope slides of the primary tumours for clinical Stage I melanoma discovered that primary lesions displayed invasion with "microscope satellites" [175]. characterized by reticular dermal and/or sub-circular nodules of tumor greater than 0.05mm in width separated from the main vertical growth phase component (Figure 3.28), are associated with lymph node metastasis, and with decreased disease-free and overall survival [157],[7] ,[176], [177], [178], [179], Clark et al.[10]

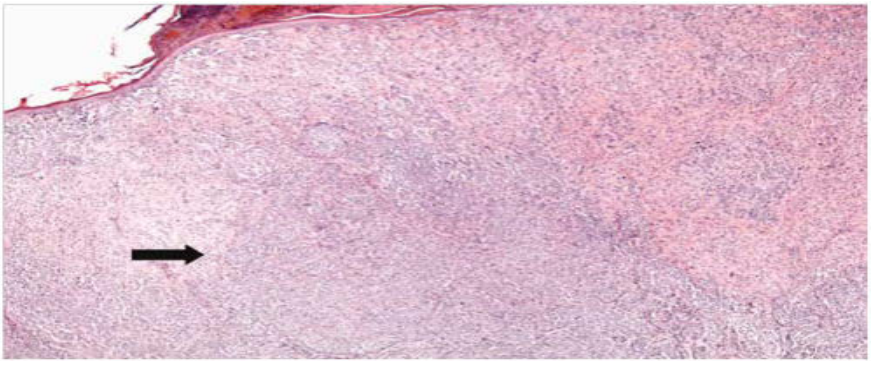


Figure 3.28 showed the Vertical growth phase melanoma

found a reduction of actuarial 8-year survival from [180] to 40% when satellites were identified. (Figure 3.29) It has been suggested that this is a difficult criterion to assess, as tumor tongues radiating from the main vertical growth phase nodule may mimic satellites due to artifacts of sectioning.[181]

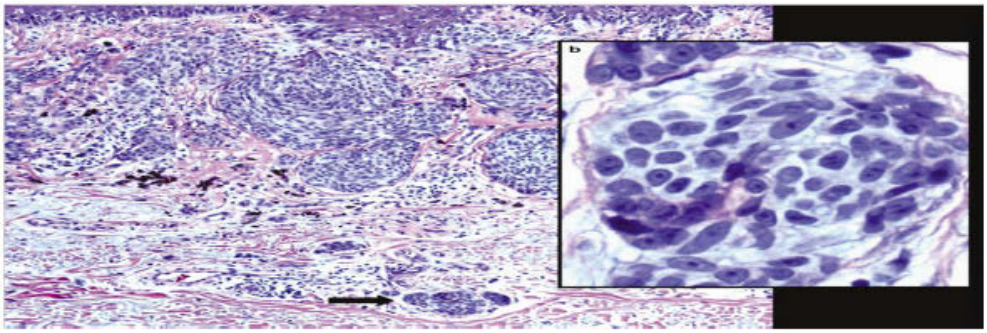


Figure 3.29 Microscopic satellites, (a). Shows neoplastic group is discontinuous (arrow) from the overlying vertical growth phase component. Shows the melanocytes must have a malignant cytomorphology (b).

➢ *3.3.6.3.9. Blood Vessel and Lymphatic Invasion*

The presence or absence of blood vessel and lymphatic invasion should be reported (Figure 3.30). Regression of over 75% of the tumor volume of a melanoma is considered a bad prognostic sign. Regression is characterized by haphazard, pattern-less fibrosis, accompanied by vertically oriented, ectatic blood vessels (straight arrow), melanophages (curved arrow) and patchy lymphocytic infiltration of the stroma. In complete, as opposed to partial regression, there is a total absence of neoplastic melanocytes in the dermis and in the overlying epidermis.

➢ *3.3.6.3.10. What are the pathologic features of a melanoma?*

- Cellular atypia: cancerous melanocytes have atypical cellular features compared to normal melanocytes, such as:
  - Increased nuclear: cytoplasm ratio
  - Found migrating throughout the layers
- Nests: atypical melanocytes are usually found in nests suggesting uncontrolled growth
- Lymphocytic infiltrate: the presence of immune cells suggests the attempt of an immune response to a malignancy

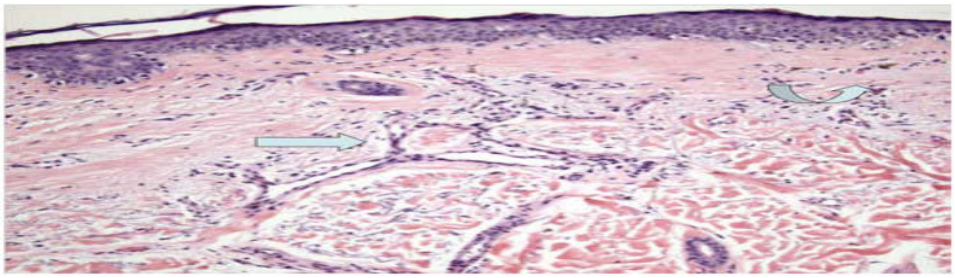


Figure 3.30 Regression: Regression of over 75% of the tumor volume of a melanoma is considered a bad prognostic sign.

The above discussion showed the essential elements of the pathology report for primary melanoma should include all of the statistically proven analytical variables.

3.3.6.4. Clinical Diagnosis using the Dermoscopy

Dermoscopy is a non-spreading method that allows in vivo evaluation of colors and microstructures of the epidermis, the dermo-epidermal junction, and the papillary dermis not visible to the naked eye. The pigmented skin lesion is covered with liquid (usually oil or alcohol) and examined under a specific optical system during a dermoscopy assessment. Recently by using polarized light in dermoscopy with LED light with polarization, immersion liquid is no longer necessary, and some of these instruments do not need direct skin contact. Researchers have reported excellent agreement for most dermoscopy colors, with the exception of blue-white veil and pink (red) color when comparing non-polarized and polarized light [182]. They concluded that the polarized light improves the visualization of red areas and vessels, especially the latter with non-contact dermoscopy [182].

Figure 3.31(a) shows an image of a lesion under the clinical (naked eye) view and 3.31(b) shows the same lesion under a dermoscopy with oil immersion. Important features are marked in the image. This methodology is reserved for experienced clinicians because of the complexity involved. See figure 3.32, for more details.

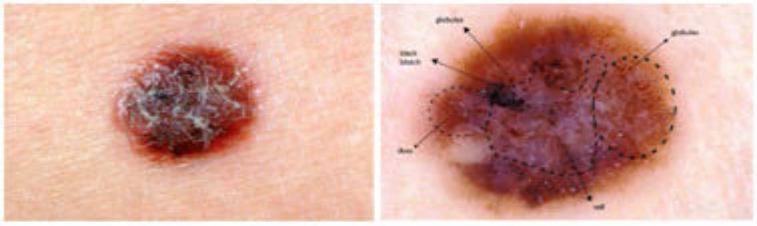


Figure 3.31: a) An image of a lesion under clinical view (naked eye). b) Shows the same lesion under a dermoscopy with oil immersion.

The four clinical diagnostic methods, which have received the largest interest among dermoscopy users, will be briefly reviewed. Initially, pattern analysis was the first dermoscopy method presented for diagnostics of pigmented skin lesions [183]. Firstly the International Dermoscopy Society (IDS) modified and refined the pattern analysis method. Then the other three methods with the largest impacts are ABCD rule, Menses method and the 7-point checklist which have been described in previous sections. All of these methods showing similar results on sensitivity for melanoma, but specificity differed slightly in favor of pattern analysis [184]. As a result, the accuracy of these simplified systems is somewhat lower than the full system, pattern analysis. For more details on the dermoscopy of pigmented skin lesions please refer to [183].

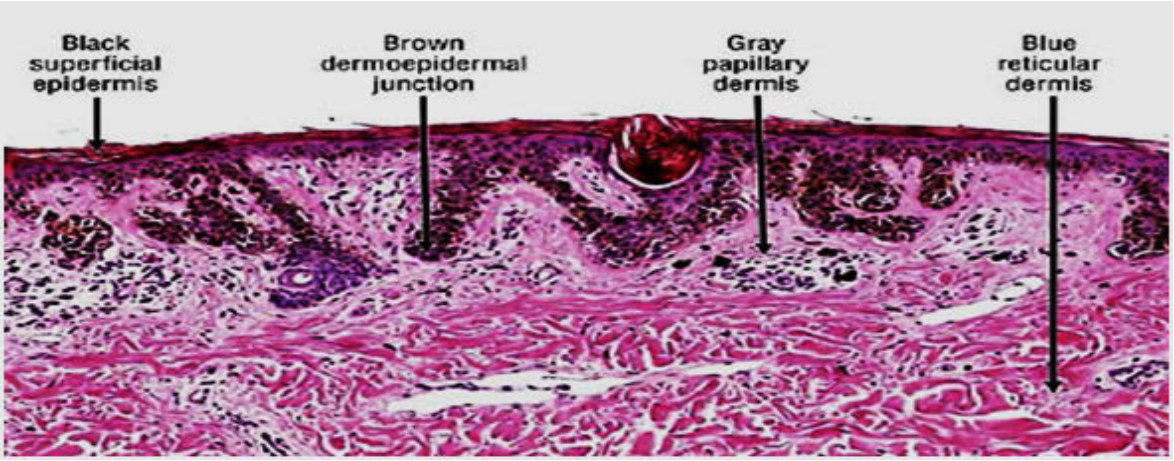


Figure 3.32: a) Colors allow the physician, to some extent, to draw conclusions about the localization of pigmented cells within the skin. Black and brown indicate pigmentation in the epidermis, while grey and blue correspond to pigmented cells within the superficial and deep dermis, respectively [185].

Figure 3.33 shows different kinds of dermoscopy and they are divided into analogue and digital types. Analogue types are more widely used however digital types are easier to take and store dermoscopy images. The magnification in all dermoscopy was identical. The authors show that in the newer dermoscopy, the image quality with regard to color and visible differential structures are distinctly improved compared to the dermoscopy (Heine Delta 10R) with only one light source[13].

Figure 3.33: Figures (a, b, c, d, and e) show analogue dermoscopy. All of them, except (a), are attachable to digital cameras to function as digital dermoscopy. Figures (f, g, and h) show DinoLite, Handy scope, and Dermoscopy which are modern digital dermoscopy.

3.3.6.5. Commercial Computer Aided Diagnostic Systems

Advances in computer science and applied mathematics are now allowing the basic technology of dermoscopy to be extended to a more complex application: forecasting whether observed spots are cancerous. Several groups are working toward computer-assisted diagnosis of melanoma using different mathematical and analytical strategies. Computer software can be used to library skin images and allow remote diagnosis and reporting by a dermatologist (digital epiluminescence microscopy, teledermoscopy, mole mapping). Mole Map NZ is such a system which is used for archiving and patient management. If we want to develop computer-aided diagnosis systems, we should address the important step of the pattern analysis approach which is the segmentation and analysis of the dermoscopy structures

3.3.6.6. Solar Scan Melanoma Monitoring

Solar Scan is a device priced for use by individual practitioners and developed by Australian startup Polartechnics, with assistance from Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO) and the Sydney Melanoma Unit[186]. Solar Scan uses image analysis, which has turned diagnostic criteria used by doctors into defined decision trees that lead toward or away from a decision to help general physicians (GPs) diagnose melanoma[187]. The device works by capturing an image of a patient's skin spot using an object shaped like a hairdryer with a built-in surface microscope. Image analysis software removes unnecessary things from the image like hairs and oil bubbles and analyses the spot's features such as its shape and color. Solar Scan then compares the features against images of melanomas and non-melanomas in a database, returning an advice to a GP. A record of the spot's status can be stored in Polartechnics' Body Map software so that it can be rechecked another time if necessary[187].

3.3.6.7. MelaFind

In figure 3.34, MelaFind is a multispectral digital dermoscopy with a specialized imaging probe and software to assist with differentiation between early melanoma and other skin lesions[188] [186]. A hand-held imaging device that shines light of 10 different, specific, wavelength bands is used to collect data. Proprietary processing software is used to extract specific features from the images [189] [190]. The software determines the edge of the lesion and generates a 10 digital image sequence. The images are then analyzed for wavelet maxima, asymmetry, color variation, perimeter changes, and texture changes, and the output is a binary recommendation of whether or not to perform a biopsy [191]. The results of a clinical trial were published in April 2008, leading to submission of the MelaFind to the FDA and recently it has been approved for use in the United States.

We believe that with new advances in dermoscopy, the texture analysis problem should be changed to an object recognition problem that involves identification, segmentation and recognition of individual shapes and structures in skin lesions and this will be employed to identify and classify specific dermoscopy structures. It is hoped that this will lead to the development of many new approaches that can be included to increase the diagnostic accuracy of automated systems.

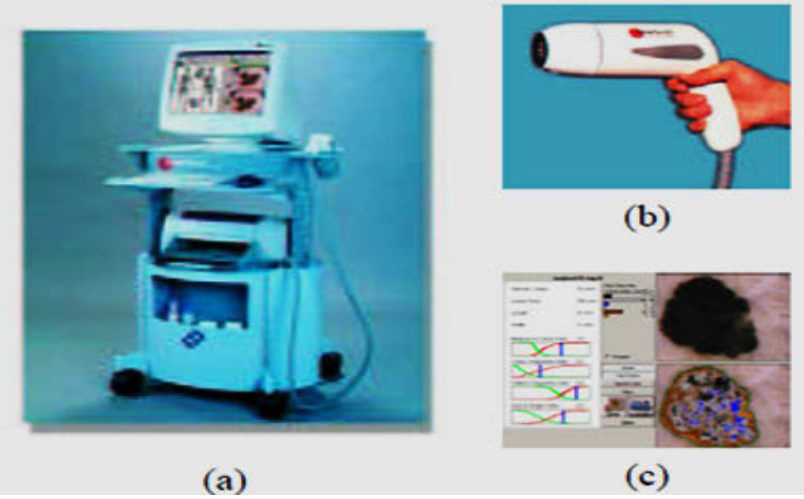


Figure 3.34. The Solar Scan instrument: (a) global appearance, (b) camera, (c) user interface [192]. Source: www.medgadget.com

Chapter summary

This chapter 'Literature Review' states that: the diagnosis of a doctor is based on visual inspection and may be reliable, although diagnosis procedure takes a lot of time and effort. These routines can be automated rather than the previous procedure saving doctors lots of time and would ensure that the diagnoses are more accurate. There are other good means of storing diagnostic information alongside computer means. These means will assist further investigations or the creation of new methods of diagnosis based on a vision system.

This chapter includes briefly reviewed sections like basics of vision system and structure of the human eye; image processing: human skin imaging; human skin biology; different types of common skin lesions, the new imaging system called digital dermoscopy and different clinical diagnostic method. Lastly the current computer aided diagnostic systems and other commercial equipment like Solar Scan Melanoma Monitoring and MelaFind are also reviewed.

The discussion in this chapter showed the essential elements of the pathology report for primary melanoma including all of the statistically proven analytical variables.

- As explained before in the modern era, clinicians and oncologists must recognize specific risk factors for metastatic disease so as to guide adjuvant chemotherapy or immunomodulatory therapy and further surgical therapy including sentinel and completion lymphadenectomy when necessary.

- It is therefore considered essential that patients receiving melanoma diagnoses be provided with accurate and comprehensive diagnostic assessment upon which these novel therapeutic strategies will be based.

CHAPTER 4

Pre-processing and Segmentation

4.1. Introduction

The rate of skin cancer (malignant melanoma [MM]) has increased dramatically over the past few decades. Researchers found that, Primary avoidance and early detection continue to be of paramount importance in addressing the public health threat of skin cancer[193] [79].

Doctor's diagnosis is reliable, but this procedure takes lots of time and effort. These routines can be automated. It could save time and could help the diagnosis to be more accurate[80]. Besides using computerized means there are good opportunity to store information with diagnostic information in order to use it for further investigations or foundation of new methods of diagnosis.

The main problem is that the diagnosis is highly dependent on individual judgment, and is just reproducible [113] [194]. Several counting systems and algorithms such as the ABCD-E rule [119] [195, 196], the seven-point checklist[197] [198], three-point checklist[199] and the Kenzie's method [200] [201] have been proposed to improve the diagnostic performance of less experienced clinicians. Although this simplification has enabled the development of these diagnostic algorithms with good accuracy but still they showed problems that have not yet been solved. The most important is that the purpose for which they were designed was not achieved, because the within- and between-observer concordance is very low, even for expert observers[119] [111] [202] [203]. Despite extensive research in investigating the varied presentations and physical characteristics of melanoma, the clinical diagnostic accuracy remains sub optimal. Thus, a growing interest has developed in the last two decades in the automated analysis of digitized images obtained by ELM techniques to assist clinicians in differentiating early melanoma from benign skin lesions.

4.1.1. Digital Image Representation

A digital image is a binary representation of a two-dimensional image. The digital representation is an array of picture elements called pixels. Each of these pixels has a numerical value which in monochromatic images represents a grey level. An image may be defined as a two-dimensional function,(x, y) where x and y are spatial coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the *intensity* of the image at that point. The term *grey level* is used often to refer to the intensity of monochrome images. Color images are formed by a combination of individual 2-D images. For example, in the RBG color system, a color image consists of three (red, green, and blue) individual component images. this is why many of the technics developed for monochrome images can be extended to color images by processing the three component images individually[27].

An image may be continued with respect to the x and y coordinates, and also in amplitude. Converting such an image to digital form requires that the coordinates, as well as the amplitude, be digitized. Digitizing the coordinate values is called *sampling*; digitizing the amplitude values is called *quantization*. Those, when $x, y,$ and the amplitude values of f are all determinate discrete quantities, we call the image a *digital image*.

4.1.2. Image Types

The image types as supported in Matlab / (IPT) Image Processing Toolbox are divided into four types of images:

- ***Intensity images***: consist of a single data matrix, each element of which denotes the intensity of the corresponding pixel (***medical imaging***).
- ***Binary images***: stored as a logical array. (i.e. each pixel is turned on or off), and is represented by a ***byte***.
- ***Indexed images***: consist of a data matrix and a couloir map matrix.
- ***RGB images***: sometimes referred to as a true-colour image, is stored in Matlab as an m-by-n-by-3 data array that defines red, green, and blue colour components for each individual.

Most monochrome image processing operations are carried out using *binary or intensity* images, so our focus is on these two image types.

1. **Intensity images**

An *intensity image* is a data matrix whose values had been scaled to represent intensities. When the elements of an intensity image are of ***class unit8***, or ***class unit16*** they have integer values in the range $(0, 255)$ and $(0, 65535)$, respectively. If the image is of ***class double***, the values are floating – point numbers. Values of scaled, class double intensity images are in the range $[0, 1]$ by convention.

2. **Binary images**

Binary images have a very specific meaning in Matlab. A *binary image* is a *logical array* of $0's$ and$1's$. Thus, an array of $0's$ and $1's$ whose values are of data class, say, unit8 is not considered a binary image in Matlab. A numeric array is converted to binary using function logical. Thus, if A is a numeric array consisting of $0's$ and$1's$, we creating a logical array B using the statement in equation 4.1:

$$B = logical\ (A) \qquad 4.1$$

If A contains elements other than $0's$ and $1's$, use of the logical function converts all nonzero quantities to logical $1's$ and all entries with value 0 to logical $0's$. Using relational and logical operators also creates logical arrays. To test if an array is logical we use the islogical function, equation 4.2:

$$islogical\ (C) \qquad 4.2$$

If C is a logical array, this function returns a1. Otherwise it returns a . Logical arrays can be converted to numeric arrays using the data class conversion functions.

4.1.3. General approach of developing a Computer Aided Design (CAD) system

Application of computational intelligence methods helps physicians as well as dermatologists in faster data process to give better and more reliable diagnoses. Studies related to the automated classification of pigmented skin lesion images have appeared in the literature as early as 1987[204]. After some successful experiments on automatic diagnostic systems for melanoma diagnosis[204] [205] [206] , utility of machine vision and computerized analysis is getting more important every year. The importance of the topic is obvious if we analyze the enormous quantity of researches related with the melanoma diagnosis. Several computerized diagnostic systems have been reported in the literature where different border detection, feature extraction, selection and classification algorithms are used. Some researchers [205] [207] [208] reviewed and tried to critically

examine image analysis techniques for diagnosis of skin cancer and compared diagnostic accuracy of expert's dermoscopists with artificial intelligence and computer aided diagnosis. More research, however, is needed to identify and reduce worries in the automatic decision support systems to improve diagnosis accuracy.

Computer aided decision-support tools are important in medical imaging for diagnosis and evaluation. Analytical models are used in a variety of medical domains for diagnostic and prognostic tasks. These models are built from experience which establishes data acquired from actual cases. The data can be pre-processed and expressed in a set of rules, such as it is often the case in knowledge-based expert systems and accordingly serve as training data for statistical and machine learning models.

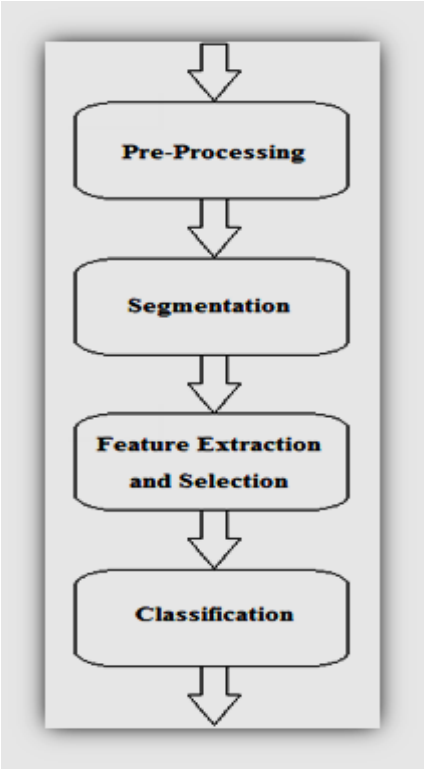
The general approach of developing a CAD system for diagnosis of skin cancer is to find the location of a lesion and also to determine an estimate of the probability of a disease. The first step in this research was to establish a standard general scheme of a CAD system for skin lesions. This scheme is shown in figure 4.1.

The inputs to the Computer Aided Systems are digital images obtained by ELM, with the possibility to add other acquisition system such as ultra-sound or confocal microscopy etc.

As showed in figure 4.1. In the first phase a pre-processing of image is done which allows us to reduce the ill effects and various objectives like hair, the true color Pathological Digital Images filterd, then followed by the detection of the lesion by image segmentation technique as shown in figure 4.2. Once the lesion is localized, different chromatic and morphological features can be quantified and used for classification.

Differentiation of malignant melanoma images demands very fast image processing and feature extraction & classification algorithms. A detailed research is necessary to make the best choice and setting the benchmarks for diagnostic system development and support. Next image processing modules consist of four section focuses on the description of the major steps that may be involved in skin cancer diagnosis, which is connected to a medical knowledge database can generate a computerized diagnosis, suggesting whether the lesion is benign or malignant.

Original Pathology Digital Image



Evaluation output

Figure 4.1 Showed: Four stages CAD system for skin lesions.

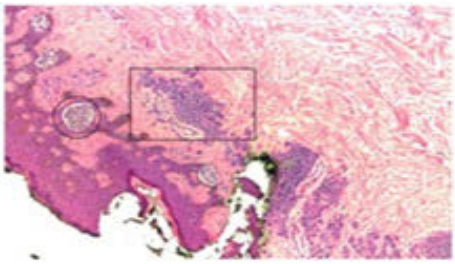


Figure 4.2 true color Pathological Digital Image.

4.2. Pre-Processing Stage

The main processing step towards a complete analysis of pigmented skin lesion is to differentiate the lesion from the healthy skin. Detection of the lesion is a difficult problem in dermatoscopic pathology images as the transition between the lesion and the surrounding skin is smooth and even for trained dermatologist, it is a challenge to distinguish accurately. It has been observed that dermoscopy images often contain objects such as rough brightness dermoscopy gel black frames, ink markings, rulers, air bubbles, as well as essential development features that can affect border detection such as blood

vessels, hairs, and skin lines and texture. These objects and unrelated elements complicate the border detection procedure, which results in loss of accuracy as well as an increase in computational time. Thus, it requires some pre-processing steps to facilitate the segmentation process by removal of unwanted objects.

Everything that might corrupt the image and consequently affect the results of image processing must be localized and then removed, masked or replaced. It may include image resizing, masking, cropping, hair removal, and conversion from RGB color to intensity grey image. It is done to reduce noise and the effect of reflection objects. It is meant to facilitate image segmentation by filtering the image and enhancing its important features.

The most straightforward way to remove these artefacts is to smooth the image using a general purpose filter such peer group filter (PGF)[209], mean filters, median filter [210] [211, 212], Gaussian filters[213] [214] or anisotropic diffusion filters (ADF). An alternative solution is to use filters that treat the pixels as vectors[215]. Another notable thing is setting of mask size proportional to the image size to manage a trade-off between smoothing of image and blurring of edges. Regardless of taking care of all the fore mentioned things, it is still not guaranteed to get an image free of all objects.

Many studies suggested their works, very few[216] [217] [217] discussed different aspects of objects together but none of them have discussed all cases of objects or the factors that complicate the detection of borders in dermoscopy images with insufficient contrast. The contrast of image is enhanced to ensure that edges of the lesion are distinction.

Recently, it has been proposed a contrast enhancement method based on independent histogram pursuit (IHP) [218]. An easy, yet powerful way to enhance the image contrast is histogram stretching, a mapping of the pixel values onto (0, 255). Another very popular technique is histogram equalization, which modifies pixel values to achieve a uniform distribution. Holomorphic filtering [219], FFT and high pass filter can be used to compensate for rough illumination or specular reflection variations in order to obtain the high contrast lesion images.

For the removal of black frames produced in the digitization process, the author in [220] [221] proposed an iterative algorithm based on the lightness component of the HSL (Hue-Saturation-Lightness) color space. In order to remove air bubbles and dermoscopy-gels, adaptive and recursive weighted median filter developed by the author in [222] that can be

utilized. This type of median filter has an edge determined capability. A method that can remove bubbles with bright edges was introduced in [223] where the authors utilized a morphological top-hat operator followed by a radial search procedure. Line detection procedure based deference of Gaussian on the 2-D results of Gaussian (DOG) [224] and exemplar-based object removal algorithm[225] can be used for removing dark lines like ruler marking. In most cases, image smoothing effectively removes the skin lines and blood vessels.

Most images are affected to some extent by noise that is unexplained variation in data: disturbances in image intensity which are either uninterpretable or not of interest. Image analysis is often simplified if this noise can be filtered out[226]. Researchers summarised the following key points about filters [226]:

• Filters re-evaluate the value of every pixel in an image. For a particular pixel, the new value is based of pixel values in a local neighbourhood, a window centered on that pixel, in order to: reduce noise by smoothing, and/or enhance edges.

• Filters provide an aid to visual interpretation of images, and can also be used as a sign to further digital processing, such as segmentation. Filters are linear or nonlinear,

– Linear filters are well understood and fast to compute, but are incapable of smoothing without simultaneously blurring edges.

– Nonlinear filters can smooth without blurring edges and can detect edges at all orientations simultaneously, but have less secure theoretical foundations and can be slow to compute.

The experiment in this thesis used three filters winner filter, Gabor Filter and adaptive median filter.

4.2.1. Wiener Filter:

Wiener filtering is a general way of finding the best reconstruction of a noisy signal [227]:

- applies in any orthogonal function basis,
- Different bases give different results,

One way of filtering is to just set components with too much noise to zero. Wiener filtering is better; it gives the optimal way of narrowing off the noisy components, so as to give the best reconstruction of the original signal. It can be applied in spatial basis (delta functions, or pixels), Fourier basis (frequency components), wavelet basis, etc. Different bases are not equivalent, because, in particular problems, signal and noise distribute differently in them. A lot of signal processing work is to find the right basis for particular problems in which signal is most concentrated. The 2-D adaptive noise-removal filtering, wiener2 low pass-filters a grayscale image that has been corrupted by constant power additive noise. Wiener2 uses a pixel wise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel. Equation 4.3 showed Matlab form [228]:

J = wiener2 (I, [m n], noise) 4.3

Where filters the image I using pixel wise adaptive Wiener filtering, using neighbourhoods of size $m - by - n$ to estimate the local image $mean$ and $standard\ deviation$. The $[m\ n]$ argument, m and n default values is 3 or 5. The additive noise (G Gaussian white noise) power is assumed to be noise.

Wiener filtering is best when the basis most cleanly separates signal from noise, and that's clear in spatial (pixel) basis, when Wiener filter is applied to the difference between an image and a smoothed image.

Spatial (pixel) basis [227]: It doesn't make sense to use the pure pixel basis, because there is no particular separation of signal and noise separately in each pixel. But a closely related method is to decompose the image into a smoothed background image (x), and then to take deviations from this as estimating signal power (S) with noise power (N), as showed in equation 4.4:

$$X_{ij} = \frac{1}{N_{hood}} \sum_{hood} x_{ij} \quad 4.4$$

'Hood' might be a 5x5 neighborhood centered on each point, as showed in equation 4.5 & 4.6:

$$\left\langle {S_{ij}}^2 + {N_{ij}}^2 \right\rangle = 1/\frac{1}{hood} \sum_{hood} (x_{ij} - X_{ij})^2 \quad 4.5$$

$$\widehat{x_{ij}} = X_{ij} + \frac{\left\langle {S_{ij}}^2 + {N_{ij}}^2 \right\rangle - \left\langle {N_{ij}}^2 \right\rangle}{\left\langle {S_{ij}}^2 + {N_{ij}}^2 \right\rangle} (x_{ij} + X_{ij}) \quad 4.6$$

From equation 4.6, you can adjust Wiener parameter part ($N_{ij}{}^{2}$).

4.2.2. Gabor Filter

One of the recent studies on mathematical modeling of visual cortical cells [229] suggested a tuned band pass filter bank structure. These filters are found to have Gaussian transfer functions in the frequency domain. Thus, taking the Invers Fourier Transform of this transfer function, it can get a filter characteristics closely resembling to the Gabor Filters. The Gabor filter is basically a Gaussian (with variances Sx and Sy along x and y-axes respectively) modulated by a complex sinusoid (with center frequencies U and V along x and y-axes respectively).

Gabor filters are used mostly in shape detection and feature extraction in image processing. The Gabor Filters have received extensive attention because the characteristics of certain cells in the visual cortex of some living thing can be approximated by these filters. In addition these filters have been shown to possess optimal localization properties in both spatial and frequency domain and thus are well suited for quality segmentation problems[230] [231]. Gabor filters have been used in many applications, such as quality segmentation, target detection, fractal dimension management, document analysis, edge detection, retina identification, image coding and image representation[231]. A Gabor filter can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope. It can be written as showed in equation 4.7:

$$h\,(x,y) = s\,(x,y)g\,(x,y) \qquad 4.7$$

Where s (x,y) is a complex sinusoid, known as a carrier, and $g\,(x,y)$ is a 2-D Gaussian shaped function, known as envelope. The complex sinusoid is defined as follows in equation 4.8,

$$s\,(x,y) = e^{-J2\omega\,(u_0 x + v_0\,y)} \qquad 4.8$$

Where (u_0, v_0) define the special frequency of the sinusoid and $w = 2\pi$.

The 2-D Gaussian function is defined as displayed in equation 4.9,

$$g\,(x,y) = {}^{1}\!/_{\sqrt{2\pi\sigma}}\, e^{-0.5\,(\frac{x^2}{\sigma_x{}^2} + \frac{x^2}{\sigma_y{}^2})} \qquad 4.9$$

Thus the 2-D Gabor filter can be written as in equation 4.10:

$$h(x,y) = e^{-0.5\left(\frac{x^2}{{\sigma_x}^2}+\frac{x^2}{{\sigma_y}^2}\right)} e^{-j2\omega(u_0 x+ v_0 y)} \qquad 4.10$$

$$= g(x,y) e^{-j2\pi (u_0 x+ v_0 y)}$$

The frequency response of the filter is:

$$H(u,v) = G(u-u_0, v-v_0)$$

$$H(u,v) = 2\pi\sigma_x\sigma_y \left[e^{-2\pi^2[(u-u_0)^2 {\sigma_x}^2+(v-v_0)^2{\sigma_y}^2]}\right] \qquad 4.11$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} e^{-0.5\left[\frac{(u-u_0)^2}{{\sigma_u}^2}+\frac{(v-v_0)^2}{{\sigma_v}^2}\right]} \quad \text{Where} \quad \sigma_u \frac{1}{2\pi\sigma_x}, \sigma_v \frac{1}{2\pi\sigma_y} \qquad 4.12$$

This is equivalent to translating the Gaussian function by (u_0, v_0) in the frequency domain. Thus the Gabor function can be thought of as being a Gaussian function shifted in frequency to position (u_0, v_0) i.e. at a distance of $\sqrt{{u_0}^2 + {v_0}^2}$ from the origin $\tan^{-1}\frac{u_0}{v_0}$. In the above equations (4.11) & (4.12), (u_0, v_0) are referred to as the Gabor filter spatial central frequency. The parameters σ_x, σ_y are the standard deviation of the Gaussian envelope along X and Y directions and determine the filter bandwidth.

Gabor Filter Bank Application

Gabor bank of filters used to get feature images. The filters used with two different values of angle and two different frequencies of cosine function in the bank. It means each image can be convolved by 4 Gabor filters as described by[232] as shown in figure 4.3.

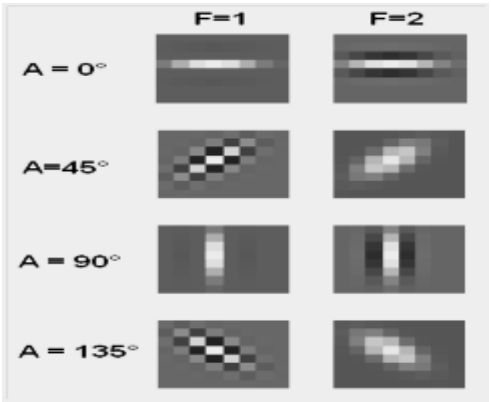


Figure 4.3. Gabor filter bank for hair detection:

8 filters of 2 frequencies F1 and F2 and for 4 angles A = 0, 45, 90 and 135 grades.

4.2.3 Median filter:

The median filter is normally used to reduce noise in an image, somewhat like the mean filter. However, it often does a better job than the mean filter of preserving useful detail in the image. It works like the mean filter; the median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.) Refer to the example in figure 4.4.

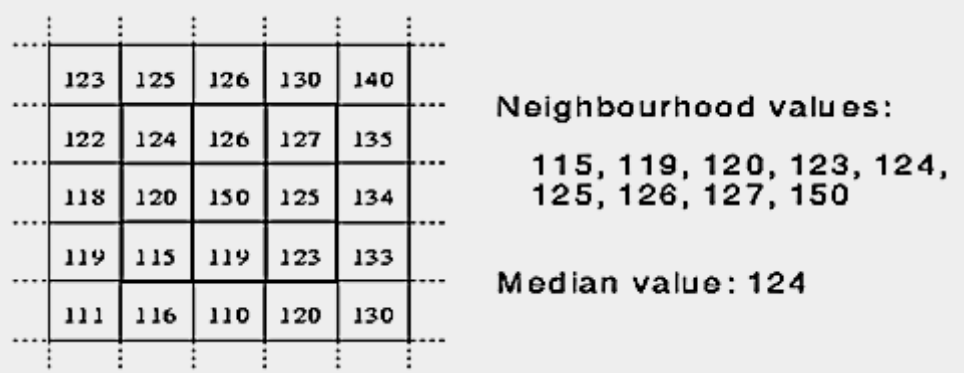


Figure 4.4, calculating the median value of a pixel neighborhood.

As can be seen, the central pixel value of 150 is rather misleading of the surrounding pixels and is replaced with the median value: 124. A 3×3 square neighborhood is used here; larger neighborhoods will produce more severe smoothing. Matlab forms median filter as displayed in equation 4.13:

B = medfilt2 (A, [m n]) 4.13

Where equation 4.11 performs median filtering of the matrix A in two dimensions. Each output pixel contains the median value in the m-by-n neighborhood around the corresponding pixel in the input image. $medfilt2$ pads the image with 0s on the edges, so the median values for the points within $[m\ n]/2$ of the edges might appear distorted.

Median filtering is a nonlinear operation often used in image processing to reduce noise. A median filter is more effective than convolution when the goal is to instantaneously reduce

noise and reserve edges. The $[m\,n]$ argumenta have the default $(3 - by - 3)$ neighborhood.

4.2.4. Adaptive Median Filter (AMF):

A new switch median filter is presented for suppression of impulsive noise in image. The proposed filter is **Modified Adaptive Center Weighted Median** (MACWM) filter with an adjustable central weight obtained by partitioning the observation vector space[233].

The framework of the MACWM filter is illustrated in Figure 4.5. It is composed of four parts: a median filter, a set of threshold by **Fuzzy C-means** (FCM), training the center weight each block by **Least Mean Square** (LMS) algorithm, and a decision as to whether noises exist or not.

At first, according to observation vector space (input image); median of input image will be calculated. We propose FCM algorithm to partition observation vector space to M block, and related weights to any blocks will be trained by using the LMS algorithm. The output value y (k) of the MACWM filter (Figure 4.6) at the processed pixel x (k) is obtained as follows in equation 4.14:

$$y(k) = \big(1 - w(k)\big)x(k) + y(k)m(k) \qquad 4.14$$

That the usual output value from a median filter is denoted as m (k) at location k in a filter window of size 2n + 1 as follows in equation 4.15:

$$m(k) = MED\{x_1(k) \dots \dots x_n(k)\} \qquad 4.15$$

Where MED means the median operation.

The MACWM filter achieves its effect through the linear combinations of the weighted output of the median filter and the related weighted input signal. Here, w (k) denotes the membership function indicating to what extent an impulsive noise is considered to be located at the pixel x (k). If w (k) = 1, an impulsive noise is considered to be located at pixel x (k), and the output value of the filter is equal to the output value of the median filter. If w (k) = 0, an impulsive noise is considered not to be located at pixel x (k), and the output value is the same as the input x (k); that is, the pixel x (k) is left unchanged. To judge whether an impulsive noise exists or not at the input pixel x (k), the membership function w (k) should take a continuous value from 0 to 1. Therefore, the major concern of

the MACWM filter is how to decide the value of the membership function w (k) at the pixel x (k).

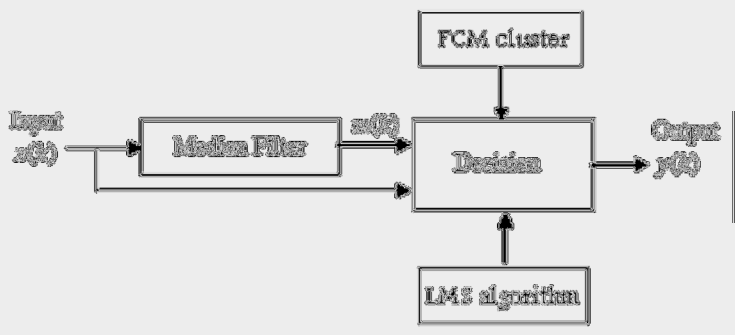


Figure 4.5. The basic structure of the MACWM filter

In general, the amplitudes of most impulses are more prominent than the fine changes of signals[234]. Thus, the following two variables can be defined to generate the observation vector[235] . (Partitioning of observation vector space)[233]

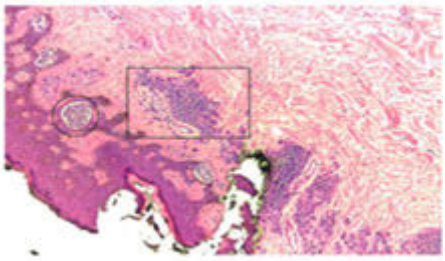


Figure 4.6. Shows the output image of Adaptive median Filtering.

4.3. Image preprocessing: enhancement

Image segmentation is one of the fundamental steps in a computer vision and is widely used in several image processing applications. Segmentation is an important step in the medical image analyses and classification for radiological evaluation or computer – added diagnoses. The main purpose of image segmentation is to partition an input image into a number of meaningful and non-overlapped regions based on some features that help in distinguishing between different image regions (object and background), by grouping together neibourhood pixels based on some pre-defined similarity principle. The similarity standard can be determined using specific properties or features of pixels representing

objects in the image. As well segmentation is a pixel classification technique that allows the formation of regions of similarities in the image [236].

a class conversion functions.

4.3.1. Image enhancement

Medical images occurred in most radiological applications are visually examined by a physician. The purpose of image enhancement methods is to process a picked image for better contrast and visibility of features of interest for visual examination as well as subsequent computer-aided analysis and diagnoses. As described in [236] different medical imaging modalities provide specific characteristics information about internal organs or biological skins. Image contrast and visibility of the features of interest depend on the imaging modality as well as the anatomical regions.

There is no unique general theory or methods for processing all kinds of medical images for feature enhancement. Specific medical imaging applications such as cardiac, neurological, breast and genital imaging, present different challenges in image are processing for feature enhancement. Medical images show characteristic information about the physiological properties of the structures and skins. But the quality and visibility of information depends on the imaging modality and the response functions of the imaging scanner. Medical images from specific modalities need to be processed using a method that is suitable to enhance the features of interest.

Image-enhancement tasks are usually characterized in two categories [237], [238], [239]:

- ***Spatial domain methods:*** These methods employ image pixel values in the spatial domain based on the distribution statistics of the full image or local regions. Histogram transformation, spatial filtering, region-growing, morphological image processing and model-based image estimation methods are some examples in this category of image and feature enhancement.
- ***Frequency domain methods:*** These methods use information in the frequency domain based on the frequency characteristics of the image. Frequency filtering, holomorphic filtering and wavelet processing methods are some examples in this category of frequency representation based image and feature enhancement.

Also model-based techniques such as *neural networks and knowledge-based systems* are also used to extract specific features for pattern recognition and classification[240, 241].

