

Computer-Aided Image Analysis for Digitized Skin Histopathology

by

Hongming Xu

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Signal and Image Processing

Department of Electrical and Computer Engineering
University of Alberta

© Hongming Xu, 2017

Abstract

Histological examination of biopsy slides is the gold standard for the diagnosis of skin cancer such as cutaneous melanoma. However, visual evaluations performed by pathologists are prone to inter- and intra-observer variations. In addition, it is very labor-intensive to manually analyze a whole biopsy slide due to the large amount of data involved. With recent advances in digital scanners and computational powers, automatic biopsy image analysis has been desired to assist pathologists in their clinical diagnosis. In this dissertation, several computerized algorithms that can aid towards building a computer-aided diagnosis (CAD) system for digitized skin biopsy slides are developed.

The main contributions of this dissertation are five fold. First, automated techniques are proposed to segment skin epidermis and dermis regions in skin whole slide images (WSIs). The proposed techniques segment the skin epidermis using a coarse-to-fine procedure based on thickness measurement and k-means clustering, and segment skin dermis based on a predefined depth of interest value. Second, two automated techniques are proposed to detect nuclei seeds in skin histopathological images. The first technique detects nuclei seeds based on ellipse descriptor analysis and an improved voting algorithm, while the second technique detects nuclei seeds by using generalized Laplacian of Gaussian (gLoG) kernels. Third, an automated technique is proposed to delineate nuclei boundaries in skin histopathological images. The technique segments nuclei boundaries by a multi-scale radial line scanning (mRLS) method, which incorporates image gradient, intensity variance and shape prior together for nuclei boundary determination. Fourth, a comput-

erized technique for melanocytic tumor classification in skin WSIs is developed. The technique analyzes both epidermis and dermis areas, and performs skin tissue classification by using a multi-class support vector machine (mSVM) with a set of cytological and textural features. These four set of techniques have been developed primarily for H&E stained images, which are widely used for clinical diagnosis. The nuclei detection and segmentation techniques work for both H&E and Ki-67 stained skin images. Finally, an automated technique is proposed for measuring the melanoma depth of invasion (DoI) in MART-1 stained skin histopathological images. Note that MART-1 is a melanoma specific stain typically used for melanoma grading. The proposed technique identifies the skin granular layer based on a Bayesian classifier, and computes the melanoma DoI using a multi-resolution framework with the Hausdorff distance measure.

Evaluations of proposed techniques have been thoroughly performed on a set of skin biopsy images provided by pathologists, which mainly includes 66 H&E stained skin WSIs, 40 Ki-67 stained and 30 MART-1 stained skin histopathological images. Experimental results indicate that the proposed techniques can provide superior performances compared to closely related methods in the literature. Due to the promising performance and relatively low complexity, the proposed techniques have the potential to be used for assisting pathologists in skin biopsy image analysis and disease diagnosis.

This thesis is dedicated to

my parents, Zipo Xu and Yulian Tong,

my wife, Miao Gao and my daughter, Angela, Yanxi, Xu,

for their endless love, support and encouragement.

Acknowledgments

I would like to start by thanking my supervisor, Prof. Mrinal Mandal, for his guidance and invaluable support throughout my PhD studies at University of Alberta (UofA). He is a knowledgeable, kind and considerate supervisor. He taught me not only in research but also in teaching, presentation and many other aspects. All my achievements have been benefited from his knowledge, inspiration and encouragement. These years of studying under his supervision will be memorable in my life.

I also would like to express my sincere gratitude to Dr. Naresh Jha, Dr. Richard Berendt and Dr. Cheng Lu, who have been my research collaborators in the past few years. They have been helping me and giving me valuable suggestions during our collaborative research. I learned and gained enormously from their knowledge and experience. My thesis would not have been done like this without their help.

I thank my committee members, Dr. Roger Zemp, Dr. Jie Chen, Dr. Vicky Zhao and Dr. Bin Han, for taking their precious time to review my candidacy report and dissertation, as well as for their valuable suggestions and guidance during my candidacy exam and dissertation defense.

Thanks also go to China Scholarship Council (CSC) and my alma mater Northwest A&F University (especially my previous supervisor Prof. Dongjian He), for awarding me the scholarship to support my study at UofA. Valuable discussion and help have come from my labmates, classmates, and friends, who are Dr. Gencheng Guo, Dr. Tao Xu, Mrs. Yue Li, Mr. Huiquan Wang, Mrs. Bingjie Ma, Mr. Salah Al-Heejawi, Mrs. Ying Wang and many others. They have made my days at UofA

more meaningful and happy. I feel so lucky to work and study with them together.

Last but far from the least, my heartiest gratitude goes to my beloved parents, Zipo Xu and Yulian Tong, my wife, Miao Gao, and my daughter, Angela Xu, for their endless love and tremendous devotion throughout my life.

Table of Contents

1	Introduction	1
1.1	Skin Cancer and Diagnosis	1
1.2	Histopathology Imaging	2
1.3	Histological Diagnosis and Melanoma Grading	4
1.3.1	Skin Structure	4
1.3.2	Histologic Diagnosis of Skin Melanoma	7
1.3.3	Histologic Grading of Skin Melanoma	8
1.4	Problem Statements and Motivation	8
1.5	Major Contribution of the Thesis	11
1.6	Organization of the Thesis	12
2	Literature Review	14
2.1	ROI Detection and Segmentation	14
2.1.1	Thresholding	14
2.1.2	Watershed	15
2.1.3	Filtering	16
2.1.4	Active Contour Model (ACM)	17
2.1.5	Clustering	18
2.1.6	Probabilistic Models	19
2.1.7	Comparison of Segmentation Techniques	19
2.2	Features for Histology Images	21
2.2.1	Textural Features	21
2.2.2	Morphological Features	22

2.2.3	Architectural Features	23
2.2.4	Features Selection	25
2.3	Classification	26
2.4	General WSIs Analysis and Classifications	27
2.5	Summary	30
3	Automated Epidermis and Dermis Segmentations	31
3.1	Introduction	32
3.2	Image Dataset	34
3.3	Proposed Epidermis Segmentation	34
3.3.1	Coarse Segmentation	35
3.3.2	Thickness Measurement	37
3.3.3	Fine Segmentation	43
3.4	Epidermis Segmentation Evaluations	44
3.4.1	Evaluation Metrics	45
3.4.2	Parameters Selection	46
3.4.3	Quantitative Results	48
3.4.4	Qualitative Results	50
3.5	Proposed Dermis Segmentation	51
3.5.1	Epidermis Segmentation & Outer Boundary Selection	53
3.5.2	Line Segments Generation	54
3.5.3	Morphological Processing	56
3.6	Dermis Segmentation Evaluations	57
3.7	Computational Complexity	58
3.8	Summary	59
4	Automated Detection of Cell Nuclei	60
4.1	Introduction	60
4.2	Proposed Voting-based Technique	63
4.2.1	Initial Segmentation	64
4.2.2	Ellipse Descriptor Analysis	65

4.2.3	Voting Detection	68
4.3	Proposed gLoG-based Technique	71
4.3.1	gLoG Kernels Construction	73
4.3.2	Response Maps Generation	77
4.3.3	Nuclei Seeds Detection	79
4.4	Performance Evaluation	83
4.4.1	Dataset Description	83
4.4.2	Parameter Configuration and Comparisons	84
4.4.3	Nuclei Detection Results	85
4.4.4	Parameter Sensitivity Evaluations	91
4.4.5	Computational Complexity and Limitations	93
4.5	Summary	94
5	Automated Segmentation of Cell Nuclei	95
5.1	Introduction	95
5.2	Proposed mRLS-based Technique	98
5.2.1	Seeds Detection	98
5.2.2	Nuclei Segmentation	99
5.3	Performance Evaluations	109
5.3.1	Datasets Description	109
5.3.2	Quantitative Evaluations	110
5.3.3	Qualitative Evaluations	115
5.3.4	Sensitivity Evaluation & Computational Complexity	117
5.4	Summary	119
6	Analysis and Classification of Skin WSI	121
6.1	Introduction	121
6.2	Proposed Technique	124
6.2.1	Epidermis & Dermis Segmentations	124
6.2.2	Epidermis Analysis	128
6.2.3	Dermis Analysis	133

6.2.4	Classification	138
6.3	Performance Evaluations	140
6.3.1	Image Dataset	140
6.3.2	Experimental Design	140
6.3.3	Classification Results	143
6.3.4	Experiments with Feature Selection	144
6.3.5	Comparisons	147
6.3.6	Computational Complexity	149
6.4	Summary	149
7	Automatic Measurement of Melanoma Depth of Invasion	151
7.1	Introduction	151
7.2	Dataset Description	154
7.3	Proposed Technique	154
7.3.1	Melanoma Detection	155
7.3.2	Epidermis Segmentation	157
7.3.3	Granular Layer Identification	159
7.3.4	Depth of Invasion Measurement	166
7.4	Performance Evaluation	167
7.4.1	Parameters Configuration	167
7.4.2	Granular Layer Detection Performance	168
7.4.3	Melanoma Depth Measurement Performance	170
7.4.4	Parameter Sensitivity Evaluation	173
7.5	Summary	174
8	Conclusions and Future Works	175
8.1	Conclusions	175
8.2	Future Works	177
	References	180

Appendix A Histology Images and Features	197
A.1 Histology images with different stains	197
A.2 Features of histology images	199
A.2.1 Textural features	199
A.3 Examples of skin WSIs	201

List of Tables

1.1	Survival rates with different melanoma depth.	8
2.1	Description of morphological features.	22
2.2	Description of architectural features.	25
2.3	Related works on various cancer diagnosis by analyzing H&E stained WSIs	28
3.1	Image dataset used for evaluating epidermis and dermis segmentations.	34
3.2	Training parameters of epidermis segmentation technique.	46
3.3	Performance evaluations of epidermis coarse segmentation in subsets of training images.	47
3.4	Quantitative evaluations of epidermis segmentation between existing techniques and proposed technique.	50
3.5	Performance evaluations of dermis segmentation.	57
4.1	Summary of nuclei detection techniques	62
4.2	Values of ellipticity parameters for several selected candidate regions shown in Fig. 4.4(a).	67
4.3	Index of gLoG kernels' configuration.	74
4.4	Index of normalized gLoG kernels with maximum responses.	77
4.5	Image dataset used for nuclei detection evaluations.	84
4.6	Comparison of nuclei detection performance ($t = 6$).	88
5.1	Cost values E_{total} for nuclei shown in Fig. 5.5.	106

5.2	Image dataset used for nuclei segmentation evaluations.	109
5.3	Comparison of nuclei segmentations.	113
5.4	Comparison of nuclei detections.	115
6.1	Skin biopsy image analysis and classification.	122
6.2	Summary of epidermis features.	132
6.3	Summary of dermis features.	139
6.4	Description of the image dataset.	141
6.5	Classification results of four test strategies.	144
6.6	Performance evaluation of melanoma diagnosis.	145
6.7	Selected epidermis and dermis features.	146
6.8	Comparisons of skin WSI classification.	147
6.9	Comparisons of skin melanoma diagnosis.	147
7.1	Bayesian information criterion (BIC) values for different models. . .	164
7.2	Summary of parameters configuration in the proposed technique. . .	168
7.3	Comparison of granular layer detection by different techniques. . .	170
7.4	Comparison of DoI measurement by different techniques.	172
7.5	Parameter sensitivity evaluations of the proposed technique.	174
A.1	Description of statistical texture features.	200
A.2	Description of Haralick features.	200

List of Figures

1.1	Examples of skin histopathological images.	4
1.2	Anatomy of normal skin.	5
1.3	An example of H&E stained skin histopathological image.	6
1.4	Breslow thickness measurement.	9
1.5	Schematic of the proposed skin melanoma diagnosis and grading system.	10
2.1	Examples of nuclei segmentations.	20
2.2	Convex hull and best fit ellipse.	23
2.3	A histological image and its graph structures	24
3.1	Skin tissue digitized slide.	32
3.2	Schematic of epidermis segmentation.	35
3.3	Two examples of epidermis coarse segmentation.	37
3.4	Illustration of morphological preprocessing.	38
3.5	Epidermis skeleton and end points.	39
3.6	Illustration of main axis identification.	40
3.7	Example of epidermis local thickness measurement.	41
3.8	Illustration of epidermis thickness measurement.	42
3.9	Illustration of k-means classification.	44
3.10	Fine segmentation result.	45
3.11	Thickness variations of coarsely segmented epidermis masks in training images.	48

3.12	Comparison of epidermis thickness based on segmentations by different techniques.	51
3.13	Comparative segmentation results on a skin WSI.	52
3.14	Schematic of dermis segmentation.	53
3.15	Example of epidermis segmentation.	53
3.16	Illustration of line segments generation.	55
3.17	Example of line segments generation.	55
3.18	Illustration of dermis mask generation.	56
3.19	Comparison of manually obtained dermis depth with the predefined segmentation depth.	58
3.20	Epidermis and dermis segmentations.	59
4.1	Nuclei in skin histopathological images.	61
4.2	Schematic of the voting-based technique.	64
4.3	An example of the initial segmentation.	65
4.4	The elliptical modeling for each connected region.	66
4.5	Variations of ellipticity parameters for isolated nuclei and nuclear clumps.	68
4.6	Illustration of voting algorithm for nuclei seed detection.	69
4.7	Illustration of voting-based nuclei detection.	72
4.8	Schematic of gLoG-based technique.	72
4.9	Nuclei in a skin image.	75
4.10	Peak responses by traditional scale normalization and log-scale normalization.	76
4.11	Example of normalized gLoG kernels.	77
4.12	Response maps with local maxima.	79
4.13	Illustration of nuclei detection.	81
4.14	Nuclei detection by different scales of gLoG kernels.	81
4.15	Nuclei scales estimation and segmentation.	83
4.16	Nuclei detection results by an existing gLoG-based technique.	86