4.3.2. Intensity transformation

The image form itself is referred to as *spatial domain*, and methods in this category are based on direct manipulation of pixels in an image. In this part we focus attention of two important categories of *spatial domain* processing: **1. *intensity (or grey-level) transformations and 2. Spatial filtering.*** The approach in our study is referred to as ***neibourhood processing, or spatial convolution***. Enhancement is highly sensitive and attracting especially to new researchers in this field. These techniques are general in choice and have uses in several other branches of digital image processing.

As noted in the spatial domain techniques operate directly on the pixels of an image. The spatial domain processes are expressed by the following equation 4.16:

$$g(x,y) = T[f(x,y)] \quad 4.16$$

Where $f(x,y)$is the input image, $g(x,y)$ is the output (processed) image, and T is an operator on$f$, defined over a specified neibourhood about point(x,y). In addition, T can operate on a set of images, such as performing the addition of K images for noise reduction. The principle approach for expressing a spatial neibourhood about a point $(x,y)$is to use square or rectangular region cantered at(x,y). The centre of the region is moved from pixel to pixel starting at the top, left corner, and as it moves, it involves different neighborhoods. Operator T is applied at each location (x,y) to yield the output g, at that location. Only the pixels in the neighborhood are used to computing a value of g at(x,y).

The simplest form of the transformation T is dependent only on the intensity values and not obviously on (x,y), intensity transformation function frequently are written in the simplified form as showed in equation 4.17

$$s = T(r) \quad 4.17$$

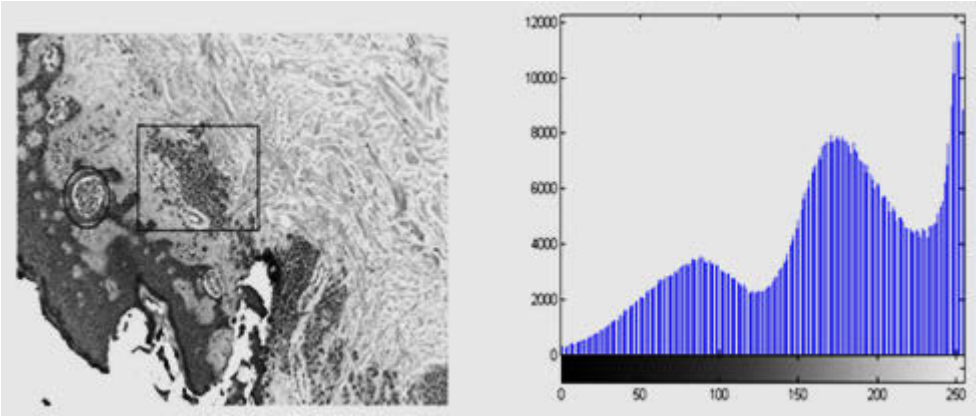
Where r denotes the intensity of $f$and s the intensity of$g$, both at any corresponding point (x,y)in the images.

4.3.3. Histogram Processing

In general, a histogram is the estimation of the probability distribution of a particular type of data. An image histogram is a type of histogram which offers a graphical representation

of the tonal obviously distribution of the grey values in a digital image. By viewing the image's histogram, it can analyse the frequency of appearance of the different grey levels contained in the image. In Figure 4.7 we can see an image and its histogram. The histogram shows that the image contains only a fraction of the total range of grey levels. In this case there are 256 grey levels and the image only use values between approximately 50-100. Therefore this image has low contrast.

Intensity transformation functions based on information extracted from image intensity histograms show a basic role in the image processing, in areas such as enhancement, compression, segmentation, and description. We focus on using histograms for image enhancement.



(a) Showed: grayscale image (b) Showed: Histogram out.

Figure 4.7 Showed: (a) grayscale image, (b) Histogram out.

The histogram of a digital image with L total intensity levels in the range $[0, G]$ is defined as the discrete function as showed in 4.18:

$$h\,(r_{\mathrm{k}}) = n_{\mathrm{k}} \qquad 4.18$$

Where rk is the kth intensity level in the interval [0, G] and nk is the number of pixels in the image whose intensity level is rk. The value of G is 255 for images of class unit 8, 65535 for images of class unit 16, and 1.0 for images of class double. Remember that indices in Matlab cannot be 0, so $r1$ corresponds to intensity level 0; $r2$ corresponds to intensity $level$ 1, and so on with rL corresponding to $level\ G$. Note that $G = L - 1$ for images of class unit8 and unit16. Generally, it is good to work with normalised histograms, obtained simply by dividing all elements of $h\,(rk)$ by the total number of pixels in the image, which we indicated by$n$, equation 4.19:

$$p(rk\ ) = \frac{h(rk)}{n} = \frac{nk}{n} \quad 4.19$$

For $k = 1, 2, \ldots, L$. From basic probability, we recognise P (r_k) as an estimate of the probability of occurrence of intensity level r_k. The core function in the Matlab toolbox for dealing with image histograms is $imhist$ which has the following basic syntax, equation 4.20:

$$h = imhist(f, b) \quad 4.20$$

Where f is the input image, h is its histogram, h (r_k) and b is the number of bins used in forming the histogram (if b is not including in the argument, b= 256 is used by default). A bin is simply a subdivision of the intensity scale. For example, if we are working with unit8 images and we let b=2, then the intensity scale is subdivided into 2 ranges: 0 to 127 and 128 to 255. The resulting histogram will have 2 values: h (1) equal to the number of pixels in the image with values in the interval [0, 127], and h (2) equal to the number of pixels with values in the interval [128, 255]. We obtain the normalized histogram simply by using equation 4.21:

$$p = imhist\ (f, b)/nume1\ (f) \quad 4.21$$

Where f gives the number of the elements in array f (i.e. the number of pixels in the image).

4.3.4. Histogram Equalization

This section describes a method of imaging processing that allows medical images to have better contrast. This is attained via the histogram of the picture, using a method that allows the areas with low contrast to gain higher contrast by spreading out the most common intensity values. When pathologies are present in an image, trying to define the area of the lesion or object of interest may be a challenge, because different structures are usually covered one over the other. For example, in the case of the chest the heart, lungs, and blood vessels are so close together that contrast is critical for achieving an accurate diagnosis.

As specified earlier, basically the histogram equalization spreads our intensity values along the total range of values in order to achieve higher contrast. This method is especially useful when an image is represented by close contrast values, such as images in which both the background and foreground are bright at the same time, or else both are dark at the

same time. For example, the result of applying histogram equalization to the image is presented and the image's contrast has been improved. The original histogram in figure 4.6 has been stretched along the full range of grey values, as can be notice in the equalized histogram in Figure 4.8.

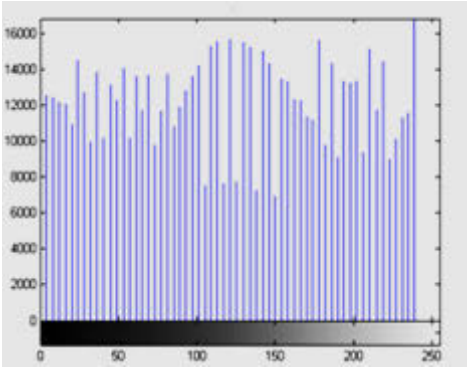


Figure 4.8. Showed: intensity image using

There are a number of different types of histogram equalization algorithms, such as:

1. Histogram expansion, (Simple and enhance contrasts of an image).
2. Local Area Histogram Equalization (LAHE) (Offers an excellent enhancement of image contrast).
3. Cumulative histogram equalization (Has good performance in histogram equalization).
4. Bar sectioning, (Easy to implement) and
5. Odd sectioning (Offers good image contrast).

As previously explained that intensity levels are continuous quantities normalized to the range $[0, 1]$, and let $pr(r)$ denote the probability density function (PDF) of the intensity levels in a given image, where the subscript is used for differentiating between the $PDF's$ of the input and output images. Suppose that we perform the following information on the input levels to obtain output (processed) intensity levels, (s) as in equation 4.22:

$$s = T(r) = \int_0^r pr(w)dw \qquad 4.22$$

Where w is a dummy variable of integration. As presented in [237]: that the probability density function of the output levels is *uniform*; that is, equation 4.23:

$$Ps(s) = \begin{cases} 1 & for\ 0 \le s \le 1 \\ 0 & 0therwise \end{cases} \qquad 4.23$$

In other words, the preceding transformation makes an image whose intensity levels is equally likely, and, in addition, covers the entire range $[0, 1]$. The net result of this intensity-level *equalization* process is an image with increased dynamite specified range, which will be likely to have higher contrast. Note that the transformation function is really nothing more than the cumulative distribution function (CDF). When dealing with discrete quantities we work with histograms and call the proceeding technique *histogram equalization*, although, in general, the histogram of the proceed image will not be uniform, due to the separate nature of the variables. Reference to $pr\ (rj), j\ =\ 1\ , 2, \dots , L,$ denote the histogram associated with the intensity levels of a given image, and the values in the normalised histogram are approximations to the probability of occurrence of each intensity level in the image. For separate quantities we work with summations, and the equalizations transformation becomes, as mentioned in equation 4.24:

$$\begin{aligned} sk &= T\ (r_k) \\ &= \sum_{j=1}^{k} p_{r\ (r_j)} \\ &= \sum_{j=1}^{k} \frac{n_j}{n} \end{aligned} \qquad 4.24$$

$For\ k\ =\ 1, 2, L,$ where s_k is the intensity value in the output (proceed) image corresponding to value r_k in the input image.

Histogram equalization is implemented in the Matlab toolbox by function $histeq,$ which has the syntax, equation 4.25:

$$g\ =\ histeq\ (f, nlev) \qquad 4.25$$

Where f the input image and $nlev$ is the number of intensity levels specified for the output image. If $nlev$ is equal to L (the total number of possible level in the input image), then $histeq$ implements the transformation function, T (r_k), directly. If levy is > L, then $histeq$ attempts to distribute the levels so that they will approximate a ***flat histogram***. Unlike $imhist$, the default value in the $histeq$ is $nlev\ =\ 64$. For the most part, we use the maximum possible number of levels (generally 256) for $nlev$ because this produces a true implementation of the histogram – equalisation just described.

4.3.5. Special filtering

As explained above under intensity transformation the neighborhood processing consists of: defining a Centre point,(x, y); performing an operation that involves only the pixels in a predefined neighborhood about that Centre point; letting the result of that operation be the "response" of the process at that point; and repeating the process for every point in the image. The process of moving the Centre point creates new neighborhoods, one of each pixel in the input image. The two principle terms used to identify this operation are ***neighborhood processing and special filtering***, with the second term being more powerful if the competitions done on the pixels of the neighborhoods are linear, the operation is ***called linear spatial filtering*** (the term *spatial convolution* also used); otherwise it's called ***nonlinear spatial filtering***.

- ***Linear spatial filtering***

The concepts of *linear filtering* have its roots in the use of the Fourier transform for signal processing in the frequency domain. Use of the term *linear spatial filtering* separates this type of process from *frequency domain filtering*.

- ***Nonlinear spatial filtering***

Generally a tool for generating nonlinear special filters in IPT is function $ordfilt2$, which generates *order-statistic filters* (also called rank filters). These are nonlinear spatial filters whose response is based on ordering (ranking) the pixels contained in an image neighborhood which are then replacing the value of the Centre pixel in the neighborhood with the value determined by the ranking results. We focus our attention on nonlinear filters generated by $ordfilt2$. The syntax of function $ordfilt2$ is, displayed in equation 4.26:

$$g = ordfilt2\,(f, order, domain) \qquad 4.26$$

The function creates the output image g by replacing each element of f by the $orderth$ element in the sorted set of neighbours specified by the non 0 elements in domain. The domain is a $m\,x\,n$ matrix of $1's$ and $0's$ that specify the pixel locations in the neighbourhood that are to be used in the computation. In this sense the domain acts like a mask. The pixels in the neighborhood that correspond to 0 in the domain matrix are not used in the computation. For example to implement a *min filter* (order 1) of size $m\,x\,n$ we use the equation 4.27

$$g = ordfilt2\,(f, 1, ones\,(m,n)) \quad 4.27$$

In this formulation the 1 denotes the $1st$ sample in the ordered set of m,n samples, and ones (m,n) creates an $m\,x\,n$ matrix of $1's$, indicating that all samples in the neighbourhood are to be used in the computation. In the terminology of statistics, a $min\,filter$ (the first sample of an ordered set) is referred to as the $0th$ percentile Similarity; the $100th$ percentile is the last sample in the ordered set, which is $mnth$ sample. This corresponds to a $max\,filter$, which is implemented using the equation 4.28:

$$g = ordfilt2\,(f, m*n, ones\,(m,n)) \quad 4.28$$

The best known order statistic filter in digital image processing is the *median filter*, which corresponds to the 50th percentile. We can use Matlab function median in $ordfilt2$ to create a median filter equation 4.29:

$$g = ordfilt2\,(f, median\,(1{:}\,m*n), ones\,(m,n)) \quad 4.29$$

Where median $(1{:}\,m*n)$ simply computes the median of the ordered sequence $1, 2, mn$. Function median has the general equation 4.30:

$$v = median\,(A, dim) \quad 4.30$$

Where v vector is whose elements are the median of A along dimension dim. Practically, the Matlab toolbox provides a specialised implementation of the $2-D\,median\,filter$, equation 4.31:

$$g = medfilt2\,(f, [m\,n], padopt) \quad 4.31$$

Where $[m\,n]$ defines a neighborhood of size $m\,x\,n$ over which the medium is computed, and $padopt$ specifies 1 of 3 possible border padding options: '$zeros$' (the default) '$symmetric$' in which f is extended symmetrically by mirror-reflecting it across its border, and '$indexed$', in which f is padded with 1's if it is of class double and with 0's otherwise. The default form of this function is, showed in equation 4.32:

$$g = medfilt2\,(f) \quad 4.32$$

This uses $(3x3)$ neighbourhoods to compute the medium and cloths the border of the input with $0's$.

We used (AMF) Adaptive Median Filter, as filtering techniques [242], this filter ignores edge effects and boundary conditions as such. The output is a collected version of the original image, as displayed above, where the amount collected is equal to the maximum window size vertically and horizontally.

All the above mentioned strategies are meant to facilitate the segmentation and feature extraction stages which consequently lead to better diagnostic results.

4.4. Image Segmentation

Segmentation subdivides an image into its principal regions or objects. The level to which the subdivision is carried depends on the problem being solved. That is, segmentation should stop when the objects of interest in an application have been isolated. Segmentation of significant images is one of the most difficult tasks in image processing. Segmentation accuracy determines the final success or failure of computerized analysis procedures. For this reason, considerable care should be taken to improve the probability of rugged segmentation. Segmentation algorithms for monochrome images generally are based on one of two basic properties of image intensity values: ***discontinuity and similarity***. In the first category, the approach is to partition an image based on immediate changes in intensity, such as edges in an image. The principle approaches in the second category are based on partitioning an image into regions that are similar according to a set of predefined criteria. Edge detection in particular has been an essential of segmentation algorithms for many years. The edge detection is followed by the introduction to thresholding techniques. Thresholding also is a fundamental approach to segmentation that enjoys a significant degree of popularity, especially in applications where ***speed*** is an important factor.

Image segmentation methods can be broadly classified into three categories:

1. ***Edge-based methods***, in which the edge information is used to determine boundaries of objects. The boundaries are then analysed and modified, if needed, to form closed regions belonging to the objects in the image.
2. ***Pixel-based direct classification methods***, in which heuristics or estimation methods derived from the histogram statistics of the image, are used to form closed regions belonging to the objects in the image.

3. ***Region-based methods,*** in which pixels are analysed directly for a region growing process based on a pre-defined similarity principle to form closed regions belonging to the objects in the image.

Once the regions are defined, features can be computed to represent regions for characterization, analysis and classification. These features may include shape and texture information of the regions as well as statistical properties, such as ***variance and mean*** of grey values.

4.4.1. Edge-based Image Segmentation

Edge-based methods use a spatial filtering method to compute the first-order or second-order gradient information of the image. ***Sobel or directional derivative masks*** can be used to compute gradient information as showed in figure 4.9. The Laplacian mask can be used to compute second-order gradient information of the image. For segmentation purposes, edges need to be linked to form closed regions. Gradient information of the image is used to track and link relevant edges. In addition, it has to deal with the worries in the gradient information due to noise and objects in the image.

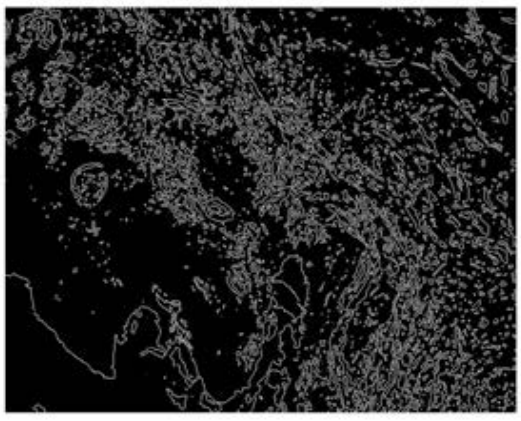


Figure 4.9. Sobel Edge Detection Image.

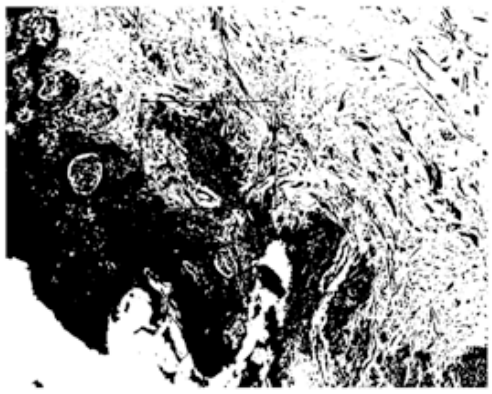


Figure 4.10.Binary Output Image Otsu's Method[243].

4.4.2. Edge Detection Operations.

The gradient magnitude and directional information from the Sobel horizontal and vertical direction masks can be obtained by convolving the respective Gx and Gy masks with the image as$[1, 2]$, equations 4.33 & 4.34:

$$Gx = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad 4.33$$

$$Gy = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad 4.34$$

$$M = \sqrt{Gx^2 + Gy^2} \approx |Gx| + |Gy| \quad 4.35$$

Where, equation 4.35 represents the magnitude of the gradient that can be approximated as the sum of the absolute values of the horizontal and vertical gradient images obtained by convolving the image with the horizontal and vertical masks, Gx and Gy .

4.4.3. Thresholding

Because of its in-built properties and simplicity of implementation, image thresholding enjoys a central position in applications of image segmentation. We discuss ways of choosing the threshold value automatically, and we consider a method for changing the threshold according to the properties of local image neighborhoods.

Suppose that the intensity histogram shown in Figure 4.11 corresponds to an image, $f(x, y)$, composed of light objects on a dark background, in such a way that object and background pixels have intensity levels grouped into two main modes. One obvious way to extract the objects from the background is to select a threshold T that separates these modes. Then any point (x, y) for which, $f(x, y) \geq T$ is called an *object point*; otherwise, the point is called a *background point*. In other words, the threshold image $g(x, y)$, is defined as equation 4.36:

$$g(x, y) = \begin{cases} 1 & if\ f(x, y) \geq T \\ 0 & if\ f(x, y) < T \end{cases} \quad 4.36$$

$Pixels\ labeled\ 1$ correspond to objects, whereas $pixels\ labeled\ 0$ correspond to the background. When T is a constant, this approach is called *global thresholding*.

As explained above, we can result with:

Histogram modality analysis: $Select\ threshold\ (s)$ in $valley\ (s)$ between peaks, as displayed on figure 4.10.

Maximal separation of classes (Otsu, 1978):

This matter explained in figures 4.11 and 4.12 below:

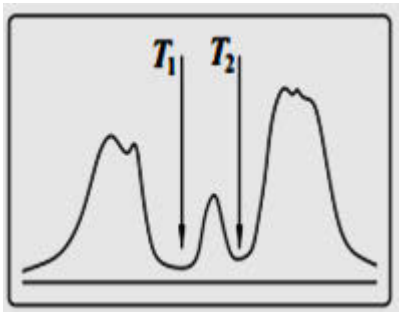


Figure 4.11, selected thresholds in valleys between peaks

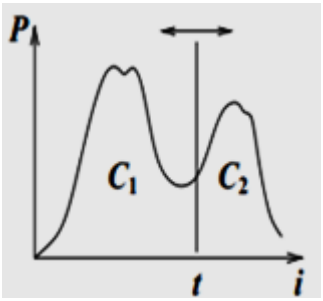


Figure 4.12, obtaining the best possible separation measure for two classes.

Improving the histogram for better peak separation: Use of gradient to improve histogram by combine intensity and gradient information for better separation of objects and background, as displayed in figure 4.13.

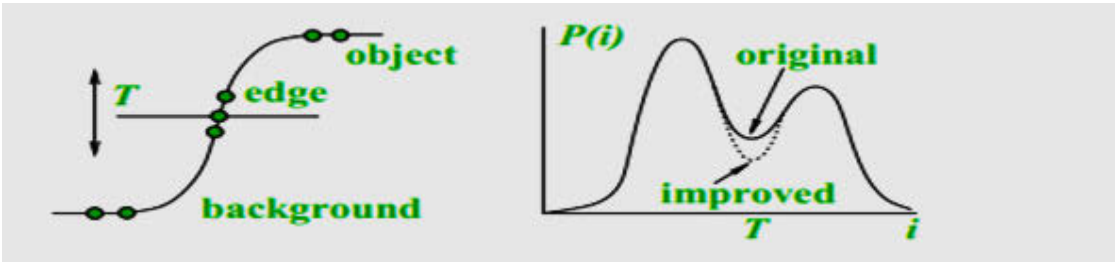


Figure 4.13, selecting a threshold by visually analyzing a bimodal histogram.

(Principle of histogram peak separation).

4.4.4. Global Thresholding

One way to choose a threshold is by visual inspection of the image histogram. The histogram clearly has two separate modes; as a result, it is easy to choose a threshold T that separates them. Another method of choosing T is by trial and error, picking different thresholds until one is found that produces a good result as judged by an observer. This is particularly effective in a communicating environment, such as one that allows the user to change the threshold using a wider threshold (graphical control) such as a ***slider*** and the result immediately.

For choosing a threshold automatically [237], describes the following significant procedure:

1. Select an initial estimate for T, (a suggested initial estimate is the mid-point between the minimum and maximum intensity values in the image).
2. Segment the image using T . This will produce two groups of pixels: $G1$, consisting of all pixels with $values \geq T$, and $G2$, consisting of pixels with $values < T$.
3. Compute the average intensity values μ_1 and μ_2 for the pixels in regions $G1$ and $G2$.
4. Compute a new threshold value: (equation 4.37)

$$T = \frac{1}{2}\ (\mu_1 + \ \mu_2\ ) \qquad 4.37$$

5. Repeat steps 2 through 4 until the difference in$T\ =\ 0$.

There is toolbox in Matlab which provides a function called ***gray thresh*** that computes a threshold using ***Otsu's method*** [243]. To examine the preparation of this histogram-based method, we start by handling the normalized histogram as an isolated probability density function, as in equation 4.38:

$$p_r\ (r_q) = \frac{n_q}{n} \qquad q = 0, 1, 2, \dots, L-1 \qquad 4.38$$

Where n the total number of pixels in the image, n_q is the number of pixels that have $level\ r_q$, and L is the total number of possible intensity levels in the image. Now suppose that a threshold k is chosen such that C_0 is the set of pixels with levels $[0, 1, \dots, k-1]$ and C_1 is the set of pixels with levels $[k, k+1, \dots, L-1]$. Otsu's method chooses the threshold value k that maximizes the *between-class variance* ${\sigma_B}^2$, which is defined as in equation 4.39:

$${\sigma_B}^2 = \ w_0\ (\mu_0 - \mu_T)^2 + w_1\ (\mu_1 - \ \mu_T)^2 \qquad 4.39$$

Where equations (34) – (39) displayed the component of ${\sigma_B}^2$ as follows:

$$w_0 = \ \sum_{q=0}^{k-1} p_q\ (r_q) \qquad 4.40$$

$$w_1 = \ \sum_{q=k}^{L-1} p_q\ (r_q) \qquad 4.41$$

$$\mu_0 = \ \sum_{q=0}^{k-1} q\ p_q\ (r_q)\ /w_0 \qquad 4.42$$

$$\mu_1 = \ \sum_{q=k}^{L-1} q\ p_q\ (r_q)\ /w_1 \qquad 4.43$$

$$\mu_T = \ \sum_{q=0}^{L-1} q\ p_q\ (r_q) \qquad 4.44$$

The gray thresh function uses Otsu's method, which chooses the threshold to minimize the infraclass variance of the black and white pixels. This gray thresh function takes an image which is used to compute its histogram, and then finds the threshold value that maximize $\sigma_B{}^2$.The threshold is returned as a normalized value between 0.0 and 1.0. The calling syntax for gray thresh is, as in equation 4.45:

$$T = graythresh\,(f) \quad 4.45$$

Where f is the input image and T computes a global threshold (level) that can be used to convert intensity image to a binary image with (im2bw), as displayed above in figure 4.9. T is a normalized intensity value that lies in the range [0, 1]. Because the threshold is normalized to range $[0, 1]$, it must be scaled to the proper range before it is used. For example, if f is of class unit 8, we multiply T by 255 before using it.

4.4.5. Optimal Global Thresholding

To determine an optimal global grey value threshold for image segmentation, parametric distribution based methods can be applied to the histogram of an image [237] [244]. Let us assume that the histogram of an image to be segmented has two normal distributions belonging to two respective classes such as background and object. Thus the histogram can be represented by a *mixture probability density function* $p\,(z)$ as in equation 4.46:

$$p(z) = P_1 p_1(z) + P_2 p_2(z) \quad 4.46$$

Where p_1 (z) and p_2 (z) are the Gaussian (normal) distributions of $class\ 1\ and\ 2$, respectively, with the class probabilities of $P1$ and $P2$ such that in equation 4.47:

$$P_1 + P_2 = 1. \quad 4.47$$

Using a grey value threshold T, a pixel in the image $f(x, y)$ can be classified to $class\ 1\ or\ class\ 2$ in the segmented image $g(\,x, y)$ as in equation 4.48:

$$g(x, y) = \begin{cases} \text{Class 1} & \text{if } f(x, y) > T \\ \text{Class 2} & \text{if } f(x, y) \leq T \end{cases} \quad 4.48$$

Let us define the error probabilities of misclassifying a pixel as in equation 4.49 and 4.50:

$$E_1(T) = \int_{-\infty}^{T} P_2(z)dz \quad 4.49$$

And

$$E_2(T) = \int_{-\infty}^{T} P_1(z)dz \quad 4.50$$

Where $E_1\ (T)$ equation" (4.49) and $E2\ (T)$ equation (4.50) are, respectively, the probability of incorrectly classifying a class 1 pixel to class 2 and a class 2 pixels to $class\ 1$. The overall probability of error in pixel classification using the threshold T is then expressed as in equation 4.51:

$$E(T) = P_2(T)E_1(T) + P_1(T)E_2(T) \quad 4.51$$

For image segmentation, the objective is to find an ***optimal threshold*** T that minimizes the overall probability of error in pixel classification. The optimization process requires the parameterization of the probability density distributions and possibility of both classes. These parameters can be determined from a model or set of training images [245] [244] [246] [247].

Let us assume σi and μi to be the ***standard deviation and mean*** of the Gaussian probability density function of the $class\ i$ ($i = 1, 2$ for two classes) such as is explained in equation 4.52:

$$p(z) = \frac{p_1}{\sqrt{2\pi\sigma_1}\ e^{-(z-\mu_1)^2/2\sigma_1^2} + \ \mathrm{P2}/\sqrt{2\pi}\,\sigma_2 \quad e^{-\frac{(z-\mu_1)^2}{2\sigma_2^2}}} \quad 4.52$$

The optimal global threshold T can be determined by finding a general solution that minimizes Equation (4.50) with the mixture distribution in Equation (4.45) and thus satisfies the following quadratic expression [248] as in equation 4.53:

$$AT^2 + BT + C = 0, \quad 4.53$$

Where equations 4.54 – 4.56 display the components displays of equation (4.53)

$$A = \sigma_1^2 - \sigma_2^2 \quad 4.54$$

$$B = 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2) \quad 4.55$$

$$C = \ \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 + 2\sigma_1^2\sigma_2^2 \mathrm{In}(\sigma_2 P_1/\sigma_1 P_2) \quad 4.56$$

If the variances of both classes can be assumed to be equal toσ^2, the optimal threshold T can be determined as in equation 4.57:

$$T = \frac{\mu_1+\mu_2}{2} + \frac{\sigma^2}{\mu_1-\mu_2}\ln\left(\frac{P_2}{P_1}\right) \quad 4.57$$

It should be noted that in the case of equal probability of classes, the above expression for determining the optimal threshold is simply reduced to the average of the mean values of the two classes.

4.4.6. Pixel Classification through Clustering

In the histogram-based pixel classification method for image segmentation, the grey values are partitioned into two or more clusters depending on the peaks in the histogram to obtain thresholds. The basic concept of segmentation by pixel classification can be extended to clustering the grey values or feature vectors of pixels in the image. In this approach, multiple parameters of interest are to be segmented or example, a feature vector may consist of grey value, contrast and local texture measures for each pixel in the image. A color image may have additional color components in a specific representation such as ***Red, Green and Blue*** components in the $R - G - B$ colour coordinate system that can be added to the feature vector. Magnetic Resonance (MR) or multi-modality medical images may also require segmentation using a ***multi-dimensional*** feature space with multiple parameters of interest.

Images can be segmented by pixel classification through clustering of all features of interest. The number of clusters in the multi-dimensional feature space thus represents the number of classes in the image. As the image is classified into cluster classes, segmented regions are obtained by checking the neighborhood pixels for the same class label. However, clustering may produce split regions with holes or regions with a single pixel. After the image data are clustered and pixels are classified, a ***post-processing algorithm*** such as region growing, pixel connectivity or rule-based algorithm is usually applied to obtain the final segmented regions [249] [250]. There are a number of algorithms developed for clustering in the literature and used for a wide range of applications [244] [251] [249] [250] [252].

Clustering is the process of grouping data points with similar feature vectors together in a single cluster while data points with different feature vectors are placed in different clusters. Thus the data points that are close to each other in the feature space are clustered together. The similarity of feature vectors can be represented by a suitable distance measure such as ***Euclidean or Mahalanobis*** disassociated with the distribution of the corresponding feature vectors of the data points in the cluster. The formation clusters is

optimized with respect to an objective function involving ***pre-specified*** distance and similarity measures along with additional constraints such as ***smoothness***.

Chapter Summary

This chapter 'pre-processing and segmentation'; includes introduction explaining extensive research in investigation of the varied presentations and physical characteristics of melanoma, the clinical diagnostic accuracy remains sub optimal. Thus, a growing interest has been developed in the last two decades in the automated analysis of digitized images obtained by ELM techniques to assist clinicians in differentiating early melanoma from benign skin lesions; image types. This chapter also covers general approach of developing a computer aided design (CAD) system; and presented the four stages of CAD system for skin lesions including all types of filters. The explanation of image segmentations are also included in all edge detection operations.

CHAPTER 5
Image Representation and Analysis

5.1. Introduction

Modern biology can benefit from the advancements made in the area of machine learning [253]Attention should be taken when judging the superiority of some machine learning approaches over other categories of methods. It is claimed [253] that the success or failure of machine learning approaches on a given problem is sometimes a matter of the quality keys used to evaluate the results, and these may vary strongly based on the expertise of the user. Of special concern with supervised applications is that all steps involved in the classifier design (selection of input variables, model training, etc.) should be cross-validated to obtain a balanced estimation for classifier accuracy. For instance, selecting the features using all available data and subsequently cross-validating the classifier training will produce a positively biased error estimate. Because of insufficient proof arrangements many studies published in the literature as successful have been shown to be overoptimistic [61]. It should be clear from the narrative examples used in this study that choice, tuning, and diagnosis of machine learning applications are far from automatic.

5.2. Machine Learning

The term machine learning refers to a set of topics dealing with the design and evaluation of algorithms that facilitate pattern recognition, classification, and calculation, based on models derived from existing data. Two facets of automation should be acknowledged when considering machine learning in broad terms. Firstly, it is intended that the classification and projection tasks can be accomplished by a suitably programmed computing machine. That is, the product of machine learning is a classifier that can be practically used on available hardware. Secondly, it is intended that the creation of the classifier should itself be highly automated and should not involve too much human input. This second facet is certainly unlimited but the basic objective is that the use of automatic algorithm construction methods can minimize the possibility that human biases could affect the selection and performance of the algorithm. The history of relations between

biology and the field of machine learning is long and complex. An early technique [254]for machine learning called the perceptron established an effort to model actual neuronal behavior, and the field of artificial neural network (ANN) design appeared from this effort. Further artificial neural network architectures such as the adaptive resonance theory (ART) [255]and neocognitron [256]were moved from the organization of the visual nervous system. In the best years, the flexibility of machine learning techniques has grown along with mathematical frameworks for measuring their reliability, and it is natural to promise that machine learning methods will improve the efficiency of discovery and understanding in the increasing volume and complexity of biological data. (More detailed information in appendix C)

5.3. Main Concepts

Two main examples exist in the field of machine learning: supervised and unsupervised learning. Both have potential applications in biology. In supervised learning, objects in a given collection are classified using a set of features. The result of the classification process is a set of rules that propose projects of objects to classes based exclusively on values of features. The goal in supervised learning is to design a system able to follow expect the class membership of new objects based on the available features. In unsupervised learning is intended to show natural groupings in the data. Thus, the two paradigms may informally be contrasted as follows: in supervised learning, the data come with class labels, and we learn how to associate labelled data with classes; in unsupervised learning, all the data are unlabeled, and the learning procedure consists of both defining the labels and associating objects with them.

We have assumed certain mathematical notations and concepts to support accurate characterization of both supervised and unsupervised machine learning methods. These to solve the problems occurred when a task is defined by a series of cases or examples rather than by predefined rules. Such problems are found in a wide variety of application domains, ranging from engineering applications in robotics and pattern recognition (speech, handwriting, face recognition, skin cancer detection), to Internet applications (text categorization) and medical applications (diagnosis, prognosis, drug discovery). Given a number of “training” examples (called data facts, samples, patterns or observations) associated with desired outcomes, the machine learning process consists of finding the relationship between the patterns and the outcomes using just the training examples. To

make it concrete, we conceder the following examples: the data points or examples are clinical observations of patient and the outcome is the health status: healthy or suffering from cancer. The goal is to expect the unknown outcome for new “test” examples, e.g. the health status of new patients. The performance on test data is called “generalisation”. To perform this task, one must build an analytical model or predictor, which is typically a function with adjustable parameters called a “learning machine”. The training examples are used to select an optimum set of parameters.

5.4. Intelligent Image Features Extraction

With the size of image databases increasing dramatically, the practicality of such information is dependent on how well it can be accessed and searched and how well knowledge can be extracted from it. Our ability to generate data currently far exceeds our ability to explore, analyze and let alone understand it.

Recently, there has been a growing and improved interest in intelligent image management systems based on domain knowledge and applications. Although over the last decade or so some researchers have developed knowledge-based approaches, they rely almost exclusively on low-level features such as texture, colour and intensity and have little high-level features interpretation capability. Also, they mainly operate in pixel domain, such that images (if compressed) need to be decompressed prior to any analysis or processing. Such an approach can slow down any application to the point of being unrealistic.

This thesis will create intelligent methods for solving many difficult high-level image feature extraction and analysis problems, in which both local and global properties as well as spatial relations must be taken into consideration. All the techniques will be implemented to work in pixel and compressed domain (avoiding the inverse transform in decompression), thus speeding up the whole process. This research will lead to a better understanding of images and development of efficient methods for other image processing, pattern recognition and computer vision problems. The research results will have a wide range of useful applications, including but not restricted to human face image recognition, medical and biological image analysis, image and video compression, and content-based image indexing and retrieval.

Initial investigation of intelligent image analysis and feature extraction techniques will be done, combining the principles from fields of multiresolution wavelet analysis, compression, computer vision, texture analysis and information retrieval. Next, by applying and extending those techniques we aim to solve a class of difficult image feature extraction problems. By combining various image analysis and signal processing techniques by develop some new high-level feature extraction methods, thus improving current state-of-the-art retrieval and classification methods. Using the resulting extracted features as a first step and input to data mining systems would lead to supreme knowledge discovery systems [257].

5.4.1. Texture:

Texture is one of the important features used in image retrieval systems. The methods of describing texture go down into two major categories: Statistical and Structural. The primary goal is to determine which texture feature or combination of texture features is most efficient in representing the spatial distribution of images. Researchers involved in analyses and comparison of texture features [258], they evaluate both Statistical and Structural texture features and presented comparison of individual texture features and combined texture features. The first-order statistics, second-order statistics, Gabor transform and 2D Wavelet transforms were considered for retrieval. The retrieval efficiency of the texture features was investigated by means of relevance. According to the results they obtained[258], it displays that it is difficult to claim any individual feature is superior to others. The performance depends on the spatial distribution of images. The test results indicated that Grey Level Co-occurrence Matrix performs well compared to other features when images are similar. In most of the image categories, Autocorrelation feature also shows similar performance[258]. It is also noted that the structural texture features are more effective than the statistical texture features. In case of combination features, combinations recorded better retrieval rate compared to the performances of those individual texture features.

The texture defined as the structural pattern of surfaces which is homogeneous in spite of fluctuations in brightness and color. Most important visual key in identifying regions.

Main categories divided to: regularity, randomness and directionality[259] as displayed in figure 5.1.

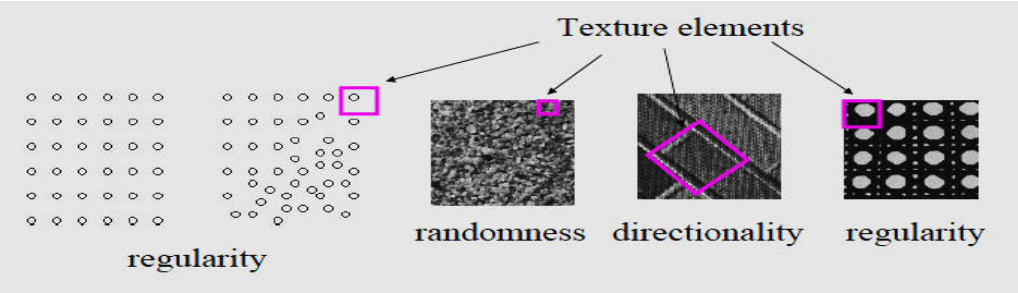


Figure 5.1 displayed the main categories: texture elements, regularity, randomness, directionality and regularity.

5.4.2. Texture features:

With the current enormous scientific image databases present, it is simple to expect any human to be able to analyses understand and extract knowledge from it. Fortunately, in spite of how logically difficult the task of knowledge discovery from such data may appear, it turns out that the process consists of some computational parts that may be automated and actually better executed by computers than by humans. This research focus on intelligent image features extraction, typically by the use of categorized or multiresolution analysis. By trying many experiments, and the aim of this research is to improve some algorithms and develop new ones to be able to generalize to various scientific image analysis problems[260]. Wavelets and other image analysis techniques have shown themselves useful in a wide array of domains such as classification of tissues in computed tomography image retrieval in medical databases.

The texture features divided to [257].

- Statistical approach
  – From first order statistics (histogram)
  – From second order statistics (grey-level co-occurrence matrix)
- Spectral approach -Fourier power spectrum
- Gabor filter
- Multiresolution wavelet

5.4.3. Grey-Level Co-occurrence Matrix and Features

The grey-level co-occurrence matrix P [i, j] is defined by specifying a displacement vector d = (dx, dy) and counting all pairs of pixels separated by d having grey levels i and j as displayed on figure 5.2.

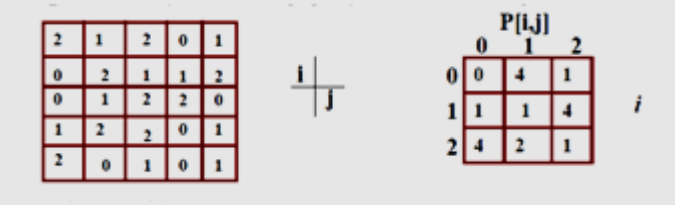


Figure 5.2, showed, (a) images (5x 5) matrix with 3 grey levels 0, 1, and 2. (b) The co-occurrence matrix for d= (1, 1).

The co-occurrence matrices can be computed for 0, 45, 90 and 135 degrees and at distances 1, 2, 3..., statistical features are computed from the co-occurrence Matrices, Entropy is a feature which measures the randomness of gay-level distribution. The angular second-moment is a measure of similarity of the image, the contrast is a measure of the amount of local variations present in the image, and the correlation is a measure of linearity in the image.

The texture matrix used was the grey-level co-occurrence matrix (GLCM). In designing the GLCM for texture representation, there are three fundamental parameters that must be defined: 1. quantization levels of the image, 2. the displacement, and 3. orientation values of the measurements.

The definition of GLCM's [261] is as follows: Suppose an image to be analyses is rectangular and has $N_x$columns and $N_y$rows. Suppose that the grey level appearing at each pixel is quantized to $N_g$levels. The texture-context information is specified by the matrix of relative frequencies P_{ij} with two neighbouring pixels separated by distance doccur on the image; one with grey level i and the other with grey level j. Such matrices of grey-level co-occurrence frequencies are a function of the angular relationship and distance between the neighbouring pixels. Let $p(i,j)$ be the(i,j)th entry in a normalized GLCM. The mean (μ_x, μ_y) and standard deviations (σ_x, σ_y) for the rows and columns of the matrix are in equations (5.1) – (5.4):

$$\mu_x = \sum_i \sum_j i.p\,(i,j), \qquad 5.1$$

$$\mu_y = \sum_i \sum_j j.p\ (i,j), \quad 5.2$$

$$\sigma_x = \sum_i \sum_j (i - \mu_x)^2\ .p\ (i,j), \quad 5.3$$

$$\sigma_y = \sum_i \sum_j (j - \mu_y)^2\ .p\ (i,j), \quad 5.4$$

As per haralick et el [261] suggests that a set of 28 textural features which can be extracted from each of the gray-tone spatial-dependence matrices. The following 5.5 equations define these features.

Notation

$p_x(i)$ ith entry in the marginal-probability matrix obtained by summing the rows of

$$p(i,j), = \sum_{j=1}^{N_g} P\ (i,j)\ . \quad 5.5$$

where N_g Number of distinct gray levels in the quantized image equations 5.6, 5.7, and 555.8:

$\sum_i and \sum_j \sum_{i=1}^{N_g} and\ \sum_{j=1}^{N_g}$, respectively.

$$p_y(j) = \sum_{i=1}^{N_g} p(i,j). \quad 5.6$$

$$P_{X+Y}(K) = \sum_{I=1}^{N_g} \sum_{J=1}^{N_g} p(i,j), \qquad K = 2,3,\dots,2N_g. \quad 5.7$$

$$P_{X-Y}(K) = \sum_{\substack{i+j=k \\ |i-j|=K}}^{N_g} \sum_{j=1}^{N_g} p(i,j), \qquad K = 0,1,\dots,N_g - 1. \quad 5.8$$

5.4.4. Gabor filter feature extraction

The Gabor Filters have received considerable attention because the characteristics of certain cells in the visual cortex (the outer layer of an internal organ) of some creatures can be approximated by these filters. In addition these filters have been shown to keep ***optimal localization properties in*** both spatial and frequency domain and thus are well suited for ***texture segmentation problems***. Gabor filters have been used in many applications, such as texture segmentation, target detection, fractal dimension management, document analysis, edge detection, retina identification, and image coding and representation. A ***Gabor filter*** can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope. It can be written as displayed on equation 5.9:

$$h\ (x,y) = s\ (x,y)\ g\ (x,y) = g\ (x,y)\ e^{-j\ 2\ \pi\ (u_0\ x + v_0\ y)} \quad 5.9$$

Where $s(x, y)$ is a complex sinusoid, known as a carrier, and g (x, y) is a $2 - D\ Gaussian$ shaped function, known as envelope. The complex sinusoid is defined as follows, equation 5.10:

$$s\ (x, y) = \ e^{-j\ 2\ \pi\ (u_0\ x + v_0\ y)} \qquad 5.10$$

In this thesis an investigation of the validity of Gabor filter banks for feature extraction in an image recognition context is presented [259]. in equation 5.11:

$$f_{a,b}(x, y) = \frac{1}{2\pi{\sigma_a}^2} \exp\left\{-\frac{x^2 + y^2}{2\ {\sigma_a}^2}\right\} \cos(2\ \pi(\omega_a\ x \cos\theta_b + \ \omega_a\ y \sin\theta_b)) \qquad 5.11$$

It showed a response from convolving image sample with filter, as displayed in equation 5.12, and figure 5.3,

$$G_{dab}\ (x, y) = \ X_d(x, y) * \ f_{ab}(x, y) \qquad 5.12$$

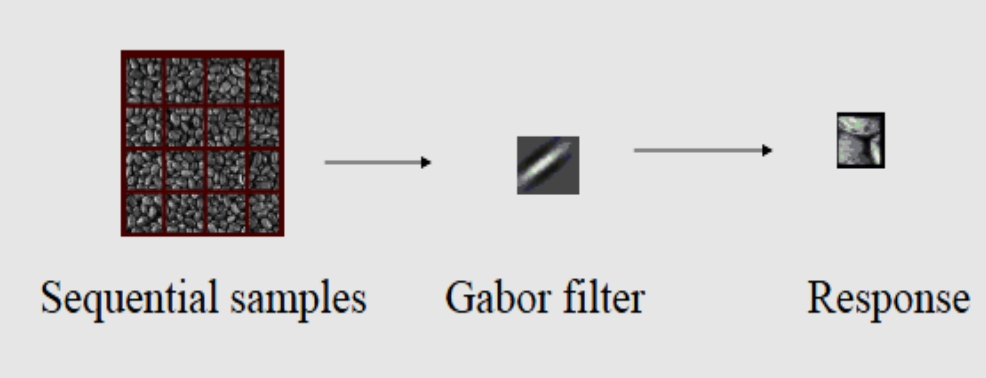


Figure 5.3, displayed the response from convolving image sample with filter.

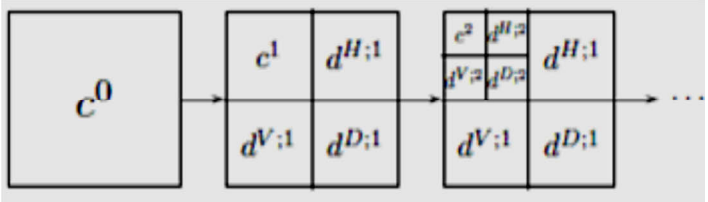
5.4.5. Wavelet decomposition of an image

Discrete wavelet transform (DWT) can be used to decomposing the image into non-overlapping sub bands as shown in Figure 5.4, a) [262], the mean and standard deviation of the decomposed image portions are calculated. These mean and standard deviation are treated as the texture features for image comparison.

The block diagram in figure 5.4, b) showed the methodology of the DWT – Gabor approach as follows. Firstly, The DWT is applied to the biomedical image to obtain its high frequency image component contains most of the desired information about the biological tissue[26]. Secondly, a bank of Gabor filters with different scales and orientations is

applied to the high-frequency image to obtain Gabor-filtered images along different spatial orientations. Thirdly, statistical features are extracted from the Gabor filtered images. Finally, the SVM is used to classify the resulting feature vector for final diagnosis.

(a)



(b)

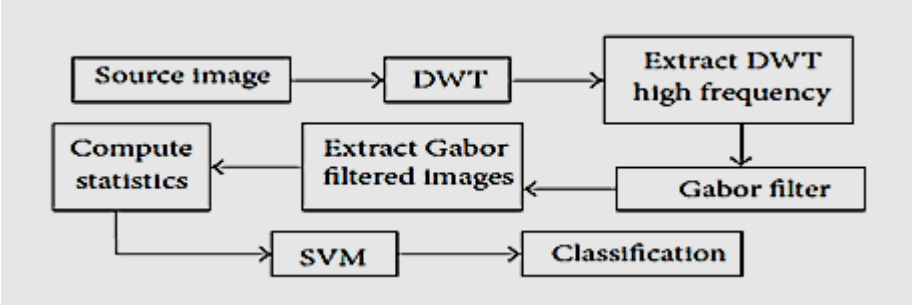


Figure 5.4, (a) Wavelet decomposition of an image (b) Block diagram of the decomposition of an image

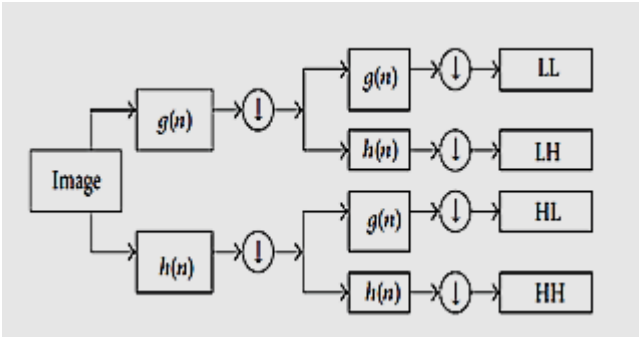


Figure 5.5 displayed a block diagram of the discrete wavelet transforms (DWT – Gabor approach).

5.4.6. Discrete Wavelet Transform

The two-dimensional discrete wavelet transform (2D-DWT) [263] [264] performs a Sub band coding of an image in terms of spectral spatial/ frequency components, using an iterative and recursive process. Figure 5.5, illustrates the case of two-level decomposition. The image is first represented by LH, HL, and HH sub bands that encode the image details in three directions and an LL sub band which provides an approximation of it. The approximation images can be decomposed again to obtain second-level detail and

approximation images, and the process can be repeated for finer analysis as each iterations doubles the image scale as shown in figure 5.4

This thesis used in total 283 features, six features extracted from first order statistics (histogram), twenty one features extracted from second order statistics (grey-level co-occurrence matrix) for each channel and 255 features extracted from Wavelet Packet Transform (WPT).

The multi-resolution wavelet features produce:

- A 2-level wavelet transform of the image is constructed using the orthonormal Daubechies filter. (Figure 5.5).
- The following features 1-3, represented by equations 5.13, 5.14 and, 5.15:

1. Wavelet energy $f_7 = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |C_{ij}|$ 5.13
2. Variance $f_8 = \sqrt{\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |C_{ij} - M|^2}$ 5.14
3. Residual energy $f_9 = = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |C_{ij} - M|$ 5.15

5.4.7. Classifier distance metric equation

It represented by equation 5.31, $D_{ij} = \sum_{r=1}^{q} \left(\frac{{}_r^i f - {}_r^j f}{\alpha\,(f_r)} \right)^2$ 5.31

Where: $\left({}_1^i f\ {}_2^i f, \dots, {}_q^i f\right)$ is the feature vector for texture class I, $\left({}_1^j f\ {}_2^j f, \dots, {}_q^j f\right)$ is the feature vector for texture class j, q is the number of features, and $\alpha(fr)$ is the std. deviation of the feature fr over the C texture classes.

5.5. Feature Extraction and Representation

A data representation (image) must be chosen [236]. Data are represented by a fixed number of features which can be binary, definite or repeats. A feature is identical to input variable or quality. It is sometimes necessary to distinguish between “raw” input variables and “features” that are variables built for the original input variables[265]. In our medical diagnosis example, the features may be indications that are a set of variables categorizing the health status of a patient.

Human expertise, which is often required to convert "raw" data into a set of useful features, can be complemented by automatic feature construction methods. Feature construction is a pre-processing stage. Pre-processing transformations may include[265]:

-Standardization: Features can have different scales although they refer to similar objects.

-Normalization: This interprets into the formula: $x = x/x$.

-Signal enhancement: The signal-to-noise ratio may be improved by applying signal or image-processing filters. These operations include baseline or background removal, de-noising, smoothing, or sharpening. The Fourier transform and wavelet transforms are popular methods. We refer to preliminary books in digital signal processing [266], wavelets[267], image processing [268], and morphological image analysis [269].

-Extraction of local features: These techniques convert problem specific knowledge into the features. It is worth mentioning that they can bring significant improvement.

-Linear and non-linear space embedding methods: When the dimensionality of the data is very high, some techniques might be used to implant the data into a lower dimensional space while retaining as much information as possible. Principal Component Analysis (PCA) and Multi-Dimensional Scaling (MDS)) [270].

-Feature discretization: Some algorithms do no handle well continuous data. it may simplify the data description and improve data understanding[271].

Texture is one of the important characteristics used in identifying objects or regions of interest in an image, whether the image is a photomicrograph, or a satellite image.

5.5.1. Methodology used in feature extraction

The literature topics in this study are shown in a thematic way rather than in a methodological way. In this section, we present a unified view of feature selection that exceeds the filter/wrapper and is inspired by the views of [271]. There are (4) four aspects of feature extraction:

1. feature construction;
2. feature subset generation (or search strategy);
3. Evaluation criterion definition (e.g. relevance index or predictive power);

4. Evaluation criterion estimation (or assessment method).

The last three aspects are relevant to feature selection and are schematically summarized in Figure 5.6. Filters (a) and wrappers (b) differ mostly by the evaluation

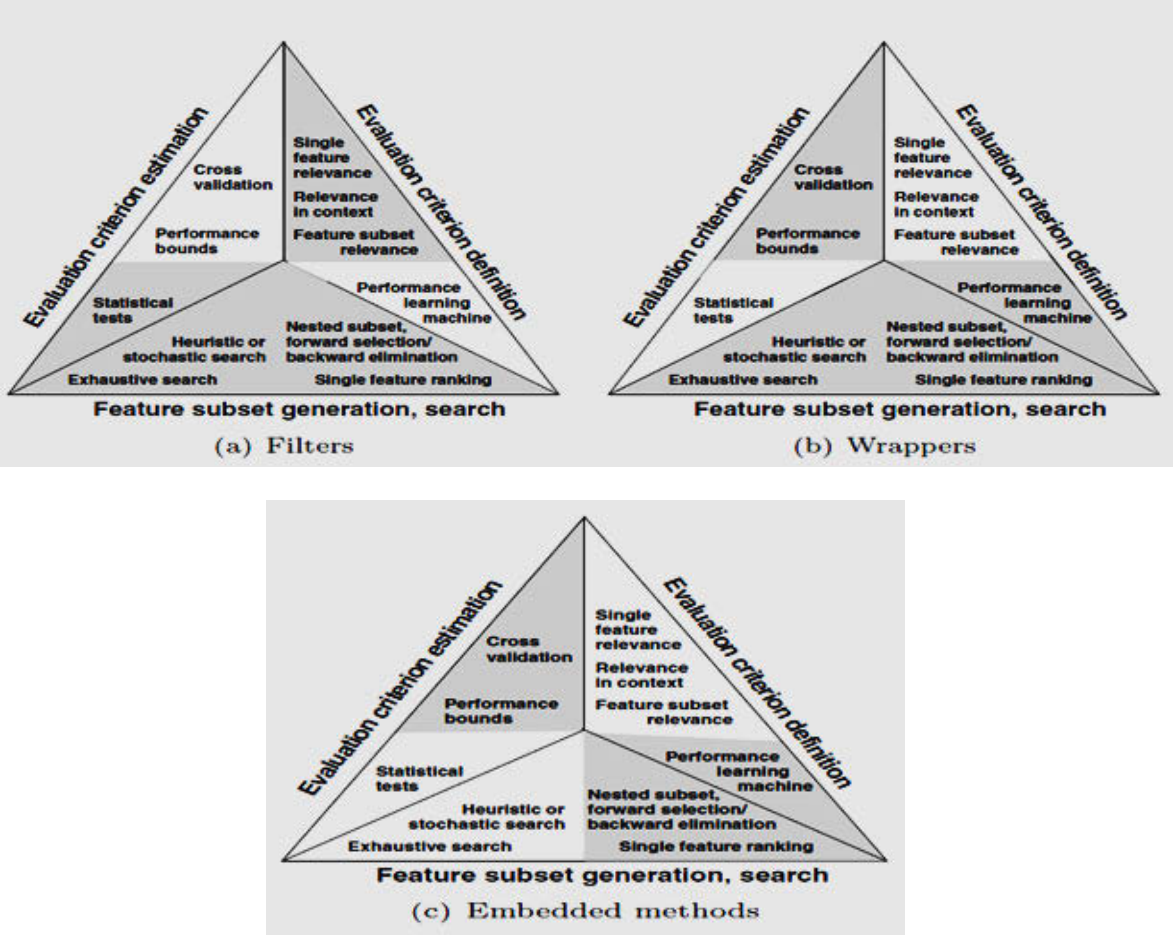


Figure 5.6, the three principal approaches of feature Mammographic e selection. The shades show the components used by the three approaches: filters, wrappers and embedded methods.

criterion. It is usually understood that: filters use criteria not involving any learning machine, e.g. a relevance index based on correlation coefficients or test statistics, whereas wrappers use the performance of a learning machine trained using a given feature subset.

5.6. Classification and Feature selection

5.6.1. Feature Selection (FS)

Researcher on [265] reported that: the problem of feature extraction rotting into (2) two steps: 1) ***feature construction:*** is one of the key steps in the data analysis process, it

compare the performances obtained with either image. 2) ***Feature selection:*** is primarily performed to select relevant and useful features. It is an important problem for pattern classification systems. We study how to select ***good features*** according to the maximal statistical need standard based on common information. We first derive an equivalent form, called ***minimal-redundancy-maximal-relevance criterion*** (***mRMR***), for first-order incremental feature selection. Then, we present a two-stage feature selection algorithm by combining mRMR and other more sophisticated feature selectors (e.g., ***wrappers***). This allows us to select a compact set of superior features at very low cost. [***Superior features = mRMR + wrappers = very low cost***].

Although Support Vector Machine ($SVMs$) as classifier directly determine the decision boundaries in the training step and the method can also provide good generalization in high-dimensional input spaces, many researchers reported that feature selection is important for $SVMs$[272]. In general, the major aims of feature selection for classification are finding a subset of variables that result in more accurate classifiers and constructing more compact models. An amount of variables can be redundant to the classification problem and feature selection can filter out the variables that are unrelated for the model. Three major types of feature selection methods are distinguished, ***namely filter, wrapper and embedded methods***[28] [273, 274] [275].

- **Filter methods**

Filter methods employ an initial analysis as a pre-processing step, where a set of features is selected as input for the classifier using a training data set[276]. Researcher in[277], proposed that the selection procedure should take the learning algorithm into account.

- **Wrapper methods**

In wrapper models [276], a subset of variables is selected based on a search by actually using the classifier. The best subset of features is obtained by estimating and comparing the performances of the various subsets of features, which are fed to the classifier. An complete search through the input space is often not feasible, therefore experimental search methods using ***backward, forward or stepwise*** variable selection are employed [278]. In addition, more sophisticated methods like ***best-first*** search are also able to navigate the space of subsets [279] [276]. For the evaluation of each subset, n-fold cross-validation or leave-one-out ***(LOO) cross-validation*** can be used. In general the wrapper models result in an ***increased accuracy*** because potential effects of the classification algorithm are taken

into account [276]. On the other hand, in principle, wrapper methods involve the higher computational cost of a search. A popular algorithm that combines SVMs with a backward search is that of ***recursive feature elimination (RFE***) [280]. In this approach the role of a feature is estimated by the change in the cost function when such a feature is removed. For this purpose, the weights in the SVM model are used while support vectors are fixed for fast feature elimination. An efficient algorithm for variable selection in high-dimensional scenarios was proposed in reference[281].

- **Embedded methods**

In contrast to filter and wrapper methods, embedded methods aim to immediately integrate the variable selection or weighting procedure in the learning algorithm. Gradient-based algorithms are used to perform feature selection by minimization of generalization error bounds. Reference [282] showed the radius-margin bound for SVM is optimized with respect to feature scaling factors, which reflect the contribution of features to the prediction rule. The work by[283]. provides derivations for several SVM generalization bounds in terms of the feature scaling factors[283]. Such formulations can serve as a methodology for multiple hyper-parameter tuning. Within this category ***automatic relevance determination (ARD),*** aims to identify informative input features as a natural result of optimizing the model selection criterion [283] [284]. Tuning parameters of $SVMs$ and Kernel and other kernel-based models usually depend on several parameters controlling the model complexity. In SVM parameters, α_i and the bias term b, are determined by solving the rounded optimization problem, while the positive penalization constant (C) and the kernel parameters remain. So as C decreases, the width of the margin increases. In the setting of $SVMs$ with $RBF\ kernel$ or $\tau\ (\geq 0)$ in polynomial kernel, determine the shape of the decision boundary as well as the dimension and complexity of the matching induced feature space. A typical a priori setting for the RBF kernel parameter σ is that of using the distance between the closest points from different classes. However, applications rely on the use of a validation set or cross-validation to choose hyper-parameter values. In practice, cross-validation-based techniques are preferred over generalization error bounds. Researcher in[285] developed an interesting benchmark study which presents an efficient methodology for hyper- parameter tuning and model building using Least squares support vector machines ($LS-SVMs$) as benchmark. The approach is based on the closed form leave-one-out (LOO) computation for $LS-SVMs$, only requiring the same computational

cost of one single $LS-SVM$ training. Performance compares favorably to bound-based techniques with the advantage of exact and efficient LOO computation.

5.6.2. Unbalanced data sets

In many binary classification problems, data sets are often skewed in favor of one class such that the contributions of false negatives and false positives to the performance assessment criterion are not balanced. In such cases the sizes of positive and negative points in the training data are not truly typical for the operational class frequencies. As mentioned in literature, appropriate prior class probabilities need to be specified for Bayesian methods and can therefore take an unbalance in the data into account. In case of standard $\boldsymbol{SVMs}$ or $\boldsymbol{LS-SVMs}$ a constant bias term correction can be performed[286]. Furthermore, different weights can be introduced for positive and negative data points, thereby balancing their contributions to the regularized loss function. Taking into account the unbalance in the data set, the objective functions for ***SVM and LS-SVM*** displayed on equations 5.34 5.35 respectively.

$$\min_{w\xi b} \tau_4 (w,\ \xi) = \frac{1}{2} w^T \mathrm{w} + \mathrm{C} \sum_{i=1}^{N} v_i \xi_i, \qquad (SVM) \tag{5.34}$$

$$\min_{wcb} \tau_5 (w,\ e) = \frac{1}{2} w^T \mathrm{w} + \frac{\gamma}{2} \sum_{i=1}^{N} v_i {e_i}^2, \qquad (LS-SVM) \tag{5.35}$$

With: $v_i = \begin{cases} \frac{N}{2N_P} & ,if\ y_{i=+1,} \\ \frac{N}{2N_N} & ,if\ y_{i=-1,} \end{cases}$,

Where N_P and N_N, represent the number of ***positive and negative data points***, respectively. Asymptotically this weighting is equivalent to resampling of the data such that there are an equal number of positive and negative data points.

Sequential Feature Selection (SFS) for Classification Sequential Feature Selection (SFS) will insert Reinforcement Learning (RL) into classification tasks, with the objective to reduce data consumption and associated costs of features during classification [287]. Feature selection with RL explained in [287]. Researcher in [288]reported that the decision process is not dependent on the internal state of the classifier, which brings their method closer to straight FS. Researchers in [289] reported that they approach reduces the number of necessary features to access to a fraction of the full input, down to 12% with RNN

classifiers. They also demonstrated that SFS is able to deal with weighted feature costs, a property that exists in plenty of real-world applications.

5.7. SVM for Classification and Regression

Support vector machines (SVM's) are currently a hot topic in machine learning community and are becoming popular in a wide variety of biological applications. A support vector machine is a computer algorithm that learns by example to assign labels to objects [290].

Researchers in [291] presented the advantage of SVM for its generalization capability. SVM utilizes procedure for reducing the training data points which participates in defining discriminant function. Thus, the participating data points, which are called support vectors, are minimized. Data points which do not participate in defining a classifier are ignored. This can reduce noise and, consequently, can improve the generalization capability. SVM has proved mostly good performance for classification in varied applications, both for real and artificial standard benchmarking data [292], including applications in medical fields. The problem of experiential data modeling is relevant to many engineering applications.

5.7.1. Linear support vector machine

SVM is a classifier which works through deciding an optimal hyperplane which optimally hyperplane which optimally separates two class data. The optimal hyperplane is also called a decision surface which is the surface optimally separating two groups of data points. Suppose there are l linearly-separable training-data xi $(i = 1,2,3, \dots, k)$ and the associated $y_i \in [-1, +1]$ is a class label $y_i = +1$ is for the class $+1$ and $y_i = -1$ is for the class –1. For example, the class $+1$ is for the data which belongs to benign, and –1 is for melanoma.

In the case of two–dimensional space, as shown in Figure 5.7 two groups of data (circles and squares) can be separated by different hyperplanes. Two possible hyperplanes are shown in Figure 5.8 (a) and (b). In SVM, the optimal hyperplane is right in the middle of the two class boundaries and the distance of the two boundaries is maximized. The optimal hyperplane promises good classification while facing unseen data and provides a good generalization. In Figure 5.7 the two parallel lines in the left and right of the hyperplane are supporting hyperplanes. The supporting hyperplanes are parallel to the hyperplane. The decision surface (or hyperplane) and the supporting hyperplanes are defined as follow:

The decision surface: $$w.x + b = 0 \quad (5.36)$$

Supporting hyperplane for class +1: $w.x + b = +1$ 5.37

Supporting hyperplane for class -1: $w.x + b = -1$ 5.38

Where w the perpendicular is distance from the hyperplane to the origin and b is called bias.

The distance between the two supporting hyperplanes is referred to as a margin(m). The optimization problem of SVM is to find the decision boundaries so that the distance between the supporting hyperplanes is as far as possible. Therefore, the optimization is to maximize the margin and to keep the decision surface equidistant from the two supporting hyperplanes. As in Figure 5.8, the margin is constrained by the points which are indicated by the filled circles and the filled squares. These points are called support vectors.

Because the optimization problem of SVM is to find the optimal hyperplane, the feasible solutions to the optimization problem are all possible hyperplanes, with their associated supporting hyperplanes. The constraints are the positions of the supporting hyperplanes in order not to cross their respective class boundaries. Considering Figure 5.8, the margin (m) between the two supporting hyperplanes can be defined as (more detailed information in ***appendix D***)

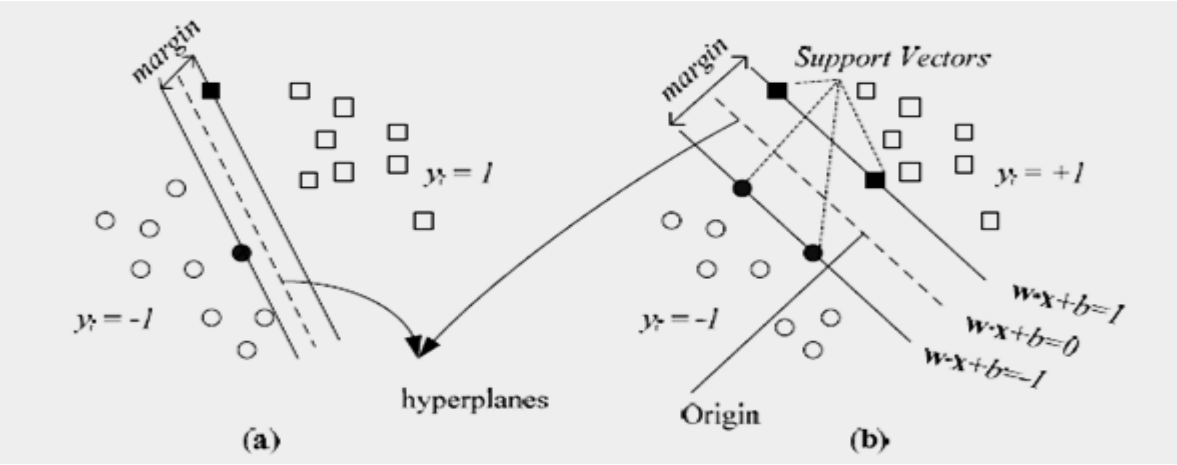


Figure 5.7: Two-out-of-many separating lines; (a) with smaller margin and (b) with larger margin

$$m = \frac{2}{|w|} \quad 5.39$$

Thus, the maximum margin $\hat{m}$ is obtained by maximizing Eq. 5.40, as

$$\hat{m} = max \frac{2}{|w|} \quad 5.40$$

$$\hat{m} = min \frac{|w|}{2} \quad 5.41$$

Eq. 5.41 can also be stated as:

$$\hat{m} = min\ |w|^2/2 \quad 5.42$$

Because optimizing over $|w|^2$ is the same as optimizing$|w|$. Eq. 5.42 can be formulated as:

$$\hat{m} = min\ ½\ w.w \quad 5.43$$

As mentioned above, the constraints of the optimization problem are the positions of the supporting hyperplanes which do not cross their respective class boundaries. Mathematically, it can be formulated as:

$w.x_1$+b$\geq +1$ for all (x_i, y_i) such that $y_i = +1$ 5.44

$w.x_1$+b≤ -1 for all (x_i, y_i) such that $y_i = -1$ 5.45

Using a more compact formula, it can be stated as:

$$y_i(w.x_i + b) \geq 1 for\ all\ (x_i, y_i) \quad 5.46$$

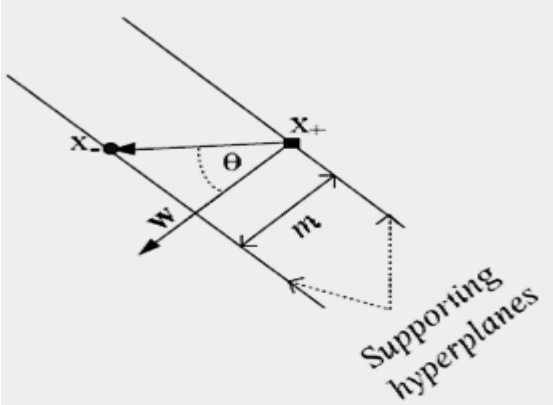


Figure 5.8: Margin m between two supporting hyperplanes.

The optimization problem in Equation 5.43 is solved using the Lagrangian. The construction of the Lagrangian for the optimization problem is:

$$L(\propto, w, b) = \frac{1}{2} w.w - \sum_{i=1}^{l} \propto_i (y_i(w.x_i - b) - 1) \quad 5.47$$

where α is a Lagrangian multiplier. This gives the Lagrangian optimization problem as:

$$\underset{\propto}{maxmin}\ L(\propto, w, b) \quad 5.48$$

Subject to

$$\propto_i \geq 0 \qquad 5.49$$

This provides the Lagrangian dual optimization problem as follow:

$$L(\propto, w, b) = \frac{1}{2} w.w - \sum_{i=1}^{l} \propto_i y_i w.x_i + b \sum_{i=1}^{l} \propto_i y_i + \sum_{i=1}^{l} \propto_i \qquad 5.50$$

and the optimal decision surface is defined as (see ***Appendix E***)

$$f(x) = sgn(w.x - b) \qquad 5.51$$

$$f(x) sgn(\sum_{i=1}^{l} \propto_i y_i x_i . x - \sum_{i=1}^{l} \propto_i y_i x_i . x_{sv} - 1) \qquad 5.52$$

5.7.2. Nonlinear support vector machine

The SVM which is mentioned above is a linear SVM. The linear SVM needs to be extended to tackle non-linearly separable data, as data in the real world are, mostly, not linearly separable. The idea of the extension is by using the kernel trick. Using a kernel function, data in the input space is transformed to a higher dimensional space, which is called a feature space. In this feature space it is possible to separate the data using a hyperplane. The transformation is illustrated in Figure 5.10. The left side of figure 5.10, the data points cannot be classified by a linear function. After the data points are transformed in three dimensional spaces, a linear function can be used to classify the data points. The right side of Figure 5.9, a plane can separate the data points; two triangles are above the plane and two squares are below the plane.

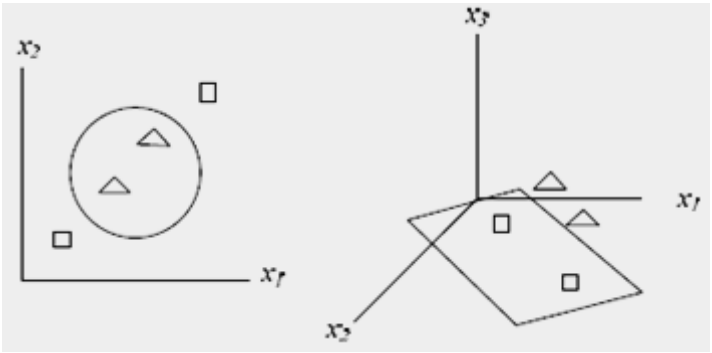


Figure 5.9: Illustration of mapping using a transform Ψ: $\mathbb{R}^2 \rightarrow \mathbb{R}^3$.

Using the kernel trick, the Lagrangian optimization problem is modified by replacing x_i using a mapping function Ψ (x) as:

$$L(\propto, w, b) = \sum_{i=1}^{l} \propto_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j \, \Psi(x_i).\Psi(x_j). \qquad 5.53$$

In the compact form Eq. 5.53 can be written as:

$$L(\propto, w, b) = \sum_{i=1}^{l} \propto_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i\alpha_j y_i y_j \; k\,(x_i . x_j). \quad 5.54$$

where k $(x_i.x_j)$ is a kernel function. Kernel functions include:

Radial basis function (RBF), $k(x_i, x_j) = \exp(-\gamma x \|x_i - x_j\|^2)$ 5.55

Sigmoid, $k(x_i, x_j)$=$(-\gamma(x_i - x_j) + 1)$ 5.56

Polynomial, $k(x_i, x_j) = x(x_i x_j + 1)^d$ 5.57

Linear, .k (x_i, x_j) =$x_i x_j$ 5.58

5.7.2.1. Soft-margin nonlinear support vector machine

A soft-margin classifier of SVM is useful to tackle imperfect data containing a noisy data point, such as that which is due to improper measurement. In this soft-marginSVM, a slack variable ξ is introduced. A slack variable measures how much of an error is introduced by allowing the supporting hyperplane to be unconstrained by a data point (Figure 5.10). By introducing a nonnegative slack variable, the corresponding modified constraint is:

$$y_i(w.x - b) + \xi - 1 \geq 0 \quad 5.59$$

and the optimization problem in 5.43 becomes

$$\hat{m} = \min(\frac{1}{2} w.w + C\sum_{i=1}^{l} \xi_i) \quad 5.60$$

The introduction of the slack variable as in the form of $C\sum_{i=1}^{l} \xi_i$ is to prevent the Construction of trivial solutions; where all training points are considered as noise during the optimization. C is a constant, called the cost, controlling the trade-off between margin size and error. By this slack variable introduction, the Lagrangian construction of the optimization problem is (see ***Appendix F***)

$$L(\propto, w, b) = \sum_{i=1}^{l} \propto_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i\alpha_j y_i y_j k(x_i . x_j) \quad 5.61$$

This Lagrangian dual for soft-margin SVM is the same as the Lagrangian dual for hard-margin SVM. Therefore the objective functions of the optimization problems for both $SVMs$ are the same. The difference is in the constraint. The constraint in the soft-margin SVM can be stated as:

$$0 \leq \propto \leq C \quad 5.62$$

In a case of an imbalanced data point number between two classes, different error weights, w_0 and w_1, are necessary to penalize more heavily the class with the smaller population[293].

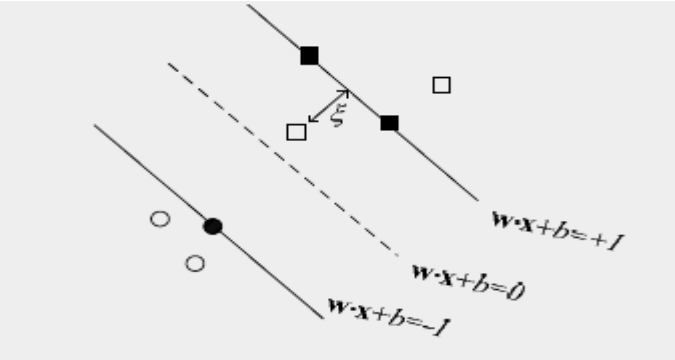


Figure 5.10: Introducing slack variable ξ in soft-margin SVM

$$\hat{m} = \min \frac{1}{2} w \,.\, w + w_0 C \sum_{i(y_i=-1)=1}^{l} \xi_i + w_1 C \sum_{i(y_i=+1)=1}^{l} \xi_i) \quad 5.63$$

In summary, the decision function of support vector machine can be represented as follow

$$f(x) = sgn(\sum_{i=1}^{l} y_i \propto_i k(x_i, x) - b) \quad 5.64$$

In which α is the solution of maximizing the following Lagrangian

$$\max_{\propto} L(\propto) = (\sum_{i=1}^{l} \propto_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \propto_i \propto_j y_i y_j k(x_i . x_j)) \quad 5.65$$

Subject to constraints

$$\sum_{i=1}^{l} \propto_i y_i = 0 \quad 5.66$$

$$0 \leq \propto_i \leq C \quad 5.67$$

The optimization of Eq. 5.65 uses the method based on sequential minimal optimization (SMO)[294, 295],[296]. In the optimization, SMO uses analytical steps, instead of numerical quadratic programming. In each iteration SMO chooses two Lagrangian multipliers and optimize these two multipliers. Two Lagrangian multipliers at every step is the smallest number in the optimization because the Lagrangian has to obey a linear equality constraint (Eq. 5.66).

The optimization which uses an SMO could be faster than the optimization which uses quadratic programming (QP). Using QP, the optimization for Eq. 5.65 involves a matrix with a number of elements that equals to the square of the number of training examples.

This number obviously needs a large memory and can result in a very slow optimization. The description of the optimization using SMO is presented in ***Appendix G*** [297].

5.7.3. Support Vector Regression

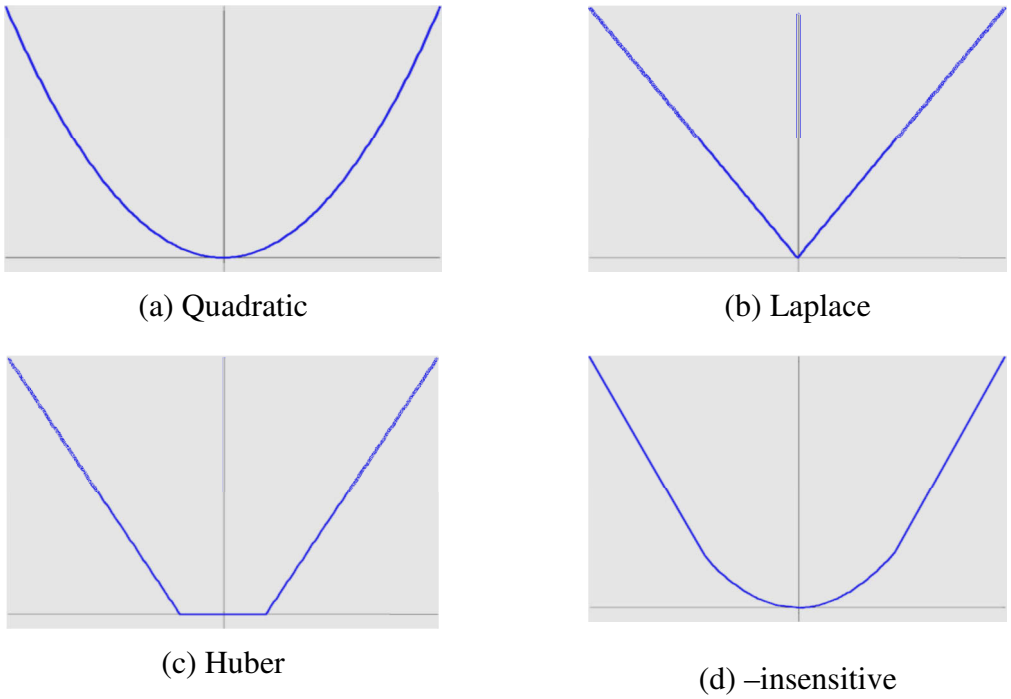
SVMs can also be applied to regression problems by the introduction of an alternative loss function [298]. The loss function must be modified to include a distance measure. Figure 5.12 illustrates four possible loss functions.

The loss function in Figure 5.11(a) corresponds to the conventional least squares error criterion. The loss function in Figure 5.11(b) is a Laplacian loss function that is less sensitive to outliers than the quadratic loss function. Huber proposed the loss function in Figure 5.11 (c) as a robust

Loss function that has optimal properties when the underlying distribution of the data is unknown. These three loss functions will produce no sparseness in the support vectors. Researcher in [299] proposed the loss function in Figure 5.11(d) as an approximation to Huber's loss function that enables a sparse set of support vectors to be obtained. The Support Vector Regression divided to: ***Linear Regression***, and ***Non Linear Regression***.

5.7.3.1. Applications in Support Vector Regression

Larger and more complex classification problems have been attacked with SVC. Researcher in [300] has applied SVC to the exacting problem of face recognition, with encouraging results. In conclusion, SVC provides a robust method for pattern classification by minimising over fitting problems by adopting the SRM principle. Use of a kernel function enables the curse of dimensionality to be addressed, and the solution implicitly contains support vectors that provide a description of the significant data for classification



(a) Quadratic (b) Laplace

(c) Huber (d) –insensitive

Figure 5.11: Loss Functions

5.8. Swarm based Support Vector Machine for Melanoma Detection

A hybrid technique of swarm-based support vector machine (SVM) is introduced for melanoma detection using the pathology images as inputs. In this technique, a particle swarm optimization (PSO) is proposed to optimize the SVM to detect melanoma.

5.8.1. Development of a Melanoma Detection based on the Swarm based Support Vector machine

The structure of the swarm based support vector machine ($SSVM$) is presented in Figure 5.12. The input was pathology image, and the output was the melanoma or benign. The parameters of SVM were optimized using PSO (the optimal parameters of SVM were found using PSO). The SVM parameters are C, w_0, w_1, γ, d. The parameter of C deals with a trade-off between maximum margin and the classification error of $SVM. SVM$ with high C provides perfect classification in the training phase which uses a training data set, but it may be poor in generalization (or bad performance in classification using a test data set). On the other hand, SVM classification with low C might give a bad performance in the training phase. Therefore, optimal C has to be chosen in order to find the optimal

classification. Likewise, the optimal values of parameters γ and d also need to be found to obtain the optimal performance.

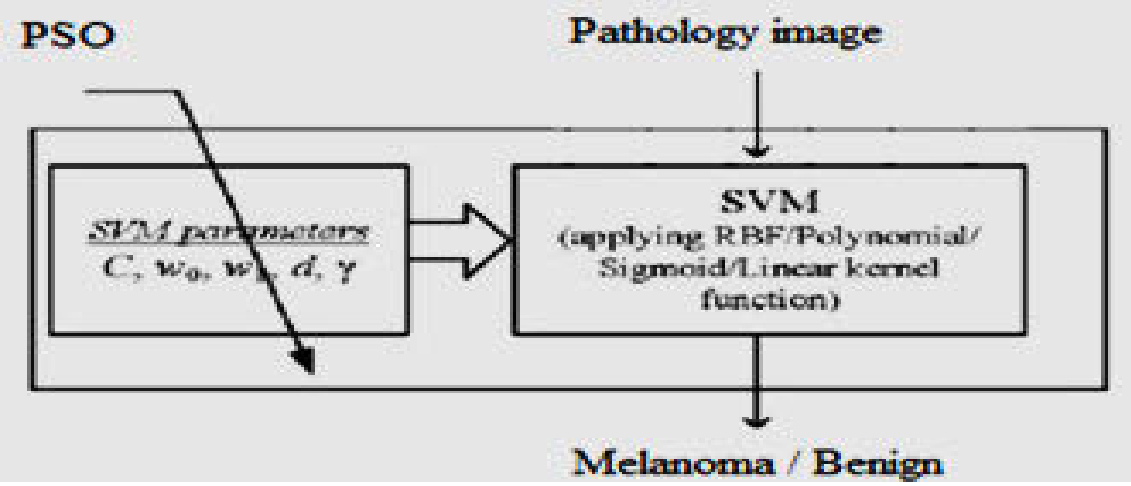


Figure 5.12: Melanoma detection using swarm based support vector machine.