4.17	Definition of true positives and false positives.	87
4.18	F-measures with different distance t values.	89
4.19	Nuclei detection in skin and breast images.	90
4.20	Nuclei detection by the proposed gLoG-based technique	91
4.21	Parameter sensitivity evaluation of the P-Voting technique.	92
4.22	Parameter sensitivity evaluation of the P-gLoG technique.	93
5.1	Schematic of the proposed mRLS-based technique.	99
5.2	gLoG kernels.	100
5.3	Nuclei detections.	100
5.4	Illustration of radial line scanning.	101
5.5	Candidate nuclei boundaries.	104
5.6	Blob segmentations.	106
5.7	Nuclei overlap processing.	108
5.8	Illustration of over-segmentation, under-segmentation and encroach- ment errors.	112
5.9	Comparative nuclei segmentations.	116
5.10	Visual examples of nuclei detections and segmentations.	118
5.11	Sensitivity evaluation of weighting parameters $(\lambda_1, \lambda_2, \lambda_3)$ on nu- clei segmentation.	119
6.1	Examples of H&E stained skin images.	124
6.2	Schematic of the proposed technique.	125
6.3	Schematic of epidermis and dermis segmentations.	125
6.4	Epidermis and dermis segmentations.	127
6.5	Image rotation and division.	127
6.6	Epidermis and dermis tiles.	128
6.7	Schematic of epidermis analysis.	129
6.8	Nuclei segmentation and melanocytes detection.	130
6.9	Illustration of melanocytes detection.	130

6.10	Division of epidermis into three sub-layers: inner layer, middle layer and outer layer.	132
6.11	Schematic of dermis analysis.	133
6.12	Example of color normalization.	134
6.13	Illustration of nuclei segmentation.	135
6.14	Nuclei detection and segmentation in dermis.	136
6.15	Architectural constructions based on detected nuclei seeds.	138
6.16	Feature selection by 10-fold cross validation.	146
6.17	An example of false diagnosis.	148
7.1	Melanoma DoI measurement in a MART-1 stained skin image.	152
7.2	Schematic of the proposed technique.	155
7.3	RGB values for manually selected melanoma pixels.	156
7.4	Melanoma detection results (indicated by blue contours).	157
7.5	Melanoma and epidermis segmentations.	159
7.6	Histogram specification for skin epidermis.	161
7.7	PDFs estimated from empirical training data.	164
7.8	Skin granular layer identification.	165
7.9	Automatic measurement of melanoma DoI.	167
7.10	Skin granular layer detection.	170
7.11	Comparison of automatically measured melanoma DoI and ground truths.	172
7.12	Normalized errors of computerized methods with ground truths as the gold standard.	172
7.13	Visual comparison of melanoma invasion measurement.	173
A.1	Skin histopathological images.	198
A.2	Examples of skin normal	201
A.3	Examples of skin nevus	202
A.4	Examples of skin melanoma	203

List of Abbreviations

Acronyms	Definition
ACM	active contours model
AE	average error
APE	average percentage of error
BIC	Bayesian information criterion
CAD	computer-aided diagnosis
CS	correct-segmenation
CT	computed tomography
DEJ	dermis-epidermis junction
DL	deep learning
DoG	Difference of Gaussian
DoI	depth of invasion
DP	dynamic programming
EM	expectation maximization
ER	encroachment errors
FP	false positive
GLCM	gray-level co-occurrence matrix
gLoG	generalized Laplacian of Gaussian
GM	granular-malpighian
GPU	graphic processing unit
H&E	hematoxylin and erosin
HPI	histopathology imaging
IHC	immunohistochemical

KNN	k-nearest neighbors
LDA	linear discriminant analysis
LoG	Laplacian-of-Gaussian
MAD	mean absolute distance
MART-1	melanoma-associated antigen recognized by T cells
mLoG	multi-scale Laplacian of Gaussian
mRLS	multi-scale radial line scanning
MRI	magnetic resonance imaging
mSVM	multi-class support vector machine
OS	over-segmentation
PCA	principal component analysis
PDFs	probability density functions
ROI	regions of interest
SD	standard deviation
SFS	sequential forward selection
SPV	single pass voting
SVM	support vector machine
TCA	tissue counter analysis
TP	true positive
US	under-segmentation
WSI	whole slide images

List of Symbols

Symbols	Definition
\mathcal{A}_{PRE}	area-based precision
\mathcal{A}_{SEN}	area-based sensitivity
\mathcal{A}_{SPE}	area-based specificity
$ \mathcal{A} $	cardinality of set \mathcal{A}
$\ A\ $	Euclidean norm of A
$\lfloor a \rfloor$	rounds number a to lower integer
$\{B_p\}_{p=1\dots Z}$	the set B which contains 1 to Z elements
b^c	the complement of mask b
b_e	epidermis mask
b_d	dermis mask
b_{e+d}	epidermis & dermis mask
\mathcal{D}_{MAD}	mean absolute distance
\mathcal{D}_{HD}	Hausdorff distance
N_{GT}	number of ground truths
N_{DS}	number of detected objects
N_{TP}	number of true positive
N_{FP}	number of false positive
N_{TN}	number of true negative
N_{FN}	number of false negative
$\partial\varphi(x, y)/\partial x$	partial derivative of the function φ with respect to x
$\partial^2\varphi(x, y)/\partial x^2$	second order partial derivative of the function φ with respect to x
$ x - y $	absolute different between x and y

(σ_x, σ_y)	scales of the gLoG kernel
θ	orientation of the gLoG kernel
$\nabla\phi$	the gradient of the function ϕ

Chapter 1

Introduction

With recent advances in medical imaging and artificial intelligence, computer-aided diagnosis (CAD) is playing an increasingly important role in the routine clinical diagnosis procedure. A wide variety of CAD schemes have been developed for detection and classification of various lesions in different medical imaging modalities, such as computed tomography (CT) [1], magnetic resonance imaging (MRI) [2] and histopathology imaging (HPI). Existing CAD systems have demonstrated their advantages in tedious tasks such as counting malignant cells and identifying tumor regions in biopsy tissues. The main goal of this thesis is to develop quantitative image analysis techniques that analyze skin histopathological images and assist pathologists in the diagnostic process. This chapter presents the motivations behind the CAD systems in skin histopathology for disease detection and classification. We briefly present the process of specimen preparation and digital imaging in histopathology. We also briefly explain skin cancer and the corresponding diagnostic features used by pathologists.

1.1 Skin Cancer and Diagnosis

Cancer is a major health problem around the world, and skin cancer is one of the most common types of cancer. Melanoma, which is a form of skin cancer, is the leading cause of mortality among all forms of skin cancers [3]. According to a

recent statistics, about 76,100 people were diagnosed with skin melanoma, and about 9,710 died from it in the United States alone in 2014 [4]. Melanoma develops from an abnormal growth of melanocytes which generate melanin, the dark pigment in skin. Cutaneous melanoma starts and grows locally, initially spreading across the surface of the skin (radial growth phase), and later invading into the dermis or even subcutaneous fat (vertical growth phase) [5]. If the melanoma is detected at early stages, it can be cured by a simple excision [3]. Therefore, an early detection and accurate prognosis of malignant melanoma will definitely help to lower the mortality from this cancer.

In the diagnosis of skin melanoma, a few diagnostic steps are performed at hospitals. The first step that doctors usually take is to do the physical exam (e.g., by naked eyes) and get patient's medical history. The patient with suspicious lesions is then referred to a dermatologist (a doctor specialized in skin diseases), who utilizes some advanced techniques (e.g., dermatoscopy) to observe the skin lesions more clearly. If a dermatologist thinks a spot might be a melanoma, a sample of skin will be excised from the suspicious area and sent to a laboratory for histopathological examinations by pathologists. Histopathological examination is a process of the microscopic examination of nuclei morphology and tissue distributions for disease diagnosis. The diagnosis from histopathology image is the gold standard in diagnosing almost all types of cancers, including skin melanoma [6].

1.2 Histopathology Imaging

Histopathology refers to the microscopic examination of tissue to study the manifestations of disease [7]. In order to identify cytological features (i.e., characteristics of cell nuclei) that indicate the presence of cancer, pathologists observe histological slides under the microscope and examine tissue components at different scales to make a judgement based on their personal experience and knowledge. Generally, a few steps are required for slide preparation before diagnostic examination. First, a biopsy sample is excised from suspicious lesion. Second, fixation by some chemical

fixatives (e.g., formalin) is performed to stop the enzyme activity, and preserve the cells and their morphological and architectural structures. Third, water is removed from the tissue, and some chemical agent (e.g., paraffin) that solidifies the tissue is embedded into the specimen to allow thin sections to be cut. Next is the sectioning where the embedded tissue sample is cut into thin sections (e.g., $4\mu m$ for light microscope) by a steel knife mounted in a microtome. Finally, these thin sections are mounted to a glass slide and stained with one or more microscopical stains.

Since cells and other extracellular structures making up most tissue specimen are colourless, staining is applied to reveal cellular components and provide contrast to the tissue [8], [9]. In clinical histology, the most-widely used stain is a combination of hematoxylin and eosin (H&E) [10]. Hematoxylin is an alkaline stain with deep purple or blue color, and eosin is an acidic stain with a red color. With H&E staining, the chromatin rich cell nuclei are stained with blue shade, while the cytoplasm and other connective tissues are stained with different shades of pink (see Fig. 1.1(a)). Although H&E stains provide a good contrast between cellular and extra-cellular structures, they are not specific to a certain type of cell nuclei. Immunohistochemical (IHC) stains are also widespread staining pigments in histopathology, which use antibodies to stain particular antigens (e.g., proteins). IHC stains are typically used to identify specific types of cells in the tissue specimen. For example, MART-1 (i.e., melanoma-associated antigen recognized by T cells) stain is commonly used to highlight the melanocytes in skin histopathological images, where melanocytes are observed as brown color and other types of cell nuclei are observed as blue color (see Fig. 1.1(b)). Ki-67 stain is widely-used to evaluate the tumor cell proliferation index [11], where immuno-positive and immuno-negative tumor cell nuclei are observed as brown and blue colors (see Fig. 1.1(c)), respectively. For other IHC stains for skin biopsy image analysis, please refer to the Appendix A.1. Using tissue specific stains, physicians diagnose a tumor as benign or malignant, and determine the stage and grade of a tumor.

After staining, the tissue specimens are covered with a glass (or plastic) coverslip to hold the specimen in place. In the past, pathologists view the slides with

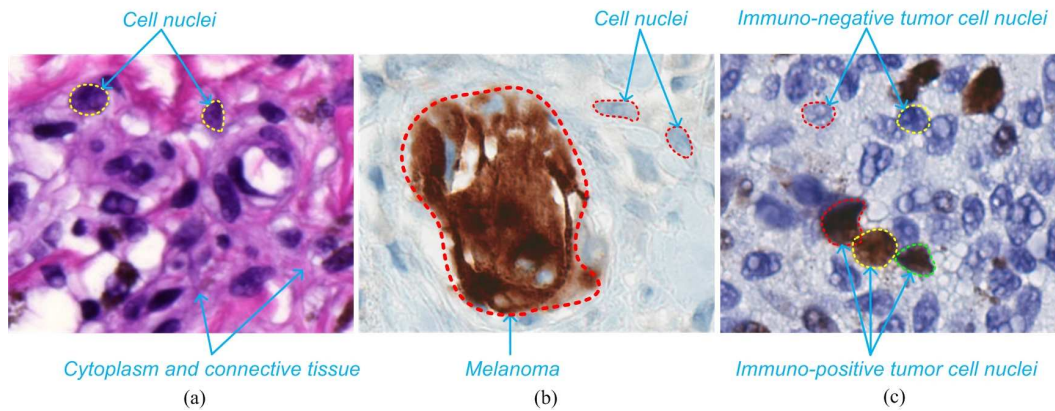


Fig. 1.1. Examples of skin histopathological images. (a) An H&E stained image. (b) A MART-1 stained image. (c) A Ki-67 stained image. Note that the images are viewed at 40X.

the specimen under a microscope, and make the diagnosis based on a series of observed features (e.g., cell types and distributions). With recent development in digital tissue slide scanners, glass slides of tissue specimen can now be digitized at high magnification to generate whole slide images (WSIs) [12]. Pathologists examine the digitized tissue slides at different scales using computer visualization softwares such as Mirax Viewer [13] and Image Scope [14] to make a diagnosis and judgement.

1.3 Histological Diagnosis and Melanoma Grading

Histopathological examination based on cytological and architectural features is currently used to confirm the presence of skin melanoma. In the following, we briefly explain the typical structures of human skin, and then illustrate how to diagnose and grade skin melanoma from histopathological images.

1.3.1 Skin Structure

The human skin is the outer covering of the body, which has many functions (e.g., thermoregulation and protection) [15]. It consists of three main layers: epidermis, dermis and subcutaneous tissue, each providing a distinct role in the overall function of the skin. Fig. 1.2 shows an anatomy of normal skin. The descriptions of each

skin layer are provided below.

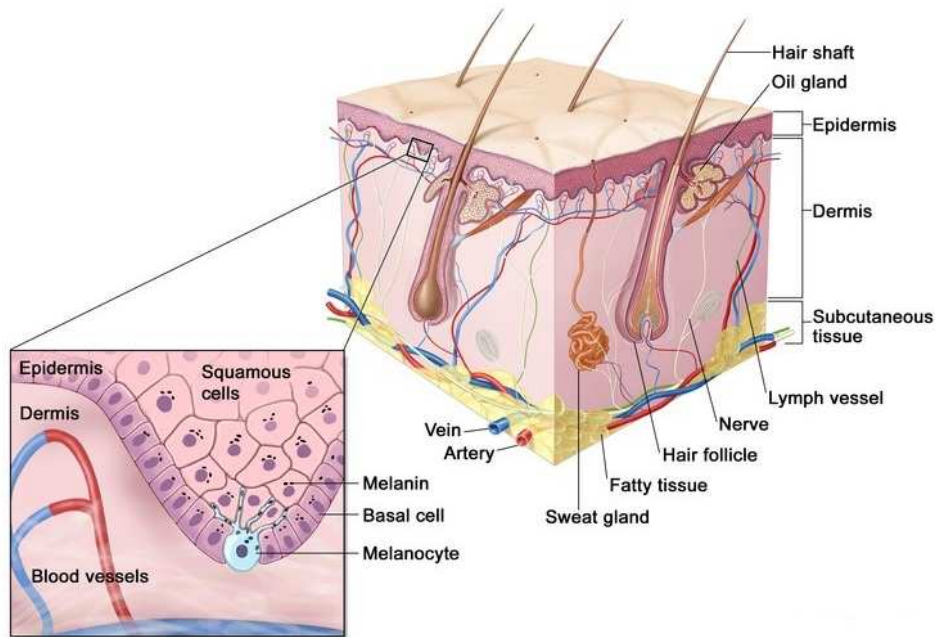


Fig. 1.2. Anatomy of normal skin (Image Credit: [16]).