5.8.2. Optimization of SVM parameters using PSO

This research utilizes the PSO algorithm to find the optimal SVM parameters according to a defined fitness function. Four SVM parameters are optimized; they are C, γ (for RBF and sigmoid kernel functions) or d (for polynomial kernel function), w_0 and w_1(weight factors for Benign and melanoma class data, respectively).

The optimization of SVM parameters using PSO can be described as in Figure 5.13. Suppose a swarm at n-th iteration is$z(n)$. It contains ω particles with τ dimensions. More clearly, in this work four parameters are used, and therefore, the τ is 4. The number of particles used in the optimization of this work is 50. Each element of the swarm is presented by$z_j{}^k Z(n)$, in which $j = 1, 2, \dots, \tau$ and$k = 1, 2, \dots, \omega$; τ is the dimension of a particle and ω denotes the number of particles in the swarm. The particles can be expressed as in the matrix in Figure 5.14.

```
Begin
    n → 1          // iteration number
    Initialize z          // z: position
    Evaluate f(z)       // f( · ): fitness function (Eq. 5.81)
    Initialize v        // v: velocity
    z̃ = z              // z̃: personal best position
    ẑ = z̃         // ẑ: global best position
while (not termination condition) do
    begin
n → n + 1
Update position of particle z_j^k (n) and velocity v_j^k (n) based on Eqs. 5.78 and
4.2, respectively.
if v(n) > vmax, v(n) = vmax end
if v(n) ≤ vmax, v(n) = −vmax end
Evaluate f(x(n))    // f(x(n)) is defined in Eq. 5.81
Update z̃ if the new position is better than the previous z̃
Update ẑ if the new position is better than the previous ẑ
    end
end
```

Figure 5.13: The pseudo of the PSO for the SVM parameter optimization

The algorithm was initialized by generating random numbers for position and velocity. At iteration n, the position ${z_j}^k\ (n)$ is determined as mentioned by [301]equation 5.68:

$${z_j}^k\ (n) = {z_j}^k\ (n-1) + {v_j}^k\ (n) \qquad 5.68$$

in which the velocity ${v_j}^k\ (n)$ is defined as[301] in equation 5.69:

$${v_j}^k\ (n) = \varphi {v_j}^k\ (n-1) + c_1 r_1 \left(\tilde{z}_j^{\,k} - {z_j}^k (n-1)\right) + c_2 r_2 \left(\hat{z}_j - {z_j}^k (n-1)\right) \qquad 5.69$$

$\tilde{z}$ is the personal best position, which is the ***best position*** of the particle at iteration n, $\hat{z}$ is the global best position which is the best position among all the particles from the first iteration until iteration n. φ is an inertia weight factor, which controls the convergence of the optimization behaviour. $c1$ and $c2$ are the acceleration constants, which control the distance, moved by a particle. $r1$ and $r2$ are the uniform random numbers between 0 to 1[302].The constants $c1$ and $c2$ are set to 2; and $r1$ and $r2$ are the random number in the range [0, 1].

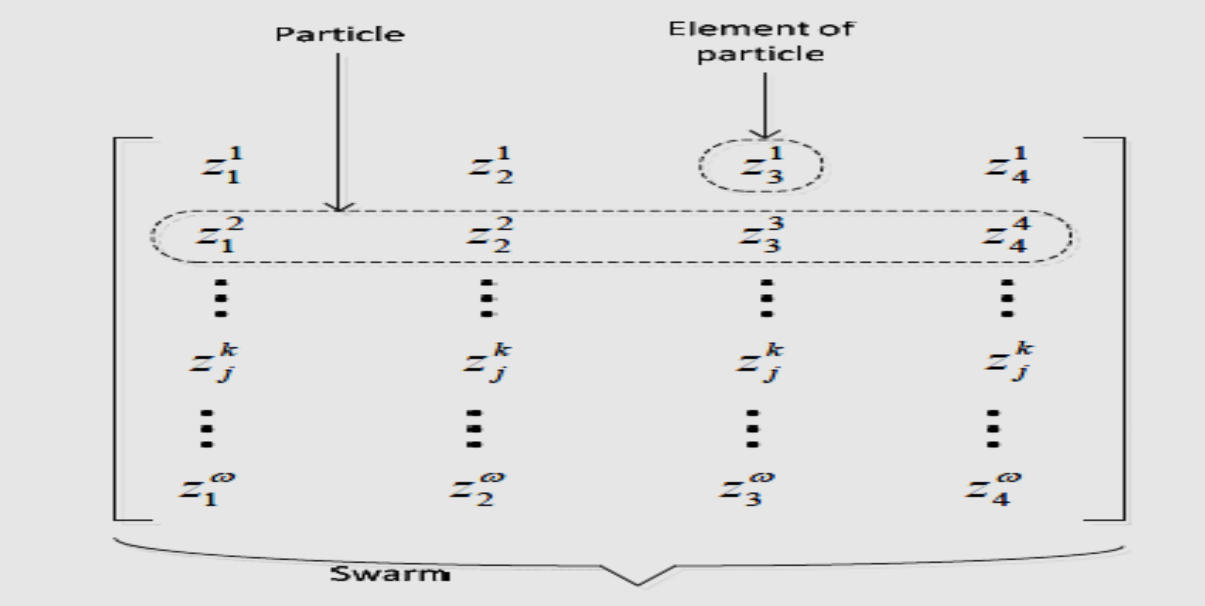


Figure 5.14: The particles of the PSO

The velocity v is limited in a certain region, which is between $-v_{max}$ and v_{max}. Thus, if v is more thanv_{max}., v is set to $v_{max}$and if v is less than$-v_{max}$, v is set to$-v_{max}$. This strategy is called velocity clamping, which is used to limit the swarm particles to a search space[303]. For this optimization, v_{max} is set to 0.2. Furthermore, the inertia weight is set as in equation 5.70:

$$\varphi = \varphi_{max} - \frac{\varphi_{max}-\varphi_{min}}{N}\ x\ n \qquad 5.70$$

where φ_{max} and $\varphi_{min}$are upper and lower inertia weights, respectively, and are set to 1.2 and 0.1, respectively, N is the total iteration number and n is the iteration number.

5.8.3. Fitness function for the optimization

The objective of the PSO optimization was to maximize the performance of the hypoglycaemia detection. The performance was measured in terms of sensitivity ϑ and specificity δ, as defined in Equations 5.45 and 5.46. Sensitivity was defined as the ratio of correct detection of melanoma to the actual number of melanoma cases. Specificity was defined as the ratio of correct detection of benign to the actual number of benign cases. Thus, the PSO optimization was essentially used to maximize sensitivity ϑ and specificity δ , and then the fitness function was defined as in equation 5.71:

$$f = -(k\vartheta_{tr} + (1-k)\delta_{tr} + k\vartheta_v + (1-k)\delta_v + \text{ℏ}) \quad 5.71$$

ϑ_{tr} and δ_{tr}were, respectively, the sensitivity and specificity of the training ϑ_v andδ_v, respectively, were the sensitivity and specificity of the validation. The training data set was used for them SVM training to find a SVM model. The validation data set was used to test the SVM model during the optimization.

The inclusion of ϑ_v and δ_v in the fitness function is to reduce the risk of overtraining[302]. Overtraining could yield a high performance in the training stage, but it might provide a low performance in the testing. In the fitness function, κ was set to 0.58 to obtain a higher sensitivity than specificity. This strategy is to prevent the risk of the low sensitivity of the melanoma detection. Furthermore, to force the detection to have high sensitivity, parameter ℏ is given by using the definition in equation 5.72:

$$\text{ℏ} = \begin{cases} 10 \; if \;\; \vartheta_{tr} > 0.7, \delta_{tr} > 0.4, \vartheta_v > 0.7, \delta_v > 0.4 \\ 0 \; otherwise \end{cases} \quad 5.72$$

By the definition of ℏ in Eq. 5.72 the detection is forced to find sensitivity and specificity of more than 70% and 40%, respectively. Considering Eq. 5.71 and Eq. 5.72, if the detection provides sensitivity and specificity of more than 70% and 40%, respectively, the fitness function has the value of more than 10. Conversely, if the detection provides sensitivity and specificity of less than 70% and 40%, respectively, the fitness function has the value of less than 2; in this formula the maximum sensitivity and specificity is 1 (or 100%).

Chapter Summary

This chapter ‘image representation and analysis’ includes introduction; machine learning; main concepts; intelligent image features extraction such as Texture feature, Grey-Level Co-occurrence Matrix, Gabor filter feature extraction, Wavelet decomposition of an image; Discrete Wavelet Transform; Classifier distance metric equation. Feature extraction and representation includes Methodology that is used in feature extraction; classification and feature selection. Other machine learning techniques for classification and regression like Linear support vector machine, Nonlinear support vector machine, Support Vector Regression; swarm based support vector machine for melanoma detection are also

discussed for the development of a Melanoma Detection based on the Swarm based Support Vector machine. Optimization of SVM parameters using PSO and Fitness function for the optimization are two other areas covered into this chapter.

CHAPTER 6

Experimental Results and Discussion

6.1. Introduction

The rate of skin cancer is rapidly increasing throughout the world including Australia [1]. Australia has one of the highest skin cancer rates in the world at nearly four times the rates in Canada, the US and the UK [76] [304]. It has been estimated 115,000 new cases of cancer are diagnosed and more than 43,000 people are expected to die from cancer this year according to the Cancer council of Australian 2010 [305], The Chair of Public Health Committee pronounced that, more than 430,000 cases treated for non-melanoma, and more than 10,300 people are treated for melanoma, with 1430 people dying each year [305]. At least two in three Australians will be diagnosed with skin cancer before 70 years of age [306], Skin cancers accounted for over 80% of all cancers diagnosed in Australia [307]each year. Skin cancer costs the health system around $300 million Australian dollar annually, the highest cost of all cancers. Melanoma has near 95% cure rate if diagnosed (detected) and treated in the early stages [305]. Visual diagnosing guide to nick-eye error since it is visually hard for medical professionals to differentiate normal from abnormal moles, and general practitioners (GP) do not have the core expertise to diagnose skin cancers. Dermatologist's data indicate that even in specialized centers diagnostic accuracy is only about 60% [308], but they are overloaded by GP referrals. In order to reduce the occurrence of skin cancer, computer- aided detection and diagnosis technologies have been developed. The extraction of the skin cancer border is a key technology in computer- aided detection technologies for skin cancer diagnosis. Previously several technologies / methods have been proposed. Several published classification systems show accuracy rates ranging from 60%- 92% [309] which coincides with the estimated rates obtained by general practitioners. These techniques were aiming to be able to provide recommendation for nonspecialized users. However the variations of diagnosis are sufficiency large and there is a lack of detail of the test methods. One commercial product, Solar Scan by Polar technics, has an accuracy rate of 92%. Solar Scan is a complex system, taking high quality Epiluminescence Light Microscopy (ELM) images and using advanced image analysis

techniques to extract a number of features for classification which it makes it unsuitable for a normal person to use. The traditional imaging is just a recording of what the human eye can see by digital camera while the Dermoscopy known as Epiluminescence Light Microscopy (ELM) images needs an experienced professional to get the required image.

This chapter is organized in 6 Experiment Based on Digital and pathological images, **Experiment 1:** These experiments investigate skin cancer segmentation using an active contour model, or snake model, driven by Gradient Vector Flow (GVF). **Experiment 2**: These experiments propose two segmentation methods to identify the normal skin cancer oriented texture and animation edges and then finding the accuracy by using the three layers back-propagation neural network classifier (BNN). **Experiment 3**: This experiment proposes an automated non-invasive system for skin cancer (melanoma) detection based on Support Vector Machine classification. The proposed system uses a number of features extracted from the Wavelet or the Curvelet decomposition of the grayscale skin lesion images from the original color images. **Experiment 4**: This experiment used a novel methodology for automatic feature extraction from histo-pathological images and subsequent classification. These techniques added Gabor filter bank to winner and adaptive median filters to improve diagnostic accuracy. Then Histogram equalization was used to enhance the contrast of the images. The extracted features were reduced by using sequential feature selection then, finally the obtained statistics was fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi. **Experiment 5**: In this experiment, 42 images were sampled from microscopic slides of skin biopsy and have been used in the current work. The algorithm used in these experiments was a 5-fold cross validation test. The experiment used LIBSVM [310] library for support vector machines, with linear, (RBF) Radial Basis Function kernel [310], and Polynomial kernel, under each magnification. The melanoma images that are confirmed by the pathologist are treated as negative images, while the nevus ones are treated as positive images. **Experiment 6**: This experiment proposed an automated system uses a wavelet packet transform (WPT) to extract more features. This thesis introduces a novel melanoma detection strategy using a hybrid particle swarm based support vector machine (SSVM) algorithm. The extracted features are reduced by using a particle swarm optimization (PSO were used to optimize the SVM parameters and finally, the obtained statistics were fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi.

Table 6.1 Sensitivity and Specificity comparison.

	Sensitivity	Specificity
PSO-SVM	94.1%	80.2%
Experts	90 %	59 %
Dermatologists	81 %	60 %
Trainees	85 %	36 %
General practitioners	62 %	63 %

The presented PSO-SVM system resulted in a sensitivity of 94.1 % a specificity of 80.2% and an accuracy of 87.1 %. The obtained sensitivity and specificity results are comparable to those obtained by Dermatologists and considerably better than those obtained by less trained doctors as seen in table 6.1 (quoted from [311]). Consequently, the proposed system can be considered as a promising method to be used by pathologists for skin cancer diagnostic. It can serve as a second opinion for the doctor in the skin cancer diagnosis process.

6.1.1. Experiment 1

This experiment investigates skin cancer segmentation using an active contour model, or snake model, driven by Gradient Vector Flow (GVF). A snake is a parameterized contour [312] that translates and deforms on the image plane according to the strength of the image's edges and the internal properties of the contour such as smoothness. This thesis implements a gradient vector flow (GVF) snake originally developed and discussed in [313]. The advantage of a GVF snake is its robustness to the initialization of the snake in [313], [314]. Although a GVF snake is robust to the initialization of the snake, it is still sensitive to image noise [315]. In order to make the snake insensitive to noise and to be able to remove the hairs, an Adaptive Filter (Wiener and Median filters) is proposed. After the noise and hairs are removed, the GVF snake will be used to segment the skin cancer region. The GVF snake extends the single direction and allows it to still be able to track the boundary of the skin cancer even if there are other objects near the skin cancer region.

The block diagram in figure 6.1 shows, stage one; the original input colour cancer image,

Stage two converts the colour image RGB into a greyscale image figure 6.2. In stage three, wiener and median filters are proposed; Wiener filtering is a general way of finding the best reconstruction of a noisy signal[227]. It gives the optimal way of narrowing off the

Figure 6.1 Block Diagram for skin cancer image segmentation

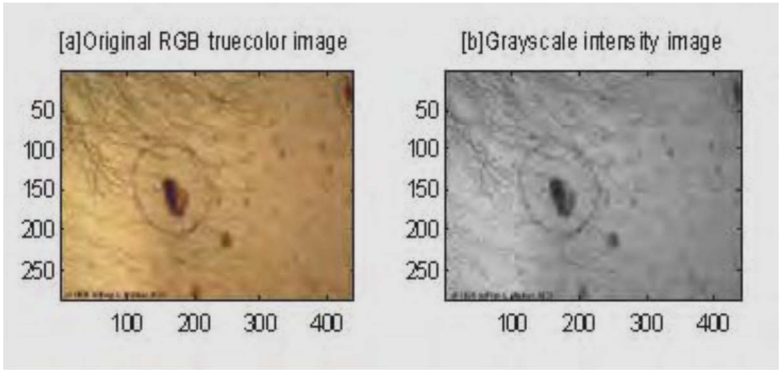


Figure. 6.2 Skin Cancer Image, (a) Original RGB true colour image

(b) Converted greyscale intensity image

noisy components, so as to give the best reconstruction of the original signal. Wiener2 uses a pixel wise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel. Equation 6.1 showed Matlab form[228]:

J = wiener2 (I, [m n], noise) 6.1

Where filters the image *I* using pixel wise adaptive Wiener filtering, using neighbourhoods of size $m - by - n$ to estimate the local image mean and standard deviation. The [m n] argument, m and n default values is 3. The additive noise (G Gaussian white noise) power is assumed to be noise. Wiener filtering is best when the basis most cleanly separates signal from noise, and that's clear when Wiener filter is applied to the difference between an image and a smoothed image.

As will The **Modified Adaptive Center Weighted Median** (MACWM) filter proposed with an adjustable central weight obtained by partitioning the observation vector space[233].

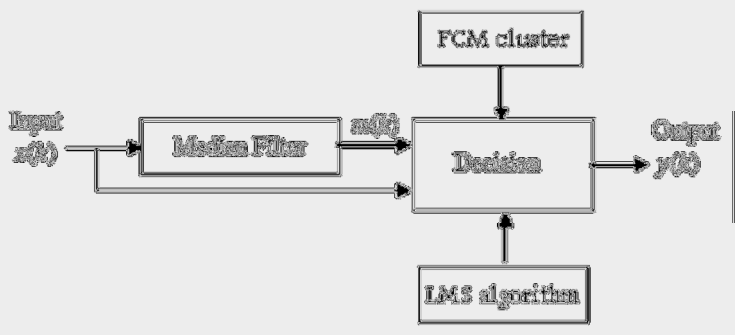


Figure 6.3. The basic structure of the MACWM filter

The framework of the MACWM filter is illustrated in Figure 6.3. It is composed of four parts: a median filter, a set of threshold by **Fuzzy C-means** (FCM), training the centre weight each block by **Least Mean Square** (LMS) algorithm, and a decision as to whether noises exist or not. At first, according to observation vector space (input image); median of input image will be calculated. We propose FCM algorithm to partition observation vector space to M block, and related weights to any blocks will be trained by using the LMS algorithm. The output value y (k) of the MACWM filter (Figure 4.6) at the processed pixel x (k) is obtained as follows in equation 6.2:

$$y(k) = \big(1 - w(k)\big)x(k) + y(k)m(k) \qquad 6.2$$

That the usual output value from a median filter is denoted as m (k) at location k in a filter window of size 2n + 1 as follows in equation 6.3:

$$m(k) = MED\{x_1(k) \dots \dots x_n(k)\} \qquad 6.3$$

Where MED means the median operation.

The MACWM filter achieves its effect through the linear combinations of the weighted output of the median filter and the related weighted input signal. Here, w (k) denotes the membership function indicating to what extent an impulsive noise is considered to be located at the pixel x (k). If w (k) = 1, an impulsive noise is considered to be located at pixel x (k), and the output value of the filter is equal to the output value of the median filter. If w (k) = 0, an impulsive noise is considered not to be located at pixel x (k), and the output value is the same as the input x (k); that is, the pixel x (k) is left unchanged. To judge whether an impulsive noise exists or not at the input pixel x (k), the membership function w (k) should take a continuous value from 0 to 1. Therefore, the major concern of the MACWM filter is how to decide the value of the membership function w (k) at the pixel x (k). The value of K set at 0.2 [316] in the experiment.

In general, the amplitudes of most impulses are more prominent than the fine changes of signals. Thus, the following two variables can be defined to generate the observation vector[235] . (Partitioning of observation vector space)[233]. In stage four, the filtered image is then treated by rough segmentation using a thresholding method developed in [298]. In stage five, Gradient Vector Flow (GVF) snake is proposed. There are two general types of active contour models in the literature today: parametric active contours[317] and geometric active contours[318] . In this paper, we focus on parametric active contours, which synthesize parametric curves within an image domain and allow them to move toward desired features, usually edges. Typically, the curves are drawn toward the edges by potential forces, which are defined to be the negative gradient of a potential function. Additional forces, such as pressure forces[319], together with the potential forces comprise the external forces. There are also internal forces designed to hold the curve together (elasticity forces) and to keep it from bending too much (bending forces).

In stage five, Gradient Vector Flow (GVF) snake is proposed, to obtain the final contour [320]. Particular advantages of the GVF snake over a traditional snake are its insensitivity to initialization and ability to move into concave boundary regions. A traditional snake is a curve $\mathrm{x(s)} = [\mathrm{x(s)}, \mathrm{y(s)}], \mathrm{s} \in [0, 1]$, that moves through the spatial domain of an image to minimize the energy functional, Equation 6.4:

$$E = \int_0^1 \frac{1}{2}\{\propto x^{-}x(s)^2 + \beta x^{=}x(s)^2\} + E_{ext}\{x(s)\}d_s \qquad 6.4$$

Where $\propto$ and β are weighting parameters that control the snake's tension and rigidity, respectively. $x^{-}(s)$ and $x^{=}(s)$ denote the first and second derivatives of $x\ (s)$ with respect to$s$. The external energy function E_{ext} is derived from the image so that it takes on its smaller values at the features of interest, such as boundaries.

To add additional flexibility to the snake model, it is possible to start from the force balance equation directly, as displayed in equation 6.5:

$$F_{int} + {F_{ext}}^2 = 0 \qquad 6.5$$

Balloon models[319] comprise an important example of this approach. In these models ${F_{ext}}^2$ is the sum of the traditional potential forces and pressure (or normal) forces, which act in a direction normal to the curve. This increases the capture range of an active contour, but also requires that the balloon be initialized to either shrink or grow.

In our first experiment, we computed the GVF field for the line drawing of Figure 6.5(a) using $\mu = 0.2$. Comparing the resulting field, shown in Figure 6.5(b) to the potential force field of Figure 6.4 (b), reveals several key differences. First, the GVF field has a much larger capture range. It is clear that in order to get this extent using traditional potential force fields, one would have to use a large σ in the Gaussian filter. But this would have the effect of blurring (or perhaps even obliterating) the edges, which is not happening in the GVF field. A second observation is that the GVF vectors are pointing somewhat downward into the top of the U-shape, which should cause an active contour to move farther into this concave region. Finally, it is clear that the GVF field behaves in an analogous fashion when viewed from the inside. That is, the vectors are pointing toward the boundary from as far away as possible and are pointing upward into the concave regions (the fingers of the U-shape) as viewed from the inside.

Figure 6.5 (c) shows the result of applying a GVF snake with parameters $\alpha = 0.6$ and $\beta = 0.0$ to the line drawing shown in Figure 5 (a) (using the external GVF field of Figure 5(b)). In this case, the snake was initialized farther away from the object than the initialization in Figure 6.4 (c), and yet it converges very well to the boundary of the U-shape. It should be noted that the blocky appearance of the U-shape results from the fact that the image is only 64 X 64 pixels. The snake itself moves through the continuum (using bilinear interpolation to derive external field forces which are not at grid points) to arrive at a sub-pixel interpolation of the boundary.

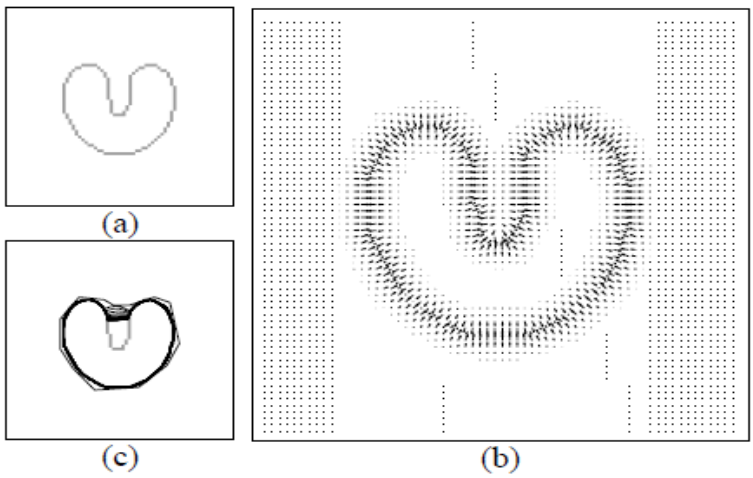


Figure 6.4. A snake with traditional potential forces cannot move into the concave boundary region.

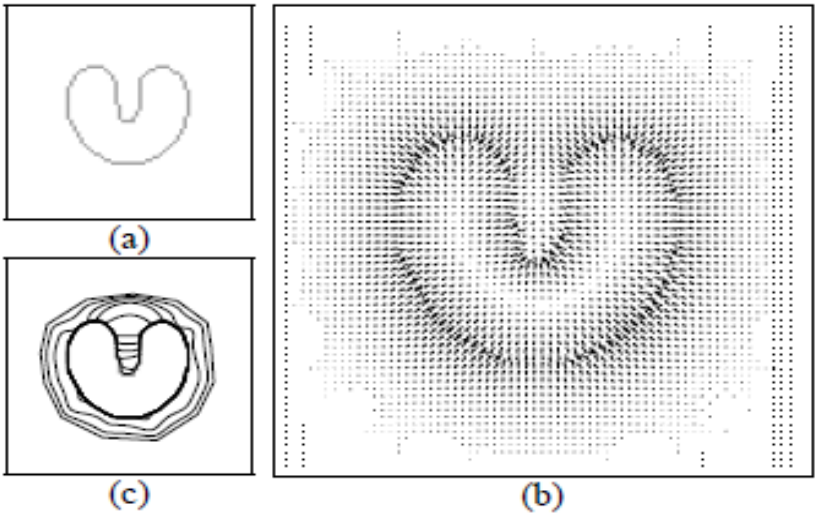


Figure 6.5 A snake with GVF external forces moves into the concave boundary region.

These experiments proposed two methods to find the border of the skin cancer. The first method is an independent search, which is used to find the seed border points. The second method is used to track the border based on the nearest border point [321]

Figure 6.6, displays the output of the skin cancer greyscale image, (a) filtered image after Wiener2 filter (b) filtered image after median filter.

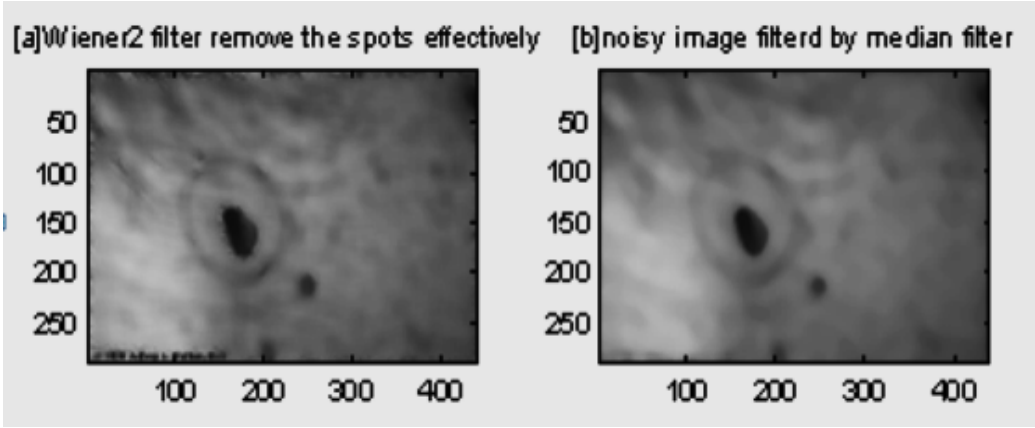
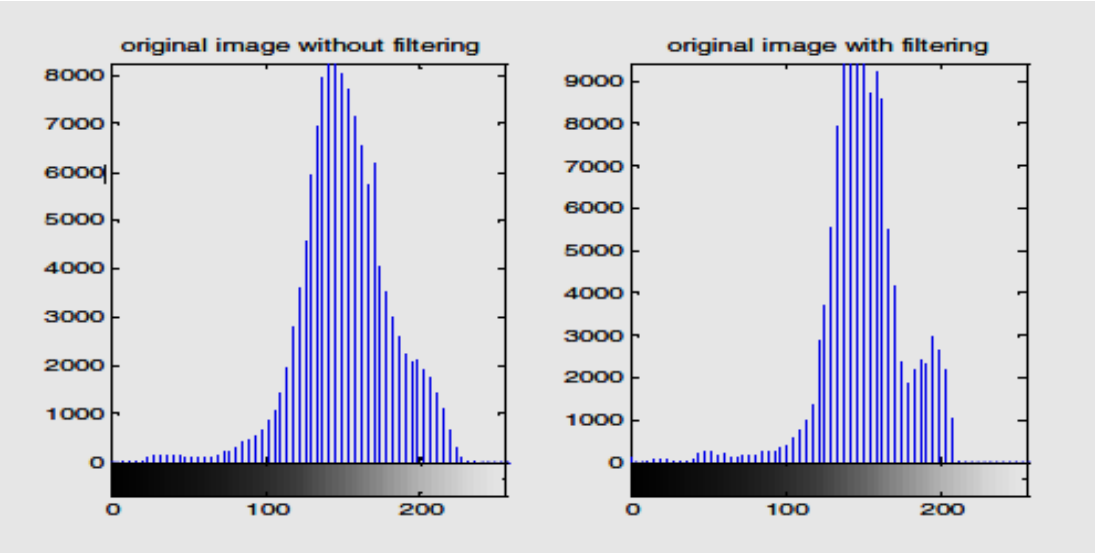


Figure. 6.6. Skin Cancer greyscale image showing: (a) Wiener2 filter has removed the spots effectively (b) noisy image filtered by the median filter

Today, health care institutions produce a large amount of image data which are used in diagnosis and treatment of patients. As the number of images produced increases, utilization, and handling of image data are becoming a difficult task for engineers, scientists, and medical physicists. Researchers say that there are two main concerns in the field of image processing and analysis applied to medical applications: *1. improving the*

quality of the image data 2. Featuring of information from medical image data in an efficient and accurate manner.

In order to demonstrate image quality, we have to start with a basic image processing overview as defined in Ref.[237]. In this experiment the process of digital images, described the brightness variation in the image using its histogram. By changing the histogram, processes that shift and scale the result (making the image brighter or dimmer, in different ways). Refer to Figure. 6.7 (a) and (b) to see image histogram stretching the pixel values onto (0, 255). Also thresholding techniques that turn an image from Gray level to binary were considered. The statistical operations can reduce noise in the image, which can benefit the feature extraction techniques to be considered later [322].



(a) Without filtering, (b) with filtering.

Figure.6.7 Skin Cancer Image Histogram,

This experiment was carried out on selected 8 (eight) skin cancer images; the images are distorted with Gaussian noise with mean = 0 and standard deviation = 0.01. The threshold value is selected to be 0.5 for all the images. The performance of the proposed operators for edge detection is evaluated and compared with gradient vector flow (GVF) snake; figure 6.8 illustrates a sample of these images).

In order to compare the segmentation performance of the proposed algorithm, the root-mean-square error [323]between the input image and the output image of same size MxN [rows and columns] and the root-mean-square of signal to noise ratio SNRrms used between the original input image and the restored output image. The root-mean-square

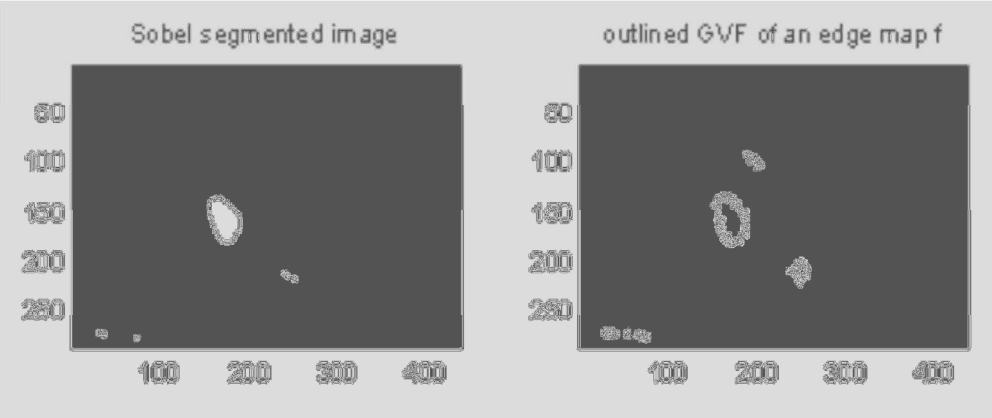
error e_{rms} [32]between the original input image $f_i\,(x, y)$ and the restored output image $f_0\,(x, y)$ of size MxN [rows and columns] is defined as in equation 6.6:

$$e_{rms} = \sqrt{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\frac{[f_0(x,y) - f_i(x,y)]^2}{MN}} \qquad 6.6$$

The root-mean-square of signal to noise ratio SNR_{rms}between the original input image f_i (x, y) and the restored output image $f_0\,(x, y)$ of size $M\ x\ N$ are defined as in equation 6.7:

$$SNR_{rms} = \sqrt{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\frac{f_0\,(x,y)^2}{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}[f_0(x,y) - f_i(x,y)]^2}} \qquad 6.7$$

The results displayed in table 6.2 for the proposed operators, the Sobel and the Prewitt operator showed the best performances (good contrast, edge map strength, and low noise content) and Robert and Laplacian have poor performance in terms of all indices].



(a) (b)

Figure. 6.8 Skin Cancer greyscale image showing: (a) Sobel segmented image (b) Gradient vector flow (GVF) segmented image

This experiment stated that the proposed operator displayed the extraction of a good quality edge map from the grey scale image, and presented a skin cancer image segmentation algorithm, which classified the most commonly used algorithms. Then these algorithms have been applied to 8 images. Table 6.2 showed a comparative way: Prewitt is more acceptable than the other operator, Laplacian is very much sensitive to noise and Canny operator shown better detection, good response and improving signal to noise ratio.

Table 6.2 Comparison between proposed operators (Sobel, Roberts, Prewitt, Laplacian, Canny and Otsu) and Gradient vector flow (GVF) segmented images.

Operators	Root mean square error	Signal to noise ratio
Sobel operator	2.3478	0.2081
Roberts operator	2.8630	0.2100
Prewitt operator	1.9687	0.3808
Laplacian operator	1.6930	0.4147
Canny operator	3.7365	0.3257
Otsu operator	1.4214	1.9052

The subjective evaluation of edge detected images shows that the proposed operators evaluated both the performance and root mean square of signal to noise ratio between the input image (operators) and the output image (GVF). The advantage of a GVF snake is its robustness to the initialization of the snake, although it is still sensitive to image noise.

6.1.2. Experiment 2

These experiments propose two segmentation methods to identify the normal skin cancer oriented texture and animation edges and then finding the accuracy by using the three layers back-propagation neural network classifier (BNN). The experiment results in table 3, displayed the best result with highest overall accuracy is 89.9%. The best BNN is three hidden layer with 40, 25 and 10 neurons for each hidden layer. The accuracy is increase with number of neuron in hidden layer. However, number of hidden layer cannot improve the result but it could reduce the probability of over-fitting [324]. (see section 6.1.2.1. Experiment Results)

This experiment used a **wavelet decomposition method** to find better feature extracting while using one of the more interesting methods, the three layer back-propagation neural network (BNN) to maintain the classification method. Wavelet decomposition allows useful cancer's feature to be extracted from the images without clinical knowledge. In this system, 2-D wavelet packet is used and the enhanced image in gray scaled as an input. Since there is lack of published articles to compare different wavelet, seven different

wavelets are used in this paper followed by same classification test to compare them and find out the best results. The 'bior5.5' wavelet found it provided the highest accuracy that achieved 88.5% as displayed in table 6.3. On the other hand, the daubechies wavelet ('db1' and 'db10') has the most stable experimental record during the testing[324].

Table 6.3 BNN Classification Test with Different Wavelet.

Wavelet	Training	Validation	Testing	Total
Db1	98.3%	48.9%	73.1%	85.6%
Db10	100%	48.1%	73.6%	86.9%
Bior1.3	100%	51.1%	68.4%	85.7%
Bior5.5	100%	53.6%	77.4%	88.5%
Coif3	100%	57.1%	69.1%	86.5%
Sym4	99%	44.4%	68.7%	84.5%
Sym7	73.7%	40.8%	53.5%	64.4%

The database images used 448 dermoscopy images collected from the Sydney Melanoma Diagnostic Centre in Royal Prince Alfred Hospital and internet websites. These images include digital malignant melanoma images and digital benign melanoma images. The images were separated into four groups to test the accuracy of each group. The database images are obtained from different sources and the size of the images is non-standard. The first step in this process is to resize the image, second apply preprocessing to remove the fine hairs, noise and air bubbles on the skin and facilitating image segmentation by using wiener and median filters. These basic operations are usually for pre-processing to improve display quality. These experiments used a powerful way to enhance the image contrast by stretching the histogram; these are a very popular technique known as histogram equalization, which alters pixel values to achieve a uniform distribution.

A post-processing method is used to enhance the shape of the image; a smart contrast enhancement technique used the histogram equalization (HE) algorithm to stretch the dynamic range of the image histogram and resulting in overall contrast improvement.

Figure 6.9 shows the result of segmentation by threshold with Gray-scale Histogram Equalization (GHE) (bottom right) and Local Histogram Equalisation (LHE). The LHE (bottom left) sharpens the lesion but also reduces the surrounding detail.

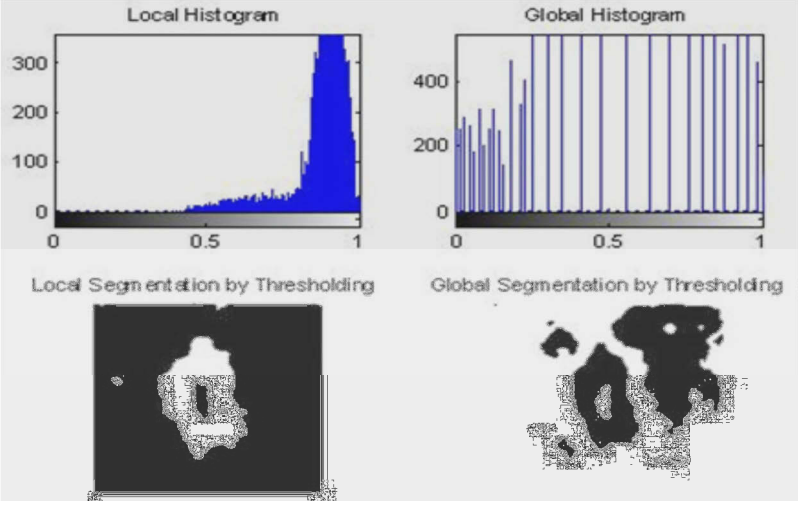


Figure 6.9. Result of segmentation by threshold with GHE (bottom right) and LHE (bottom left), Top: shows histogram Local and Global results

The segmentation methods used the region of interest (ROI) to extract the interesting features of melanoma within the border since most of the cancer cells are lump structures. The border structure provides vital information for accurate diagnosis. Many clinical features including asymmetry and border irregularity are calculated from the border. Threshold segmentation is an advanced method for segmentation to compute the intensity value from grey images. Figure.6.10 shows the result of threshold segmentation where the white color represents the region area of interest (ROI).

Figure 6.10.Segmentation by Thresholding (ROI)

This experiment used the segmentation to remove the healthy skin from the image and find the region of interest, usually the cancer cell that remains in the image. A Statistical region merging (SRM) algorithm is based on region growing and merging. The performance of SRM has a better segmentation result compared to other popular segmentation methods [221]. The theory of region merging starts at a seed point and compares its four neighbor points or pixels. The region is grown when the neighbors pixel share the same homogenous properties. SRM usually works with a statistical test to decide the merging of regions. Figure. 6.11 show the result of SRM segmentation where the white color represents the

area of interest, which is more accurate in segmentation. It includes more useful areas and does not put the outer region as part of the area of interest.

Figure 6.11. Segmentation by SRM

Compare two segmentation method, SRM has a better and more accuracy in segmentation. It includes more useful area and does not put the outer region as part of the area of interest. However, the accuracy for of neural network as classifier could not support that, since by using Backpropagation neural network (BNN), it found that the threshold has 0.5% higher accuracy than SRM in the classification test as illustrate in table 6.4.

Table 6.4. BNN Classification Test for SRM & Thresholding.

Method	Training	Validation	Testing	Total
Threshold	98.3%	48.9%	73.1%	85.6%
SRM	97%	46.6%	72.6%	84.7%

Feature extraction is the procedure that takes out the hidden properties or the raw data from the image that is further used in the classifier. These experiments used a modern method of feature extraction, wavelet, and curvelet transformation .These approaches, such as discrete wavelet and discrete curvelet, transform the frequency information at different scales and locations are obtains [325].

Wavelet decomposition, as explained in sections 5.4.6, can be a repeated process occurring over several levels. The original signals are decomposed into approximations feeding the next level of decomposition, creating a decomposition tree. The approach of this feature extraction is similar to the work in [324]; the image energy is distributed according to the resolution. Each of the nodes represents the represent one feature, which then can be used as an input for classification stage. The experiments in the above mentioned work stated that the second level decomposition can generate 16 nodes or features. Two-dimensional wavelet packet returns the coefficients in a 2 dimensions matrix. The features are

calculated by their mean, maximum, minimum, median, standard deviation and variance. Therefore, 144 features for each image (16 nodes x 9 features) for 448 images are produced to be the input for the classification system. This would increase the computation time without any improvement to the performance [324]. Therefore principal component analysis (PCA) is used before the training process, the PCA factor chosen = 0.002, to reduce the 144 features to 60 features. Figure 6.12. Shows the original image and its transform (wavelet coefficients). This method allows useful cancer features to be extracted from the images without clinical knowledge. In these experiments, 2-D wavelet packet is used to enhance the image in grey scaled as an input.