The epidermis is the outermost layer of the skin, which serves as the first barrier to infection. Generally, the epidermis can be broadly divided into four layers: cornified layer, granular layer, spinous layer and basal layer (see Fig. 1.3). It is mainly composed of keratinocytes (95%) but also contains melanocytes, Langerhans cells and Merkel's cells. The thickness of epidermis varies in different regions of body but is usually around 0.1mm thick.

The dermis is a layer of skin between the epidermis and subcutaneous tissues, which cushions the body from stress and strain. It consists of two layers: papillary dermis and reticular dermis. The papillary layer includes the ridges and papillae. The reticular layer is composed of coarse interlacing collagenous fibers and an elastic network. Hair follicles, sebaceous glands, sweat glands, receptors and blood vessels are also located in the reticular layer. The thickness of dermis may vary greatly ranging from 0.3mm (e.g., eyelids) to 4.0mm (e.g., palms and soles of feet), but most parts of human body generally have a thickness about 2mm .

The subcutaneous tissue is the lowermost layer of human skin, which blends

indistinctly with the dermis. It contains loose connective tissue and lobules of fat that holds large blood vessels and nerves. Fig. 1.3 shows a skin WSI with annotated tissue components.

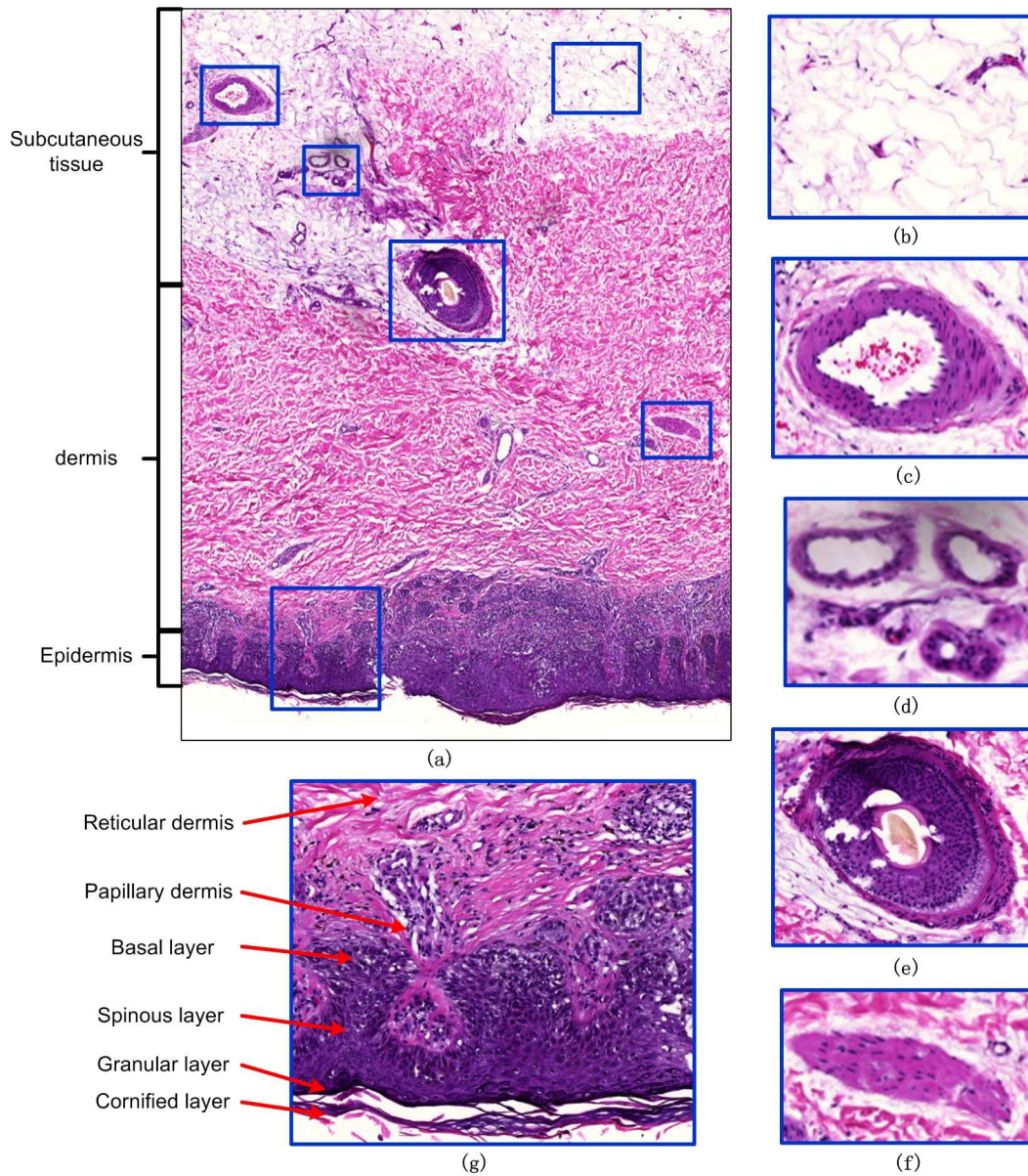


Fig. 1.3. An example of H&E stained skin histopathological image. (a) Skin WSI. (b) Sebaceous fat. (c) Blood vessel. (d) Sweat glands. (e) Hair follicle. (f) Muscle. (g) Selected epidermis and dermis areas. Note that (b)(c)(d)(e)(f)(g) are the magnifications of selected image patches in (a).

1.3.2 Histologic Diagnosis of Skin Melanoma

In the diagnosis of skin melanoma by analyzing histopathological images, pathologists are mainly interested in the epidermis and dermis-epidermis junction (DEJ) areas as most histologic diagnosis criteria are observed in these regions [17]. In the following, we list the commonly-used features to identify skin melanomas in H&E stained images.

1.3.2.1 Epidermis features

There are a series of architectural and cytological features within epidermis that are related to malignant melanoma. Generally, architectural changes seen within the epidermis in melanoma include broader lesion breadth, poor circumscription of melanocytes, single melanocytes predominating over nests of melanocytes, asymmetrical changes in epidermis thickness, the presence of melanocytes above the basal layer and the irregular configuration of epidermis [17], [18]. Besides architectural features, cytological features are also important in indicating the presence of melanoma. These features include large size of melanocytes, irregular contours of nuclei, prominent or multiple nucleoli and severe mitosis.

1.3.2.2 Dermis features

In the dermis area, the architectural changes indicating the presence of melanoma mainly include the irregular shape of dermis base contour (close to the DEJ region), the lack of maturation and asymmetrical host inflammatory response [17]. Just like the cytological features in the epidermis area, the severe mitotic activity, the highly nuclear pleomorphism (e.g., irregular nuclei shape and size), prominent or multiple nucleoli and hyperchromatism (i.e., excessive nuclei staining) also indicate the presence of melanoma in the dermis area.

More explanations and examples of histologic features for skin melanoma diagnosis can be found in [17], [18]. However, it should be mentioned that it is difficult to make a 100% correct judgement based on the above listed features,

even by an experienced pathologists. In practice, the experience and knowledge of pathologists are quite important in disease diagnosis, and cannot be replaced by any CAD systems in near future.

1.3.3 Histologic Grading of Skin Melanoma

If a skin lesion is diagnosed with malignant melanoma, pathologists will further examine the biopsy to determine the state (extent) of the disease in order to effectively plan patient’s treatment. The Breslow thickness, which is defined as the maximum distance between malignant cells and the top of the granular layer [19], is one of the most important prognostic factors to stage the skin melanoma. The deeper is the Breslow thickness, the greater is the possibility of metastases and the worse is the prognosis for the patient. For example, a study in [20] has indicated that a 5-year survival rate decreases continuously with the increase of tumor depth, which is illustrated in Table 1.1. Since MART-1 stain is specific to the melanocytes, it is usually used to determine the depth of melanoma invasion by pathologists. Fig. 1.4 shows an example of measuring Breslow thickness in a MART-1 stained skin image. Note that in Fig. 1.4 the melanoma (malignant cells) is stained with dark brown color, while other normal cells are stained with light blue color. As observed in Fig. 1.4, the depth of melanoma invasion is $351.7\mu m$.

TABLE 1.1
SURVIVAL RATES WITH DIFFERENT MELANOMA DEPTH.

Melanoma Depth	<1mm	1-2mm	2.1-4mm	>4mm
Survival Rates (~5 years)	95-100%	80-96%	60-75%	50%

1.4 Problem Statements and Motivation

Although many non-invasive techniques have been developed for melanoma diagnosis, e.g., epiluminescence microscopy [3] and confocal microscopy [21], which can provide initial diagnosis, the histopathological examination of a skin WSI by

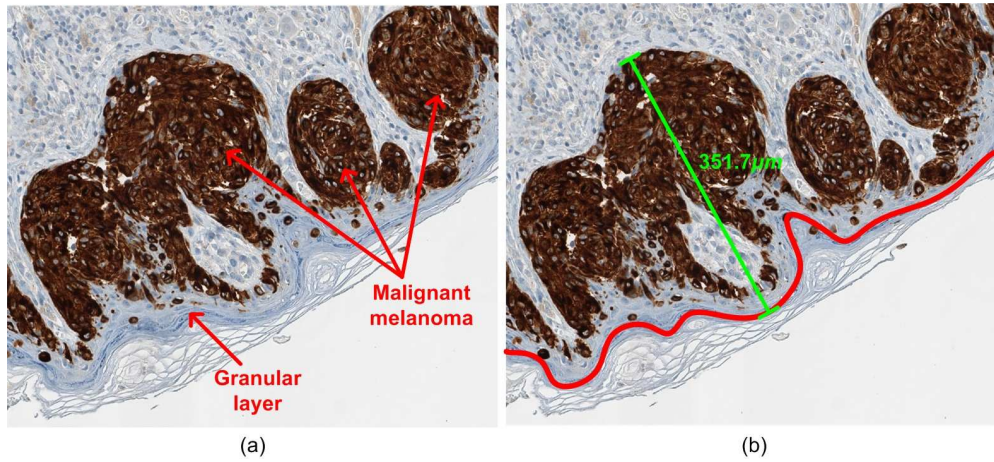


Fig. 1.4. Breslow thickness measurement. (a) A skin image stained by MART-1. (b) Breslow thickness.

pathologists remains the gold standard for the diagnosis as the histopathology slides provide a cellular level view of the disease [22]. Traditionally, the histopathological slides are examined under a microscope, and pathologists make the diagnosis based on their personal experience and knowledge. However, the diagnosis by pathologists are typically subjective and often lead to intra- and inter-observer variability [23], [24]. For example, it has been reported that in the diagnosis of melanoma (on a randomly selected set of 62 suspicious pigmented skin lesions), the inter-observer variation of diagnosis sensitivity ranges from 55% to 100% between 20 pathologists (overall sensitivity 87%) [25]. Additionally, the manual analysis of the WSI with high resolution is labor intensive due to the large volume of the data to be analyzed [26]. To address these problems, computer-aided image analysis which can provide reliable and reproducible results is desirable.

Besides bringing benefits in objectivity and labour-saving, computerized analysis of digitized histological slides also has advantages in other aspects. For example, it usually can provide a wide variety of micro and macro prognostic cues, which may not be easily observed by a qualitative visual examination. Using quantitative histopathology, it is possible to process the entire WSI instead of only examining representative image regions, and hence the sampling bias by manual examination is prevented. Providing quantitative tools for characterization of tissue changes and

structure is also important for basic biological research to understand the biological mechanisms of disease progression [3].

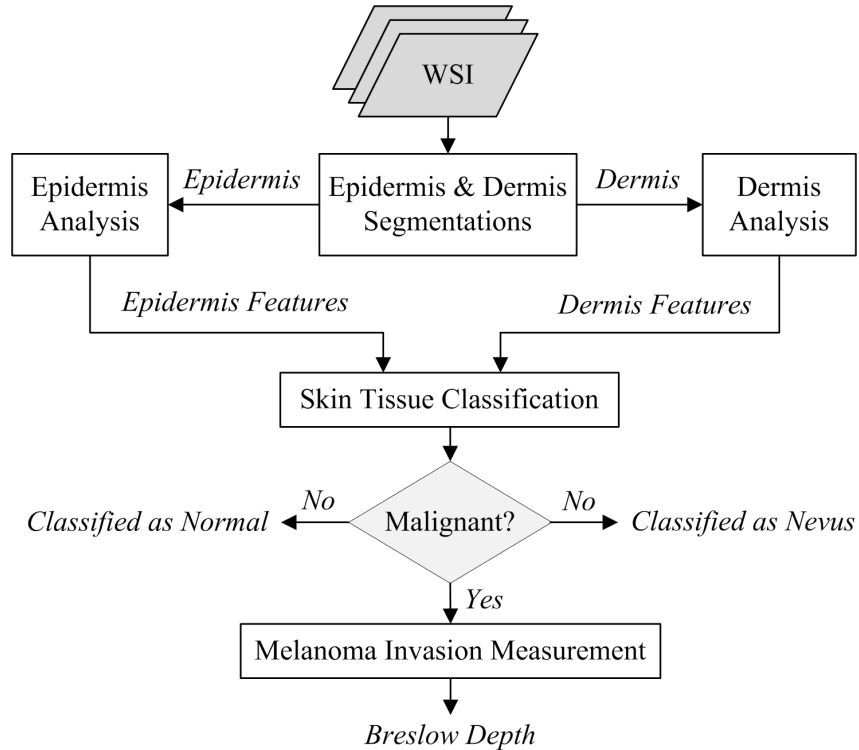


Fig. 1.5. Schematic of the proposed skin melanoma diagnosis and grading system.

The aim of this thesis is to develop robust and accurate image analysis algorithms for computer-aided interpretation of skin histopathology imagery. A CAD system for skin histopathological image classification based on the features from both epidermis and dermis areas will be developed. This CAD system will operate on reliable quantitative measures (from pixel-level to object-level), and will provide objective and reproducible information that can be used by pathologists for diagnosis. For the specimens identified with melanoma, the depth of invasion (DoI) will be measured to determine the severity of skin melanoma. Fig. 1.5 shows the framework of the proposed CAD system for skin melanoma diagnosis and grading. As observed in Fig. 1.5, the proposed CAD system includes five main modules: epidermis & dermis segmentations, epidermis analysis, dermis analysis, skin tissue classification and melanoma invasion measurement. The techniques proposed in this thesis are used by this CAD system for skin biopsy image analysis, but they are

also expected to be applied into other CAD applications of biopsy image analysis.

1.5 Major Contribution of the Thesis

Parts of this thesis have been published in the following journals and conferences:

Refereed Journals

- [1] H. Xu and M. Mandal, “Epidermis segmentation in skin histopathological images based on thickness measurement and k-means algorithm”, *EURASIP Journal on Image and Video Processing*, vol.2015, no.1, pp.1-14, 2015. (Chapter 3)
- [2] H. Xu, C. Lu, N. Jha and M, Mandal, “An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm”, *IEEE Journal of Biomedical and Health Informatics*, vol.18, no.5, pp.1729-1741, 2014. (Chapter 4)
- [3] H. Xu, C. Lu, R. Berendt, N. Jha and M. Mandal, “Automatic nuclei detection based on generalized Laplacian of Gaussian filters”, *IEEE Journal of Biomedical and Health Informatics*, vol.21, no.3, pp.826-837, 2017. (Chapter 4)
- [4] H. Xu, C. Lu, R. Berendt, N. Jha and M. Mandal, “Automatic nuclear segmentation using multi-scale radial line scanning with dynamic programming”, *accepted by IEEE Transactions on Biomedical Engineering*, January 2017. (Chapter 5)
- [5] H. Xu, R. Berendt, N. Jha and M. Mandal, “Automatic measurement of melanoma depth of invasion in skin histopathological images”, *Micron*, vol.97, pp.56-67, 2017. (Chapter 7)

Refereed Conference Proceedings

- [1] H. Xu and M. Mandal, “Efficient segmentation of skin epidermis in whole slide histopathological images”, *Proc. of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp.3855-3858. (Chapter 3)
- [2] H. Xu, C. Lu and M. Mandal, “Automated segmentation of regions of interest in whole slide skin histopathological images”, *Proc. of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp.3869-3872. (Chapter 3)
- [3] H. Xu, H. Wang, R. Berendt, N. Jha and M. Mandal, “Automated nuclear segmentation in skin histopathological images using multi-scale radial line scanning”, *Proc. of IEEE-NIH 2016 Special Topics Conference on Healthcare Innovations and Point-of-Care Technologies*, 2016, pp.175-178. (Chapter 5)
- [4] H. Xu, R. Berendt, N. Jha and M. Mandal, “Automated diagnosis of melanoma from skin biopsy images”, *Proc. of BHI-2017 International Conference on Biomedical and Health Informatics*, February 2017. (Chapter 6)
- [5] H. Xu, H. Wang, R. Berendt, N. Jha and M. Mandal, “Computerized measurement of melanoma depth of invasion in skin biopsy images”, *Proc. of BHI-2017 International Conference on Biomedical and Health Informatics*, February 2017. (Chapter 7)

1.6 Organization of the Thesis

The rest of this thesis is organized as follows: Chapter 2 presents the broad literature review on image processing techniques used for histology analysis as well as existing whole slide image analysis systems. Chapter 3 presents the proposed techniques for epidermis and dermis segmentations in skin WSIs. Chapter 4 presents two proposed techniques for nuclei detections (e.g., identify nuclei locations) in skin histopathological images. Chapter 5 presents the proposed nuclei segmentation

technique (e.g., delineate nuclear boundaries) in skin histopathological images. Chapter 6 presents the proposed technique for automatic analysis and classification of melanocytic tumor on skin WSIs. Chapter 7 presents the proposed technique for measuring Breslow depth in MART-1 stained skin histopathological images. Future work and conclusion for the research are presented in Chapter 8, which is followed the bibliography and appendix.

Chapter 2

Literature Review

In this chapter we present a brief overview of computer-aided image analysis techniques for histopathological images. The literature review covers a range of commonly used image processing algorithms related to regions of interest (ROI) segmentation, features construction and classification. The state of the art techniques related to WSIs analysis with different types of cancer including skin cancer are also included.

2.1 ROI Detection and Segmentation

In the diagnosis of various cancers by analyzing microscopic images, the ROI detection and segmentation are usually the first step before any subsequent quantitative analysis, as the architectural and morphological features of these ROI often provide the most important diagnostic clues in cancer identification. Numerous works have been conducted to detect and segment different tissue components from microscopic images.

2.1.1 Thresholding

Thresholding is an intuitive and simple method in image segmentation, which converts an intensity image into a binary image by assigning all pixels to the value one

or zero as follows:

$$B(x, y) = \begin{cases} 1 & \text{if } I(x, y) \geq T \\ 0 & \text{otherwise} \end{cases} \quad (2.1)$$

where (x, y) represents a pixel coordinate in the image I or B , and T is a pre-defined threshold. Many researchers [24], [27]–[30] have used thresholding techniques for ROIs segmentation. Gurcan *et al.* [27] proposed a thresholding based method for nuclear segmentation in neuroblastoma images, which first performs image pre-processing by morphological top-hat transform (i.e., subtracting an opened image from the original), and then image binarization by the hysteresis thresholding followed by watershed segmentation. Korde *et al.* [28] proposed to segment cell nuclei in bladder and skin images by a global threshold. The technique tries to find the valley of the image histogram as the threshold. If no distinct valleys are detected in the image, the darkest 20% of the pixels in the image are assumed to be nuclear pixels. Lu *et al.* [30] proposed a locally adaptive threshold selection technique for nuclear segmentation in skin histopathological images. This technique first performs image binarization by an adaptive thresholding [31], and then determines an optimal threshold for abnormally large nuclear regions (i.e, nuclear clumps) based on prior knowledge about nuclear size and shape. Although these thresholding techniques have advantages of efficiency and simplicity, it usually results in under segmentations or missed segmentations when there exist considerable intensity variations in the image.

2.1.2 Watershed

Watershed is another popular segmentation technique in image processing field. It considers the grayscale image as a topographic surface, where pixels intensity are interpreted as heights. The “water” starts from the local minima in the image and gradually floods the surrounding regions (i.e., catchment basin), and the watershed lines are formed when “waters” from different basins meet together. The whole process of “water” flooding stops when the water level has reached the highest peak in the topographic surface, and thus the image is divided into different meaningful

regions [31]. The watershed transform can be represented as:

$$L(x, y) = \mathcal{W}(I(x, y)) \quad (2.2)$$

where $\mathcal{W}(\cdot)$ represents the watershed transform operation, and $L(x, y)$ is the segmented image. The watershed segmentation can be conducted by combining with the Euclidean distance transform or image gradient maps, but it tends to produce severe over-segmentations when there are too many local minima in the image. In order to tackle this problem, a number of marked watershed algorithms [32]–[36] have been proposed. These algorithms differ in the strategies of finding the markers of interested objects. For instance, Yang *et al.* [32] proposed a method that detects nuclei markers for segmentation by an iteratively coarse and fine erosion process. Jung *et al.* [34] formulated the marker extraction based on H-minima transform as an optimization problem, which determines an optimal depth value (corresponds to the minimum fitting error) to suppress the undesired local minima in the Euclidean distance map. George *et al.* [37] proposed to detect nuclei markers by a circular Hough transform, followed by watershed segmentation. The accuracy of these marked watershed segmentation depends critically on the accuracy and reliability of detected markers.

2.1.3 Filtering

In order to identify the locations and numbers of objects in an image (e.g., cell nuclei), filtering based techniques which can enhance the signals around object centers may provide a good performance. Laplacian of Gaussian (LoG) filter, as defined below, has been widely used in histological images:

$$LoG(x, y; \sigma) = \frac{\partial^2 G(x, y; \sigma)}{\partial x^2} + \frac{\partial^2 G(x, y; \sigma)}{\partial y^2} \quad (2.3)$$

where $G(x, y; \sigma)$ is a Gaussian kernel with zero mean and scale σ , as defined below:

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2.4)$$

Byun *et al.* [38] proposed a nuclear detector in retinal images that is designed based on the LoG filter at a fixed scale (set empirically). This method has advantages such as simple implementation and low computational complexity, but it fails to provide a good performance when the image has nuclei with different size (corresponding to different scales). Al-Kofahi *et al.* [39] proposed a method that detects nuclei based on multi-scale Laplacian of Gaussian (mLoG) filters, which first extracts image foreground with a graph-cuts-based binarization, and then detects nuclear seeds by mLoG filtering with a distance-map-constrained adaptive scale selection. The mLoG filter is efficient in locating the centers of 2-D near-circular blobs but has limitations in detecting blobs with general elliptical shapes. Kong *et al.* [40] proposed a generalized Laplacian of Gaussian (gLoG) filter based technique, which tries to detect general elliptical blob structures (like cell nuclei) in images. Although a good performance is reported in [40], the technique is highly sensitive to the image noise and often misses low intensity nuclei in the image.

2.1.4 Active Contour Model (ACM)

A widely-used approach for object segmentation is based on ACM or snakes. The basic idea of these algorithms is to evolve a continuous curve in a 2-D plane to delineate structures in an image. The evolution of the ACM is subject to constraints from the given image by seeking to minimize an energy function based on gradient or intensity information. For example, A classical snake model, acting as an edge-detector, is proposed in [41], which evolves the curve by minimizing the following energy function:

$$E_{snake} = \alpha \int_0^1 |C'(s)|^2 ds + \beta \int_0^1 |C''(s)|^2 ds - \lambda \int_0^1 |\nabla I(C(s))|^2 ds \quad (2.5)$$

where $C(s)$, $s \in [0, 1]$ represents a parametric curve, α , β and λ are the positive weight parameters. The first two terms on the right side control the smoothness of the contour (i.e., internal energy), while the third term attracts the contour toward object boundaries in the image (i.e., external energy). Several works have applied the ACM for ROI segmentation in microscopic images. These works are related to cell segmentation [42], [43], nuclear and glandular structures segmentation [44], [45]. Although these techniques are reported to provide a good performance, the results are very sensitive to the initialization of ACM and local intensity variations in images. In addition, the ACM has a high computational complexity.

2.1.5 Clustering

Clustering is a process of grouping objects into different clusters, where the objects in the same group (i.e, a cluster) are more similar to each other than those in other groups. The k-means algorithm is a popular clustering technique that divides a set of observations (i.e., feature vectors) into k sets $S = \{S_1, S_2, \dots, S_k\}$ by minimizing the within-cluster variances, i.e.,

$$S = \arg \min_S \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (2.6)$$

where x represents the observation and μ_i is the mean of feature vectors in S_i . Many works have utilized clustering analysis techniques for objects segmentation and identification in medical image applications. Sertel *et al.* [46] designed a system for follicular lymphoma grading, which utilizes k-means clustering algorithm on the lymphoma microscopic images to segment cell nuclei, cytoplasm and extra-cellular material based on Lab* color features. Lu *et al.* [22] applied mean-shift clustering algorithm to smooth the image and make nuclear regions homogenous such that nuclei can be accurately segmented by adaptive thresholding in skin histopathological images. Qi *et al.* [43] proposed a cell seed detection algorithm that uses mean-shift clustering on a voting map and considers cluster centres as cell seeds.

2.1.6 Probabilistic Models

Based on the prior knowledge (e.g., color distribution of detected objects), probabilistic models can be used to classify the objects in an image into different categories. The well-known Bayesian theorem has been used in many medical applications. Using the Bayesian theorem [31], the posterior conditional probability $P(\omega_j|X)$ can be computed as:

$$P(\omega_j|X) = \frac{p(X|\omega_j)P(\omega_j)}{\sum_{j \in \{1 \dots K\}} p(X|\omega_j)P(\omega_j)} \quad (2.7)$$

where X is a feature vector (e.g., pixel intensities), ω_j is the index of classes and K is the number of classes. $p(X|\omega_j)$ and $P(\omega_j)$ are the likelihood function and prior probabilities, respectively. The evidence $\sum_{j \in \{1 \dots K\}} p(X|\omega_j)P(\omega_j)$ is the scale factor. By using the Bayesian theorem, Naik *et al.* [47] segmented the image into three classes including lumen, cytoplasm, and nuclei for prostate cancer grading. Sertel *et al.* [48] detected the mitosis and karyorrhexis cells in neuroblastoma histological images. Basavanhally *et al.* [49] detected the lymphocytes in breast cancer images by the Bayesian model. Lu *et al.* [50] segmented mitotic cells from multi-spectral images based on Bayesian classification. Since the likelihood and the prior terms are generally determined from training samples, the performance of Bayesian theorem based detection techniques are highly relied on the training database.