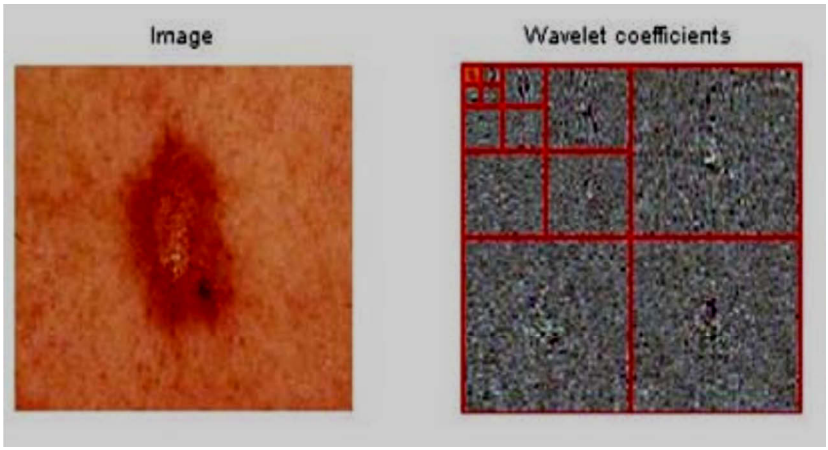


Figure 6.12. Display of the image and its transform (wavelet coefficients)

As explained in [304], Curvelet are waveforms which are highly anisotropic at fine scales, with effective support obeying the parabolic principle (length width2). Just as for wavelets, there is both a continuous and a discrete curvelet transform as shown in figure.6.13

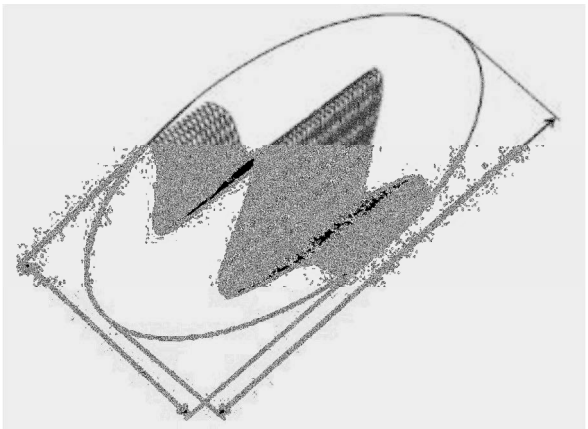


Figure. 6.13: A single curvelet with width 2^{-1} and length $2^{-\frac{1}{2}}$.

curvelet have useful geometric features that set them apart from wavelets and combine the frequency responses of the curvelet at different scales and orientations, a rectangular frequency tiling that covers the whole image in the spectral domain as shown in Figure 6.14. Thus, the curvelet spectra completely cover the frequency plane and there is no loss of spectral information as there is with the Gabor filters.

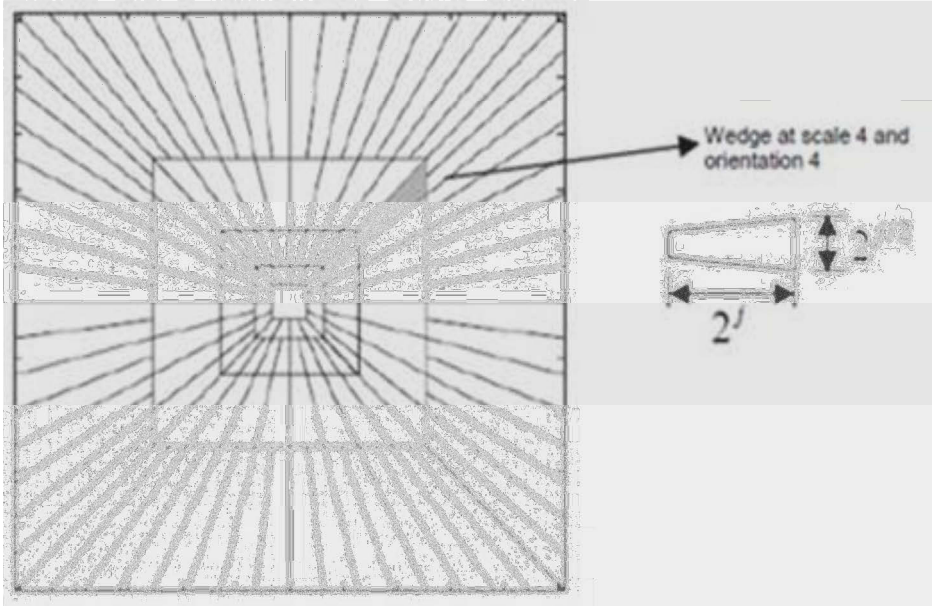


Figure 6.14. Rectangular frequency (basic digital) tiling of an image with 5 level curvelet.

To achieve higher level of efficiency, the curvelet transform is usually implemented in the frequency domain. The product is then inverse Fourier transformed to obtain the curvelet coefficients. The process can be described as:

Curvelet transform = IFFT [FFT (Curvelet) x FFT (Image)], and the product from the multiplication is a wedge.

These experiments calculate the mean and standard deviation for the first half of the total cells at each scale. The test on melanoma image using the curvelet method showing the original, noisy and the output of curvelets is displayed in Figure 6.15.

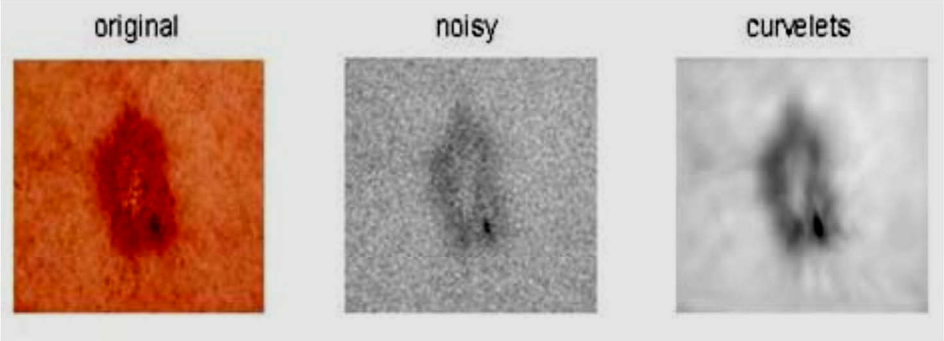


Figure 6.15. Curvelet demising image that contains oriented texture and cartoon edges (a) Original (b) Noisy, and (c) Curvelet.

Classification system used to identify if the image is the image if it cancer or not. The classification system is usually supported by an intelligent classifier, such as a neural network or support vector machine.

6.1.2.1. Experiment Results

This experiment used 60% of the total 448 images for training, 20% for testing and 20% for validation. The best result with highest overall accuracy is 70.44% for curvelet with Back-Propagation Neural Network (BNN) with three layers [40 25 10 neurons] as shown in table 6.6. The accuracy for neural network as a classifier is increased with the number of neurons in the hidden layer. However, the number of hidden layers could reduce the probability of over-fitting; it was found that the curvelet method has 14.5% higher accuracy than wavelet in the classification test as illustrated in table 6.5.

Table 6.6. Displaying the comparison between wavelet and curvelet features results**.**

Method	No of Neurons	No of Layers	Training	Validation	Testing
Wavelet	40 25 10	3	76.29	55.67	55.94
Curvelet	40 25 10	3	79.76	68.06	70.44

6.1.3. Experiment 3

This experiment proposes an automated non-invasive system for skin cancer (melanoma) detection based on Support Vector Machine classification. The proposed system uses a number of features extracted from the Wavelet or the Curvelet decomposition of the grayscale skin lesion images from the original color images. The dataset used include both digital images and Dermoscopy images for skin lesions that are either benign or malignant. The recognition accuracy obtained by the Support Vector Machine classifier used in this experiment is 87.7% for the Wavelet based features and 83.6 % for the Curvelet based

ones. The proposed system also resulted in a sensitivity of 86.4 % for the case of Wavelet and 76.9% for the case of Curvelet. It also resulted in a specificity of 88.1% for the case of Wavelet and 85.4% for the case of Curvelet. The obtained sensitivity and specificity results are comparable to those obtained by dermatologists.

6.1.3.1. Experiment Evaluation:

The aim was to evaluate the ability of the proposed system to classify digital images of skin lesions as benign or malignant. The dataset used in the evaluation was a collection of images obtained from the Sydney Melanoma Diagnostic Centre in Royal Prince Alfred Hospital and internet websites. In total, the dataset included 448 images divided into four groups as follows: Group A contains 93 digital images for malignant skin lesions obtained from the internet; group B contains 142 dermoscopy images for malignant skin lesions obtained from the Sydney Melanoma Diagnostic Centre; group C and D contain 121 and 92 digital images for benign skin lesions obtained from the internet.

Four experiments were conducted; the data groups for each experiment are varied for the training and testing or the feature extraction method. The classification technique used for all of the experiments was the Support Vector Machines (SVM) with Radial Basis Function Kernel using LIBSVM [326]. I trialed different kernels during my experiments, in each case, 10-fold cross validation was used to pick the best RBF kernel parameters, namely, C and γ. because it demonstrated the best results. For each experiment, 0.8 of the data was used for training and 0.2 was used for the purpose of testing. The experiments conducted are described below. The data used for the computation of the specificity and sensitivity is based on dataset. In the experiment, used 60% of the total images for training, 20% for testing and 20% for validation.

Test 1 included dataset of 448 images both digital skin images and dermoscopy images and was used for the evaluation purposes and Wavelets ‘db4’were used for the feature extraction step. Three evaluations were conducted based on the number of features extracted from the second level wavelet decomposition which resulted in 16 two-dimensional matrices. The first used 9 features per matrix (calculated as a combination of the mean, maximum, minimum, median, standard deviation and variance of each two dimensional matrix). The second used the same 9 features together with L1 and L2 norms. The third used the previously mentioned 11 features and the variances of the R, the G and the B layers of the image to include the color effect. The results of ***test 1*** are given in table

6.7. It can be seen from this table that the cross validation accuracy increases with the number of features and that including the colour features improved the cross validation accuracy by 6.7%.

Test 2, the dermoscopy data was excluded. Only digital skin images (305 images) were used for the evaluation purposes. Similar to test 1, Wavelets 'db4'were used for the feature extraction step. The wavelet function Daubechies 'db4' was used because it has the most stable experimental record during the testing.Three evaluations were conducted based on the number of features extracted from the second level wavelet decomposition which resulted in 16 two-dimensional matrices. The first used 9 features per matrix (calculated as a combination of the mean, maximum, minimum, median, standard deviation and variance of each two dimensional matrix). The second used the same 9 features together with L1 and L2 norms. The third used the previously mentioned 11 features and the variances of the R, the G and the B layers of the image to include the colour effect. The results of ***test 2*** are shown in table 6.4. It can be seen from table 6.7 that the cross validation accuracy increased with the number of features. It can be also noticed that the specificity, sensitivity and the testing accuracy increased dramatically with the inclusion of the colour features. The testing accuracy resulting when including the colour features is 87.7 % with around 11% increase compared with the same case in table 6.8 with the dermoscopy images included in the dataset. The trend of improved testing accuracy with color features observed in test 2 was not noticed in test 1 as the image set used included two types on images with different characteristics namely the digital skin images and the dermoscopy images.

Test 3 used the 448 images which include both digital skin images and dermoscopy images for the evaluation. Curvelet were used for the feature extraction step. Two evaluations were conducted based on the number of features obtained from the Curvelet decomposition of the image. The first used 52 features (calculated using the mean and standard deviation). The second used the same 52 features together with the variances of the R, the G and the B layers of the image to include the colour effect. The results of test 3 are given in table 6.9. This table showed that the cross validation accuracy increases with the number of features and that including the colour features improved the cross validation accuracy by 4.4%.

Test 4, the dermoscopy data was excluded. Only digital images (305 images) were used for the evaluation purposes. Similar to test 3, curve lets were used for the feature extraction

step. Two evaluations were conducted based on the number of features used. The first used 52 features (calculated from the mean and standard deviation of the Curvelet coefficients). The second used the same 52 features together with the variances of the R, the G and the B layers of the image to include the colour effect. The results of test 4 are given in table 6.10. This table showed that the cross validation accuracy increased with the number of features. It can be also being noticed that the specificity, sensitivity and the testing accuracy increased with the inclusion of the colour features. The resulting testing accuracy with the inclusion of with the inclusion of the colour features is 83.6 % with around 17.6 % increase compared with the same case in table 3 with the dermoscopy images included in the dataset. The dramatic improvement in the specificity, sensitivity and the testing accuracy for test 4 compared to test 3 can be justified by understanding that the dataset used for test 4 was more homogenous as it included only digital images for the skin lesions as opposed to test3 where the dataset included two types on images with different characteristics namely, the normal digital images and the dermoscopy images. The inclusion of two types on images with different characteristics reduced the ability of the SVM to generalize as is clear from the fact that the cross validation accuracy results were considerably higher than the testing accuracy for test 3.

In general, it can be said that the inclusion of the colour features improved the testing accuracy and that the Wavelet extracted features gave better performance (87.7 %) compared with those found by Curvelet extracted features (83.6 %).

Table 6.7. svm classification for all the dataset (including dermoscopy images) with features extracted from wavelets

	Specificity	Sensitivity	Testing Accuracy	10-fold Cross Validation Accuracy	RBF Kernel Parameters
Wavelet 9 Features	70.8	78.0	74.2	72.4	C=2 $\gamma = 1.10\ x\ 10^{-2}$
Wavelet 11 Features	77.1	85.4	80.9	74.6	C=4.757 $\gamma = 3.05\ x\ 10^{-5}$
Wavelet and Colour Features	72.9	80.5	76.4	81.3	C=861.1 $\gamma = 1.75\ x\ 10^{-4}$

Table 6.8. svm classification excluding dermoscopy images with features extracted from wavelets

	Specificity	Sensitivity	Testing Accuracy	10-fold Cross Validation Accuracy	RBF Kernel Parameters
Wavelet 9 Features	76.6	64.7	74.1	78.6	C=362.0 $\gamma = 8.63 \times 10^{-5}$
Wavelet 11 Features	81.0	65.2	76.5	84.8	C=4.757 $\gamma = 5.52 \times 10^{-3}$
Wavelet and Color Features	88.1	86.4	87.7	85.3	C=1024 $\gamma = 8.63 \times 10^{-5}$

Table 6.9. svm classification for all the dataset (including dermoscopy images) with features extracted from curvelets

	Specificity	Sensitivity	Testing Accuracy	10-fold Cross Validation Accuracy	RBF Kernel Parameters
Curvelet Features	**68.6**	**63.0**	**65.2**	**76.9**	**C=13.45** $\gamma = 6.25 \times 10^{-2}$
Curvelet and Color Features	**76.9**	**61.9**	**66.3**	**81.3**	**C=11.31** $\gamma = 2.21 \times 10^{-2}$

Table 6.10. svm classification excluding dermoscopy images with features extracted from curvelets

	Specificity	Sensitivity	Testing Accuracy	10-fold Cross Validation Accuracy	RBF Kernel Parameters
Curvelet Features	83.3	69.2	80.3	76.6	C=2 $\gamma = 1.25 \times 10^{-1}$
Curvelet and Color Features	85.4	76.9	83.6	78.6	C=861.1 $\gamma = 1.95 \times 10^{-3}$

The presented system resulted in a sensitivity of 86.4 % for the case of wavelet and 76.9% for the case of curvelet. It also resulted in a specificity of 88.1% for the case of Wavelet and 85.4% for the case of curvelet. The obtained sensitivity and specificity result are comparable to those obtained by dermatologists as seen in table 6.11 from [310]. Consequently, the proposed system can be considered as a promising non-invasive method to be used for skin cancer diagnosis that does not require well trained practitioners as for the case of Dermoscopy images. It can serve as a second opinion for the doctor in the skin cancer diagnosis process.

Table 6.11. sensitivity and specificity statistics [310]

	Sensitivity	**Specificity**
Experts	90 %	59 %
Dermatologists	81 %	60 %
Trainees	85 %	36 %
General Practitioners	62 %	63 %

The next part based on pathological images is; melanoma is a pathology clinical lesion is irregularly shaped, an asymmetrical lesion with varying colour with a history of recent change in size, shape, colour or sensation. Melanoma may arise within an existing benign lesion. Most melanomas have an initial radial growth phase within the epidermis and sometimes within the papillary dermis, which may be followed by a vertical growth phase with deeper extension. The experiments 1, 2 and 3 concludes that there are some possible factors to improve the accuracy of detecting malignant melanoma by having a higher number of images for training of the SVM network.

6.1.4. Experiment 4

The complexity and the parameter adjustment can be a serious problem in several industrial computer vision applications; it had been proposed a spatially adaptive median filter (SAMF) would be more efficient for image processing [327]. This filter (SAMF) applied widely as an advanced method compared with standard median filtering is used [328]. Color-based segmentation is one of the most widely investigated research areas in pathological image analysis. Fundamentally, there are several objects and each object has several features and this algorithm has been applied to classify the objects based on the features [110].

One of the methods had been used in this thesis to find the features is by generalizing the k-mean clustering into n features, and to define the centroid as a vector where each component is the average value of that component. Each component represents one features [329] .

The proposed system used a number of features extracted from histo-pathological images of skin lesions through image processing techniques which consisted of a spatially adaptive color median filter for filtering and a k-means clustering for segmentation. The extracted features were reduced by using sequential feature selection and then classified by using

support vector machine (SVM) to diagnose skin biopsies from patients as either malignant melanoma or benign nevi.

The results on figure 6.16 show, original image, image after adaptive median filter, gray scale image, histogram out, adjust image intensity, and output image using histogram equalization.

This experiment is a new application based on histo-pathological images of skin lesions that required finding out new features and the correlation of the reduced numbers of features and getting better accuracy. It was required to modify many of the mentioned techniques to make them work for such application. The higher accuracy of diagnosis proposed work is calculated and displayed.

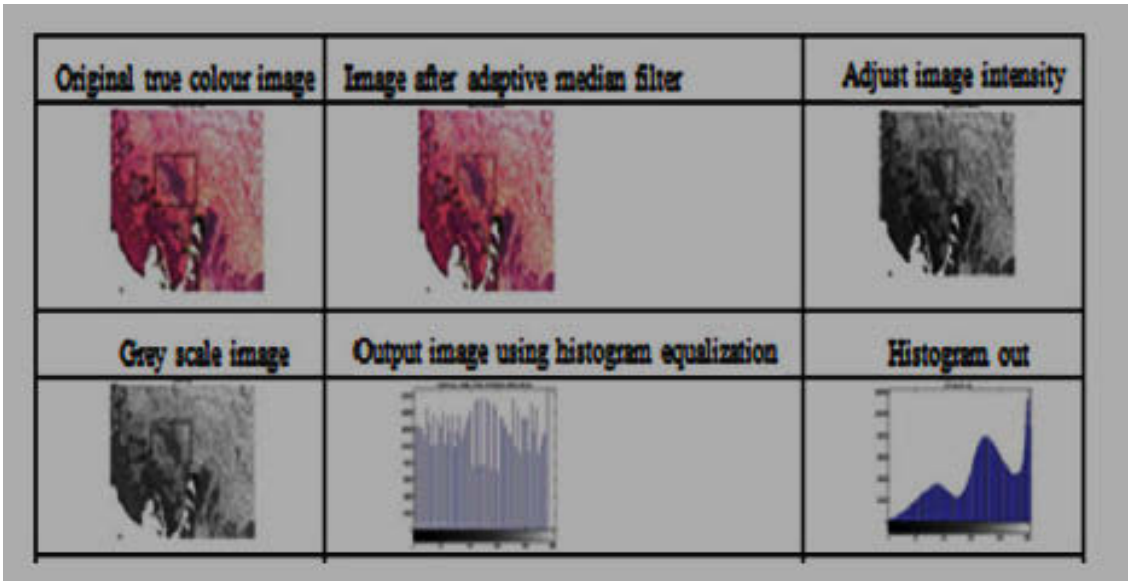


Figure. 6.16. The results showed, original image, image after median filter, gray scale image, histogram out; adjust image intensity, and output image using histogram equalization

6.1.5.1. Results and discussion-Experiment 4

In this experiment, 42 images were sampled from microscopic slides of skin biopsy and have been used in the current work.

The algorithm used in these experiments was a 5-fold cross validation test. The experiment used LIBSVM [310] library for support vector machines, with linear, (RBF) Radial Basis Function kernel [310], and Polynomial kernel, under each magnification. The melanoma

images that are confirmed by the pathologist are treated as negative images, while the nevus ones are treated as positive images. Researchers, in [330] classifying breast tissue found 67.1% accuracy in labeling nuclei as benign or malignant; the classification of histopathology images presented in [310] resulted in 88.7% accuracy in the diagnosis of lung cancer, 94.9% accuracy in the typing of malignant mesothelioma, and 80.0% to 82.9% accuracy in the diagnosis of malignant mesothelioma for Fulgent stained lung sections. Keenan et al. [310] reported accuracies of 62.3% - 76.5% in the grading of H&E stained cervical tissue. Skin cancer is a fast developing disease of modem society, reaching 20% [310] increasing of diagnosed cases every year. Dermatoscopic methods used for diagnostics are non-invasive and require a great deal of experience to make correct diagnosis. As reported in [310], only experts arrive at 90% sensitivity and 59% specificity in skin lesion diagnosis; while for less trained doctors, these figures show a significant drop to till around 62%- 63% for general practioners, as mentioned on Table 6.3.

In this experiment, the images are sampled and split into a training set and a test set using MATLAB software. The testing process is done for 42 images (28 benign and & 14 melanoma). These 42 images were are fed to the proposed SVM network In the experiment, used 60% of the total images for training, 20% for testing and 20% for validation. The performance of the algorithm was evaluated by computing the percentages of Sensitivity (SE), Specificity (SP) and Accuracy (AC) [331]. From our experimental results CPU time faster than experiment 3 [22], from our experimental results CPU time = 9 min., which is 17 times faster than the old experiment[19] ,(CPU time = 156 min) and produce more accurate classification results, than physicians do.

Table 6.13 shows the average resulting SE, SP and AC of 5 fold validation of the proposed networks by using LIBSVM. This experiment used RBF kernel with gamma=0.l25 and C=1. The results showed this to be one of the best results comparing with expertise / physicians results displayed on table 6.12. The obtained accuracy of the system is 88.9 % whereas the sensitivity and specificity were found to be equal to 87.5 % and 100% respectively. By comparing the results on tables 6.12 and 6.13, it is concluded that the proposed system gives faster and more accurate classification of skin tumors than physicians do. This experiment concludes that there were some possible factors to improve the accuracy of detecting malignant melanoma by having a higher number of images for training of the SVM network.

Table 6.12. Described sensitivity and specificity

	Sensitivity	Specificity
Experts	90%	59%
Dermatologists	81%	60%
Trainees	85%	36%
General practitioners	62%	63%

Table 6.13. show result after training or the (svm) network. using (sfs) technique

No of images	Sensitivity	Specificity	Accuracy
42	87.5	100	88.9

6.1.5. Experiment 5

6.1.5.1 Evolution of Melanoma Diagnosis

Early detection of malignant melanoma remains the key factor in reducing mortality from skin cancer. Recognizing the importance of this issue in the 1990s, dermoscopy enabled the recognition of new subsurface features to differentiate between malignant and benign pigmented lesions. During the last decade, new computer-based technologies have improved diagnostic sensitivity and specificity and may result in optimizing lesion selection for biopsy and pathology review. Even with all of the advances in melanoma diagnosis, timely recognition, detection, and rapid treatment of melanoma remain critical. Although pathologic examination remains the gold standard for diagnosis, this cancer has the potential to be diagnosed through non-invasive approaches because of its affecting the skin location. From the development of the ABCDs through current efforts that use complex computer algorithms and genetic markers, a clinician's ability to detect melanoma in its earliest form has been improved. However, a "good clinical eye" is still fundamental to selecting the lesions for evaluation among the sea of those that are widespread. As current approaches are refined and new techniques are developed, the improved ability to diagnose this cancer will hopefully enhance reaching the goal of reducing melanoma mortality[191].This experiment used a novel methodology for automatic feature extraction from histo-pathological images and subsequent classification. These techniques added Gabor filter bank to winner and adaptive median filters to improve diagnostic accuracy. Then Histogram equalization was used to enhance the contrast of the images. The extracted features were reduced by using sequential feature selection then; finally the obtained

statistics was fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi.

This experiment used forty two (42) images sampled from microscopic slides of skin biopsy. MATLAB R2013b software was used; the features were carried out to generate training and testing of the proposed SVM. , this 42 images (28 benign images and 14 melanoma images) were sampled and split into a training set (mixed 28 images) and test set mixed 14 images to be used as input to the proposed SVM network. The performance of the algorithm was evaluated by computing the percentages of Sensitivity (SE), Specificity (SP) and Accuracy (AC). Table 2 shows the average resulting SE, SP and AC of 2 fold cross validation of the proposed networks by using L1BSVM. In this study used RBF kernel with gamma=0.125 and C=10. We did many trials to improve our experimental results; these experiments present a new application based on histo-pathological images of skin lesions that required finding out new features and the correlation of the reduced numbers of features and getting better accuracy. Measuring the many of techniques to make them work for such an application. The results which are displayed on table 6.14 in which the SFS technique was used technique produces one of the good and acceptable results compared with dermoscopy and pathologic diagnosis results table 6.16 and also with the results displayed on table 6.15, which shows the results without using (SFS) technique.

Table 6.14. the result after training of the (svm) network with using (sfs) technique.

No of images	Sensitivity	Specificity	Accuracy
42	76	100	81

Table 6.15. the result after training of the (svm) network without using (sfs) technique.

No of images	Sensitivity	Specificity	Accuracy
42	55.6	80.6	68.1

Table 6.16 Comparison of different algorithms for dermoscopic and pathologic diagnosis

DERMOSCOPY DIAGNOSTIC ALGORITHM	SENSITIVITY	SPECIFICITY	DIAGNOSTIC ACCURACY
Pattern analysis[332]	85%	79%	71%
ABCD [332]	84%	75%	76%
7-Point checklist[332]	78%	65%	58%
CASH[333]	98%	68%	—
Menzies[334]	85%	85%	81%
SolarScan[335]	91%	68%	—
SVM using sfs technique[23]	87.5%	100%	88.9%

Several different algorithms for dermoscopy diagnosis are used to differentiate melanoma from benign melanocytic lesions: pattern analysis; ABCD rule of dermoscopy; 7-point checklist; colour, architecture, symmetry, and homogeneity (CASH); and Menzies method (Table 6.16)[332] [336] [333]. Based on the results of the Consensus Net Meeting on Dermoscopy, sensitivity varied between 82.6% and 85.7%, and specificity varied between 70% and 83.4% for distinguishing melanomas from benign lesions[337]. Although pattern analysis had the highest sensitivity, specificity, and diagnostic accuracy for melanoma detection when performed by dermatologists, the Menzies method proved better when used by nondermatology physicians[334]. Clinicians may be discouraged from using dermoscopy because they feel it is too time consuming.

SolarScan is an automated instrument (Polartechnics, Sydney, Australia), SolarScan was found to give a sensitivity of 91% and specificity of 68% for melanoma, which was comparable to that of expert clinicians[338]

The pathologic results obtained in this experiment are comparable to those obtained by different algorithms for dermoscopic as seen in table 6.16 consequently; the proposed system can be considered as a promising method to be improved by increasing the number of images to be used later by pathologists for skin cancer diagnostic.

6.1.6. Experiment 6

This experiment proposed an automated system uses a wavelet packet transform (WPT) to extract more features. This thesis introduces a novel melanoma detection strategy using a hybrid particle swarm based support vector machine (SSVM) algorithm. The extracted features are reduced by using a particle swarm optimization (PSO were used to optimize the SVM parameters and finally, the obtained statistics were fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi.

The support vector machine (SVM) classifier is used widely in bioinformatics, due to its high accuracy, ability to deal with high-dimensional data and in this syntax diverse sources of data [310] [339] [340]. It is possible that the data found from the RBF mapping is more likely to be correctly classified by SVM than from the other kernel function mappings. Furthermore, SVM-linear could be a special case of SVM-RBF [339] [341], which means

that SVM-RBF has more possibilities to obtain a better performance than SVM-linear. The better performance of the SVM based system which employs RBF is confirmed in another application [310] [342].

This thesis used in total 283 features (six features extracted from histogram, twenty one features extracted from co-occurrence matrix for each channel and 255 features extracted from Wavelet Packet Transform (WPT) see chapter 5). WPT was used to implement the feature extraction process. Variability appears to be what most separates malignant melanoma from benign nevi, therefore the best approach at feature extraction would be to retain as much of the data variability as possible [308]. Wavelet analysis looks at these changes over different scales which should detect whole lesion changes such as texture, color, and local changes like granularity. The WPT is a generalized version of the wavelet transform: the high-frequency part also splits into a low and a high frequency part; this produces a decomposition tree as shown in the Figure. 6.17. The WPT provides a high dimensional feature vector thus providing more information about the images. However, the WPT complicates the analysis process as the high dimensionality of the feature vector causes an increase in the learning parameters of the pattern classifier, and the convergence of the learning error deteriorates. Consequently, dimensionality reduction played an important role before applying the feature vector to the pattern classifier. The extracted features were reduced by using a particle swarm optimization (PSO) [343]. This was used to optimize the SVM parameters as a feature selection and finally, the obtained statistics were fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi. The obtained classification accuracies show better performance in comparison to similar approaches for feature extraction.

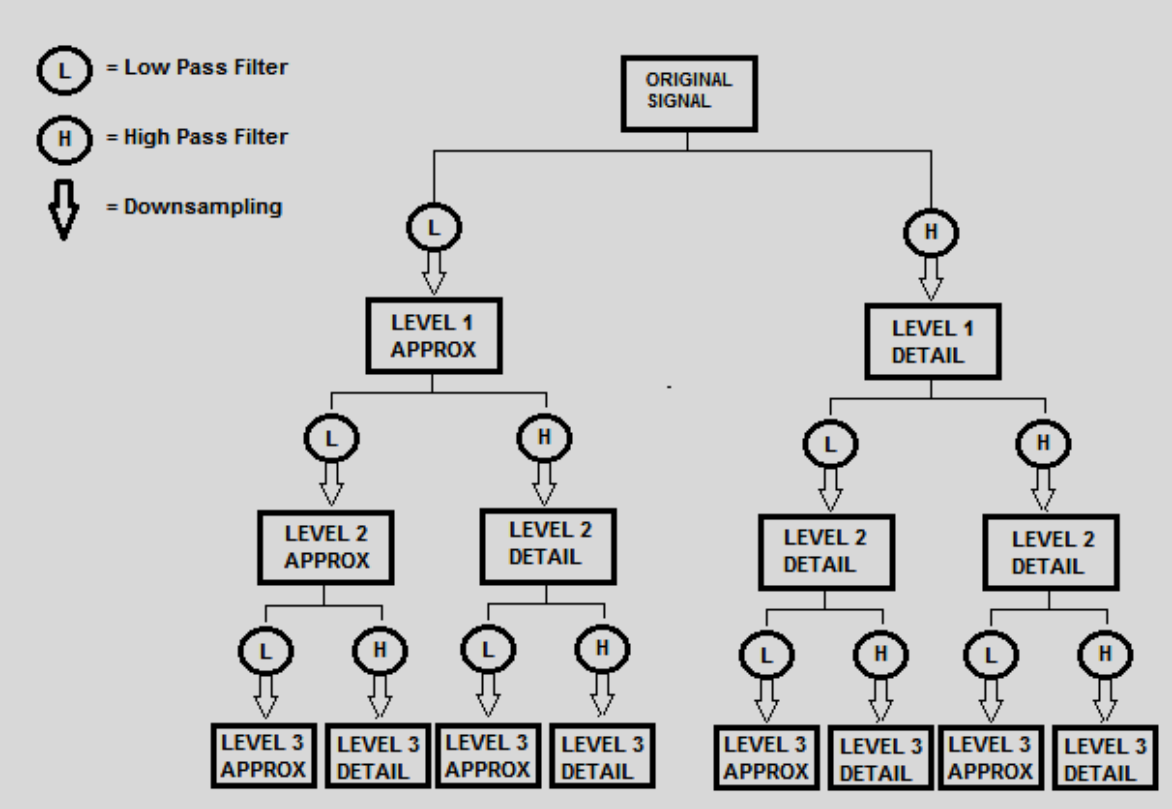


Figure. 6.17. - A Wavelet Packet decomposition tree

6.1.6.1. Results and Discussion for Experiment 6

The proposed image processing approach for computer-aided melanoma detection has been presented. Thesis employs the SVM algorithm and the input of pathology images. SVM has been developed with an output of +1, which means that a melanomas event is happening, and -1, which means that there is no melanoma (benign).

One of the main parameters of SVM is C that constrains the vector margin as explained previously in section 5.6. The incensement of C does not always provide improvement of the performances of the SVM algorithms in the melanoma detection. The reason could be that theoretically an increasing C value allows more points to cross the supporting hyperplanes of SVM. This could reduce the generalization performance of the detection algorithm. The increase in C value from 10^4 to 10^6 decreases the testing sensitivities in terms of geometric means. On the other hand, this increase in C raises the training performances in all terms (sensitivities, specificities and geometric means). The increase of the training performances (while the testing performance decreases) means that the generalization of the algorithm becomes worse.

The other thing is the modification in the weight factor. The modification of w_0 or w_1 is similar to modification of C in the associated class. Modification of w_0 means modification of C for the data points in the nonmelanomic (benign) class, and modification of w_1 means modification of C for the data points in the melanomic class. The optimal SVM parameters have been found in the optimization stages. The optimal $w_1$is higher than w_0. In the optimization, the weight factor is automatically tuned to prevent a low sensitivity. In other words, the automatic selection of the weight factor performs well.

Our aim is to evaluate the ability of our proposed system to classify Histo-pathological images of skin lesions as benign or malignant. The dataset used in our evaluation is obtained from the Southern Pathology Laboratory in Wollongong NSW, Australia. It includes 79 Histo-pathological images (29 benign images and 50 melanoma images).

Two experiments were conducted. They differ from each other in the feature selection method employed and the way the optimal SVM parameters are picked (C and γ). The classification technique used for all of the experiments is SVM with Radial Basis Function Kernel using LIBSVM [344]The algorithms were implemented using MATLAB R2013b. For each experiment, 0.8 of the data was used for training and 0.2 was used for the purpose of testing.

The melanoma images that are confirmed by the pathologist are treated as negative images, while the nevus ones are treated as positive images. The performance of the algorithms is evaluated by computing the percentage Sensitivity (SE), Specificity (SP) and Accuracy (AC) using equations 6.8, 6.9, 6.10, respectively:

$$SE = \frac{TP}{(TP+FN)} \text{ X } 100 \tag{6.8}$$

$$SP = \frac{TN}{(TN+FP)} \text{X} 100 \tag{6.9}$$

$$AC = \frac{(TP+FN)}{(TN+TP+FN+FP)} \tag{6.10}$$

Where TP is the number of true positives, TN is the number of true negatives, FN is the number of false negatives, and FP is the number of false positives.

In the first experiment, SFS was used to choose the most relevant features that result in the best performance of the SVM. 5-fold cross validation was used to pick the best RBF kernel parameters (C=10 and γ=0.125). The results of experiment 1 are shown in table 6.17.

PSO-SVM arrangement was used in the second experiment to choose the most relevant features and the optimal RBF kernel parameters that optimize the performance of the SVM. 5-fold cross validation was used within the PSO-SVM arrangement. The resulting RBF kernel parameters are C=3525.0051 and γ=0.0084732). The results of experiment 2 are shown in table 6.18.

Comparing the results in table 6.17 with those in table 6.18, it can be seen the performance of the PSO-SVM arrangement for feature selection and SVM parameter optimization is considerably better than that obtained when using SFS for feature selection. An improvement of 10 % can be noticed in the values of the Sensitivity, the Specificity and the Accuracy. This highlights the success of the PSO-SVM arrangement.

The presented PSO-SVM system resulted in a sensitivity of 94.1 % a specificity of 80.2% and an accuracy of 87.1 %. The obtained sensitivity and specificity results are comparable to those obtained by Dermatologists and considerably better than those obtained by less trained doctors as seen in table 6.19 (quoted from [345])Consequently, the proposed system can be considered as a promising method to be used by pathologists for skin cancer diagnostic. It can serve as a second opinion for the doctor in the skin cancer diagnosis process.

Table 6.17 The Result after Training of the (SVM) Network, with using (SFS)Technique.

No of images	Sensitivity	Specificity	Accuracy
79	83.6	70.7	77.4

Table 6.18 The Result after Training of the (SVM) Network. with using (SVM+WLG+PSO)

No of images	Sensitivity	Specificity	Accuracy
79	94.1	80.2	87.1

Table 6.19 Described Sensitivity and Specificity [31]

	Sensitivity	Specificity
Experts	90 %	59 %
Dermatologists	81 %	60 %
Trainees	85 %	36 %
General practitioners	62 %	63 %

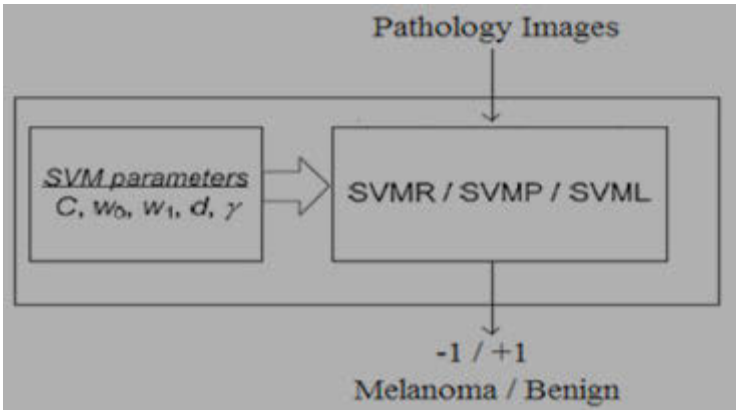


Figure. 6.18 shows the melanoma detection employing SVM (SVMR, SVMP, and SVML); the input is pathology images and the output is melanoma (-1) or benign (+1).

$$Geometric\ mean = \sqrt{Sensitivity\ xSpecificity} \qquad 6.3$$

where:

- TP (true positive) is the number of inputs which correspond to melanoma classified as melanoma.
- FP (false positive) is the number of inputs which correspond to benign classified as melanoma
- TN (true negative) is the number of inputs which correspond to benign classified as benign.
- FN (false negative) is the number of inputs that correspond to melanoma classified as benign.

The geometric mean (gm) is suitable for indicating the performance of a case with imbalanced data [346]. The imbalanced data means that the data point number of a class is significantly more than that of another class, as is the data in this research.

The performance of the detection algorithms was measured in terms of sensitivity, specificity and geometric mean, instead of a root mean square error or RMSE [347]. Using sensitivity the correct detection of melanomic class data can be known, and using specificity the correct detection of nonmelanomic (benign) class data can be seen. On the other hand, RMSE only provides one value for the correct detection of both classes, instead of one value for each class.

The final results in experiment 6 displayed melanoma detections using the swarm based SVM algorithms. $SSVMs$ with different kernel functions have been compared. Furthermore, the detection using $SSVM$ has also been compared with the swarm based multiple regression and the SVM algorithms without the swarm optimization.

The performance of $SSVMR$ ($SSVM\ with\ RBF$ kernel function) is better compared to the other three $SSVMs$, which employ polynomial, sigmoid and linear kernel functions. It is possible that the data found from the RBF mapping is more likely to be correctly classified by SVM than from the other kernel functions mappings. Furthermore, $SVM - linear$ could be a special case of $SVM - RBF$ [296], which means that $SVM - RBF$ has more possibilities to obtain a better performance than $SVM - linear$. The better performance of the SVM based system which employs RBF is confirmed in another application [263].

The proposed swarm optimization $SSVM$ could be suitable to prevent a low sensitivity in melanoma detection. A high sensitivity of a disease detector means it has a high true positive. The high true positive in melanoma detection means that the melanomas events can be detected properly. (more detailed information in Appendix H: Abbreviation)

In the fitness function for the optimization section 5.8.3 and (Eq. 5.53), κ was set to 0.58. This κ value was determined through several experiments. The important thing of ***κ value is that the coefficient for the sensitivity is set higher than the coefficient for the specificity.*** By this setting a low sensitivity could be avoided. This setting is necessary because the data number of the melanomas class is far lower than the nonmelanomic (benign) class.

This difference has the risk of low sensitivity. However, if κ is too high, the sensitivity could be very high and the specificity could be very low. Therefore, ***the κ value should be more than 0.50, but not too high***[348]. (refer to section 5.8.3, fitness function for the optimization).

The computational time required by the optimization using PSO depends on the iteration number provided and the time consumed in every iteration. Every iteration creates n SVM training to make n SVM models; $\boldsymbol{n}$ is the number of particles of $\boldsymbol{PSO}$. In this thesis, the ***iteration number is 200 and the particle number is 50***.

Hence, after 200 iterations the optimization terminates. Another termination principle is the ***rises*** of the ***global best*** during $\boldsymbol{m\ iteration}$; if the rises are less than a defined value, the optimization terminates. It means that if the optimal values do not change significantly during ***m iteration***, the optimization terminates. (More detailed information in applications E).

The proposed swarm based SVM outperforms the swarm based multiple regression (SMR) as showen in experiment 6. The optimization methods of these two algorithms are the same, in which their parameters are optimized using thePSO. One difference of the two algorithms in finding the decision function is that SVM considers maximizing the margin between the two nearest points of the two classes in SVM, whereas SMR does not. The maximization of the margin could result in better generalization.

Chapter Summary

This chapter 'Experimental Results and Discussion' is organized and structured in describing the experiments based on digital and pathological images; experiment 1 investigates skin cancer segmentation using an active contour model, or snake model, driven by Gradient Vector Flow (GVF); experiment 2 proposed two segmentation methods to identify the normal skin cancer oriented texture and animation edges followed by finding the accuracy by using the three layers back-propagation neural network classifier (BNN); experiment 3 proposed an automated non-invasive system for skin cancer (melanoma) detection based on support vector machine classification. The proposed system uses a number of features extracted from the wavelet and the curvelet of decomposed greyscale skin lesion images from the original color images; experiment 4 proposed a system used a number of features extracted from histo-pathological images of skin lesions through image processing techniques which consisted of a spatially adaptive color median filter for filtering and a k-means clustering for segmentation. The extracted

features were reduced by using sequential feature selection and then classified by using support vector machine (SVM) to diagnose skin biopsies from patients as either malignant melanoma or benign nevi; experiment 5 proposed a novel methodology for automatic feature extraction from histo-pathological images and subsequent classification. These techniques added Gabor filter bank to winner and adaptive median filters to improve diagnostic accuracy. Then Histogram equalization was used to enhance the contrast of the images. The extracted features were reduced by using sequential feature selection then. Finally the obtained statistics was fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi. Experiment 6 proposed an automated system using a wavelet packet transform (WPT) to extract more features. This thesis introduces a novel melanoma detection strategy using a hybrid particle swarm based support vector machine (SSVM) algorithm. The extracted features are reduced by using a Particle Swarm Optimization (PSO) were used to optimize the SVM parameters and finally, the obtained statistics were fed to a support vector machine (SVM) binary classifier to diagnose skin biopsies from patients as either malignant melanoma or benign nevi.