2.1.7 Comparison of Segmentation Techniques

Figs. 2.1(a)(b) show an H&E stained color skin image and its corresponding gray-scale image, respectively. Figs. 2.1(c)(d)(e)(f)(g)(h) separately show the nuclei segmentations by using Otsu's thresholding method [51], marked watershed algorithm [34], LoG filter [38], ACM [41], k-means clustering algorithm [52] and Bayesian classification technique [50]. Note that in Fig. 2.1(f), the initializations (i.e., yellow circles in the image) for ACM are obtained based on a voting algorithm [53]. As observed in Fig. 2.1, the global thresholding, clustering and probabilistic model

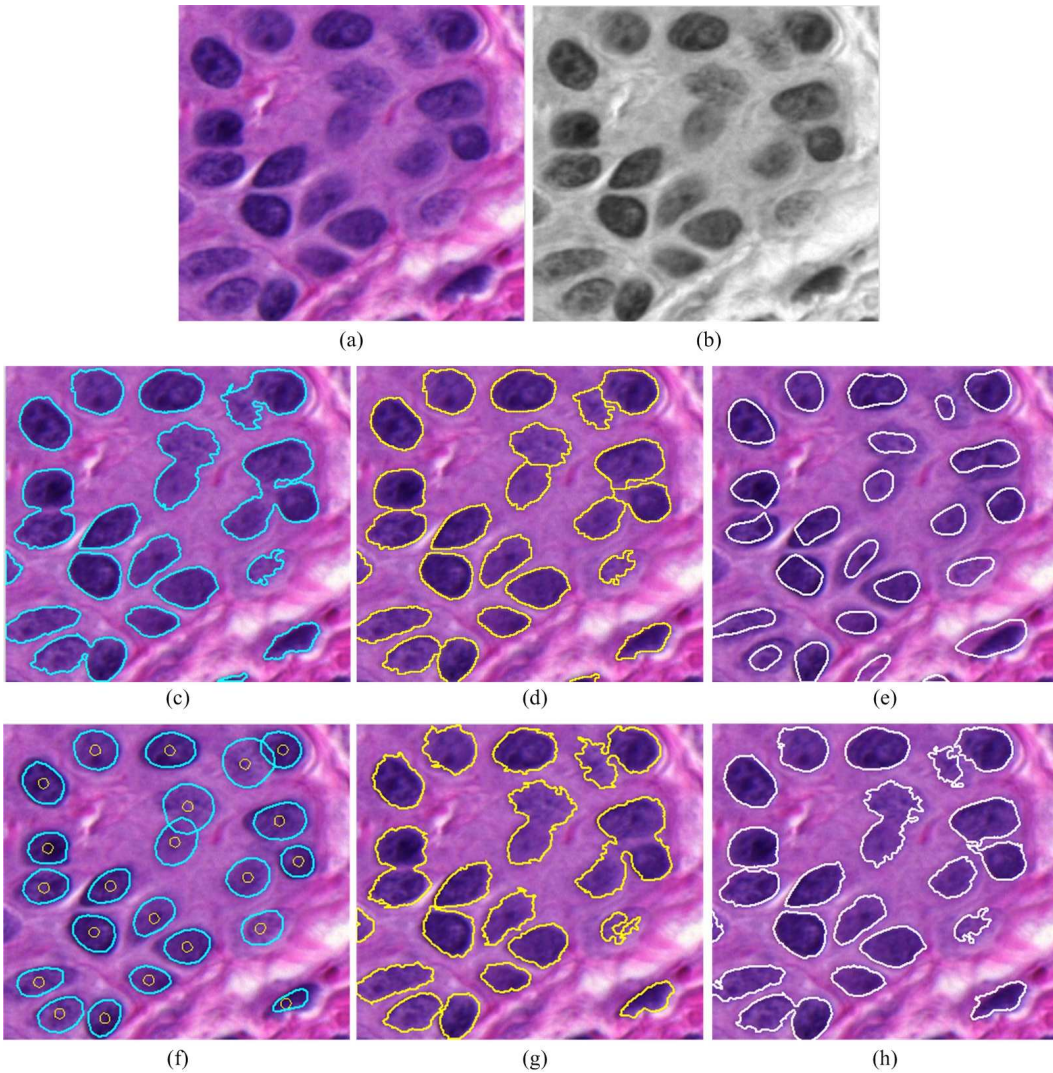


Fig. 2.1. Examples of nuclei segmentations by different techniques. (a) Original color image. (b) Gray-scale image. (c) Otsu's thresholding. (d) Watershed algorithm. (e) LoG filtering. (f) ACM. (g) K-means clustering. (h) Bayesian classification. Note that in (f) the yellow circles around nuclear centers are the initializations for ACM.

(see Fig. 2.1(c)(g)(h)) generally results in under-segmentations for the touching or overlapping cell nuclei, as these techniques perform segmentations mainly based on pixel intensities. The marked watershed and LoG filters can generally split clustered nuclei (Fig. 2.1(d)(e)), but the obtained nuclear boundaries are usually not very accurate due to the color variations within nuclei regions. In comparison, the ACM can provide much smoother nuclear boundaries and achieve more accurate segmentations (see Fig. 2.1(f)). However, the performance of ACM is quite sensitive to the initializations and local intensity variations in the image.

2.2 Features for Histology Images

There have been numerous research works that characterize histological images with different features. These features are usually related to the texture, morphology and architecture of segmented ROI.

2.2.1 Textural Features

Textural features are important in histology analysis since different tissue components (e.g., melanoma and normal cells) are stained with different colors and present different textures [54]. A frequently-used approach for texture analysis is based on statistical properties of the intensity histogram. Given a set of pixels from a segmented ROI, the textural features from intensity histogram including the mean, standard deviation, smoothness, third moment, uniformity and entropy [31] can be extracted. For more explanation of these features, please see the Appendix A.2.1.1 These feature descriptors have been used in tissue counter analysis for distinguishing benign nevi and malignant melanoma [55]. In [26], they were used as texture features to segment squamous epithelium by the SVM classifier.

Haralick texture features [56] are another category of popular features used in histological image analysis. These features are derived from the gray level co-occurrence matrix \mathcal{C} ,

$$\mathcal{C} = \begin{bmatrix} p(1, 1) & p(1, 2) & \cdots & p(1, N_g) \\ p(2, 1) & p(2, 2) & \cdots & p(2, N_g) \\ \vdots & \vdots & \ddots & \vdots \\ p(N_g, 1) & p(N_g, 2) & \cdots & p(N_g, N_g) \end{bmatrix} \quad (2.8)$$

where N_g is the total number of gray levels, $p(i, j)$ is the probability that a pixel with gray level i is adjacent to a pixel with gray level j . The co-occurrence matrix records the spatial dependency between the gray-level pairs. Although there are many Haralick features [56], [57], only some of them are commonly used in histological

applications. These features include energy, contrast, correlation, homogeneity, dissimilarity and entropy. For more explanations about Haralick texture features, please see the Appendix A.2.1.2. Sertel *et al.* [58] proposed an automatic technique to classify the stromal development in neuroblastoma images, which utilizes Haralick features for classification. Lu *et al.* [59] proposed a technique for nuclear atypia classification in breast cancer images, which is designed based on Haralick features.

There are also other kinds of texture features that have been applied for histological image analysis such as local binary pattern [58], Gabor features [60], [61] and Tamura features [62]. For more details about these features, please refer to the provided references.

2.2.2 Morphological Features

During the manual analysis of digitized biopsies, pathologists usually describe many gland or cell characteristics in terms of morphology, and hence morphological features play a significant role in automatic histopathological image analysis. Based on the segmented ROI, a set of morphological features can be calculated. Table 2.1 lists the common morphological features used in histology.

TABLE 2.1
DESCRIPTION OF MORPHOLOGICAL FEATURES.

Name	Description
Regional features	The features are derived directly from the segmented ROI, and usually include the area, perimeter and center.
Convex hull features	The features are derived from the convex hull [63] of ROI. These features may include the solidity, convex deficiency, convex area and concavity.
Elliptical features	The features are derived from the best fit ellipse [64] of ROI. These features may include the major and minor axis and the ratio between them, eccentricity, orientation, and ellipticity.

Fig. 2.2(a) shows the convex hull of a ROI (i.e., lymph node). The solidity is the ratio of ROI pixels (e.g., green pixels in Fig. 2.2(a)) over the total pixels in the convex hull, while the convex deficiency is the ratio of pixels in concavity

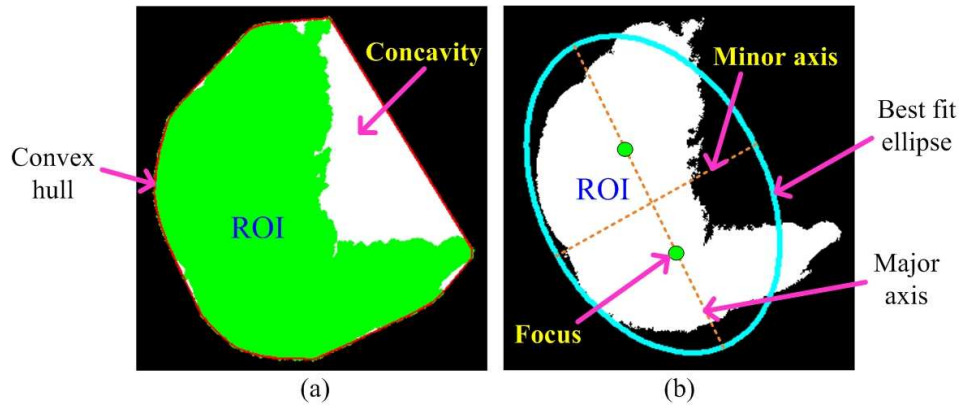


Fig. 2.2. Convex hull and best fit ellipse. (a) The convex hull for a lymph node. (b) The best fit ellipse for a lymph node.

regions (e.g., white regions in Fig. 2.2(a)) over the total pixels in the convex hull. Fig. 2.2(b) shows the best fit ellipse of a ROI. The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length, while the ellipticity is the ratio of ROI pixels within the ellipse over the total pixels of the ellipse. For more explanations about different morphological features, please refer to [31]. The morphological features are usually combined with other features (e.g., textural features) for disease diagnosis and grading. The typical applications include follicular lymphoma grading [46], cervical intraepithelial neoplasia diagnosis [26] and melanocytic histology classification [65].

2.2.3 Architectural Features

Architectural features measure the spatial arrangement and structural information of ROI in the histological image. The graphs can generally be used for encoding the spatial-relation features. Formally, a simple graph is defined as $G = (V, E)$, with V and E being the node and edge set. To capture the spatial distribution of segmented ROI, the centers of ROI can be considered as the nodes for constructing graphs. There are many graph structures that can be used to encode spatial information, and correspondingly many features can be selected for classification [54]. The typical graphs include Voronoi Tessellation, Delaunay Triangulation and Minimum Spanning Tree. Let us assume that an image region I contains m objects of interest

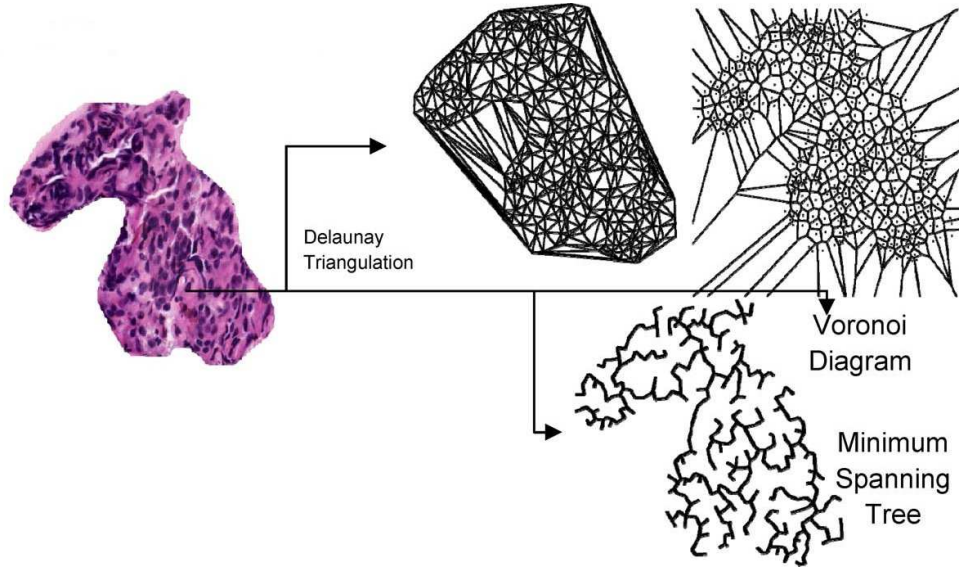


Fig. 2.3. A histological image and its graph structures with nuclear centers as nodes (Image Credit: [54]).

with centroids c_1, c_2, \dots, c_m . The graphs are defined as follows:

- **Voronoi Tessellation:** The Voronoi diagram V is a set of polygons $P = \{P_1, P_2, \dots, P_m\}$. For any polygon P_i ($1 \leq i \leq m$), it is comprised of pixels x ($x \in I$) if $d(x, c_i) = \min_j \{\|x - c_j\|\}$ where $i, j \in \{1, 2, \dots, m\}$ and $d(a, b)$ is the Euclidean distance between any two pixels $a, b \in I$.
- **Delaunay Triangulation:** The Delaunay triangulation D is constructed on top of the Voronoi diagram V , where any two centroids c_i and c_j with $i, j \in \{1, 2, \dots, m\}$ are connected by an edge if P_i and P_j share a side in V .
- **Minimum spanning tree:** The minimum spanning tree MST is a non-cyclic subgraph that connects all nodes in D . Note that a single graph D may have many different spanning trees, but the spanning tree with the minimum length (i.e., edge length) is selected as the MST .

Fig. 2.3 shows an example of Voronoi tessellation, Delaunay triangulation and Minimum spanning tree for the nuclei in the histopathological image. The features derived from these graphs are provided in Table 2.2. Many works on histological image classification have used architectural features derived from graphs.

TABLE 2.2
DESCRIPTION OF ARCHITECTURAL FEATURES.

Graph structure	Typical features
Voronoi tessellation	number of nodes, number of edges, average polygon area, area disorder and roundness factor
Delaunay Triangulation	number of nodes, average edge length, maximum edge length, number of edges and average triangle area
Minimum spanning tree	number of nodes, average edge length, total edge length, edge length disorder and number of neighbors

These works includes lymphocytic infiltration grading in breast histological images, colon tissue classification [66], prostate cancer grading [60] and cervical cancer grading [67]. Besides features from graphs, some other architectural features are also used for measuring spatial arrangement of objects in the image. For example, Petushi *et al.* [24] proposed to measure nuclear density (i.e., number of nuclei per square window) to assist breast cancer grading. Sertel *et al.* [46] proposed a histopathological image analysis system for follicular lymphoma grading, which uses measurements such as number of neighboring nuclear components and average distance to neighboring nuclear components to encode architectural features of the image.

2.2.4 Features Selection

In real applications, it is usually difficult to perform the diagnosis and grading of tissue diseases based on one category of features, and hence textural, morphological and architectural features are usually combined together for tissue classification. The combination of different categories of features may result in a large dimension of feature vector, including redundant features without contributing much to the discrimination power. In fact, a large set of features may possibly degrade the overall classification performance, a phenomenon known as “the curse of the dimensionality” [54]. To avoid this problem, the well-known dimensionality reduction methods such as principal component analysis (PCA) [31] and linear discriminant analysis (LDA) [68] can be used such that the feature dimensionality is reduced.

However, such kind of features projection will “hide” some feature elements that contribute (although not much) to the classification performance. Feature selection methods which aim at determining the optimal subset of features and providing the best classification performance have been widely-studied and used in histological applications [54]. Among many different methods, the sequential forward selection (SFS) method is well known and widely used in practice due to its good performance and relatively low computational cost [69]. The SFS algorithm starts with selecting the individual feature that maximizes an optimality criterion \mathcal{F} (e.g., the classification accuracy). Each consequent stage of the algorithm augments the set of already selected features by including a new feature such that the resulting feature subset maximizes \mathcal{F} . The process of adding new features is continued until \mathcal{F} reaches the maximum. The SFS algorithm has been used for features selection before classification of prostate cancer biopsies in [70].

2.3 Classification

After constructing a number of feature vectors, the classification of histopathological images is performed to judge the tissue biopsy and divide them into different categories. The classification can be performed by many classifiers such as k-nearest neighbors (KNN) classifier [58], LDA classifier [67], Bayesian classifier [46], SVM [26], [60], [66], [71], Neural network [24], Decision tree [24], [55], or combinations of different classifiers [72]. Among these classifiers, SVM is one of the most widely used classifiers mainly due to its excellent performance and efficiency. Let $\{x_i, y_i\}_{i=1}^N$ be the training data where $x_i \in R^n$ is a feature point and $y_i \in \{-1, 1\}$ is the corresponding label. The optimal hyperplane $w \cdot \phi(x) + b = 0$ for separating two classes by SVM can be obtained by minimizing the following function:

$$\min \frac{1}{2} \|w\|^2, \quad s.t., \quad y_i [w \cdot \phi(x_i) + b] \geq 1, \quad 1 \leq i \leq N \quad (2.9)$$

where w and b are the coefficients for the hyperplane, and $\phi(\cdot)$ is the kernel function (e.g., linear kernel $\phi(x_i) = x_i$). Once the optimal hyperplane is determined by Eq. (2.9), the input data points x_i are mapped to one of two classes by using the binary decision function $f(x_i) = \text{sgn}[w \cdot \phi(x_i) + b]$. By using the SVM classifier, Wang *et al.* [26] developed a computer assisted system to diagnose cervical intraepithelial neoplasia based on cytological features. Altunbay *et al.* [66] proposed to diagnose and grade colon cancers by using SVM classifier with tissue component structures computed from color graphs. Ali *et al.* [73] proposed to use SVM classifier with spatially aware cell cluster graphs to distinguish oropharyngeal p16+ tumors.

2.4 General WSIs Analysis and Classifications

The WSI is a high magnification digital image of an entire microscopic slide, which can provide global information of tissue specimen for quantitative image analysis. With recent development in computational powers, several systems on different cancers of diagnosis and grading based on the multi-resolution framework have been proposed. Table 2.3 shows the state-of-the-arts literatures about digitized WSIs analysis and classifications.

Petushi *et al.* [24] proposed a technique for grading breast cancer from histology images. They first partitioned each breast WSI into 64 sections, and then partitioned each of these section-images into 25 sub-images. For each sub-image, they employed gray scale conversion, adaptive thresholding and morphological operations to segment cell nuclei, and classified nuclei into three different types (e.g., lymphocytes, epithelial cells and cancer cells) using supervised learning. The nuclear densities and the number of segmented tubules in images are used as the features for classification by quadratic classifier. The classification accuracy of 75% has been reported on a database of 24 H&E stained WSIs.

Mete *et al.* [71] proposed a block-based supervised technique for detection of malignancy in head and neck histopathological images. The proposed technique first divides the training WSI into half-overlapping sub-images (e.g., 128×128

TABLE 2.3
RELATED WORKS ON VARIOUS CANCER DIAGNOSIS BY ANALYZING H&E STAINED WSIs

Reference	Dataset	Feature	Classifier	Accuracy
[24]	24 Breast, 40X	Architectural features of nuclei and tubules	Quadratic	75%
[71]	7 Head and Neck, 20X	Color features of HSV channels	SVM	96.25%
[26]	31 Cervical, 40X	Morphological and architectural features of nuclei	SVM	79.74%
[58]	43 Neuroblastoma, 40X	Textural features of image blocks	KNN	88.4%
[72]	33 Neuroblastoma, 40X	Textural features of cytoplasm and neuropil	Multiple*	87.88%
[61]	100 Prostate, 40X	Textural features	Bayesian	74%
[59]	66 Skin, 40X	Morphological and architectural features	SVM	89.07%

*Multiple refers to a combination of different classifiers such as KNN, Bayesian and SVM.

pixels), and then extracts prominent colors from malignant and normal training samples, respectively. These extracted colors are clustered into several groups that are used to train the SVM. For malignancy detection in a candidate image, the trained SVM is applied to classify the divided sub-images into different categories. A classification accuracy of 96.25% has been achieved on 7 H&E stained WSIs.

Wang *et al.* [26] developed a computer-aided system for the diagnosis of cervical intraepithelial neoplasia. The system first segments the squamous epithelium using texture features in conjunction with SVM, and then generates the perpendicular lines in the segmented epithelium (with respect to the medial axis). Based on the computed features (e.g., nuclear size and density) in image blocks along perpendicular lines, the SVM is employed to perform the classification. This CAD system achieves an overall 79.74% accuracy for four-class tissue classification on 31 H&E stained WSIs.

Sertel *et al.* [58] developed a multi-resolution CAD system for classification of stromal development in neuroblastoma. The system uses texture features extracted from co-occurrence statistics and local binary patterns of divided image tiles for classification. At each resolution level, the most discriminative features

are determined by the feature selection method that minimizes the leave-one-out classification error on the training set. A modified KNN classifier is employed to determine the confidence level of the classification at each resolution. The classification starts from the low resolution and switches to high resolution when necessary. Experimental results on 43 H&E stained WSIs show a 88.4% classification accuracy.

Kong *et al.* [72] developed a multi-resolution CAD system for grading of neuroblastic differentiation in WSIs. This system first divides the WSI into a set of non-overlapping image tiles (e.g., 500×500) with different resolutions, and then segments each image tile into multiple cytological components (e.g., nuclei, cytoplasm, neuropil) using an expectation maximization (EM) based approach. The statistical and Haralick features derived from the pre-segmented cytological components are fed to a multi-classifier combiner for training. The trained classification system is tested on 33 images and provides a 87.88% accuracy.

Doyle *et al.* [61] proposed a technique for detection of prostate cancer from digitized needle biopsies. The technique starts from the low resolution image, where the texture features such as Haralick and Gabor features are extracted. The extracted features are fed into a set of boosted Bayesian classifiers to generate a mask with cancerous tissues as foreground. The segmented cancerous tissues are then mapped to high resolution for another round of classification. This process repeats until image resolution is sufficient for application of advanced region-based grading systems. Experiments on 100 H&E stained WSIs show a 74% classification accuracy.

Lu *et al.* [59] proposed a CAD system for classification of skin histopathological images. The system first down-samples the WSI and segments the skin epidermis area at the low resolution level. The segmented epidermis is then mapped to the high resolution such that nuclei and melanocytes within epidermis are segmented. Based on the morphological features and architectural features of melanocytes in the epidermis area, the skin image is classified into different categories (e.g., melanoma, nevi, normal) by SVM. A classification accuracy of 89.07% has been

obtained based on 10-fold cross validation on 66 H&E stained WSIs.

2.5 Summary

In this chapter, we presented a brief review of automated image analysis techniques in histopathology, which includes ROI detection and segmentation, feature extraction and selection, and image classification. Several representative works on WSIs analysis and classification with different types of cancers (including skin cancer) are also provided.

Chapter 3

Automated Epidermis and Dermis Segmentations

Automatic segmentation of epidermis and dermis areas in skin histopathological images is an essential step for computer-aided diagnosis of various skin cancers including skin melanoma. However, it is a challenging problem to accurately segment epidermis and dermis regions from digitized WSIs due to the presence of many different structures within skin tissues and the large volume of image size. In this chapter, we first propose an epidermis segmentation technique that segments skin epidermis based on thickness measurement and k-means classification. We then propose a dermis segmentation technique that segments skin dermis based on morphological operations and the predefined depth of interest. Experiments and comparisons indicate that the proposed techniques provide a good performance in epidermis and dermis segmentations.

Part of this chapter has been published in: Xu *et al.* “Epidermis segmentation in skin histopathological images based on thickness measurement and k-means algorithm”, *EURASIP Journal on Image and Video Processing*, vol.2015, no.1, pp.1-14, 2015. Xu *et al.* “Automated segmentation of regions of interest in whole slide skin histopathological images”, *Proc. IEEE Engineering in Medicine and Biology Society (EMBC)*, pp.3855-3858, August 2015.

3.1 Introduction

Fig. 3.1 shows a skin WSI stained with H&E. As observed in Fig. 3.1, a typical digitized skin slide can be divided into three main parts: epidermis, dermis and subcutaneous tissues. The epidermis area has dark purple color due to the large amount of absorption of the haematoxylin by cell nuclei. On the other hand, the dermis area is mainly presented the pink color due to the large absorption of eosin whereas the subcutaneous tissues have net-like pattern with white color background. The diagnosis of skin melanoma can generally be made by analyzing the architectural and morphological features of atypical cells in the epidermis and DEJ areas [18]. In addition, epidermis segmentation helps in identifying the relative positions between carcinoma cells and epidermis boundaries. The invasion depth of carcinoma cells into the skin tissue can be measured, which is a critical indicator for skin cancer grading and therapy [21]. Therefore, automatic segmentation of epidermis and dermis areas is an important step in melanoma diagnosis by analyzing histopathological images.

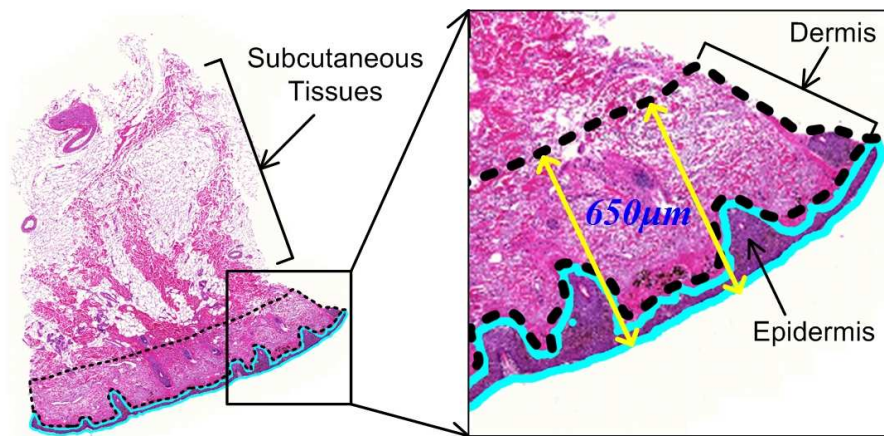


Fig. 3.1. Skin tissue digitized slide. Note that a selected image patch is zoomed in for better viewing. The epidermis is manually labeled with cyan contours, while the dermis with a depth of $650\mu m$ is labeled with dashed black contours.

Recently a few techniques for epidermis segmentation have been proposed. Lu *et al.* [74] proposed a Global Thresholding and Shape Analysis based technique (henceforth referred to as the GTSA technique) that segments epidermis area in skin

histopathological images. The GTSA technique first down-samples WSIs with 40X magnification by a factor of 16, and then performs epidermis segmentation on the red channel of the down-sampled image. Haggerty *et al.* [75] presented a Contrast Enhancement and Thresholding based technique (henceforth referred to as the CET technique) for epidermal tissue segmentation in WSIs with 10X magnification. Unlike GTSA technique, the CET technique performs global thresholding on a contrast enhanced composite image (which is a combination of Y and b^* channel). Both GTSA and CET techniques assume that there are small number of cell nuclei present in the dermis area, and they eliminate noisy regions by shape and area analysis. Mokhtari *et al.* [21] developed a system for measuring melanoma depth of invasion in microscopic images, which includes the segmentation of epidermal layer. The epidermal layer is segmented by a Morphological Closing and Global Thresholding based technique (henceforth referred to as the MCGT technique). The MCGT technique makes an assumption that the morphological closing operation can remove all low intensity components in the dermis area (e.g., cell nuclei and other skin components). However, it is usually difficult to define an appropriate structuring element for closing operation to remove all low intensity components in the dermis area and keep epidermis unchanged when dealing with the WSIs. Note that since all existing epidermis segmentation techniques are mainly relied on thresholding, they usually fail to provide a high precision when different dark skin components (e.g., cell nuclei, hair follicles) are present in the dermis area.

In this chapter, we propose an epidermis segmentation technique, which has overcome the limitations (i.e., low segmentation precision) of existing techniques. Based on epidermis segmentation, we further propose a dermis segmentation technique, which is, to our knowledge, the first technique to segment skin dermis in skin WSIs. The details of the proposed techniques and experimental results are provided in following sections.

3.2 Image Dataset

The dataset used for evaluating epidermis and dermis segmentation performance includes histopathological images from formalin-fixed paraffin-embedded tissue blocks of skin biopsies. The sections prepared are about $4\mu m$ thick and are stained with H&E using automated stainer. The original digital WSIs were captured under 40X magnification on Carl Zeiss MIRAX MIDI Scanning system. All images were obtained from University of Alberta hospital. For examples of skin WSIs, please refer to the Appendix A.3. Since the original WSIs have a large volume size (e.g., several gigabytes) and are difficult for real time processing, these images were down-sampled by a factor of 16 and saved into TIFF format using MIRAX Viewer software. Overall the image dataset consists of 66 different skin WSIs with the resolution between $2,500 \times 3,000$ and $6,000 \times 10,000$ pixels. Table 3.1 lists the distributions of different tissue biopsies included in the dataset.

TABLE 3.1
IMAGE DATASET USED FOR EVALUATING EPIDERMIS AND DERMIS SEGMENTATIONS.

Tissue biopsies	No.	Pct.
Normal	17	25.7%
Nevus	17	25.7%
Melanoma	32	48.6%

3.3 Proposed Epidermis Segmentation

The schematic of the proposed technique for epidermis segmentation is shown in Fig. 3.2. The technique has three modules. In the first module, the epidermis coarse segmentation is performed based on thresholding and shape analysis. In the second module, the thickness of coarsely segmented epidermis area is measured using line segments perpendicular to the main axis of the epidermis mask. The coarsely segmented result is evaluated based on the measured epidermis thickness. In the third module, a second-pass fine segmentation by an unsupervised clustering

algorithm is performed on the epidermis region with the poor quality segmentation result. The three modules are now presented in details in the following.