CHAPTER 7

Conclusion and Future Work

7.1. Conclusion

Skin cancer is increasing and affects many people in different countries in the world including Australia. Malignant melanoma is the deadliest type of skin cancer that can be treated successfully if it is detected early. Early intervention will lead to a better survival rate. Since the clinical observation of a melanoma is subject to human error, early detection can be enhanced by utilising an automated process. In this thesis, an automatic skin cancer detection system had been developed to assist doctors in the accurate detection of skin cancers while reducing the computational time. The database images were enhanced through an image processing procedure. This study intended to show that it was possible, with some common utilization hardware – a simple digital camera – and some well-structured software, to reach results that would increase the possibilities of accurate detection of malignant skin lesions and, consequently, increase the possibilities of recovery from skin cancer, thus, also increasing life expectancy of the people affected by this disease.

Dermatology centers are very specialized work environments that are not always available in small clinical institutions. Up to now, all the available systems rely on rather sophisticated and expensive equipment, not directed or even accessible to the general public. Regardless of the cost associated with the available systems, most of them are nothing but simple image storage solutions. Those few that implement classification methods, reached results that are much less than satisfactory.

These considerations made it obvious for the author that, along the research process, various difficulties would arise. This became a fact, namely in were image capturing conditions were concerned. These varied environmental conditions create several problems to the feature extraction system, particularly in the detection of edges, which is, in fact, the basis for the classification process.

The aim of this thesis is to develop an automated process for the rapid and accurate detection of a benign (safe) or dangerous (melanoma) within a smaller perimeter of error than that which is achievable by clinical observation. A high performance computer aided diagnostic system is developed to assist physicians and avoid misdiagnosis.

The common approaches to skin lesion early detection include different stages of pre-processing, segmentation, feature extraction and selection, and classification. Since the output of each stage is the input of the next stage, all stages have an important role to avoid misdiagnosis.

Image enhancement is one of the most important issues in low-level image processing. Its purpose is to improve the quality of low contrast images. Therefore, the underlying principle of the enhancement is to enlarge the intensity difference between objects and surroundings, so that the image is not apparently distorted.

The histogram which has been used in the thesis provides much evidence to represent the characteristic of the images. The proposed multi-peak Generalized Histogram Equalization (GHE) method described in described in chapter 4.3.6 is very effective not only in enhancing the entire image but also in enhancing the image texture. It makes the change of the order of grey levels of the original image completely controllable.

Image segmentation is a critical and essential component and is one of the most difficult tasks in image processing, pattern recognition, and determines the quality of the final analysis. Segmentation subdivides an image into its principal regions or objects. Segmentation it an essential issue in digital image processing which is used for image description and classification.

Hundreds of features might be derived from an image. However not all of the features are suitable for mass classification. Too many irrelevant features will not only make the classifier complicated, but will also reduce the accuracy of the classification. The most important issue is to select features that are able to represent the characteristics of masses in the skin cancer images, and based on these features, the malignant mass can be significantly discriminated from the benign masses by the classifier. The lump will be separated from normal skin and useful information will be extracted from the image. Then, it will pass through classifiers for training and testing to define whether it is malignant melanoma or benign.

Support Vector Machines are a suitable approach to data modelling. They combine generalisation control with a technique to address the problem of dimensionality.

The detection of melanoma using a swarm-based SVM (SSVM) algorithm has been conducted. In SSVM, a particle swarm optimization (PSO) was used to optimize the SVM parameters. A fitness function was defined in the optimization to find a high performance of the melanoma detection, especially in the sensitivity.

Different kernel functions were investigated to find the suitable kernel function for the SSVM which incomes a good performance in the melanoma detection. The SSVM which used the RBF kernel function performs well with 94.1 %, 80.2 % and 87.1 % in terms of sensitivity, specificity and geometric means. The thesis outcome results are very significant they are one of the best results in this type of research as seen in table 6.18 describing sensitivity and specificity in chapter 6, section 6.1.6.1, pg. 179. results and discussion for experiment 6. By Comparing the results in table 6.17 with those in table 6.18, it can be seen the performance of the PSO-SVM arrangement for feature selection and SVM parameter optimization is considerably better than that obtained when using SFS for feature selection. An improvement of 10 % can be noticed in the values of the Sensitivity, the Specificity and the Accuracy. This highlights the success of the PSO-SVM arrangement

7.2. Future Research

Future research aims to develop hybrid artificial intelligence techniques employing the following genetic algorithms; hybrid of fuzzy inference system (FIS), Ant Colony Optimization (ACO), called PSO-ACO, Ontology, Contourlet Transform and others in multifunction pattern recognition systems.

Looking forward to conduct further experiments and to test this algorithm on a large dataset having different pathologies and also with different skin colours.

Future work will propose the implementation of this approach and at the end, with strategic selection of different fusions tested within this report, we can better give global diagnostics for any given pathology.

Chapter Summary

This chapter conclusion and future research has presented the contributions to melanoma detection technology. The models of melanoma detection have been developed and examined. In addition, the hybrid algorithms with the base of SVM were introduced and tested for melanoma detection. The main algorithms are support vector machine (SVM), and a swarm optimization technique (PSO).

The detection of melanoma using a swarm-based SVM (SSVM) algorithm has been conducted. In the SSVM, all the possible combinations of the SVM parameters are tested for the input. The SSVM performs better than SVM. In the SSVM, and SVM are combined and the SVM parameters are optimized using a swarm technique. In terms of computational time, the two algorithms differ in the off-line Optimization stage. The computational time can be controlled so that the time could still be acceptable. In the experiment, the SSVM performs better than the SVM algorithm, with the sensitivity, specificity and geometric mean of 94.1 %, 80.2 % and 87.1 %, respectively.

we plane to conduct further experiments and to test this algorithm on a large dataset with selection of different fusions tested within this report, to present better diagnostics for any given pathology.

Appendices

A. *Appendix A: Glossary of Pathology Terminology*

- Actin: a protein found especially in microfilaments (as those comprising myofibrils) and it is active in muscular contraction, cellular movement, and maintenance of cell shape.
- Adenocarcinoma: a malignant tumour originating in glandular epithelium.
- Fine-needle aspiration cytology (FNAC): a diagnostic procedure used to investigate superficial (just under the skin) lumps or masses. In this technique, a thin, hollow needle is inserted into the mass for sampling of cells. These cells are examined under a microscope after being stained. Figure A1 & A2

Figure A1, showed a physician's hands are seen performing a needle biopsy to determine nature of lump either fluid-filled cyst or solid tumour

[This image was released by the National Cancer Institute, an agency part of the National Institutes of Health, with the ID 1973 (image)]

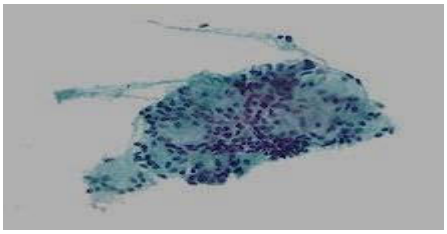


Figure A2, the diagnostics showed: Micrograph of a needle aspiration biopsy specimen of a salivary gland showing adenoid cystic carcinoma. (Source: Pap stain. MeSH D044963).

- Benign: not causing death / disease without cancer/ not cancerous. Figure A3

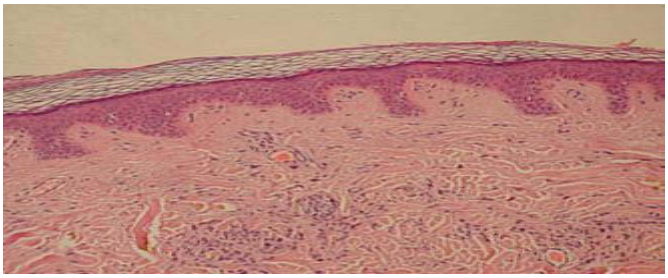


Figure A3, showed a Normal Epidermis and Dermis with Intradermal Nevus 10x-cropped. (Source, Kilbad).

- Brightfeld microscopy: This is the simplest of a range of techniques used for illumination of samples in light microscopes. It is a simple technique and that made it a quite popular. The typical appearance of a bright-field microscopy image is a dark sample on a bright background. Figure A4.

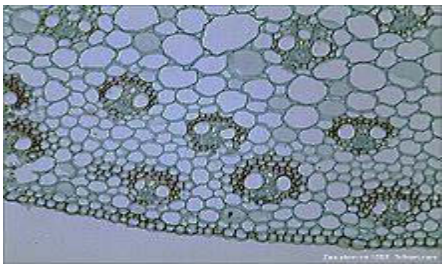


Figure A4, this image shows a crossection of the vascular tissue in a plant stem. Showed as an example for bright field micrograph. (Source: Wikipedia).

- Carcinoma: is a type of cancer.
- Cellularity: is the state of a tissue or other mass with regard to the number of constituent cells.
- Confocal microscopy: is an optical imaging technique used to increase optical resolution and contrast of a micrograph by using point illumination and a spatial pinhole to eliminate out-of-focus light in specimens that are thicker than the focal plane[349]. Figure A5.

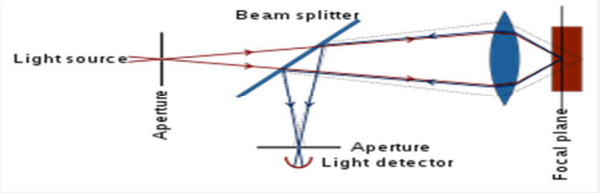


Figure A5, Principle of confocal microscopy

- Cytology: the study of cells as described in e Medicine Dictionary. It is branch of life science, which deals with the study of cells in terms of structure, function and chemistry.
- Cytopathology: the study of cellular disease and the use of cellular changes for the diagnosis of disease.
- Cell biology: the study of (normal) cellular anatomy, function and chemistry.
- Densitometry: is the quantitative measurement of optical density in light-sensitive materials, such as photographic paper or photographic film, due to exposure to light.
- Ductal carcinoma: is the most common type of breast cancer in women. It comes in two forms: invasive ductal carcinoma (IDC), ductal carcinoma in situ (DCIS), (lactiferous ducts), where breast cancer most often originates.
- Dysplasia: is an ambiguous term used in pathology to refer to an abnormality of development or an epithelial anomaly of growth and differentiation, as explained at Dorland's Medical Dictionary.
- The endothelium: This is the thin layer of cells that lines the interior surface of blood vessels and lymphatic vessels as explained at Dorland's Medical Dictionary.
- Fourier transform spectroscopy: is a measurement technique whereby spectra are collected based on measurements of the coherence of a radiative source, using time-domain or space-domain measurements of the electromagnetic radiation or other type of radiation[350].
- Glycogen: is made and stored primarily in the cells of the liver and the muscles. It functions as the secondary long-term energy storage (with the primary energy stores being fats held in adipose tissue)[351].

- Histology: It is commonly performed by examining cells and tissues by sectioning and staining which is followed by examination under a light microscope or electron microscope.
- Histopathology: in clinical medicine, histopathology refers to the examination of a biopsy or surgical specimen by a pathologist, after the specimen has been processed and histological sections have been placed onto glass slides.
- In situ: is a Latin phrase that translates literally to 'in position', as described in Collins Latin Dictionary & Grammar, it is used in many different contexts.
- in vivo: (Latin) are those in which the effects of various biological entities are tested on whole, living organisms usually animals including humans, and plants as opposed to a partial or dead organism.
- Karyometry: the measurement of the nucleus of a cell.
- Malignant: Many cancer treatments work by coaxing malignant cells to commit suicide, a process known to biologists as apoptosis.
- Metastasis: is the spread of a cancer from one organ or part to another non-adjacent organ or part.
- Microarray: A microarray is a multiplex lab-on-a-chip. It is a 2D array on a solid substrate (usually a glass slide or silicon thin-film cell) that assays large amounts of biological material using high-throughput screening miniaturized, multiplexed and parallel processing and detection methods.
- Nucleolus: is a structure found in the nucleus of cells. It forms around specific chromosomal regions in the nucleus of eukaryotic cells, and is made up of proteins and ribonucleic acids. Figure A6.

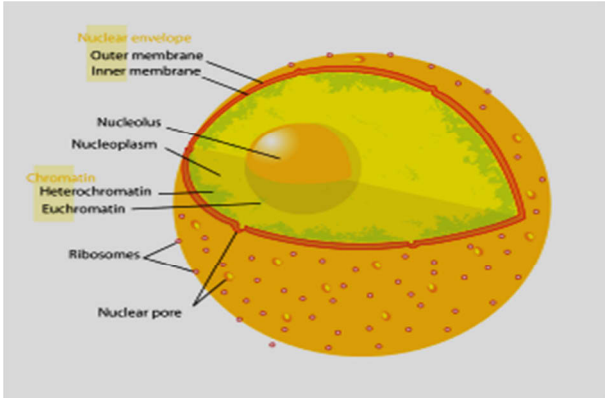


Figure A6 showed the nucleolus is contained within the cell nucleus. (Source: Wikimedia Foundation, Inc.)

- Optical Density (OD): is a convenient tool to describe the transmission of light through a highly blocking optical filter (when the transmission is extremely small). OD is defined as the negative of the logarithm (base 10) of the transmission, where the transmission varies between 0 and 1. Figure A7

$$OD = -\log_{10}(T) \text{ or } T = 10 - OD$$

- Pathology: is a significant component of the causal study of disease and a major field in modern medical practice and diagnosis. It is a study of disease in general, "general pathology," it includes a number of distinct but inter-related medical specialties which diagnose disease mostly through the analysis of tissue, cell, and body fluid samples.

Figure A8.

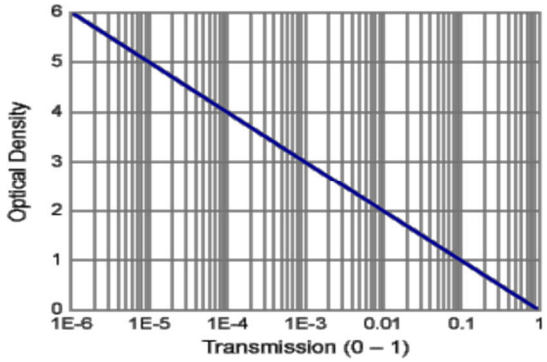


Figure A7 displayed the negative of the logarithm (base 10) of the optical Density (OD). (Source: semrock@idexcorp.com)

Figure A8, showing a pathologist examines a tissue section for evidence of cancerous cells while a surgeon observes. (Source: wikipedia.org)

- Premalignancy: is a generalized state associated with a significantly increased risk of cancer. If it is left untreated, these conditions may lead to cancer.
- Stroma: the connective, functionally supportive framework of a biological cell, tissue, or organ.

- Transmission electron microscopy (TEM): is a microscopy technique in which a beam of electrons is transmitted through an ultra-thin specimen, An image is formed from the interaction of the electrons transmitted through the specimen which is magnified and focused onto an imaging device, such as a fluorescent screen, on a layer of photographic film, or to be detected by a sensor such as a CCD camera. Figure A9.

Optical density (OD) provides a linear relationship between image intensity and staining density which can be used for transmission microscopy of stained tissue. In order to get this, firstly, the intensity of light transmitted through a specimen is calculated using Lambert- Beer's law for the atmosphere is usually written in equation B1[352]:

$$I = I_o * 10^{-A*c*l} \quad \text{B.1}$$

Where I is the intensity of the observed light, I_o the intensity of incident light, A is the amount of stain, c is the absorption factor of the stain, and l is the distance travelled through the sample [353] [354, 355]. After this, OD reduces this relationship to a linear one equation B2,[353] [354] [355]:

$$OD = -\frac{1}{l}\log_{10}(\frac{I}{I_o}) = A * c \quad \text{B.2}$$

OD Is calculated on a channel-by-channel basis. There are many different applications of OD as well other than stained medical imagery. In general, the equations are generally specified in a similar fashion whenever the absorption of light by a material is of importance.

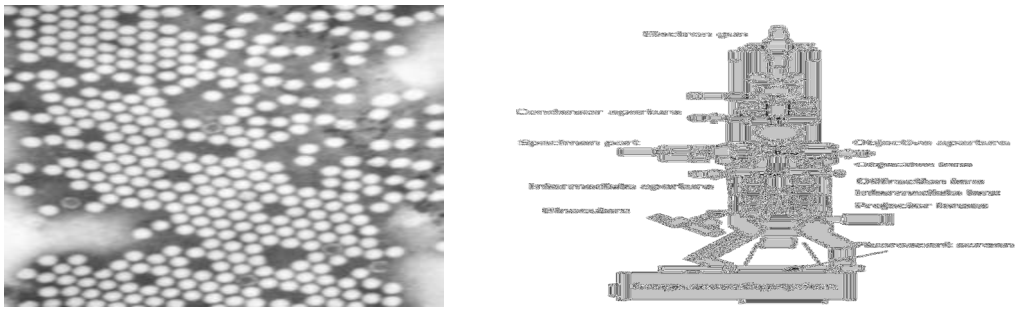


Figure A9, showed: a) A TEM image of the polio virus is 30 nm in size, b) Layout of Optical components in a basic TEM. (Source: Wikimedia)Appendix B: Optical Density of Transmission Microscopy Images

For instance, RGB image and the corresponding OD image can be seen in Figure B1 (a). The incident light intensity $I_o$for each channel is estimated by the brightest pixel. It is

noteworthy that, in the OD image, brighter pixels correspond (linearly) to larger amounts of stain. There are two stains present in the image presented in Figure B1, i.e. hematoxylin and eosin. Each stain has a different absorption factor c, left over from the original formulation of Beer- Lambert's law. One cannot directly separate the individual contributions of each stain without images of singly-stained tissue, but will instead get an estimate of the total stain present.

However, the analysis performed in[354], presents relative contributions of hematoxylin and eosin to the $R, G,$ and B channels as [0.1 0.2 0.08] and [0.01 0.13 0.01], respectively.

It can be observed that these relative contributions are qualitatively expressed in the optical density image of Figure B1. The individual channels of Figure B1 (b) are shown in B1(c)-(e) as reference. The hematoxylin (which stains nuclei) is represented in both the red and green channels, eosin (which stains cytoplasm and fibres) is represented mostly in the green channel, and the blue channel has very little response for either of the stain; this corroborates with the quantitative responses presented in[353].

B.*Appendix B. Glossary of Pathology Terminology*

While the optical density corresponding to a 3-channel RGB image (Figure B 1) is being presented here, the same computation can also be performed for each channel of a multispectral image stack. This operation is widely used in the histopathology analysis.

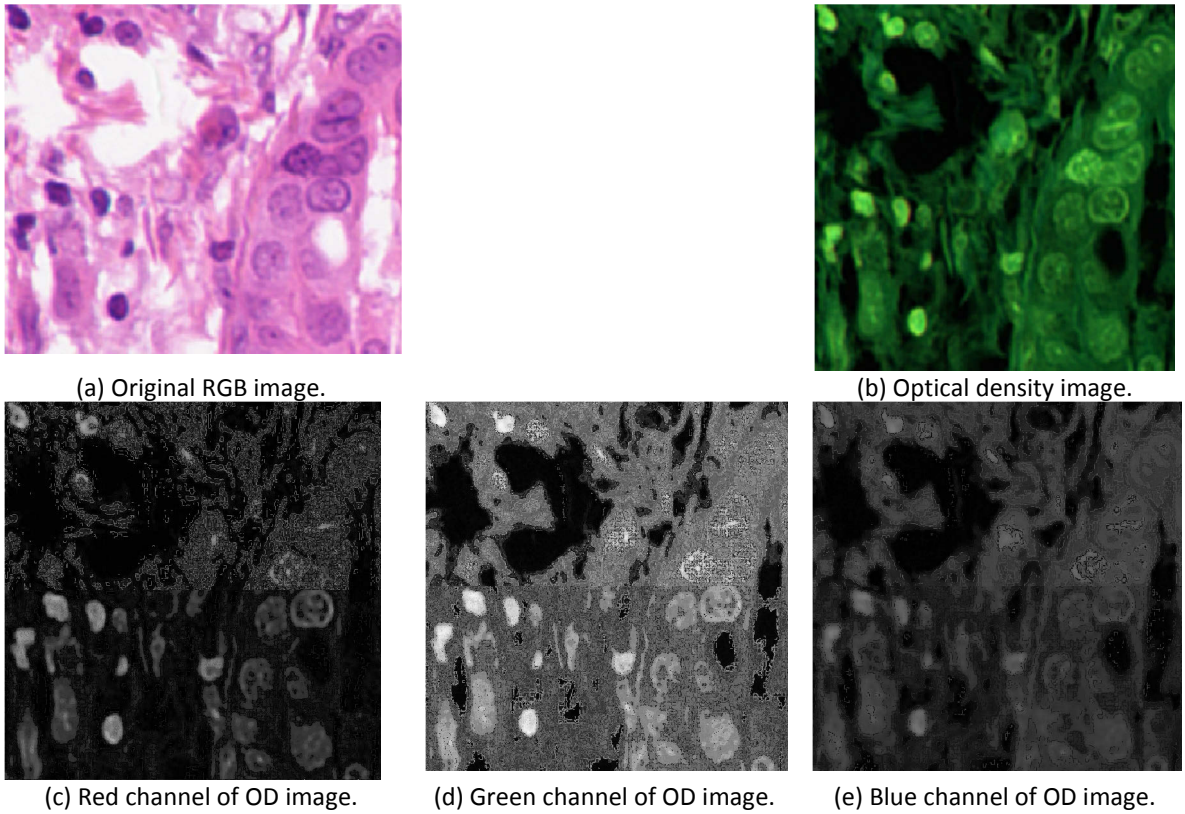


Figure B1. Example optical density image.

C. *Appendix C: Glossary of Machine Learning and Computer Vision Terminology*

- **Accuracy:** In the fields of science, engineering, industry, and statistics, the accuracy of a measurement system is the degree of closeness of measurements of a quantity to the actual (true) value of that quantity[356]
- **Classification:** Classification can be referred to categorization. It is the process in which ideas and objects are recognized, differentiated, and understood.
- **In graph theory:** a connected component or component: is a sub graph (undirected graph) in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in the subparagraph, as displayed in figure C1.

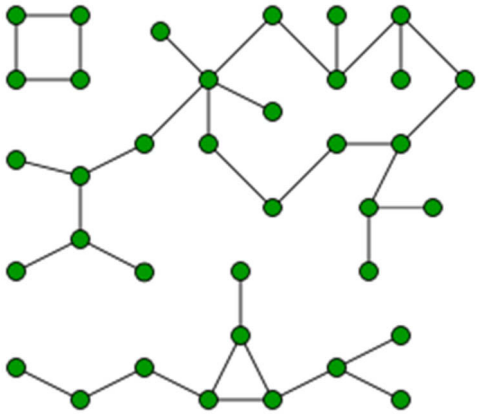


Figure C1, showed a graph with three connected components.

- **Cycle**: is a simplifiedsimplicial chain with 0 boundaries.
- **Densitometry**: is the quantitative measurement of optical density due to exposure to light in light-sensitive materials, such as photographic paper or photographic film.
- **Detection rate:** is considered the most important characteristic of a Satellite-AIS system.
- **Eigen decomposition:** In the mathematical discipline of linear algebra, Eigen-decomposition is the factorization of a matrix into a canonical form, whereby the matrix is represented in terms of its eigenvalues & eigenvectors. Only diagonalizable matrices can be factorized in this way [357].

- **Entropy:** Its originates from machine engineering, where in a process involving heat it is a measure of the portion of heat becoming unavailable for doing work, e.g. reversal of the process just mentioned. Entropy is also frequently used with emphasis of order or disorder properties of dissimilar areas.
- **Error:** It is a measure of the estimated difference between the observed or calculated value of a quantity and its respective true value.
- **False alarm rate:** In statistics, performing multiple comparisons, the term false positive ratio, also known as the false alarm ratio, usually refers to the probability of falsely rejecting the null hypothesis for a particular test.
- **Feature selection:** In machine learning figure C2 and statistics, feature selection, known as variable selection, attribute selection or variable subset selection, is the process of selecting a subset of relevant features for use in the model construction. A feature selection technique is based on the central assumption that the data contains many redundant or irrelevant features.

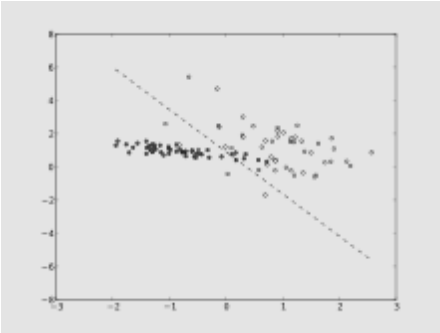


Figure C2, displayed the machine learning and data mining which used to solve problem in the areas of Classification, Clustering, Regression, Anomaly detection, Association rules, Reinforcement learning, Structured prediction, Feature learning, Online learning, Semi-supervised learning, and Grammar induction.

- **Fractal dimension:** A fractal dimension is a ratio providing a statistical index of complexity comparing how much detail in a pattern (fractal pattern) changes with the scale at which it is measured. It has also been characterized as a measure of the space-filling capacity of a pattern that tells how a fractal scales differently from the space it is embedded in; a fractal dimension does not have to be an integer[358] [359] [360].
- **Gabor filter:** In image processing, a Gabor filter, named after Dennis Gabor, is a linear filter that is used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system.

They have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave.

- **Gabor wavelet:** Gabor filters are directly related to Gabor wavelets, since they can be designed for a number of dilations and rotations. However, in general, expansion is not applied for Gabor wavelets, because this requires computation of bi-orthogonal wavelets, which may be very time-consuming. Therefore, usually, a filter bank consisting of Gabor filters with various scales and rotations is created. The filters are convolved with the signal, resulting in a so-called Gabor space. This process is closely related to processes in the primary visual cortex[361]. The Gabor space is very useful in image processing applications such as optical character recognition, iris recognition and fingerprint recognition. Relations between activations for a specific spatial location are very distinctive between objects in an image. Furthermore, important activations can be extracted from the Gabor space in order to create a sparse object representation. This is an example implementation in MATLAB/Octave:

Function gb =gabor_fn (sigma, theta, lambda, psi, gamma)

sigma_x = sigma; sigma_y = sigma/gamma;

- **Graph:** is a representation of a set of objects where some pairs of objects are connected by links. The interconnected objects are represented by mathematical abstractions called vertices, and the links that connect some pairs of vertices are called edges[362]. Typically, a graph is depicted in diagrammatic form as a set of dots for the vertices, joined by lines or curves for the edges. Graphs are one of the objects of study in discrete mathematics.
- **Ground truth:** In machine learning, the term "ground truth" refers to the accuracy of the training set's classification for supervised learning techniques. This is used in statistical models to prove or disprove research hypotheses. The term "ground trothing" refers to the process of gathering the proper objective data for this test.
- **In-sample:** makes it easy to expect medical costs and health outcomes in your patient population.

- **Optical density:** is a logarithmic ratio of the falling radiation to the transmitted radiation through a material. It is also referred as a fraction of absorbed radiation at a particular wavelength.
- **Out-of-sample:** is a popular way to test the expected accuracy of an estimating method. It is balanced since it's unprotected of all adjustments and filters.
- **K-means:** is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells.
- **Laplacian of Gaussian:** is a 2-D isotropic measure of the 2nd spatial derivative of an image. The Laplacian of an image highlights regions of rapid intensity change and is therefore often used for edge detection. The Laplacian is often applied to an image that has first been smoothed with something approximating a Gaussian smoothing filter in order to reduce its sensitivity to noise, and hence the two variants will be described together here. The operator normally takes a single graylevel image as input and produces another graylevel image as output.
- **Margin:** in machine learning, it is the distance between a decision boundary and a data point.
- **Mark-up**: A (document) mark-up language is a modern system for annotating a document in a way that is syntactically distinguishable from the text.
- **Over segmentation techniques:** For some images it is not possible to set segmentation process parameters, such as a threshold value, so that all the objects of interest are extracted from the background or each other without over segmenting the data. Over segmentation is the process by which the objects being segmented from the background are they segmented or fractured into subcomponents.
- **Pattern spectrum:** the application of pattern spectra algorithms to images of materials of different types, for the purpose of pattern classification. As materials are often best characterized by their texture, pattern spectra constitute a very important tool for texture analysis. Researchers by experiments discussed pattern spectra and they compare the algorithms in terms of speed and accuracy. From the

results it is evident that the slope pattern spectra algorithm is a fast and robust alternative to granulometric-based techniques.

- **Principal Components Analysis (PCA):** is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to (i.e., uncorrelated with) the preceding components. Principal components are guaranteed to be independent if the data set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables[363].
- **Receiver Operating Characteristic (ROC) curve:** In signal detection theory, a receiver operating characteristic (ROC), or simply ROC curve is a graphical plot which illustrates the performance of a binary classifier system as its discrimination threshold is varied. It is created by plotting the fraction of true positives out of the total actual positives (TPR = true positive rate) vs. the fraction of false positives out of the total actual negatives (FPR = false positive rate), at various threshold settings. TPR is also known as sensitivity or recall in machine learning. The FPR is also known as the fall-out and can be calculated as one minus the more well-known specificity. The ROC curve was first developed by electrical engineers and radar engineers during World War II for detecting enemy objects in battlefields and was soon introduced to psychology to account for perceptual detection of stimuli. ROC analysis since then has been used in medicine, radiology, biometrics, and other areas for many decades and is increasingly used in machine learning and data mining research [364].
- **Segmentation:** Image segmentation; mean the partitioning of a digital image into two or more regions.
- **Under segmentation:** A outcome case to the over segmentation problem [365] is that of under segmentation. In under segmentation, segmenting processes are applied to extract objects of interest from an image as overlapping objects, rather than separate components. The best solution to the under segmentation problem is

to avoid it. For example, the techniques of region growing[365] can be used as a segmentation method to prevent under segmentation. However, even these methods might produce under segmented results with some classes of data.

- **Watershed transform:** used different approaches for image segmentation, such as: Local minima of the gradient of the image may be chosen as markers, in this case an over-segmentation is produced and a second step involves region merging. And marker based watershed transformation make use of specific marker positions which have been either explicitly defined by the user or determined automatically with morphological operators or other ways[366].

D.*Appendix D: Margin between two hyper planes*

Figure D1 suppose the margin is *m*, it can be computed as equation D.1:

$$m = \frac{w.(x_- - x_+)}{|w|} \quad \text{D.1}$$

This is the projection of $(x_- - x_+)$ in the direction of w. The equation above can be rewritten as equation D.2:

$$m = \frac{w.x_- - w.x_+}{|w|} \quad \text{D.2}$$

Considering Equations. 3.2 and Equation. D.2 can be written as in equation D.3 & D.4:

$$m = \frac{(b+1)-(b-1)}{|w|} \quad \text{D.3}$$

$$= 2/|w| \quad \text{D.4}$$

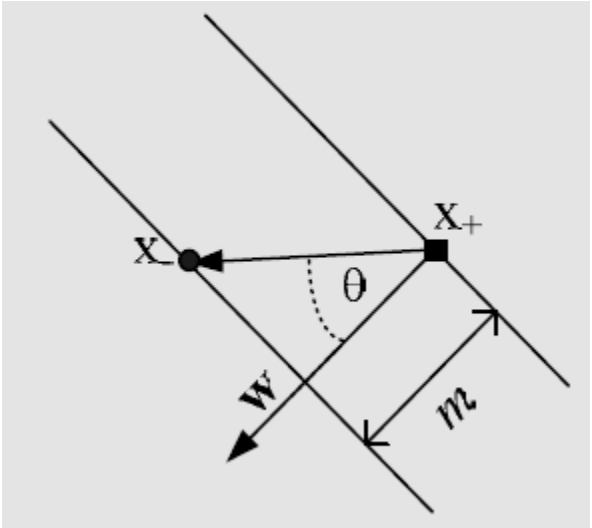


Figure D1: Margin *m* between two supporting hyperplanes

E. *Appendix E: Lagrangian dual optimization*

Considering Eq. 3.13, the construction of the Lagrangian for the optimization problem is displayed in equation E.1 & E.2:

$$L\ (\propto, w, b) = \frac{1}{2}\ w.w - \sum_{i=1}^{l} \propto_i \quad (y_i(w.x_i - b) - 1) \qquad \text{E.1}$$

$$(\propto, w, b) = \frac{1}{2}\ w.w - \sum_{i=1}^{l} \propto_i\ y_i\ w.x_i + b \sum_{i=1}^{l} \propto_i\ y_i + \sum_{i=1}^{l} \propto_i \qquad \text{E.2}$$

This gives the Lagrangian optimization problem as displayed on equation E.3:

$$\max_{\propto} \min_{w,b} L(\propto, w, b) \qquad \text{E.3}$$

Subjects to:

$$\propto_i \geq 0 \qquad \text{E.4}$$

Since the objective function is rounded and the restraint is linear, the solution for $\propto$, w and b has to satisfy the following KKT conditions, equations E.5 – E.9:

Gradient condition $\frac{\partial \mathrm{L}}{\partial \mathrm{w}}\ (\propto, \mathrm{w}, \mathrm{b}) = 0$ E.5

$$\frac{\partial \mathrm{L}}{\partial \mathrm{b}}\ (\propto, \mathrm{w}, \mathrm{b}) = 0 \qquad \text{E.6}$$

Orthogonally condition $\propto_{\mathrm{i}}\ (\mathrm{y_i}\ (\mathrm{w}.\mathrm{x_i} - \mathrm{b}) - 1 \geq 0$ E.7

Feasibility condition $\mathrm{y_i}\ (\mathrm{w}.\mathrm{x_i} - \mathrm{b}) - 1 \geq 0$ E.8

Non-negative condition $\propto_{\mathrm{i}} \geq 0$ E.9

The Lagrangian optimization can be solved through the partial derivative as displayed in the following equations E.10 – E.12:

$$\frac{\partial \mathrm{L}}{\partial \mathrm{w}}\ (\propto, \mathrm{w}, \mathrm{b}) = \mathrm{w} - \sum_{\mathrm{i}=1}^{\mathrm{l}} \propto_{\mathrm{i}} \mathrm{y_i} \mathrm{x_i} = 0 \qquad \text{E.10}$$

$$\mathrm{w} - \sum_{\mathrm{i}=1}^{\mathrm{l}} \propto_{\mathrm{i}} \mathrm{y_i} \mathrm{x_i} \qquad \text{E.11}$$

$$\frac{\partial \mathrm{L}}{\partial \mathrm{b}}\ (\propto, \mathrm{w}, \mathrm{b}) = \sum_{\mathrm{i}=1}^{\mathrm{l}} \propto_{\mathrm{i}} \mathrm{y_i} = 0 \qquad \text{E.12}$$

The Lagrangian dual optimization problem, displayed as in equation E.13:

$$\mathrm{L}\ (\propto, \mathrm{w}, \mathrm{b}) = \sum_{\mathrm{i}=1}^{\mathrm{l}} \propto_{\mathrm{i}} - \frac{1}{2} \sum_{\mathrm{i}=1}^{\mathrm{l}} \sum_{\mathrm{j}=1}^{\mathrm{l}} \propto_{\mathrm{i}} \propto_{\mathrm{j}} \mathrm{y_i} \mathrm{y_j}\ \mathrm{x_i}.\mathrm{x_j} \qquad \text{E.13}$$

Bias b can be determined as in equation E.14:

$$\mathrm{b} = \mathrm{w}.\mathrm{x_{sv}} - 1 = \sum_{\mathrm{i}=1}^{\mathrm{l}} \propto_{\mathrm{i}} \mathrm{y_i} \mathrm{x_i}.\mathrm{x_{sv}} - 1 \qquad \text{E.14}$$

Finally, the optimal decision surface can be defined as displayed in equation E.15 and E.16:

$$\hat{f}\ (x) = \text{sgn}\ (w.x - b) \qquad \text{E.15}$$

$$\hat{f}\,(x) = \text{sgn}\left(\sum_{i=1}^{l} \propto_i y_i x_i . x - \sum_{i=1}^{l} \propto_i y_i x_i . x_{sv} - 1\right) \qquad \text{E.16}$$

F. *Appendix F: Soft-margin nonlinear support vector machine*

$$m^* = \min \frac{1}{2} w.w + \sum_{i=1}^{l} \xi_i \quad \text{F.1}$$

The Lagrangian (is named after Italian-French mathematician, it's a function that descripts the state of a dynamic system in terms of position coordinates and their time derivatives and that is equal to the…) structure of the optimization problem is modified as displayed in equation F.2:

$$L(\alpha, \beta, \xi, w, b) = \frac{1}{2} w.w + C \sum_{i=1}^{l} \xi_i - \sum_{i=1}^{l} \alpha_i (y_i (w.x_i - b) + \xi_i - 1) \quad \text{F.2}$$

This gives the Lagrangian optimization problem as displayed in equation F.3:

$$\max_{\alpha} \min_{w,b} L(\alpha, \beta, \xi, w, b) \quad \text{F.3}$$

Subject to the limits:

$$\propto_i \geq 0 \quad \text{F.4}$$

$$\beta_i \geq 0 \quad \text{F.5}$$

Since the objective function is rounded and the limit is linear, the solution for α, β, w, ξ, b has to satisfy the first order necessary conditions for a solution in mathematics, the Karush-Kuhn-Tucker (KKT) conditions (also known as the Kuhn-Tucker) conditions as displayed in equations F.6 – F.14:

$$\frac{\partial L}{\partial w} \mathrm{L}(\propto, \beta, \xi, w, b) = 0 \quad \text{F.6}$$

$$\partial \mathrm{L}/\partial \xi \ \mathrm{L}(\propto, \beta, \xi, w, b) = 0 \quad \text{F.7}$$

$$\partial \mathrm{L}/\partial b \ \mathrm{L}(\propto, \beta, \xi, w, b) = 0 \quad \text{F.8}$$

$$\alpha_i (y_i (w.x_i - b) + \xi - 1) = 0 \quad \text{F.9}$$

$$\beta_i \xi_i = 0 \quad \text{F.10}$$

$$y_i (w.x_i - b) + \xi - 1 \geq 0 \quad \text{F.11}$$

$$\propto_i \geq 0 \quad \text{F.12}$$

$$\beta_i \geq 0 \quad \text{F.13}$$

$$\xi_i \geq 0 \qquad \text{F.14}$$

The partial derivatives are displayed in equations F.15 - F.19:

$$\frac{\partial L}{\partial w} L\ (\propto, \beta, \xi, w, b) = w - \sum_{i=1}^{l} \propto_i y_i x_i = 0 \qquad \text{F.15}$$

$$w = \sum_{i=1}^{l} \propto_i y_i x_i \qquad \text{F.16}$$

$$\frac{\partial L}{\partial b} L\ (\propto, \beta, \xi, w, b) = \sum_{i=1}^{l} \alpha_i y_i = 0 \qquad \text{F.17}$$

$$\partial L/\partial \beta\ L\ (\propto, \beta, \xi, w, b) = \text{C} - \alpha_i - \beta_i \qquad \text{F.18}$$

$$C = \alpha_i - \beta_i \qquad \text{F.19}$$

Substituting partial derivatives above to the Lagrangian of the optimization problem gives the Lagrangian dual as displayed in equation F.20:

$$L\ (\alpha, w, b) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j x_i \cdot x_j \qquad \text{F.20}$$

This Lagrangian dual for soft-margin SVM is same with the Lagrangian dual for hard-margin SVM. Therefore the objective functions of the optimization problems for both SVMs are same. The difference is in the limitation. In the soft-margin SVM, the constraint can be stated as in equation F.21:

$$0 \leq \propto \leq C \qquad \text{F.21}$$

G.*Appendix G: Sequential minimal optimization (SMO) for SVM*

The sequential minimal optimization (SMO) is used to find the Lagrangian multiplier α in Equation 3.30, which is the solution displayed on equation G.1:

$$\max(L\ (\propto)) \quad \text{G.1}$$

Where

$$L\ (\propto) = \sum_{i=1}^{l} \propto_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{i=1}^{l} \propto_i \propto_j y_i y_j\ k\ (x_i . x_j) \quad \text{G.2}$$

subjects to:

$$0\ \leq \propto \leq C \quad \text{G.3}$$

And

$$\sum_{i=1}^{l} y_i \propto_i = 0 \quad \text{G.4}$$

This maximization problem is equivalent to a minimization problem by multiplying it by-1, so that it becomes the minimization problem as displayed in equations G.5 & G.6:

$$\max_{\propto} L\ (\propto) \equiv\ \min_{\propto}(-L\ (\propto)) \quad \text{G.5}$$

$$= \min_{\propto}(\ell\ (\propto)) \quad \text{G.6}$$

Where

$$\ell\ (\propto) = -L\ (\propto) \quad \text{G.7}$$

Or

$$\ell\ (\propto) = (\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \propto_i \propto_j y_i y_j\ k\ (x_i . x_j) - \sum_{i=1}^{l} \propto_i) \quad \text{G.8}$$

Thus, to find α, minimization of $\ell\ (\alpha)$ in Eq. G.8 is used. Can be written in another form as displayed in equation G.9:

$$\ell\ (\propto) = \frac{1}{2} x \propto^T\ Q \propto\ - q^T\ \propto \quad \text{G.9}$$

where:

- $\boldsymbol{q} = x[1, \ldots, 1]^T$ is the vector of all ones.