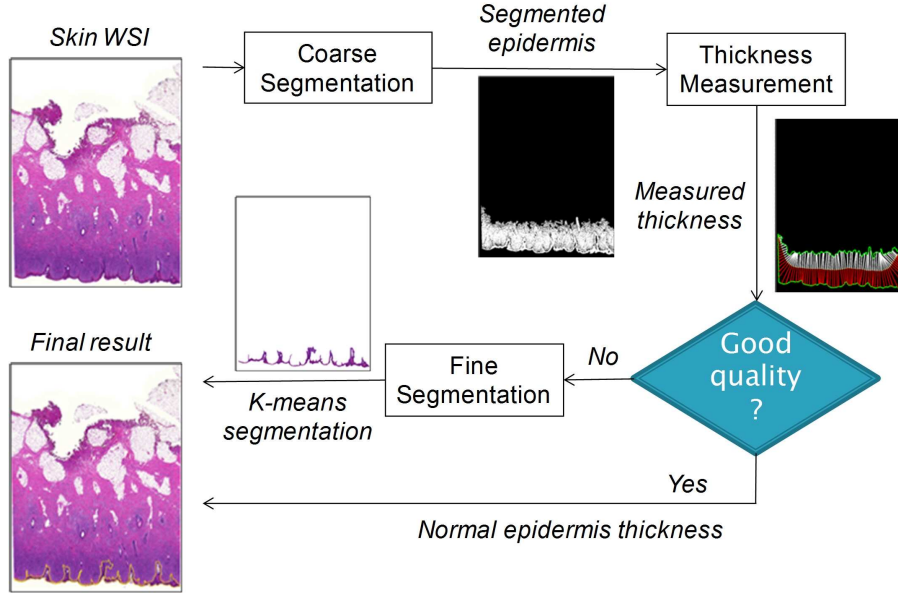


Fig. 3.2. Schematic of epidermis segmentation.

3.3.1 Coarse Segmentation

Given an RGB image I_l , the red channel R_l is selected for the epidermis coarse segmentation, since the red channel of the H&E stained skin histopathological image provides good distinguishable information [74]. With the red channel image R_l , the epidermis coarse segmentation is performed as follows:

(1) Removing white background pixels: In this step, we empirically select a threshold τ_1 (e.g., $\tau_1 = 240$) to separate skin tissues from the background (which are typically white). The pixels in R_l with gray values smaller than τ_1 are classified as the foreground. Let the foreground pixels be denoted by $\{F_k\}_{k=1\dots M}$, where M is the number of pixels.

(2) Applying global thresholding: The Otsu's thresholding technique [51] is applied to group the pixels $\{F_k\}_{k=1\dots M}$ into two classes. A binary mask b_0 is

generated as follows:

$$b_0(i, j) = \begin{cases} 1 & \text{if } F_k \leq \tau_2 \\ 0 & \text{if } F_k > \tau_2 \end{cases} \quad (3.1)$$

where (i, j) is the 2-D coordinate of the pixel F_k in R_l , τ_2 is the threshold obtained by the Otsu's technique.

(3) Eliminating false regions: We label all the regions in the binary mask b_0 via 8-connected criterion. Let the 8-connected regions in b_0 be denoted by $\{C_k\}_{k=1 \dots O}$ where O is the number of connected regions. For each region C_k , we calculate the area C_{area} , the major axis length r_{maj} and minor axis length r_{min} of the best fit ellipse [64]. A binary mask b_1 with epidermis regions as the foreground is determined as follows:

$$b_1(C_k) = \begin{cases} b_0(C_k), & \text{if } (C_{area} > T_{area}) \wedge (r_{maj}/r_{min} > T_{ratio}) \\ 0, & \text{otherwise} \end{cases} \quad (3.2)$$

where $b_0(C_k)$ represents the pixels of the region C_k in b_0 , \wedge is the AND operation, T_{area} and T_{ratio} are the predefined thresholds. Note that T_{area} is used to remove small noisy regions in b_0 , while T_{ratio} is used to select the epidermis region that has a long and narrow shape after global thresholding [74][75]. In this work, T_{area} and T_{ratio} values are determined based on the domain prior knowledge and experiments on training images. Specially, we set the thresholds as $T_{area} = 0.006M$, and $T_{ratio} = 3$. For more details, please refer to parameters selection in the section of epidermis segmentation evaluations.

Fig. 3.3 shows two examples of both intermediate and final coarse segmentation results. Figs. 3.3(d)(h) show the segmented epidermis regions (b_1) corresponding to Figs. 3.3(a)(e), respectively. Note that Fig. 3.3(d) shows a good quality segmentation, whereas Fig. 3.3(h) shows a poor quality (incorrect) segmentation where the false positive region is highlighted by the manually labeled (red) contour.

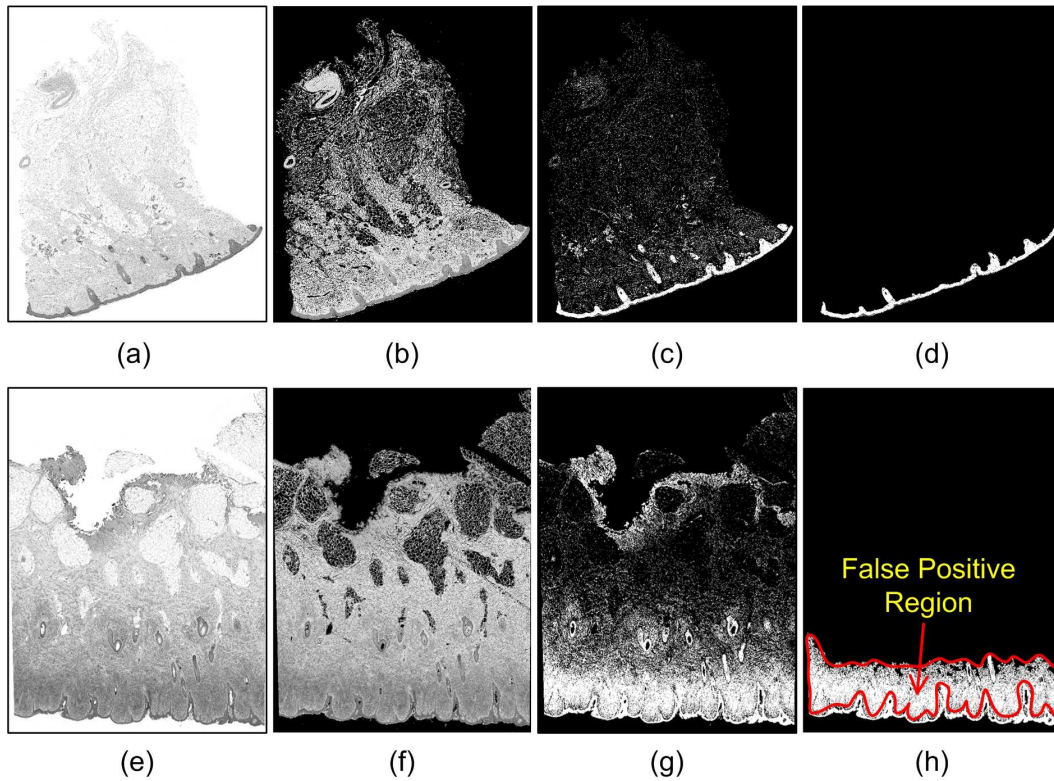


Fig. 3.3. Two examples of epidermis coarse segmentation. (a) & (e) Red channel images. (b) & (f) Images after removing background pixels. (c) & (g) Binary images after global thresholding. (d) & (h) Final binary masks. Note that white regions in (d) and (h) correspond to segmented epidermis areas.

3.3.2 Thickness Measurement

It is observed in Fig. 3.3 that coarse segmentation module may result in both good and poor quality segmentations. With a pixel resolution of $3.72 \mu m/\text{pixel}$, the segmented epidermis as shown in Fig. 3.3(d) on average has a thickness of 52 pixels (or 0.19 mm), whereas the segmented epidermis as shown in Fig. 3.3(h) has a thickness of 276 pixels (or 1.03 mm). The skin epidermis varies in thickness in different parts of a body but should be within a limited range [76]. In our database, the epidermis of skin histopathological images roughly has a thickness of $0.1\text{-}0.4 \text{ mm}$, and hence a second-pass segmentation can be carried out based on thickness measurement. In this module, we measure the thickness of the coarsely segmented result to classify it as good or poor quality segmentation. The steps of thickness measurement are detailed below.

(1) Morphological preprocessing: In order to smooth the boundaries of the epidermis area, the morphological closing operation is first performed on the mask b_1 as follows:

$$b_2 = b_1 \bullet S \quad (3.3)$$

where \bullet is the morphological closing operator, and S is the structuring element. In this work, a disk shape structuring element with a radius of 10 pixels is empirically selected for the closing operation. Next, the holes within the mask b_2 are filled by performing the morphological reconstruction operation:

$$b_3 = [\mathfrak{S}(b_2^c, b_m)]^c \quad (3.4)$$

where \mathfrak{S} is the morphological reconstruction operator [31], b_2^c is the complement of b_2 , and b_m is the marker image which is set to be 0 everywhere except on the image border, where it is set to be b_2^c . Fig. 3.4(a) shows a mask b_1 [cropped from Fig. 3.3(h)], and Figs. 3.4(b)(c) show the corresponding b_2 and b_3 .

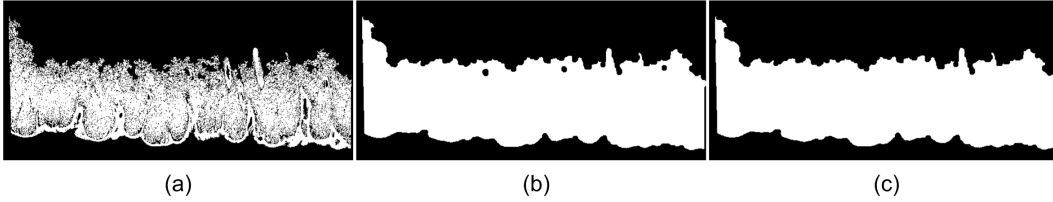


Fig. 3.4. Illustration of morphological preprocessing. Epidermis mask (a) b_1 . (b) b_2 . (c) b_3 . Note that figures (a)-(c) are cropped from the whole slide image and zoomed up for clear illustration here.

(2) Thinning of epidermis mask: This step reduces the epidermis area in the mask b_3 to a connected stroke (a thin line) that is only a single pixel wide. The connected stroke can be considered as the skeleton of the epidermis area. To obtain the connected stroke, the parallel thinning algorithm [77] is performed on the mask b_3 . The algorithm is executed in a number of iterations until the generated mask b_4 stops changing. Fig. 3.5(a) shows the generated epidermis skeleton in the mask b_4 superimposed on the mask b_3 .

(3) End points extraction: After generating the mask b_4 , the end points of the epidermis skeleton are detected by a lookup table (LUT) technique [31]. A LUT

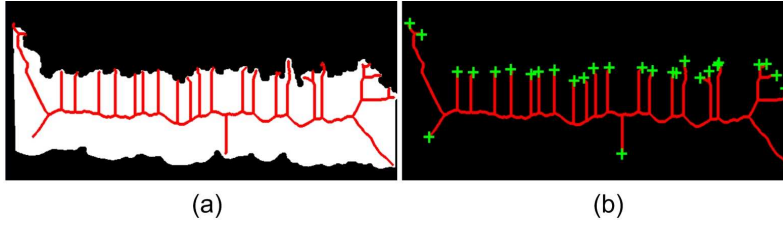


Fig. 3.5. Epidermis skeleton and end points. (a) Skeleton b_4 superimposed on epidermis mask b_3 . (b) Epidermis skeleton with end points marked by + symbols.

is first constructed based on the observation that an end point (in the epidermis skeleton) has exactly one foreground neighbor. The mask b_4 is then processed by using the generated LUT to extract end points of the epidermis skeleton. Let the end points be denoted by $\{E_k\}_{k=1\dots N}$, where N is the number of end points. In Fig. 3.5(b) the end points are marked with + symbols.

(4) Main axis identification: It is observed in Figs. 3.5(a)(b) that there are many branches in the epidermis skeleton. The longest path joining two end points on the skeleton reflects the main axis of the mask b_3 . In this step, we calculate all paths joining each possible pair of end points on the skeleton, and select the longest path as the main axis. Given two arbitrary end points E_i and E_j , let the geodesic distance (i.e., the number of pixels on the shortest path connecting E_i and E_j) be denoted by D_{ij} . The main axis is calculated as follows:

Step 1: Calculate all possible D_{ij} based on the geodesic distance transform [78], where $1 \leq i, j \leq N$.

Step 2: Select the longest geodesic distance among all possible D_{ij} and consider the corresponding constrained path as the main axis.

Step 3: Smooth the main axis by using a moving average filter of length 200 pixels.

Note that there are usually a large number of end points, and hence it may be computationally expensive to calculate all possible D_{ij} in Step 1. As observed in Fig. 3.5(b), the pair of end points corresponding to the longest constrained path usually has a relatively long Euclidean distance. In order to speed up the main axis identification, we calculate the Euclidean distance between all possible end points,

and select a short list of pairs (e.g., 10 pairs) based on (large) Euclidean distance. The main axis identification can then be efficiently performed by applying steps 1-3 on the selected pairs.

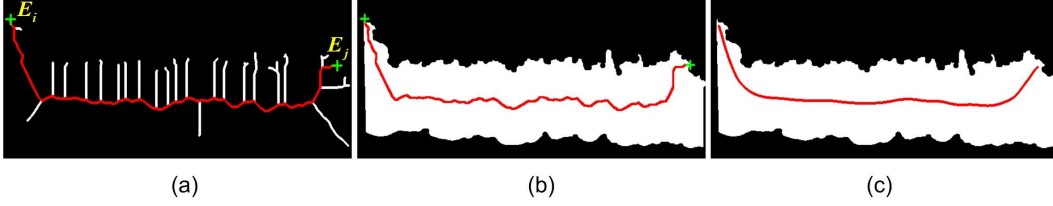


Fig. 3.6. Illustration of main axis identification. (a) A constrained path joining E_i and E_j . Epidermis mask b_3 with main axis (b) before smoothing. (c) after smoothing.

Let the obtained main axis be denoted by the points set $\{Z_k\}_{k=1\dots Q}$ where Q is the number of points on the main axis. Fig. 3.6 illustrates the main axis identification with an example. Fig. 3.6(a) shows a constrained path (the red line) joining points E_i and E_j . Figs. 3.6(b)(c) show the epidermis main axis before smoothing and after smoothing, respectively, superimposed on the mask b_3 .