- $\boldsymbol{\alpha} = [\propto_1, \propto_1, \ldots, \propto_l]^T$
- Q is a $l\ x\ l$ symmetric matrix (l is the number of training data, with components:

- $Q_{ij} = y_i y_j\ k\ (x_i x_j)$ G.10

Furthermore, the constraint defined in Equation 3.31, $\sum_{i=1}^{l} \propto_i y_i = 0$, can be written in the other form as displayed in equation G.11:

$$\sum_{i=1}^{l} \propto_i y_i = Y \propto \quad \text{G.11}$$

Where Y is the vector $1\ x\ l$ with components y_i,

It's reported that the difficulty in the minimization of $\ell\ (\alpha)$ is that $\boldsymbol{Q}$ is a thick matrix which crops a problem of memory and is time consuming[367]. SMO solves this problem by decomposition such that at each step two Lagrangian multipliers $\propto$ are optimized and update the SVM at the end of each optimization. Researchers in[212] described the algorithm as follows:

a) Defining $\propto^1$ for the initial solution (iteration n = 1).
b) If $\propto^n$is a stationary point of the minimization of$\ell(\alpha)$, stop. Otherwise, a two–element working set $W = [i, j]$ is determined.
c) Defining $V \equiv [1 \ldots l]W$ (the set V contains all those elements of $[1 \ldots l]$ which are not in$W$).
d) Let $\propto_w{}^n$ and $\propto_v{}^n$ be the sub-vectors of $\propto^n$corresponding to W and$V$, respectively.
e) Considering to Equation G.9, solve for the following sub-problem with the variable $\propto_w$:

$$\min_{\propto B} \ell_{\propto B}$$

Where

$$\ell_{\propto B} = \frac{1}{2}[\propto_w{}^T\ (\propto_v{}^n)^T]\begin{bmatrix} Q_{ww} & Q_{wv} \\ Q_{vw} & Q_{vv} \end{bmatrix}\begin{bmatrix} \propto_w \\ \propto^n{}_v \end{bmatrix} - [q_w{}^T \quad q_N{}^T]\begin{bmatrix} \propto_w \\ \propto^n{}_v \end{bmatrix} \quad \text{G.12}$$

$$= \frac{1}{2}[\propto_w{}^T\ Q_{ww}\ \propto_w + \propto_w{}^T\ Q_{wv}\ \propto_v{}^n + \ (\propto_v{}^n)^T Q_{vw}\ \propto_v + \ (\propto_v{}^n)^T Q_{vv}\ \propto_v{}^n] - [q_w{}^T\ \propto_w + \quad q_N{}^T\ \propto_v{}^n] \quad \text{G.13}$$

$$= \frac{1}{2}[\propto_w{}^T\ Q_{ww}\ \propto_w + \propto_w{}^T\ Q_{wv}\ \propto_v{}^n + (\propto_v{}^n)^T Q_{vw}\ \propto_v - q_w{}^T\ \propto_w] + \ ] \quad (\propto_v{}^n)^T Q_{vv}\ \propto_v{}^n - q_N{}^T\ \propto_v{}^n \quad \text{G.14}$$

$$= \frac{1}{2}\propto_w{}^T\ Q_{ww}\ \propto_w + (-\ q_w + Q_{wv}\ \propto_v{}^n)^T\ \propto_w + constant \quad \text{G.15}$$

$$= \frac{1}{2}[\propto_i \propto_j]\begin{bmatrix} Q_{ii} & Q_{ij} \\ Q_{ji} & Q_{jj} \end{bmatrix}\begin{bmatrix} \propto_i \\ \propto_j \end{bmatrix} + (-q_w + Q_{wv} \propto_v{}^n)^T \begin{bmatrix} \propto_i \\ \propto_j \end{bmatrix} + constant \qquad \text{G.16}$$

f) Set $\propto_B{}^{n+1}$ to be the optimal solution of the sub-problem Equation G.16 and $\propto_B{}^{n+1} \equiv \propto_w{}^n$.Set $n \leftarrow n + 1$ and go to Step b.

The index w is updated at every iteration, but for simplicity, w is used instead ofw^n.

The stopping criteria consider that the feasible solution of α is a stationary point of Equation G.9 (or in another form, Equation G.10 if and only if there is a number $\mathscr{b}$ and two non-negative $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ such that reported in [212].

$$\nabla \ell(\propto) + \mathscr{b} y = \lambda - \mu \qquad \text{G.17}$$

With the limitations:

$$\lambda_i \propto_i = 0, p\,(C - \alpha)_i = 0. \lambda_i \geq 0, \rho_i \geq 0 \quad i = 1, \dots, l \qquad \text{G.18}$$

Where $\nabla l(\propto) \equiv Q\alpha - q$ is the gradient of$\ell(\propto)$. Equations G.17 and G.18 result displayed in equation G.19:

$$\nabla \ell \propto_i + \mathscr{b} y_i \begin{cases} \leq 0. \ if \ \propto_i < C \\ \leq 0, \ if \ \propto_i > C \end{cases} \qquad \text{G.19}$$

Because of that ($y_i \in [+1 - 1]$), considering in Equation G.19, $\mathscr{b}$ can be expressed as in equation G.20:

$$m(\propto) \leq \mathscr{b} \leq M(\propto) \qquad \text{G.20}$$

Where

$$m(\propto) \equiv \max_{i \in I_b(\propto)} - y_i \nabla \ell(\propto) \quad \text{And} \quad M(\propto) \equiv \max_{i \in I_b(\propto)} - y_i \nabla \ell(\propto) \qquad \text{G.21}$$

And,

$$I_a(\propto) \equiv [i | \propto_i < C, y_i = 1 \ or \ \propto_i < C, y_i = -1],$$
$$I_b(\propto) \equiv [i | \propto_i < C, y_i = -1 \ or \ \propto_i > C, y_i = 1], \qquad \text{G.22}$$

- Therefore, a feasible $\propto$ is a stationary point of problem in Equation G.10 if and only if:

$$m(\propto) \leq | M(\propto) \qquad \text{G.23}$$

and the stopping principles can be defined as in equation G.24:

$$m(\propto) - M(\propto) \leq \varepsilon_s \qquad \text{G.24}$$

where ε_s is tolerance; typically$\varepsilon_s = 10^{-5}$,

H.*Appendix H: Glossary and Abbreviations*

ABCD-E:		Asymmetry, Border, Colour, Diameter, Evaluation
AC	:	Accuracy
ACO	:	Ant Colony Optimization
ADC	:	Analog to digital Convertor.
ADF	:	Anisotropic Diffusion Filters
AFE	:	Automated Feature Extraction
ALOS	:	Average Length Of Stay
AMF	:	Adaptive Median Filter
AMPCo	:	Australian Medical Publishing Company
ANN	:	Artificial Neural Network
ANS	:	Autonomic Nervous System
ART	:	Adaptive Resonance Theory
BCC	:	Basal Cell Carcinoma.
BI	:	Binary Images.
B: M	:	Benign to malignant ratio
BNN	:	Backpropagation neural network
CA	:	Concavity Alignment
CAD	:	Computer-aided design.
CDF	:	Cumulative Distribution Function.
CBIR	:	Content–Based Image Retrieval System.
CCD	:	Charge Coupled Device
CLSM	:	Confocal laser scanning microscopy
CMOS	:	Complementary Metal-Oxide Semiconductor.
CNS	:	Central nervous system
CSE	:	Clinical skin examination
CSIRO	:	Commonwealth Scientific and Industrial Research Organization

CT : CAT scan
DALY : Disability-Adjusted life year
DCIS : Ductal Carcinoma In Situ
DIA : Digital Image Acquisition
DIR : Digital Image Representation
DNA : Deoxyribonucleic Acid.
DOG : Difference Of Gaussian
DR : Detection Rate
DWT : Discrete Wavelet Transform
2D-DWT : Two-Dimensional Discrete Wavelet Transform
EEG : Electroencephalogram Test to Measure Brain Electrical: Activity
ELM : Epiluminescence Light Microscopy
EM : Electron microscope
EM : Expectation Maximization
EP : Extra Pixels
FAR : False Alarm Rate
FCM : Fuzzy C-means
FFT : Fast Fourier transforms.
FIM : Feature Intensity Map
FIS : Fuzzy Inference System
FL : Fuzzy Logic
FN : False Negative
FNAC : Fine-Needle Aspiration Cytology
FP : False Positive
FPR : False Positive Rate
FS : Feature Selection
GA : Genetic Algorithm
GF : Gabor Filter.
GFBA : Gabor Filter Bank Application.
GHE : Gray-scale Histogram Equalization
GM : Geometric Mean
GP : General Practitioner Physicians
GT : Ground Truth

GVF : Gradient Vector Flow (Snakes)
H.E. : Histogram Equalization.
H&E : Haematoxylin & Eosin
HIS : Hybrid Intelligent System
HPF : High Pass Filter
HP : Histogram Processing.
HSI : Hue, Saturation, Intensity
HSL : Hue-Saturation-Lightness.
HSV : Hue, Saturation, Value
IDC : Invasive Ductal Carcinoma
IDS : International Dermoscopy Society
IHP : Independent Histogram Pursuit
II : Intensity Images.
IPT : Image Processing Toolbox.
JSEG : Japan Society of Engineering geology.
KKT : Karush-Kuhn-Tucker
K-NN : k-Nearest Neighbour
LAHE : Local Area Histogram Equalization.
LHE : Local Histogram Equalisation
LM : Light Microscope
LMS : Least Mean Square
LoG : Laplacian of Gaussian
LMS : Least Mean Square
LOO : Leave-one-out
LSVM : Linear Support Vector Machine
MACWM: Modified Adaptive Centre Weighed Median
MACWM : Modified Adaptive Canter Weighted Median
MDS : Multi-Dimensional Scaling
MED : Minimum Euclidean Distance
MF : Median Filter.
.MI : Medical Imaging
MIR : Mortality-to-Incidence Ratio
MM : Malignant Melanoma.

MOD : Mean Optical Density
MRI : Magnetic Resonance Imaging
MRF : Markov Random Field
MRMR : Minimal-Redundancy-Maximal-Relevance
NHMRC : National Health and Medical Research Council
NLSVM : Nonlinear Support Vector Machine
NMSC : Non-Melanoma Skin Cancer.
NN : Neural Networks
OD : Optical Density
OM : Optical Microscope
OS-FCM : Orientation Sensitive Fuzzy C-Means.
ORSCSS : Oxford Recognition Skin Cancer Screen System
ORL : Oxford Recognition Limited
P : Performance
PCA : Principal Components Analysis
PDF : Probability Density Function.
PNS : Peripheral nervous system
PSO : Particle Swarm Optimization
RACGP : Royal Australian College of General Practionrs
RBF : Radial Basis Function
RCM : Reflectance Confocal Microscopy.
RFE : Recursive Feature Elimination
RGB : Red, Green, Blue
RGBI : Red, Green and Blue images
RL : Reinforcement Learning
ROC : Receiver Operating Characteristics
ROI : Region Of Interest
RTD : Relative tumour density
SAQ : Self-administered questionnaire
SBS : Sequential Backward Selection
SCC : Squamous Cell Carcinoma
SE : Sensitivity
SEM : Scanning Electron Microscope

SF : Spatial Filtering.
SP : Specificity
SFS : Sequential Forward Selection
SIT : Skin Imaging Techniques
SMO : Sequential Minimal Optimization
SMR : Swarm Based Multiple Regression
SN : Spinal nerves
SR : Segmentation Regions
SRM : Statistical Region Merging
SSE : Skin self-examination
SSVM : Swarm Based Support Vector Machine
SVM : Support Vector Machine
SSVM : Swarm Optimization Support Vector Machine
TDV : Total Dermoscopy Value
TEM : Transmission Electron Microscope
TLM : Transillumination Light Microscopy.
TN : True Negative
TP : True Positive
TPR : True Positive Rate
US : United States of America
UV : Ultraviolet Radiation
UVR : Ultraviolet radiation
VGP : Vertical Growth Phase
WF : Wiener Filter.
WLG Wavelet Gabor
WPT : Wavelet Packet Transform
YLD : Years of healthy life lost due to disability
YLL : Years of Life Lost

Glossary [77]

Aboriginal or Torres Strait Islander: A person of Aboriginal and/or Torres Strait Islander descent that identifies as an Aboriginal and/or Torres Strait Islander. See also *Indigenous*.

Additional diagnosis: A condition or complaint either coexisting with the principal diagnosis or arising during the episode of care.

Administrative databases: Observations about events that are routinely recorded or required by law to be recorded. Such events include births, deaths, hospital separations and cancer incidence. Administrative databases include the Australian Cancer Database, the National Mortality Database and the National Hospital Morbidity Database.

Admitted patient: A person who undergoes a hospital's formal admission process to receive treatment and/or care. Such treatment or care can occur in hospital and/or in the person's home (as a 'hospital-in-home' patient).

Age-specific rate: A rate for a specific age group. The numerator and denominator relate to the same age group.

Age-standardisation: A method of removing the influence of age when comparing populations with different age structures. This is usually necessary because the rates of many

diseases vary strongly (usually increasing) with age. The age structures of the different populations are converted to the same 'standard' structure; then the disease rates that would have occurred with that structure are calculated and compared.

Average length of stay (ALOS): The average (mean) number of patient days for admitted patient episodes. Patients admitted and separated on the same date are allocated a length of stay of 1 day.

Benign: Non-cancerous tumours that may grow larger but do not spread to other parts of the body.

Burden of disease and injury: Term referring to the quantified impact of a disease or injury on an individual or population, using the *disability-adjusted life year* (DALY) measure.

Cancer (malignant neoplasm): A large range of diseases in which some of the body's cells become defective begin to multiply out of control, can invade and damage the area around them, and can also spread to other parts of the body to cause further damage.

Chemotherapy: The use of drugs (chemicals) to prevent or treat disease, with the term being applied for treatment of cancer rather than for other uses.

Cohort method: A method of calculating *survival* that is based on a cohort of people diagnosed with cancer in a previous time period and followed over time.

Disability-adjusted life years (DALYs): A year of healthy life lost, either through premature death or equivalently through living with disability due to illness or injury. It is the basis unit used in *burden of disease and injury* estimates.

Death due to cancer: A death where the underlying cause is indicated as cancer.

Expected survival: A measure of *survival* that reflects the proportion of people in the general population alive for a given amount of time. Expected survival estimates are crude estimates calculated from *life tables* of the general population by age, sex and calendar year.

Histology: The microscopic characteristics of cellular structure and composition of tissue.

Incidence: The number of new cases (of an illness or event, and so on) in a given period.

Indigenous: A person of Aboriginal and/or Torres Strait Islander descent who identifies as an Aboriginal and/or Torres Strait Islander. See also *Aboriginal or Torres Strait Islander.*

Length of stay: Duration of hospital stay, calculated by subtracting the date the patient was admitted from the day of separation. All leave days, including the day the patient went on leave, are excluded. A same-day patient is allocated a length of stay of 1 day.

Life tables: Tables of annual probabilities of death in the general population.

Limited-duration prevalence: The number of people alive at a specific time who have been diagnosed with cancer over a specified period (such as the previous 5 or 25 years).

Malignant: A tumour with the capacity to spread to surrounding tissue or to other sites in the body. 186 Cancer in Australia: an overview, 2012

Median: The midpoint of a list of observations that have been ranked from the smallest to the largest.

.**Mortality due to cancer:** The number of deaths which occurred during a specified period (usually a year) for which the underlying cause of death was recorded as cancer.

Mortality-to-incidence ratio (MIR): The ratio of the age-standardised mortality rate for cancer to the age-standardised incidence rate for cancer.

Neoplasm: An abnormal (‘neo’, new) growth of tissue. Can be ‘benign’ (not a cancer) or ‘malignant’ (a cancer). Also known as a tumour.

Non-Indigenous: People who have declared that they are not of Aboriginal or Torres Strait Islander descent.

Observed survival: A measure of *survival* that reflects the proportion of people alive for a given amount of time after a diagnosis of cancer. Observed survival estimates are crude estimates calculated from population-based cancer data.

Other Australians: Includes people who have declared that they are not of Aboriginal or Torres Strait Islander descent as well as those who have not stated their Indigenous status.

Overnight patient: An admitted patient who receives hospital treatment for a minimum of 1 night (that is, is admitted to, and separates from, hospital on different dates).

Patient days: The number of full or partial days of stay for patients who were admitted for an episode of care and who underwent separation during the reporting period. A patient who is admitted and separated on the same day is allocated one patient day.

Period method: A method of calculating *survival* that is based on the survival experience during a recent *at-risk* or *follow-up* time period.

Primary cancer: A tumour that is at the site where it first formed (see also *Secondary cancer*).

Principal diagnosis: The diagnosis listed in hospital records to describe the problem that was chiefly responsible for the patient's episode of care in hospital.

Procedure: A clinical intervention that is surgical in nature, carries a procedural risk, carries an anaesthetic risk, requires specialised training and/or requires special facilities or equipment available only in the acute care setting.

Relative survival: The ratio of *observed survival* of a group of persons diagnosed with cancer to *expected survival* of those in the corresponding general population after a specified interval following diagnosis (such as 5 or 10 years).

Risk factor: Any factor that represents a greater risk of a health disorder or other unwanted condition or event. Some risk factors are regarded as causes of disease, others are not Cancer in Australia: an overview, 2012 187 necessarily so. Along with their opposites, namely protective factors, risk factors are known as 'determinants'.

Same-day patient: A patient who is admitted to, and separates from, hospital on the same date.

Secondary cancer: A tumour that originated from a cancer elsewhere in the body. Also referred to as a metastasis.

Separation: An episode of care for an admitted patient which may include a total hospital

stay (from admission to discharge, transfer or death) or a portion of a hospital stay that begins or ends in a change of type of care (for example, from acute to rehabilitation). In this report, separations are also referred to as hospitalisations.

Statistical significance: An indication from a statistical test that an observed difference or association may be significant or 'real' because it is unlikely to be due just to chance. A statistical result is usually said to be 'significant' if it would occur by chance only once in 20 times or less often (see Appendix B for more information about statistical significance).

Stage: The extent of a cancer in the body. Staging is usually based on the size of the tumour, whether lymph nodes contain cancer, and whether the cancer has spread from the original site to other parts of the body.

Survival: A general term indicating the probability of being alive for a given amount time after a particular event, such as a diagnosis of cancer.

Symptom: Any indication of a disorder that is apparent to the person affected.

Underlying cause of death: The disease or injury that initiated the sequence of events leading directly to death.

Valid FOBT test: Only faecal occult blood test (FOBT) results that are either positive or negative are classified as valid results. Inconclusive results are excluded from analysis.

Years of healthy life lost due to disability (YLD): For each new case of cancer, YLD equals the average duration of the cancer (to remission or death) multiplied by a severity weight for cancer (which depends upon its disabling effect over the disease duration).

Years of life lost (YLL): For each new case, YLL equals the number of years between premature death and the standard life expectancy for the individual.

Bibliography

1. Curado, M.P., et al., *Cancer incidence in five continents*. 2007, Intl Agency for Research on Cancer: Lyon,. p. 1-895.
2. Government, A. *Cancer in Australia statistics*. 2014; 1-3]. Available from: (http://canceraustralia.gov.au).
3. *"Cancer Facts & Figures 2002"*, in *American Cancer Society*. 2002: Atlanta, Georgia.
4. Aziz, N. and R.J. . *Trends and advances in cancer survivorship research: Challenge and opportunity*. in *Seminars in Radiation Oncology*. 2003.
5. Aziz, N., *Cancer survivorship research: Challenge and opportunity.* Journal Nutr., 2002. **132**(Suppl.11): p. 3494S-503S.
6. Criscione, V. and M. Weinstock, *Melanoma Thickness Trends in the United States, 1988-2006.* Journal of Investigative Dermatology, 2009.
7. Balch, C.M., S. Soong, and M.I. Ross, *Long-term results of a multi-institutional randomized trial comparing prognostic factors and surgical results for intermediate thickness melanomas (1.0 to 4.0 mm).* Intergroup Melanoma Surgical Trial. Ann Surg Oncol, 2000. **7**: p. 87-97.
8. Barnhill, R.L., J.A. Fine, and G.C. Roush, *Predicting fiveyear outcome for patients with cutaneous melanoma in a population-based study. Cancer.* Cancer, 1996. **78**: p. 427-432.
9. Argenziano G, S.H., Chimenti S, Talamini R, Corona R, Sera F, *Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet.* J Am Acad Dermatol, 2003. **48**(5): p. 679-93.
10. Jablonski, N.G., Chaplin, G. , *Human skin pigmentation as an adaptation to UV radiation.* Proceedings of the National Academy of Sciences, 2010. **107**.
11. Stratigos, A.J., et al., *Euromelanoma: a dermatology-led European campaign against nonmelanoma skin cancer and cutaneous melanoma. Past, present and future.* British Journal of Dermatology, 2012. **167**: p. 99-104.
12. Blackledge, J.M. and D.A. Dubovitskiy, *Object Detection and Classification with Applications to Skin Cancer Screening.* ISAST TRANSACTIONS ON INTELLIGENT SYSTEMS, 2008. **1**(2): p. 1-12.
13. Blum, A., Jaworski,S., *Clear differences in hand-held dermoscopes.* JDDG, 2006. **4**: p. 1054–1057.
14. Boucheron, L.E., *Object- and Spatial-Level Quantitative Analysis of Multispectral Histopathology Images for Detection and Characterization of Cancer*, in *Electrical*

and Computer Engineering. 2008, UNIVERSITY of CALIFORNIA: Santa Barbara. p. 321.

15. Abu Mahmoud, M.K. and A. Al-Jumaily, *A Hybrid System for Skin Lesion Detection: Based on Gabor Wavelet and Support Vector Machine*, in *CISP, BMEI 2014*, IEEE, Editor. 2014, IEEE: Dalian, China.
16. (NRHA), N.R.H.A. *Skin Cancer in Australia: awareness, early diagnosis and management*. National Rural Health Conference Australian Journal of Rural Health 2014; Available from: http://ruralhealth.org.au/sites/default/files/documents/nrha-policy-document/submissions/sub-skin-cancer-inquiry-21-march-2014.pdf.
17. Rigel DS, C.J., *Malignant melanoma: prevention early detection, and treatment in the 21st century.* CA Cancer J Clin, 2000. **50**: p. 215-36.
18. Authors, A., *Cancer Facts & Figures 2002.* American Cancer Society, 2002: p. 1-44.
19. Chiem, A., A. Al-Jumaily, and R.N. Khushaba. *A Novel Hybrid System for Skin Lesion Detection*. 2007. IEEE.
20. Abu Mahmoud, M.K., Al-Jumaily,A., *Segmentation of Skin Cancer Images Based on Gradient Vector Flow (GVF) Snake*, in *2011 IEEE International Conference on Mechatronics and Automation ICMA*. 2011. p. 216-220.
21. Abu Mahmoud, M.K., Al-Jumaily, A., Takruri, M., *Wavelet and Curvelet Analysis for Automatic Identification of Melanoma Based on Neural Network Classification*, in *International Journal of Computer Information Systems and Industrial Management (IJCISIM)*. 2013, IJCISIM: MIR Labs, USA. p. 606-614.
22. Abu Mahmoud, M.K., et al., *Classification of Malignant Melanoma and Benign Nevi from Skin Lesions based on Support Vector Machine*, in *Computational Intelligence, Modelling and Simulation (CIMSim), 2013 Fifth International Conference*. 2013: Seoul, Korea (South). p. 236 - 241.
23. Abu Mahmoud, M.K. and A. Al-Jumaily, *Novel feature extraction methodology based on histopathalogical images and subsequent classification by Support Vector Machine.*, in *IEEE*, I.C. Publications, Editor. 2014. p. 1 - 6
24. Takruri, M., A. Al-Jumaily, and M.K. Abu Mahmoud, *Automatic Recognition of Melanoma Using Support Vector Machines: A Study Based on Wavelet, Curvelet and Color Features*, in *IEEE*. 2014.
25. Abu Mahmoud, M.K. and A. Al-Jumaily, *A Hybrid System for Skin Lesion Detection: Based on Gabor Wavelet and Support Vector Machine*, in *IEEE*. 2014, IEEE.
26. Jamarani, S.M.H., G. Rezai-rad, and H. Behnam, *A novel method for breast cancer prognosis using wavelet packet based neural network*, in *27th Annual International Conference of the Engineering in Medicine and Biology Society* 2005, (IEEE-EMBS '05: Shanghai, China.
27. Gonzalez, R.C., R.E. Woods, and S.L. Eddins, *Digital Image Processing Using Matlab*, ed. 109876543. 2004, Upper Saddle River, New Jersey, USA: Pearson, Prentice Hall.

28. Dietterich, T.G., *Approximate statistical tests for comparing supervised classification learning algorithms.* Neural Computing & Applications, 1997. **10**: p. 1895.

29. Abu Mahmoud, M.K., A. Al-Jumaily, and M. Takruri, *Wavelet and Curvelet Analysis for Automatic Identification of Melanoma Based on Neural Network Classification*, in *International Journal of Computer Information Systems and Industrial Management (IJCISIM)*. 2013, IJCISIM: MIR Labs, USA. p. 606-614.

30. Takruri, M., A. Al-Jumaily, and M.K.A. Mahmoud. *Automatic recognition of melanoma using Support Vector Machines: A study based on Wavelet, Curvelet and color features*. in *Industrial Automation, Information and Communications Technology (IAICT), 2014 International Conference on*. 2014. IEEE.

31. Abu Mahmoud, M.K. and Al-Jumaily.A., *Novel feature extraction methodology based on histopathalogical images and subsequent classification by Support Vector Machine*, in *(ICCVIA2014) International Conference on Computer Vision & Image Analysis, Ras Al Khaimah, UAE*. 2014.

32. Abu Mahmoud, M.K., et al., *Classification of Malignant Melanoma and Benign Nevi from Skin Lesions based on Support Vector Machine*, in *IEEE 5th International conference on Computational Intelligence, Modelling and Simulation*, I.c. socity, Editor. 2013, CIMSim2013: South Korea. p. 236-241.

33. Abu Mahmoud, M.K., A. Al-Jumaily, and M. Takruri, *The Automatic Identification of Melanoma by Wavelet and Curvelet Analysis: Study Based on Neural Network Classification*, in *11th International Conference on Hybrid Intelligent Systems (HIS)*. 2011, IEEE. p. 680-685.

34. Abu Mahmoud, M.K. and A. Al-Jumaily, *Segmentation of Skin Cancer Images Based on Gradient Vector Flow (GVF) Snake*, in *2011 IEEE International Conference on Mechatronics and Automation ICMA*. 2011. p. 216-220.

35. Young, B., et al., *Wheater's Functional Histology*, ed. t.e. 2006. 2006: Elsevier.

36. Ross, M.H., G.I. Kaye, and W. Pawlina, *Histology: a text and atlas with cell and molecular biology*. 2003.

37. Kelly, J. and D. Van Essen, *Cell structure and function in the visual cortex of the cat.* The Journal of physiology, 1974. **238**(3): p. 515-547.

38. Lodish, H.F., et al., *Molecular cell biology*. Vol. 4. 2000: WH Freeman New York.

39. Albrecht-Buehler, G., *Daughter 3T3 cells. Are they mirror images of each other?* The Journal of cell biology, 1977. **72**(3): p. 595-603.

40. Young, B., *Wheater's Functional Histology: A and Colour Atlas (with student consult online access),* . Functional Histology Wheater's, ed. 6e. 2013.

41. Stolz W, B.-F.O., Bilek P, Landthaler M *Color alas of drmatoscopy*. Cognetta AB. Vol. 2nd ed. 2002, Berlin: Burgdorf WHC.

42. Norton, M.E., N.C. Rose, and P. Benn, *Noninvasive prenatal testing for fetal aneuploidy: clinical assessment and a plea for restraint.* Obstetrics & Gynecology, 2013. **121**(4): p. 847-850.

43. L., G. and G.W.G. Kumar *Special Stains and H & E Second Edition.* Pathology education guide, 2010.

44. TutorVista, *Skin Cell Structure.* tutoring sessions.

45. Michael, M., *Human Evolutionary Biology.* Cambridge University Press, 2010: p. 192–213.

46. Lucas, R., McMichael,T., Smith,W.,Armstrong,B., *Solar Ultraviolet Radiation: Global burden of disease from solar ultraviolet radiation.* Environmental Burden of Disease Series, 2006. **13**.

47. Christe, D., *Biology.* Biology, 2012,2013: p. 1-9.

48. McGrath, J., R. Eady, and F. Pope, *Anatomy and organization of human skin.*

49. Giovannucci, E.L., Y., et al., *Prospective study of predictors of vitamin D status and cancer incidence and mortality in men.* JNCI Journal of the National Cancer Institute, 2006. **98**(7): p. 451.

50. Staples, M., et al., *Non-melanoma skin cancer in Australia.* Medical Journal Australia, 2006. **198**(184).

51. Foundation, S.C. *Skin Cancer Foundation*. Available from: http://www.skincancer.org.

52. Diepgen, T.L. and V. Mahler, *The epidemiology of skin cancer.* British Journal of Dermatology, 2002. **146**(s61): p. 1-6.

53. Staples, M., R. Marks, and G. Giles, *Trends in the incidence of non-melanocytic skin cancer (NMSC) treated in Australia 1985-1995: are primary prevention programs starting to have an effect?* International Journal of Cancer, 1998. **78**: p. 144-148.

54. Terrill, P., Fairbanks, S., Bailey, M., *Is there just one lesion? The need for whole body skin examination in patients presenting with non-melanocytic skin cancer.* ANZ Journal of Surgery, 2009. **79**(10): p. 707-712.

55. Welfare, A.I.o.H.a., *Australian Association of cancer Registratries*, in *23*, AIHW, Editor. 2003, AIHW: canberra.

56. Randell, P., Cole,J., Quirk, C., Von Nida,J., Yiasemides, E., and Queensland Health *The ABCD of melanoma detection.* Cancer Council Western Australia, 2009.

57. Morton, D.L., et al., *Sentinel node biopsy for early-stage melanoma: accuracy and morbidity in MSLT-I, an international multicenter trial.* Annals of surgery, 2005. **242**(3): p. 302.

58. Feng, J., et al., *Metabolites Studies of Secondary Melanoma on C57BL/6J Mouse Liver Using 1H NMR Metabolomics.* 2013. **3**(4): p. 1011-1035.

59. Jerant AF, J.J., Sheridan CD, Caffrey TJ *"Early detection and treatment of skin cancer".* Am Fam Physician, 2000. **62**(2): p. 357–68, 375–6, 381–2.

60. Armstrong, B., et al., *Challenges in health and health care for Australia.* Medical Journal of Australia, 2007. **187**(9): p. 485.

61. Ferlay, J., Shin HR, Bray, F, Forman, D, Mathers, C, Parkin, DM *Cancer Incidence and Mortality Worldwide: IARC Cancerbase No.10.* GLOBOCAN 2008 & International Agency for Research on Cancer, 2008. **2**.

62. Gies, P., C. Roy, and P. Udelhofen, *Solar and ultraviolet radiation.* In Prevention of Skin Cancer, 2004: p. 21-54.

63. Hershkovitz L, S.J., Treves AJ, Besser MJ *"Focus on adoptive T cell transfer trials in melanoma".* Clin. Dev. Immunol. 2010: 260267, 2010.

64. Wurm EM, S.H., *Scanning for melanoma.* Australian Prescriber 2010. **33**: p. 150–5.

65. Hocker, T. and H. Tsao, *Ultraviolet radiation and melanoma: a systematic review and analysis of reported sequence variants.* Human mutation, 2007. **28**(6): p. 578-588.

66. James, W.D.B., Timothy G., *Diseases of the Skin: clinical Dermatology.*, ed. Andrews. 2006: Saunders Elsevier. 694–9.

67. Wisconsin, M.C.o., *"Malignant Melanoma: staging".* Collaborative Hypertext of Radiology, 2006.

68. Balch C, B.A., Soong S, Atkins M, Cascinelli N, Coit D, Fleming I,, Houghton A, Kirkwood J, McMasters K, Mihm M, Morton D, Reintgen D, Ross M, Sober A, Thompson J, Thompson J, *"Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma".* Gershenwald J and J Clin Oncol, 2001. **19**(16): p. 3635–48.

69. Weedon, D., *Skin pathology, 2nd Edition* 2002, Sydney: Churchill-Livingstone.

70. Breslow, A., *Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma.* Ann Surg, 1970. **172**: p. 902-908.

71. Aziz, N., Rowland JH. . . *Trends and advances in cancer survivorship research: Challenge and opportunity*. in *Seminars in Radiation Oncology*. 2003.

72. Jemal, A., Murray, T., Ward, E, Samuels, A., Tiwari, RC, Ghafoor, A., Feuer, EJ, Thun, MJ, *Cancer Statistics 2005.* CA Cancer Journal Clin., 2005. **55**: p. 10-30.

73. Rowland, J., Mariotto, A.., Aziz, N., Tesauro, G., Feuer E. , *Cancer survivorship United States 1971-2001.* MMWR, 2004. **53**: p. 526-9.

74. Statistics, A.B.o., *National Health Survey: Summary of Results 2004-2005 2006*. 2008, Australian Bureau of Statistics.

75. **Cancer Council Australia**. *Screening and early detection of skin cancer*. 2007; Available from: **www.cancer.org.auCancer**

76. Cancer, I.A.f.R.o., *Cancer incidence in five continents*, in *International Agency for Research on Cancer*. 2002: Lyon.

77. Welfare, A.I.o.H.a., *Cancer in Australia: An overview, 2012*, in *AIHW*. 2012: Canberra, Australia.

78. Brady, M.S., *Melanoma and Other Skin Cancers.* 2013.

79. A., N., et al., *Skin cancer-related prevention and screening behaviors: a review of the literature.* J Behav Med, 2009. **32**: p. 406-428.

80. Valiukevičienė, S., *Melanomų diagnostika ir gydymas.* Kaunas, KMU, 2002: p. 85p.

81. Blackledge, J., Fellow, IET, Dubovitskiy, DA, *Object Detection and Classification with Applications to Skin Cancer Screening.* Intelligent Systems, 2008. **1**(1): p. 34-45.

82. Wyszecki, G., Stiles, W.S., *Color Science, Second Edition*. McGraw-Hill, NY. 1982.

83. Stockham, T.G., *Image Processing in the context of a Visual Model.* Jr. *Proceedings* of *IEEE*, 1972. **60**(7): p. 828-842.

84. Demirkaya, O., M.H. Asyali, and P.K. Sahoo, *Image Processing with Matlab: Applications in medicine and biology*. 2009: CRC Press.

85. Pham, D.T., Alcock, R., *Smart Inspection Systems: Technicuqes and Applications of Intelligent Vision*. Academic Press, Oxford. 2003.

86. Lissesand, T.M., Kiefer, *R.W.* , *Remote Sensing and Image Interpretation*. 1999: 4th Edition, John Wiley and Sons.

87. Jensen, J.R., *Remote Sensing of the Environment: An Earth ResourcePerspective*. 2000: Prentice Hall.

88. Suetens, *P.*, *Fundamentals of Medical Imaging*. Cambridge University Press. 2002.

89. Kupinski, A.A., Giger, h4.. *Automated Seeded Lesion Segmentation on Digital Mammograms.* IEEE Trans. Med. Image, 1998. **17**: p. 510-51 7.

90. Mitra, S.S. and T. Acharya, *Data Mining: Multimedia. Soft Computing, and Bioinformatics*. Vol. *NJ*. 2003: *Wiley, Hoboken*.

91. Acharya, T. and A.K. Ray, *Information Technology: Principles and Applications*. prentice Hall of India. 2004, New Delhi, India.

92. Cucchiara, R., et al., *The Sakbot System for Moving Object Detection and Tracking. Video-Based Surveillance Systems- Computer Vision and Distributed Processing*, 2001: p. 145-157.

93. Davies, E.R., *Machine Vision: Theory, Algorithms, Practicalities*. 1997, Academic press: London.

94. Freeman, J.H., *Machine vision. Algorithms, Architectures, and Systems*, in *Academic press*. 1988, Academic press: London,.

95. Louis, J., Galbiati, JR, *Machine vision and digital image processing fundamentals.* State University of New York, New-York, 1990.

96. Snyder, W.E. and H. Qi., *Machine Vision*. 2004, Cambridge University Press: England.

97. Rosenfeld, A. and A.C. Kak, *Digital Picture Processing*. Second Edition, Academic Press. Vol. 1. 1982.

98. Bracewell, *R.N.*, *Two-Dimensional Imaging.* Prentice Hall, Englewood Cliffs, NJ, 1995.

99. Toennies, K.D., *Guide to Medical Image Analysis: Methods and Algorithms*. 2012, London: Springer-Verlag.

100. Hall, C.F., Hall, E.L., *A Nonlinear Model for the Spatial Characteristics of the Human Visual System. IEEE Trans. Systems, Man, and Cybernetics,* SMC, 1977. **7**(3): p. 161-170.

101. Bayer, B.E., *Color Imaging Array.* Eastman Kodak Company, US *Patent*, 1976. **3**: p. 971,065.

102. Sakamoto, T., C. Nakanishi, and T. Hase, *Software Pixel Interpolation for Digital Still Cameras Suitable for **A** 32-bit MCU. IEEE Transactions on Consumer Electronics*, 1998. **44**(4): p. 1342-1352.

103. wilhelm , P., Enzo,B., Peter,E., Maibach,I., , *"Bio Engineering of Skin: Skin Imaging and Analysis", Dermatology: Basic science series.* Dermatology: Basic science series., 2008.

104. Marozas, M., *Review on skin lesion imaging, analysis and automatic classification.* R.J. Biomedical Engineering Institute, 2006.

105. DS., R., *Malignant melanoma: perspectives on incidence and its effects on awareness, diagnosis, and treatment.* CA Cancer J Clin, 1996. **46**: p. 195-8.

106. Australia, A.G.C., *Melanoma of the skin statistics.* Information Publication Scheme, 2013.

107. Bafounta ML, B.A., Aegerter P, Saiag P, *Is dermoscopy (epiluminescence microscopy) useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests.* Arch Dermatol, 2001. **137**

p. 1343-50.

108. Rigel DS, F.R., Kopf AW, *The incidence of malignant melanoma in the United States: Issues as we approach the 21st entury.* J Am Acad Dermatol, 1996. **34**: p. 839-47.

109. Stolz W, B.-F.O., Bilek P, Landthaler M, *Blackwell Wissenschafts-Verlag; Cognetta AB. Color alas of drmatoscopy. 2nd ed*. 2002, Berlin: Burgdorf WHC.

110. Binder, M., Puespoeck-Schwarz, Steiner, A., Kittler, H., Muellner, M., Wolff, A.,Pehamberger, H. , Austria, A., *Epiluminescence Microscopy. "A useful tool for the diagnosis of pigmented lesions for formally trained dermatologists"*. Journal of the American Academy of Dermatology, 1997. **36**(2): p. 197-202.

111. Binder M, S.M., Winkler A, Steiner A, Kaider A, Wolff K,, *Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists.* Arch Dermatol, 1995. **131**: p. 286-91.

112. Kittler H, P.H., Wolff K, Binder M, *Diagnostic accuracy of dermoscopy.* Lancet Oncol, 2002. **3**: p. 159-65.

113. Pehamberger, H., Steiner,A., Wolff,K., *In vivo epiluminescence microscopy of pigmented skin lesions. i. pattern analysis of pigmented skin lesions.* Journal of the American Academy of Dermatology, 1987. **17**(4): p. 571–583.

114. Tennessee, U.o., *Comparison of OM, TEM and SEM. Electron Microscopy*, ed. D.o.M.S.a.E. (ed.). Vol. Topics 3b,c. University of Tennessee.

115. Fink-Puches, R.H.-W.R., J. Smolle, and H. Kerl, *Confocal laser scanning microscopy: a new optical microscopic technique for applications in pathology and dermatology.* Cutan Pathol, 1995. **22**(3): p. 252-9.

116. Chu, M., N. Kollias, and M.A. Luedtke, *Confocal Scanning Laser Microscopy: Applications for Imaging Dynamic Processes* Skin In Vivo, 2010.

117. **A., S.**, et al., *Statistical evaluation of epiluminescence microscopy criteria for melanocytic pigmented skin lesions. J. Am. Acad. Derm.*, 1993. **29**(4): p. 581-588.

118. Steiner A, P.H., Wolff K, *In vivo epiluminescence microscopy of pigmented skin lesions. II. Diagnosis of small pigmented skin lesions and early detection of malignant melanoma.* J Am Acad Dermatol, 1987 **17** p. 584-91.

119. Nachbar F, S.W., Merkle T, Cognetta AB, Vogt T, Landthaler M, *The ABCD rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions.* J Am Acad Dermatol, 1994. **30**: p. 551-9.

120. Robert, C. and M.D. Burton, *Malignant Melanoma in the Year 2000.* CA Cancer J Clin 2000, 2000. **50**(4): p. 209-213.

121. Menzies, S.W.C.K. and I. C., *An atlas of surface microscopy of pigmented skin lesions*, ed. M.W. 2003. 2003, Sydney: McGraw-Hill Book Company.

122. Argenziano, G., et al., *Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet.* J Am Acad Dermatol, 2003. **48**(5): p. 679-93.

123. Menzies SW, I.C., McCarthy WH. , *A sensitivity and specificity analysis of the surface microscopy features of invasive melanoma.* Melanoma Res 1996 **6**: p. 55-62.

124. Braun RP, R.H., Oliviero M, Kopf AW, Saurat JH., *Pattern analysis: a two step procedure for the dermoscopic diagnosis of melanoma.* Clin Dermatol, 2002. **20**: p. 236-9.

125. Simon, J.D., et al., *Current challenges in understanding melanogenesis: bridging chemistry, biological control, morphology, and function.* Pigment cell & melanoma research, 2009. **22**(5): p. 563-579.

126. Hansen, C., Wilkinson, D., Hansen, M., Argenziano, G., *How good are skin cancer clinics at melanoma detection?Number needed to treat variability across a national clinic group in Australia.* American Acadimic Dermatol 2009(Am Acad Dermatol,).

127. Kreusch J, R.G., *Standardisierte auflichtmikroskopische Unterscheidung melanozytischer und nichtmelanozytischer Pigmentmale.* Hautarzt, 1991(42): p. 77-83.

128. Stolz W, R.A., Cognetta AB, *ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma.* Eur J Dermatol, 1994: p. 4:521-527.

129. Staples, M.P., Elwood, M., Burton, RC., *Nonmelanoma skin cancer in Australia: the 2002 national survey and trends since 1985.* Medical Journal Austustralia, 2006. **184**: p. 6-10.

130. Anthony, F. and M.D. Jerant, *Early detection and Treatment of Skin cancer.* Am Fam Physician, 2000. **62**(2): p. 357-368.

131. MenziesS.W., et al., *The performance of solarscan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma.* Archives of dermatology, 2005. **141**(11): p. 1388–1396.

132. Zalaudek, I., et al., *Three-point checklist of dermoscopy: an open internet study.* British journal of dermatology, 2006. **154**(3): p. 431-437.

133. Celebi, M.E., *Development of Algorithms for Dermoscopy Image Analysis*. 2006: ProQuest.

134. Sydney, B.D., *Skin lession boundary tracing algorithm.* Melanoma-skin cancer reviewed, 2006.

135. Celebi, M.E., et al., *Unsupervised border detection in dermoscopy images.* Skin Research and Technology and Blackwell Munksgaard, 2007. **13**: p. 454-462.

136. Zouridakis, G., M. Doshi, and N. Mullani, *Early diagnosis of skin cancer based on segmentation and measurement of vascularization and pigmentation in nevoscope images*, in *26th Annual International Conference of the IEEE EMBS*. 2004.

137. Winter, A. and C. Nastar, *Differential feature distribution maps for image segmentation and region queries* image databases, 1999: p. 1-5.

138. Ruzon, M.A., *Segmentation.* 1997.

139. Ma, W. and B. Manjunath, *Texture features and learning similarity.* Computer Vision and Pattern Recognition, 1996. **5**: p. 425-430.

140. Zarita, R. and S. Lelandais, *Wavelets and high order statistics for texture classification.* SCIA, 1997. **8**: p. 95-102.

141. Jain, A.K., Murty, M. N., Flynn, P. J. , *Data clustering: a review*, in *American Computer Magazin Computing Surveys (CSUR)*. 1999. p. 1-60.

142. Squire, D., *Data mining - clustering techniques and association rule discovery.* 2005.

143. Chaira, T. and A.K. Ray, *Fuzzy Image Processing and Applications with MatLab*. 2010, U.S.: Taylor & Francis Group.

144. Wolfram, *fuzzy clustering.* 2005. **1** p. 12

145. Schmid, P., *Segmentation and symmetry measure for image analysis: application to digital dermatoscopy.*, in *PRÉSENTÉE AU DÉPARTEMENT D'ÉLECTRICITÉ*. 1999, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE: Lausanne, EPFL. p. 192.

146. Erkol, B., et al., *Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes.* Skin Research and Technology, 2004. **11**: p. 17-26.

147. Geman, S. and D. Geman, *Stochastic relaxation and gibbs distributions and the bayesian restoration of images.* IEEE-PAMI, 1984. **6**: p. 721-741.

148. Deng, H. and D.A. Clausi, *Unsupervised image segmentation using a simple mrf model with a new implementation scheme.* Pattern Recognition, 2004. **37**: p. 2323-2335.

149. Crowson, A.N., C.M. Magro, and M.C. Mihm, *Prognosticators of melanoma, the melanoma report, and the sentinel lymph node.* Modern Pathology, 2006. **19**: p. S71-S87.

150. Clark Jr, W.H., *A classification of malignant melanoma in man correlated with histogenesis and biologic behavior.* Advances in the Biology of the Skin, 1967. **8**: p. 621-647.

151. Clark Jr WH, From L, and B. EA., *Histogenesis and biologic behavior of primary human malignant melanomas of the skin.* Cancer Res, 1969. **29**: p. 705-726.

152. Crowson, A., et al., *The pathology of melanoma.* Cutaneous Melanoma, 4th edn. Quality Medical Publishers Inc.: St Louis, 2003: p. 171-206.

153. Starfield, B., et al., *The effects of specialist supply on populations' health: assessing the evidence.* Cancer, 2005. **103**(20.93): p. 23.18.

154. Babu, M.N., et al., ***Histo-pathological Image Analysis using OS-FCM and Level Sets***. 2010.

155. Demir, C. and B. Yener, *Automated cancer diagnosis based on histopathological images: a systematic survey*, in *TR-05-09*. 2009, Rensselaer Polytechnic Institute, Department of computer science.

156. Mulrane, L., et al., *Automated image analysis in histopathology: a valuable tool in medical diagnostics.* 2008.

157. Clark Jr WH, D.E. Elder, and D. Guerry, *Model predicting survival in stage I melanoma based on tumor progression.* Natl Cancer Inst, 1989. **81**: p. 1893-1904.

158. Clark, W.H. and M.C. Mihm, *Lentigo maligna and lentigo maligna melanoma.* Am J Pathol, 1969. **55**: p. 39-54.

159. Reed, R.J., *Acral lentiginous melanoma. In: New Concepts in Surgical Pathology of the Skin.* Wiley: New York, 1976: p. 89-90.

160. Koh, H.K., E. Michalik, and A.J. Sober, *Lentigo maligna melanoma has no better prognosis than other types of melanoma.blogsome.com.* J Clin Oncol 1984. **2**: p. 994-1001.

161. Vollmer, R.T., *Malignant melanoma. A multivariate analysis of prognostic factors.* Pathol Annu, 1989. **24**: p. 383-407.

162. Weinstock, M.A. and A.J. Sober, *The risk of progression of lentigo maligna to lentigo maligna melanoma.* Br J Dermatol, 1987. **116**: p. 303-310.

163. Fidler, I.J., *Tumor heterogeneity and the biology of cancer invasion and metastasis.* Cancer Res, 1978. **38**: p. 2651-2660.

164. Balch, C.M., S.J. Soong, and J.E. Gershenwald, *Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system.* J Clin Oncol, 2001. **19**: p. 3622-3634.

165. Balch, C.M., T.M. Murad, and S.J. Soong, *A multlifactorial analysis of melanoma: prognostic histopathological features comparing Clark's and Breslow's staging methods.* Ann Surg, 1978. **188**: p. 732-742.

166. Balch, C.M., S.J. Soong, and T.M. Murad, *A multifactorial analysis of melanoma. II. Prognostic factors in patients with stage I (localized) melanoma.* Surgery, 1979. **86**: p. 343-351.

167. Cascinelli, N., A. Morabito, and R. Bufalino, *Prognosis of stage I melanoma of the skin.* Int J Cancer, 1980. **26**: p. 733-739.

168. Sondergaard, K. and G. Schou, *Survival with primary cutaneous malignant melanoma, evaluated from 2012 cases. A multivariate regression analysis.* Virchows Arch A Pathol Anat Histopathol, 1985. **406**: p. 179-195.

169. Rigel, D.S., R.J. Friedman, and A.W. Kopf, *Factors influencing survival in melanoma.* . Dermatol Clin, 1991. **9**: p. 631-642.

170. Stidham, K.R., J.L. Johnson, and H.F. Seigler, *Survival superiority of females with melanoma. A multivariate analysis of 6383 patients exploring the significance of gender in prognostic outcome.* Arch Surg, 1994. **129**: p. 316-324.

171. Garbe, C., P. Buttner, and J. Bertz, *Primary cutaneous melanoma. Identification of prognostic groups and estimation of individual prognosis for 5093 patients.* Cancer, 1995. **75**: p. 2484-2491.

172. Retsas, S., K. Henry, and M.Q. Mohammed, *Prognostic factors of cutaneous melanoma and a new staging system proposed by the American Joint Committee on Cancer (AJCC): validation in a cohort of 1284 patients.* Eur J Cancer, 2002. **38**: p. 511-516.

173. Ronan, S.G., M.C. Han, and T.K. Das Gupta, *Histologic prognostic indicators in cutaneous malignant melanoma.* Semin Oncol, 1988. **15**: p. 558-565.

174. Vollmer, R.T. and H.F. Seigler, *A model for pretest probability of lymph node metastasis from cutaneous melanoma.* Am J Clin Pathol, 2000. **114**: p. 875-879.

175. DAY, C.L., et al., *Malignant Melanoma: Prognostic Significance of "Msicroscopic Satellites" in the Reticular Dermis and Subcutaneous Fat.* MICROSCOPIC SATELLITES IN MELANOMA PRIMARIES, 1981. **194**(1): p. 108-112.

176. Day, C.L., T.J. Harrist, and F. Gorstein, *Malignant melanoma. Prognostic significance of 'microscopic satellites' in the reticular dermis and subcutaneous fat.* Ann Surg, 1981. **194**: p. 108-112.

177. Harrist, T.J., D.S. Rigel, and J.C.L. Day, *'Microscopic satellites' are more highly associated with regional lymph nodes metastases than is primary melanoma thickness.* . Cancer, 1984. **53**: p. 2183-2187.

178. Leon, P., J.M. Daly, and M. Synnestvedt, *The prognostic implications of microscopic satellites in patients with clinical stage I melanoma.* Arch Surg, 1991. **126**: p. 1461-1468.

179. Rao, U.N., J. Ibrahim, and L.E. Flaherty, *Implications of microscopic satellites of the primary and extracapsular lymph node spread in patients with high-risk melanoma: pathology corollary of Eastern Cooperative Oncology Group Trial E1690.* J Clin Oncol, 2002. **20**: p. 2053-2057.

180. Mascaro, J.M., M. Molgo, and T. Castel, *Plasma cells within the infiltrate of primary cutaneous malignant melanoma of the skin. A confirmation of its histoprognostic value.* Am J Dermatopathol, 1987. **9**: p. 497-499.

181. Piepkorn, M., *Prognostic factors in cutaneous melanoma.* Busam KJ (eds). ed. Pathology of Melanocytic Nevi and Malignant Melanoma, ed. n.e.N. York. 2004: Springer, Barnhill RL. 372-394.

182. Benvenuto-Andrade, C.D., S.W.; Agero, A.L.C.;Scope,A.; Rajadhyaksha, M.; Halpern,A.C.; Marghoob,A.A., *Differences Between Polarized Light Dermoscopy and Immersion Contact Dermoscopy for the Evaluation of Skin Lesions* Arch Dermatol., 2007. **143**: p. 329-338.

183. Argenziano, G., et al., *Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet.* Journal of the American Academy of Dermatology, 2003. **48**(5): p. 679-693.

184. Kendler, K.S., et al., *Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins.* American Journal of Psychiatry, 2003. **160**(4): p. 687-695.

185. Clark, W.H., et al., *The histogenesis and biologic behavior of primary human malignant melanomas of the skin.* Cancer research, 1969. **29**(3): p. 705-727.

186. McGovern, V., *Digital diagnosis: new tool for detecting skin cancer.* Environ Health Perspect, 2003. **111**(14): p. McGovern, V.

187. MenziesS.W, B., H.L., Talbot, A. Gutenev,T.A.,Avramidis M. ,Wong,, Lo,S.K., Mackellar,g., and V. Skladnev, McCarthy, W., *The performance of solarscan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma.* Archives of dermatology, 2005. **141** (11): p. 1388–1396.

188. Elbaum M, K.A., Rabinovitz HS, Langley RG, Kamino H, Mihm MC Jr, Sober AJ, Peck GL, Bogdan A, Gutkowicz-Krusin D, Greenebaum M, Keem S, Oliviero M, Wang S., *Automatic differentiation of melanoma from melanocytic nevi with multispectral digital dermoscopy: a feasibility study.* J Am Acad Dermatol, 2001. **44**(2): p. 207-18.

189. Gutkowicz-Krusin D, E.M., Jacobs A, Keem S, Kopf AW, Kamino H, Wang S, Rubin P, Rabinovitz H, Oliviero M, *Precision of automatic measurements of pigmented skin lesion parameters with a MelaFind(TM) multispectral digital dermoscope.* Melanoma Res, 2000. **10**(6): p. 563-70.

190. Kopf, A.W., ElbaumM.,Provost,N., *The use of dermoscopy and digital imaging in the diagnosis of cutaneous malignant melanoma.* Skin Research and Technology (impact factor: 1.71). 10/ /j.1600-0846.1997.tb00152.x, 2006. **3**(1): p. 1 - 7.

191. Rigel DS, R.J., Friedman R., *The evolution of melanoma diagnosis: 25 years beyond the ABCDs.* CA Cancer J Clin., 2010. **60**(5): p. 301-16.

192. Talbot, H., Bischof,L., *An overview of the Polartechnics SolarScan melanoma diagnosis algorithms.* 2003.

193. Philippa, H.Y., Peter, D. Baade, Monika, Janda, Christopher, B. Del Mar, et al.,, *Diagnosing skin cancer in primary care: how do mainstream general practitioners compare with primary care skin cancer clinic doctors?* Medical Journal of Australia, 2007. **187**(4): p. 215.

194. Steiner, A., H. Pehamberger, and K. Wolff, *Improvement of the diagnostic accuracy in pigmented skin lesions by epiluminescent light microscopy.* ANTICANCER RESEARCH, 1987. **7(3 Pt B)**: p. 433-434.

195. Abbasi, N.R., et al., *Early diagnosis of cutaneous melanoma: Revisiting the abcd criteria.* JAMA: The Journal of the American Medical Association, 2004. **292**(22): p. 2771-2776.

196. Robinson Jk, T.R., *SKills training to learn discrimination of abcde criteria by those at risk of developing melanoma.* Archives of Dermatology, 2006. **142**(4): p. 447-452.

197. Johr, R.H., *Dermoscopy: alternative melanocytic algorithms—the ABCD rule of dermatoscopy , menzies scoring method, and 7-point checklist.* Clinics in dermatology, 2002. **20**(3): p. 240-247.

198. Healsmith, M.F., *An evaluation of the revised seven-point checklist for the early diagnosis of cutaneous malignant melanoma.* British Journal of Dermatology, 1994. **130**(1): p. 48-50.

199. Zalaudek, I.A.G., et al., *The Dermoscopy Working, Group, Three-point checklist of dermoscopy: an open internet study.* British Journal of Dermatology, 2006. **154**(3): p. 431-437.

200. Vestergaard, M.E.M. and S. W., *Automated Diagnostic Instruments for Cutaneous Melanoma.* Seminars in Cutaneous Medicine and Surgery, 2008. **27**(1): p. 32-36.

201. Menzies SW, I.C., C. K., and M. W.H., *Frequency and morphologic characteristics of invasive melanomas lacking specific surface microscopic features.* Archives of Dermatology, 1996. **132**: p. 1178-1182.

202. Carli, P., *Reliability and inter-observer agreement of dermoscopic diagnosis of melanoma and melanocytic naevi.* European Journal of Cancer Prevention, 1998. **7**(5): p. 397-402.

203. Stanganelli, I., *Intraobserver agreement in interpretation of digital epiluminescence microscopy.* Journal of the American Academy of Dermatology, 1995. **33**(4): p. 584-589.

204. Cascinelli, N., *A possible new tool for clinical diagnosis of melanoma: The computer.* Journal of the American Academy of Dermatology, 1987. **16**(2, Part 1): p. 361-367.

205. Hall, P.N., E. Claridge, and D.M. Smith, *Computer screening for early detection of melanoma—is there a future?* . British Journal of Dermatology, 1995. **132**(3): p. 325-338.

206. Cristofolini, M., et al., Diagnosis of cutaneous melanoma: accuracy of a computerized image analysis system (Skin View). Skin Research and Technology, 1997. **3**(1): p. 23-27.

207. Rajpara, S.M., *Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma.* British Journal of Dermatology, 2009. **161**(3): p. 591-604.

208. Rosado B, M.S., *Harbauer A et al., Accuracy of computer diagnosis of melanoma: a quantitative meta-analysis.* Archives of Dermatology, 2003. **139**: p. 361-367.

209. Celebi, M.E., Y.A. Aslandogan, and P.R. Bergstresser, *Unsupervised border detection of skin lesion images.* ITCC International Conference onInformation Technology: Coding and Computing, 2005, 2005.

210. Sookpotharom, S. *Border Detection of Skin Lesion Images Based on Fuzzy C-Means Thresholding*. in *3rd International Conference on Genetic and Evolutionary Computing, 2009. WGEC '09*. 2009.

211. Silveira, M. and Marques.J.S. *Level set segmentation of dermoscopy images*. in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2008*. 2008.

212. Celebi, M., Aslandogan, YA, Stoecker, WV, Iyatomi, H, Oka, H, Chen, X, *Unsupervised border detection in dermoscopy images.* Skin Research and Technology, 2007. **13**(4): p. 454-462.

213. Maglogiannis, I., et al., *Intelligent Segmentation and Classification of Pigmented Skin Lesions in Dermatological Images*, in *Advances in Artificial Intelligence,*, S. Berlin, Editor. 2006: Heidelberg. p. 214-223.

214. Bulent Erkol, R.H.M., et al., *Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes.* Skin Research and Technology, 2005. **11**: p. 17-26.

215. Celebi, M.E., H.A. Kingravi, and Y.A. Aslandogan, *Nonlinear vector filtering for impulsive noise removal from color images.* Journal of Electronic Imaging, 2007. **16**(3): p. 033008-033008.

216. Celebi, M.E., G.S. Hitoshi Iyatomi, and W.V. Stoecker, *Lesion Border Detection in Dermoscopy Images.* Comput Med Imaging Graph, 2009. **33**(2): p. 148-153.

217. Abbas, Q., *Lesion border detection in dermoscopy images using dynamic programming.* Skin Research and Technology, 2011. **17**(1): p. 91-100.

218. Gomez, D.D., *Independent Histogram Pursuit for Segmentation of Skin Lesions.* IEEE Transactions on Biomedical Engineering, 2008. **55**(1): p. 157-161.

219. Adelmann, H.G., *Butterworth equations for homomorphic filtering of images.* Computers in Biology and Medicine, 1998. **28**(2): p. 169-181.

220. Celebi, M.H.K. and Lee.J., *Fast and accurate border detection in dermoscopy images using statistical region merging.* Proc. SPIE Medical Imaging., 2007.

221. Celebi, M.E., et al., *Border detection in dermoscopy images using statistical region merging.* Skin Research and Technology Journal, 2008. **14**: p. 1-7.

222. Sweet, M.R., *adaptive and recursive median filtering*.

223. Butler, J.K., *Methods for improved light microscope microtomy.* Biotechnic & Histochemistry, 1979. **54**(2): p. 53-69.

224. QinzLi, *Dark line detection with line width extraction*, in *15th IEEE International Conference on Image Processing, 2008. ICIP 2008*. 2008.

225. Criminisi, A., P. Perez, and K. Toyama, *Region filling and object removal by exemplar-based image inpainting.* IEEE Transactions on Image Processing, 2004. **13**(9): p. 1200-1212.

226. *Filters*.

227. Press, W.H., *Wiener Filtering (and some Wavelets)*, in *Computational Statistics with Application to Bioinformatics*, T.U.o.T.a. Austin, Editor. 2008: Texas at Austin. p. 1-26.

228. Lim, J.S., *Two-Dimensional Signal and Image Processing*, in *Two-Dimensional Signal and Image Processing*. 1990, Englewood Cliffs, NJ, Prentice Hall. p. p. 548, equations 9.26, 9.27, and 9.29.

229. J.J. Kulikowski, J.J., S. Mar6elja, and P.O. Bishop, *Theory of Spatial Position and Spatial Frequency Relations in the Receptive Fields of Simple Ceils in the Visual Cortex.* Biol. Cybern, 1982. **43**: p. 187-198.

230. Jain, A. and S. Bhattacharjee, *Address block location on envelopes using gabor filters.* 1992. **25**(12).

231. Jain, A., N. Ratha, and S. Lakshmanan, *Object detection using gabor filters.* Pattern Recognition, 1997. **30**: p. 295-309.

232. Utah, E. *Gabor Filters*. 2002; Available from: www.cs.utah.edu/~arul/report/node13.html.

233. Yazdi, H.S. and F. Homayouni, *Impulsive Noise Suppression of Images Using Adaptive Median Filter* International Journal of Signal Processing, Image Processing and Pattern Recognition, 2010. **3**(3): p. 1-12.

234. Arakawa, K., *Median filter based on fuzzy rules and its application to image restoration.* Fuzzy Sets and Systems 1996. **77**: p. 3-13.

235. Lin, T.-C. and P.-T. Yu, *Partition fuzzy median "lter based on fuzzy rules for image restoration.* Fuzzy Sets and Systems 2004. **147**: p. 75-97.

236. Dhawan, A.P., *Medical Image Analysis*, ed. I. press. Vol. ch. 8. 2003, United States of America: Wiley-Interscience.

237. Gonzalez, R.C. and R.E. Woods, *"Digital Image Processing", 2nd ed.* Upper Saddle River, 2002.

238. Jain, A.K., *Fundamentals of Digital Image Processing*. Englewood Cliffs, 1989.

239. Russ, J.C., *The Image Processing Handbook, 2nd Edition*. Boca Raton. 1995: CRCPress.

240. Jain, R., R. Kasturi, and B.G. Schunck, *Machine Vision*. 1995, New York: McGraw-Hill.

241. Rosenfeld, A. and A.V. Kak, *Digital Picture Processing, 2nd ed.* Vol. 1 & 2. 1982, Orlando: Academic Press.

242. Czerwinski, R.N., D.L. Jones, and W.D. O'Brien Jr. *Ultrasound speckle reduction by directional median filtering*. in *Image Processing, 1995. Proceedings., International Conference on*. 1995. IEEE.

243. Otsu, N., *A Tlreshold Selection Method from Gray-Level Histograms.* IEEE T. Syst. Man Cy, 1979. **9**: p. 62-66.

244. Dawant, B.M., et al., *Image segmentation, in Handbook of Medical Imaging,*. Vol. 2: Medical Image Processing and Analysis. 2000, Paress, Bellingha, 2000: SPIE.

245. Dhawan, A.P. and L. Arata, *Segmentation of medical images through competitive learning* Comp, Methods and Prog. in Biomed, 1993. **40**: p. 203-215.

246. Cline, H.E., et al., *Three-dimentional segmentation of MR images of the head using probabilityand connectivity.* Comp. Assis. Tomog., 1990. **14**: p. 1037-1045.

247. Zavaljevski, A., et al., *Multispectral MR brain image classification.* Comp. Med. Image. Graphics Image Proc., 2000. **24**: p. 87-98.

248. Reibman, A.R. and L. Nolte, *Optimal detection and performance of distributed sensor systems.* Aerospace and Electronic Systems, IEEE Transactions on, 1987(1): p. 24-30.

249. Hall, L.O., et al., *A comparison of neural network and fuzzy clustering techniques in segmentation magnetic resonance images of the brain.* IEEE Trans. Neural Networks, 1992. **3**: p. 672-682.

250. Loncaric, S., et al., *3-D image analysis of intracerebral brain hemorrhage.* Comp. Methods Prog. Biomed, 1995. **46**: p. 207-216.

251. Clarke, L., et al., *MRI: Stability of three supervised segmentation techniques.* Mag. Res. Image, 1993. **11**: p. 95-106.

252. Beni Xuanli, G., *A validity measure for fuzzy clustring.* IEEE Trans Patt. Analy. Machine Intell., 1991. **133**(8).

253. Tarca, A.L., et al., *Machine learning and its applications to biology.* PLoS computational biology, 2007. **3**(6): p. e116.

254. Yao, X., *Evolving artificial neural networks.* Proceedings of the IEEE, 1999. **87**(9): p. 1423-1447.

255. Marchiori, S.C., et al., *Neural network based on adaptive resonance theory with continuous training for multi-configuration transient stability analysis of electric power systems.* Applied Soft Computing, 2011. **11**(1): p. 706-715.

256. Fukushima, K., *Neocognitron: A hierarchical neural network capable of visual pattern recognition.* Neural networks, 1988. **1**(2): p. 119-130.

257. Page, I. *Intelligent Image Features Extraction in Knowledge Discovery Systems*. 2007-2015; Available from: www.imagefeatures.org/research/.

258. Selvarajah, S. and S.R. Kodituwakku, *Analysis and Comparison of Texture Features for Content Based Image Retrieval.* International Journal of Latest Trends in Computing (E-ISSN: 2045-5364) 108 2011. **2**(1): p. 108-113.

259. Manian, V. *Image feature extraction and segmentation*. Pattern Recognition; Available from: www.ece.uprm.edu/~manian/ImageAnalysis4.

260. *Intelligent Image Features Extraction in Knowledge Discovery Systems*. Available from: http://www.imagefeatures.org/research/.

261. Haralick, R.M., K. Shanmugan, and I.H. Dinstein, *Textural features for image classification.* IEEE Trans. Syst., 1973. **SMC-3**: p. 610–621.

262. Selvarajah, S. and S.R. Kodituwakku, *Analysis and Comparison of Texture Features for Content Based Image Retrieval.* International Journal of Latest Trends in Computing, 2011. **2**(1): p. 108-113.

263. Li, B. and M.Q.H. Meng, *Texture analysis for ulcer detection in capsule endoscopy images.* Image and Vision Computing, 2009. **27**(9): p. 1336–1342

264. Chui, C.K., *An Introduction to Wavelets*, in *Academic Press*. 1992: San Diego, Calif, USA.

265. Guyon, I. and A. Elisseeff, *An introduction to feature extraction*, in *Feature Extraction*. 2006, Springer. p. 1-25.

266. Lyons, R.G., *Understanding Digital Siginal Processing*. 2010, Upper Saddle River: Prentice Hall. 984.

267. S., J. and Walker, *A Primer on Wavelets and Their Scientific Applications.* Paperback, 1999.

268. Gonzalez, R.C., *image processing.* 1992.

269. Soille, *morphological image analysis.* 2004.

270. Kruskal and Wish, *Multidimensional Scaling (MDS).* 1978.

271. Liu and H. Motoda, *Feature Extraction, Construction and Selection: A Data Mining Perspective.* Kluwer Academic, 1998.

272. Abe, S., *Support Vector Machines for Pattern Classification.* Springer, New York, 2005.

273. Kohavi, R. and G.H. John, *Wrappers for feature subset selection.* Artif. Intell. Med, 1997. **97**: p. 273–324.

274. D. Wettschereck, D., D.W. Aha, and T. Mohri, *A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms.* Artif. Intell. Med, 1997. **11**: p. 273–314.

275. Bluma, A.L. and P. Langley, *Selection of relevant features and examples in machine learning.* Amficial Intelligence, 1997. **97**: p. 245-271.

276. Kohavi, R., John, G.H., *Wrappers for feature subset selection.* Artificial Intelligence 97, 1997: p. 273-324.

277. G.H. John, G.H., R. Kohavi, and K. Pfleger. *Irrelevant features and the subset selection problem*. in *Proceedings of the International Conference on Machine Learning*. 1994. New Brunswick.

278. Neter, J., M.H. Kutner, and W. Wasserman, *Applied Linear Statistical Models.* C.J. Nachtsheim, 1996.

279. Ginsberg, M.L., *Essentials of Artificial Intelligence.* Morgan Kaufmann, Palo Alto, 1993.

280. I.M. Guyon, I.M., et al., *Gene selection for cancer classification using support vector machines.* Mach. Learn, 2002. **46**: p. 389–422.

281. F. Ojeda, F., J.A.K. Suykens, and B. De Moor, *Low rank updated LS-SVM classifiers for fast variable selection.* Neural Netw., 2008. **21**(2–3): p. 437–449.

282. Weston, S., et al. *Feature selection for SVMs*. in *Proceedings of the Conference on Neural Information Processing Systems*. 2000. Denver.

283. chapelle, O., et al., *Choosing Multiple Parameters for Support Vector Machines.* Machine Learning, , , , 2002. **46**: p. 131–159.

284. Rasmussen, C.E. and C.K.I. Williams, *Gaussian Processes for Machine Learning.* MIT Press Cambridge, 2006.

285. Cawley, G.C. *Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs*. in *Proceedings of the International Joint Conference*. 2006. Neural Networks, Vancouver.

286. Suykens, J.A.K., et al., *Least Squares Support Vector Machines.* World Scientific, Singapore,, 2002.

287. Duda, R.O., P.E. Hart, and D.G. Stork, *Pattern classification*. 2012: John Wiley & Sons.

288. Ooi, C.H., M. Chetty, and S.W. Teng, *Differential prioritization in feature selection and classifier aggregation for multiclass microarray datasets.* Data Mining and Knowledge Discovery, 2007. **14**(3): p. 329-366.

289. Rückstieß, T., C. Osendorfer, and P. van der Smagt, *Minimizing data consumption with sequential online feature selection.* International Journal of Machine Learning and Cybernetics, 2013. **4**(3): p. 235-243.

290. Gunn, S.R., *Support vector machines for classification and regression.* ISIS technical report, 1998. **14**.

291. Duin, R.P., *Classifiers in almost empty spaces*, in *In Pattern Recognition*, P.t.I. Conference, Editor. 2000, IEEE. p. 1-7.

292. K. Hoffmann, K., et al., *Diagnostic and neural analysis of skin cancer (DANAOS). A multi-centre study for collection and computeraided analysis of data from pigmented skin lesions using digital dermoscopy.* Br. J. Dermatol., 2003. **149**: p. 801–809.

293. BATUWITA, R. and V. PALADE, *FSVM-CIL: fuzzy support vector machines for class imbalance learning.* IEEE Transactions on Fuzzy Systems, 2010. **18**: p. 558-571.

294. Platt, J.C. *Fast training of support vector machines using sequential minimal optimization*. In Advances in kernel methods 1999 1999, February]; 185-208].

295. KEERTHI, S.S., et al., *Improvements to Platt's SMO algorithm for SVM classifier design.* Neural Computation, 2001. **13**: p. 637-649.

296. FAN, R.E., P.H. CHEN, and C.J. LIN, *Working set selection using second order information for training support vector machines.* The Journal of Machine Learning Research, 2005. **6**: p. 1889-1918.

297. Chiang, M., *Convex Optimization and Lagrange Duality*. 2007.

298. Smola, A.J., *Regression estimation with support vector learning machines*. 1996, Technische Universit: M¨unchen.

299. Vapnik, V., et al., *Support vector method for function approximation, regression estimation, and signal processing.* Advances in Neural Information Processing Systems, 1997. **9**: p. 281– 287.

300. Osuna, E., et al., *An improved training algorithm for support vector machines.* J. Principe 1997: p. 276 – 285.

301. DEL VALLE, Y., et al., *Particle swarm optimization: basic concepts, variants and applications in power systems.* IEEE Transactions on Evolutionary Computation, 2008. **12**: p. 171-195.

302. ASTION, M.L., et al., *Overtraining in neural networks that interpret clinical data.* Clinical Chemistry, 1993. **39**: p. 1998-2004.

303. SIDDIQUI, A.B., et al., *Block- Based Pixel Level Multi-Focus Image Fusion Using Particle Swarm Optimization.* International Journal of Innovative Computing, Information and Control, 2011. **7**: p. 3583 - 3596.

304. Cand`es, E.J. and D.L. Donoho, *New tight frames of curvelets and optimal representations of objects with piecewise-C 2 singularities.* Comm. on Pure and Appl. Math, 2002. **57**: p. 219-26.

305. Australian, *C.C. Skin Cancer Facts and Figures*. 2010; Available from: <http://www.cancer.org.au/cancersmartlifestyle/SunSmart/Skincancerfactsandfigures.htm>.

306. Welfare, A.c.n.H.A.I.o.H.a. *Health system expenditures on cancer and other neoplasms in Australia, 2000–01*. 2005; Available from: (http://www.aihw.gov.au).

307. Australian Institute of Health and Welfare, C., *Cancer in AUSTRALIA: an overview 2010*, in *AIHW cat no. CAN 56*, s.n. 60, Editor. 2010: Canberra.

308. Sikorski, J. *Identification of malignant melanoma by wavelet analysis*. in *Student/Faculty Research Day*. 2004. CSIS, Pace University.

309. Zhang, Z., Stoecker,W., Moss, R. , *Border detection on digitized skin tumor images.* IEEE Trans. Med. Imaging 2000. **19**(11): p. 1128-1143.

310. Scholkopf, B., Tsuda, K., Vert, J.P., *kernel Methods in computational biology.* MIT Press, 2004.

311. Dolianitis, C., et al., *Comparative performance of 4 dermoscopic algorithms by nonexperts for the diagnosis of melanocytic lesions.* Archives of dermatology, 2005. **141**(8): p. 1008-1014.

312. Kass, M., Witkin,A.,Terzopolous, D., *Snakes: active contour models.* Int. J. Comput, 1987. **1**(4): p. 321-331.

313. Xu, C. and *J.L.* Prince, *Snakes, Shapes, and Gradient Vector Flow.* IEEE TRANSACTIONS ON IMAGE PROCESSING, 1998. **7**(3): p. 359.

314. Tang, J., et al., *Surface extraction and thickness measurement of the articular cartilage from MR images using directional gradient vector flow snake.* IEEE Trans. Biomed. Eng, 2006. **52**(5): p. 896-907.

315. Tang, J., Acton,S. , *Vessel boundary tracking for intravital microscopy via multi-scale gradient vector flow snakes.* IEEE Trans. Biomed. Eng., 2004. **51**(2): p. 316-324.

316. Arakawa, K., *Median filter based on fuzzy rules and its application to image restoration.* Fuzzy Sets and Systems, 1996. **77**: p. 3-13.

317. Kass, M., Witkin,A.,Terzopolous, D., *Snakes: active contour models.* Int. J. Comput, 1988. **1**(4): p. 321-331.

318. Caselles, V., et al., *A geometric model for active contours.* Numerische Mathematik, 1993. **66**: p. 1-31.

319. L.D., C., *On active contour models and balloons.* CVGIP: Image Understanding, 1991. **53**(2): p. 211–218.

320. Xu, C. and J.L. Prince, *Gradient Vector Flow: A New External Force for Snakes.* IEEE Proc. Conf. on Comp. Vis. Patt. Recog. (CVPR'97), 1997: p. 66-71.

321. Tang, J., *A multi-directionGVFsnakeforthesegmentationofskincancerimages.* Pattern Recognition Lett, 2009. **42**: p. 1172 -- 1179.

322. Nixon, M. and A. Aguado, *Feature Extraction and Image Processing*. 2002, London, UK: Academic Press is an imprint of Elsevier.

323. Levinson, N., *The Wiener RMS (root mean square) error criterion in filter design and prediction.* 1947.

324. Lau, H.T. and A. Al-Jumaily, *Automatically Early Detection of Skin Cancer: Study Based on Nueral Netwok Classification*, in *International Conference of Soft Computing and Pattern Recognition*. 2009.

325. Sumana, I.J., *Image Retrieval Using Discrete Curvelet Transform*, in *Gippsland School of Information Technolog*. 2008, Monash University: Australia. p. 84.

326. Takruri, M., A. Al-Jumaily, and M.K. Abu Mahmoud, *Automatic Recognition of Melanoma Using Support Vector Machines: A Study Based on Wavelet, Curvelet and Color Features*, in *2014 International Conference on Industrial Automation, Information and Communications Technology (IAICT 2014)*. 2014, IEEE Indonesia Communications Society Chapter and RAS/CSS (Robotics & Automations Society and Control System Society) Joint Chapter: Grand Inna Kuta Bali Indonesia.

327. Robert-Inacio, F., Dinet, E., *An adaptive median filter for colour image processing.* Proc. 3rd CGIV, 2006: p. 205-210.

328. Dinet, E. and F. Robert-Inacio, *Color Median Filtering: a Spatially Adaptive Filter.* Proceedings of Image and Vision Computing, 2007: p. 71-76.

329. Teknomo, K., *K-Means Clustering Tutorial*. 2007. p. What is K-Means Clustering?

Simply speaking it is an algorithm to classify or to group your objects based on attributes/features into K number of group. K is positive integer number. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid. Thus, the purpose of K-mean clustering is to classify the data.

330. Van De Wouwer, G., Weyn, B., Scheunders, P., Jacob, W., Van Marck, E., Van Dyck, D.. *Wavelets as chromatin texture descriptors for the automated identification of neoplastic nuclei.* PubMed - indexed for MEDLINE, 2000. **197**(Pt1): p. 25-35.

331. Ali, S., Smith-Miles, K. A., *Improved Support Vector Machine Generalization Using Normalized Input Space.* © Springer-Verlag Berlin Heidelberg 2006, 2006: p. 362-371.

332. Annessi, G., et al., *Sensitivity, specificity, and diagnostic accuracy of three dermoscopic algorithmic methods in the diagnosis of doubtful melanocytic lesions: the importance of light brown structureless areas in differentiating atypical melanocytic nevi from thin melanom*

J Am Acad Dermatol, 2007. **56**: p. 759-767.

333. Henning, J.S., S.W. Dusza, and S.Q. Wang, *The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy.* J AmAcad Dermatol, 2007. **56**: p. 45-52.

334. Dolianitis, C., et al., *Comparative performance of 4 dermoscopic algorithms by nonexperts for the diagnosis of melanocytic lesions.* Arch Dermatol, 2005. **141**: p. 1008-1014.

335. Menzies SW1, et al., *The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma.* Arch Dermatol, 2005. **141**(11): p. 1388-96.

336. Burroni, M., P. Sbano, and G. Cevenini, *Dysplastic nevus vs. in situ melanoma: digital dermoscopy analysis.* J Dermatol, 2005. **152**: p. 679-684.

337. Argenziano, G., et al., *Dermoscopy of pigmented skin lesions: results of a consensus meeting via the Internet.* J Am Acad Dermatol, 2003. **48**: p. 679-693.

338. Menzies, S.W., L. Bischof, and H. Talbot, *The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma.* Arch Dermatol 2005. **141**: p. 1388-1396.

339. Kecman, V., *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models (Complex Adaptive Systems).* 2001, Cambridge, MA, 2001: The MIT Press.

340. Wang, L.P. and X.J. Fu, *Data Mining with Computational Intelligence*. 2006, Berlin: Springer. 288.

341. Sloin, A. and D. Burshtein, *Support Vector Machine Training for Improved Hidden Markov Modeling.* Signal Processing, IEEE Transactions, 2008. **56**(1): p. 172-188.

342. Chu, F. and L.P. Wang, *Applications of support vector machines to cancer classification with microarray data.* International Journal of Neural Systems, 2005. **15**(6): p. 475-484.

343. del Valle, Y., et al., *Particle Swarm Optimization: Basic Concepts, Variants and Applications in Power Systems.* IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, 2008. **12**(2): p. 171-195.

344. Hsu, C.-W., C.-C. Chang, and C.-J. Lin, *A practical guide to support vector classification*. 2003.

345. Wallace, V., et al., *Classification of reflectance spectra from pigmented skin lesions, a comparison of multivariate discriminant analysis and artificial neural networks.* Physics in medicine and biology, 2000. **45**(10): p. 2859.

346. GEORGOULAS, G. and C.D. STYLIOS, *Predicting the risk of metabolic acidosis for newborns based on fetal heart rate signal classification using support vector machines.* IEEE Transactions on Biomedical Engineering, 2006. **53**: p. 875-884.

347. DEEB, O. and M. GOODARZI, *Exploring QSARs for Inhibitory Activity of Non-peptide HIV-1 Protease Inhibitors by GA-PLS and GA-SVM.* Chemical Biology and Drug Design, 2010. **75**: p. 506-514.

348. Lin, S.-W., et al., *Particle swarm optimization for parameter determination and feature selection of support vector machines.* Expert Systems with Applications, 2008. **35**(4): p. 1817-1824.

349. Pawley, J., *Handbook of biological confocal microscopy*. 2010: Springer.

350. Infrared spectroscopy from Wikipedia, t.f.e.; Available from: http://en.wikipedia.org/wiki/Infrared_spectroscopy.

351. Tyne, U.o.N.u. *On-line medical dictionary*. Available from: http://cancerweb.ncl.ac.uk/omd.

352. Mehta, A., *Limitations and Deviations of Beer–Lambert Law.*

353. Brey, E.M., et al., *Automated Selection of DAB-labeled Tissue for Immunohistochemical Quantification.* The Journal of Histochemistry & Cytochemistry, 2003. **51**(5): p. 575–584.

354. Ruifrok, A.C., Johnston, D.A., *Quantification of histochemical staining by color deconvolution.* Analytical and quantitative cytology and histology, 2001. **23**(4): p. 291-299.

355. Beer-Lambert. *law from Wikipedia*. Available from: http://en.wikipedia.org/wiki/Beer-Lambert_law.

356. Bipm, I., et al., *The international vocabulary of metrology—basic and general concepts and associated terms (VIM).* JCGM, 2008. **200**.

357. Meyer, C.D., *Matrix Analysis and Applied Linear Algebra.* SIAM, 2000.

358. Falconer and Kenneth, *Fractal Geometry*. 2003, New York: Wiley.

359. Sagan and Hans, *Space-Filling Curves*. 1994, Berlin: Springer-Verlag.

360. Vicsek and Tamás, *Fractal growth phenomena*. 1992, Singapore New Jersey: World Scientific.

361. Daugman, J.G., *Two-dimensional spectral analysis of cortical receptive field profiles.* Vision Res, 1980. **20**(10): p. 847–56.

362. Trudeau and Richard, *Introduction to Graph Theory (Corrected, enlarged republication. ed.)*, in *New York: Dover Pub*. 1993, Retrieved 8 August 2012. p. 19.

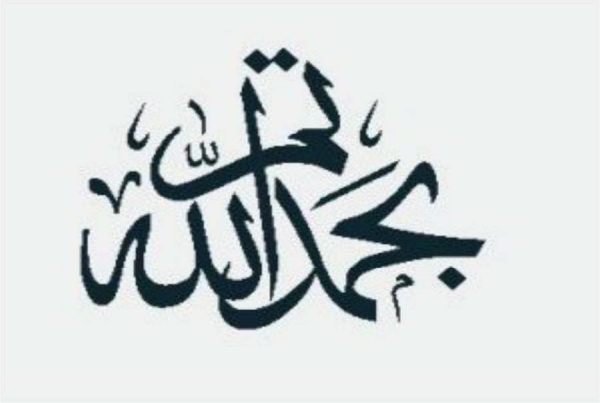
363. Jolliffe, I.T., *Principal Component Analysis, second edition*. 2002: Springer.

364. Swet and A. John, *Signal detection theory and ROC analysis in psychology and diagnostics : collected papers.* NJ, 1996.

365. *Vision Systems Design.* 1999 **4**(2): p. 21.

366. Beucher, S. and Meyer.F., *The morphological approach to segmentation: the watershed transformation*. In Mathematical Morphology in Image Processing (Ed. E. R. Dougherty). 1993. 433–481.

367. Platt, J.C., *Fast Training of Support Vector Machines using Sequential Minimal Optimization*. 2000, Redmond, WA 98052, USA: Microsoft Way.



END