(5) Epidermis thickness calculation: In this step, we first calculate the gradient image of the mask b_3 (obtained by Eq. 3.4). The epidermis boundary points correspond to the points with non-zero gradients (on the binary mask). Let this boundary points set be denoted by $\{A_k\}_{k=1\dots W}$ where W is the number of points. We then calculate the epidermis thickness based on the epidermis main axis and epidermis boundary points. Note that there are Q points on the main axis. In order to reduce the computational complexity, we calculate the epidermis thickness using selected points on the main axis. In this work, a set of r points, $\{Z_k\}_{k=h,2h,\dots,rh}$ where $r = \lfloor \frac{Q}{h} \rfloor$, is selected. Fig. 3.8(a) shows the epidermis contour with r selected points on the main axis, where $h = 3$. To calculate the epidermis thickness, a perpendicular line for each selected point on the main axis is defined. Given a point $Z_k(x_k, y_k)$ (see Fig. 3.7), the steps to calculate the local thickness are as follows.

Step 1: Let f_k denote the directed line passing through points $Z_{k-1}(x_{k-1}, y_{k-1})$ and $Z_{k+1}(x_{k+1}, y_{k+1})$. Note that the direction is from the point Z_{k-1} to Z_{k+1} . The

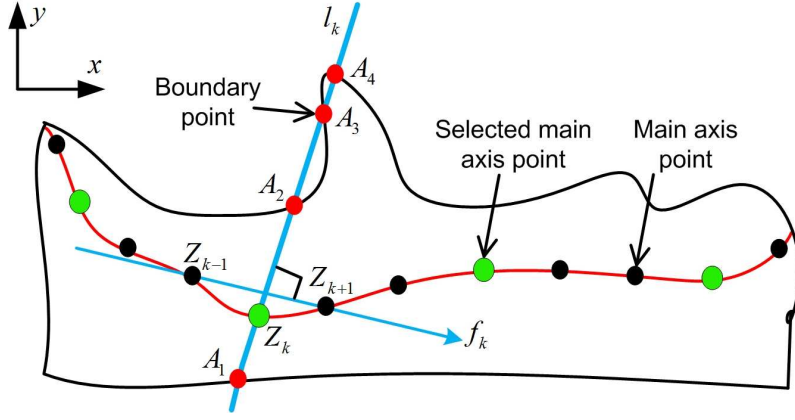


Fig. 3.7. Example of epidermis local thickness measurement (with $h = 3$).

slope s_k of the line l_k perpendicular to f_k is computed as follows:

$$s_k = \begin{cases} 0, & \text{if } x_{k+1} = x_{k-1} \\ \infty, & \text{if } y_{k+1} = y_{k-1} \\ \frac{x_{k-1} - x_{k+1}}{y_{k+1} - y_{k-1}}, & \text{otherwise} \end{cases} \quad (3.5)$$

Step 2: The intersection points between the line l_k and the epidermis boundary $\{A_k\}_{k=1 \dots W}$ are calculated. A boundary point $A_k(u_k, v_k)$ is considered to be on the perpendicular line l_k if it satisfies the following condition:

$$\left| \tan^{-1}(s_k) - \tan^{-1}\left(\frac{v_k - y_k}{u_k - x_k}\right) \right| \leq \eta \quad (3.6)$$

where η is a small positive number (e.g., $\eta = 0.05$) to allow for a small error in intersection points calculation. Note that for an arbitrary point Z_k , there will be two or more intersection points. For example, in Fig. 3.7 the line l_k intersects with the epidermis contour at four points A_1, A_2, A_3 and A_4 .

Step 3: The directed line f_k divides the intersection points (e.g., A_1, A_2, A_3 and A_4) into two groups: right side points (RSP) and left side points (LSP). Note that RSP and LSP are seen from the direction of the line f_k . The position of a point $A_k(u_k, v_k)$ with respect to the line f_k is determined by the following equation:

$$\varphi = u_k \alpha + v_k \beta + \gamma \quad (3.7)$$

where $\alpha = y_{k+1} - y_{k-1}$, $\beta = x_{k-1} - x_{k+1}$, $\gamma = x_{k+1}y_{k-1} - x_{k-1}y_{k+1}$. If $\varphi < 0$, the point belongs to RSP (i.e., A_k is located on the right side of f_k); if $\varphi > 0$, the point belongs to LSP; if $\varphi = 0$, the point is on the f_k . In Fig. 3.7, the points A_2, A_3, A_4 are in the LSP, whereas the point A_1 is in the RSP.

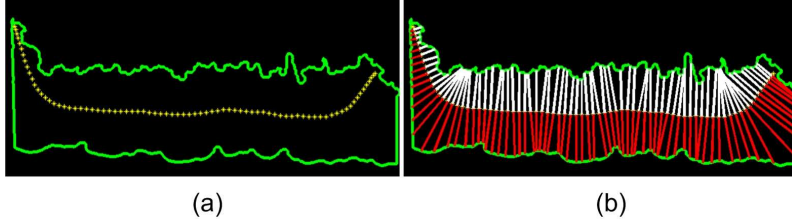


Fig. 3.8. Illustration of epidermis thickness measurement. (a) Epidermis contour with selected points on the main axis ($h = 20$). (b) Line segments measuring epidermis thickness.

Step 4: The local thickness t_k for a point Z_k is computed as follows:

$$t_k = \min \{ \|A_i - A_j\| \}, A_i \in RSP, A_j \in LSP \quad (3.8)$$

where $\|\cdot\|$ is the 2-D Euclidean distance between points A_i and A_j . In Fig. 3.7, the Euclidean distance between points A_1 and A_2 is computed as the local thickness t_k .

Likewise, the local thicknesses $\{t_k\}_{k=h,2h,\dots,rh}$ for all selected points on the main axis are calculated by using steps 1-4. Fig. 3.8(b) shows the line segments measuring epidermis thickness.

(6) Segmentation quality evaluation: The quality of coarse segmentation result is evaluated based on the average value of measured epidermis thickness, which is as follows:

$$\rho = \begin{cases} 1 & \text{if } \bar{t} < \tau_3 \\ 0 & \text{otherwise} \end{cases} \quad (3.9)$$

where $\bar{t} = \frac{1}{r} \sum_{k=1}^r t_{kh}$, τ_3 is a threshold value and ρ is a parameter to indicate the coarse segmentation quality. Note that the threshold τ_3 , which corresponds to the maximum epidermis thickness, is determined based on experiments on training images (please see parameters selection in the section of epidermis segmentation evaluations). In this work, we set the threshold τ_3 as 150 pixels (or 0.55 mm). For a good quality segmentation, $\rho = 1$, whereas for a poor quality segmentation,

$\rho = 0$, which needs to be enhanced by the fine segmentation module.

3.3.3 Fine Segmentation

The coarse segmentation results are classified into good and poor quality segmentations based on thickness measurement. In this module, we consider the poor quality segmentation for further analysis in order to obtain a more accurate segmentation.

When $\rho = 0$, it is likely that some dermis pixels are incorrectly classified as epidermis pixels. In order to obtain a more accurate segmentation, it is necessary to conduct a second-pass fine segmentation to divide the pixels into two classes (e.g., epidermis and dermis pixels). To obtain a robust performance, we perform the second-pass fine segmentation by using $\{R, G, B\}$ color channels. Due to the possible variations in the color spectrum between different digitized slides, k-means classification [52], which is an unsupervised classification algorithm, is selected to perform the fine segmentation. The $\{R, G, B\}$ values of the pixels that are binary true in the coarsely segmented epidermis area (e.g., the mask b_1) are taken from the image I_l and used as clustering attributes. The k-means algorithm divides the pixels into 2 classes based on their attributes (e.g., $\{R, G, B\}$ color values) by iteratively minimizing the following cost function:

$$J = \sum_{j=1}^2 \sum_{i=1}^{n_j} \|x_i^j - c_j\|^2 \quad (3.10)$$

where n_j is the number of pixels in the class j , x_i^j is the i th pixel in the class j , and c_j is the centroid of the class j . Note that the number of classes is set as 2 that corresponds to dermis and epidermis respectively.

Fig. 3.9(a) reproduces the coarse segmentation result shown in Fig. 3.4(a) in color. Figs. 3.9(b)(c) show two classes of pixels obtained by the k-means algorithm. It is observed that the class with epidermis pixels has relatively darker color (i.e., lower R,G,B values) than the class with dermis pixels. The two classes can be

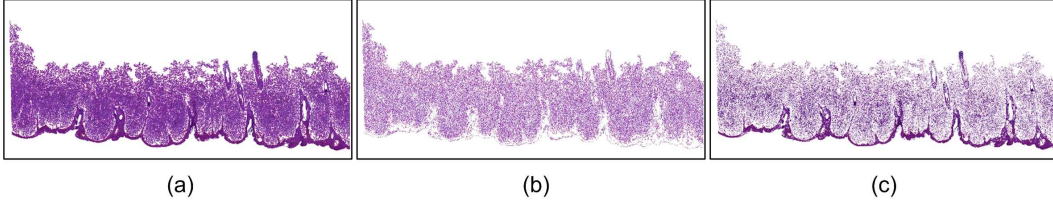


Fig. 3.9. Illustration of k-means classification. (a) Coarse segmentation result Fig. 3.4(a) in color. (b) Class with dermis pixels ($k^* = 2$). (c) Class with epidermis pixels ($k^* = 1$).

identified as follows:

$$k^* = \begin{cases} 1 & \text{if } (\overline{R}_1 + \overline{G}_1 + \overline{B}_1) < (\overline{R}_2 + \overline{G}_2 + \overline{B}_2) \\ 2 & \text{otherwise} \end{cases} \quad (3.11)$$

where $(\overline{R}_1, \overline{G}_1, \overline{B}_1)$ and $(\overline{R}_2, \overline{G}_2, \overline{B}_2)$ are the centroids of the two classes. Note that for the class with epidermis pixels, $k^* = 1$, while for the class with dermis pixels, $k^* = 2$;

The foreground pixels shown in Fig. 3.9(c) are considered to be epidermis pixels according to the Eq (3.11). However, it is observed in Fig. 3.9(c) that a number of low intensity pixels in the dermis area are classified as epidermis pixels. Note that most of the false positive pixels (belonging to dermis area) are isolated pixels, or correspond to regions with smaller area. Therefore, false positive pixels can easily be eliminated by the area opening operation. Regions that have areas below the threshold T_{area} (see the coarse segmentation module) are removed. Finally, the morphological closing operation with a disk shape structuring element (with a radius of 5 pixels) is performed to smooth the epidermis area, and the holes within the epidermis area are filled by the reconstruction operation. Fig. 3.10(a) shows the finally obtained epidermis region. Fig. 3.10(b) shows the epidermis contour on the original image.

3.4 Epidermis Segmentation Evaluations

In this section, we present the performance of the proposed technique and compare it with the performance obtained by existing techniques.

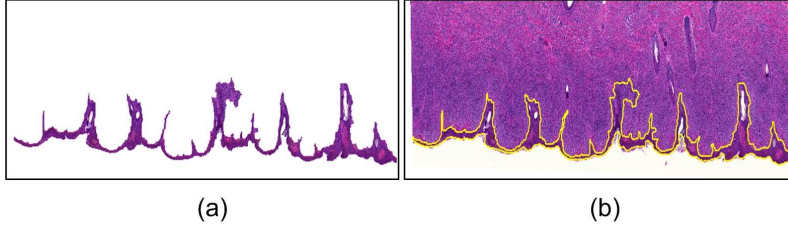


Fig. 3.10. Fine segmentation result. (a) Finally segmented epidermis area. (b) Epidermis contour on the original image.

3.4.1 Evaluation Metrics

The automatic epidermis segmentation results are compared to the ground truth segmentations obtained by visual inspection. The evaluations are performed by computing area based metrics [74] namely precision (\mathcal{A}_{PRE}), sensitivity (\mathcal{A}_{SEN}) and specificity (\mathcal{A}_{SPE}), and boundary based metrics [79] namely Hausdorff distance (\mathcal{D}_{HD}) and mean absolute distance (\mathcal{D}_{MAD}). We denote the manually obtained boundary as $g = \{c_i^g | i \in (1, 2, \dots, m)\}$, and the boundary of the automatic segmentation as $s = \{c_j^s | j \in (1, 2, \dots, n)\}$, where m and n are the numbers of the ground truth and automatically segmented boundary points, respectively. The area based metrics are defined as follows:

$$\mathcal{A}_{PRE} = \frac{|\mathfrak{R}(s) \cap \mathfrak{R}(g)|}{|\mathfrak{R}(s)|} \times 100\% \quad (3.12)$$

$$\mathcal{A}_{SEN} = \frac{|\mathfrak{R}(s) \cap \mathfrak{R}(g)|}{|\mathfrak{R}(g)|} \times 100\% \quad (3.13)$$

$$\mathcal{A}_{SPE} = \frac{|\mathfrak{R}^c(s) \cap \mathfrak{R}^c(g)|}{|\mathfrak{R}^c(g)|} \times 100\% \quad (3.14)$$

where $\mathfrak{R}(\cdot)$ is the area of the closed boundary, $|\cdot|$ is the cardinality operator, \cap is the intersection operation and $\mathfrak{R}^c(\cdot)$ is the complement of $\mathfrak{R}(\cdot)$. To evaluate the automatically segmented boundary contours, we calculate the distance of every point in g from all points in s . The boundary based metrics are defined as follows:

$$\mathcal{D}_{HD} = \max_i \left[\min_j \|c_j^s - c_i^g\| \right] \quad (3.15)$$

$$\mathcal{D}_{MAD} = \frac{1}{m} \sum_{i=1}^m \left[\min_j \|c_j^s - c_i^g\| \right] \quad (3.16)$$

Note that the Hausdorff distance (\mathcal{D}_{HD}) measures the worst possible disagreement between two contours. The mean absolute distance (\mathcal{D}_{MAD}) estimates the disagreement averaged over the two contours.

3.4.2 Parameters Selection

There are 66 different skin histopathological images in the whole dataset, which are provided with ground truth segmentations by pathologists. The 66 WSIs consist of three categories: 17 normal skins, 17 melanocytic nevi and 32 skin melanomas. Note that there are three parameters that should be selected appropriately in the proposed technique, which includes T_{area} , T_{ratio} (thresholds for eliminating false positive regions in Section 3.3.1), and τ_3 (threshold to determine the coarse segmentation quality in Section 3.3.2). To determine the values of these parameters, we randomly selected 6 normal skins, 6 melanocytic nevi and 8 skin melanomas as training images. The 20 training images were randomly selected from three categories to avoid any bias. The other 46 images were taken as testing images, which were used as an independent validation set. The values of training parameters are shown in Table 3.2. We explain the process of determining parameter values in the following.

TABLE 3.2
TRAINING PARAMETERS OF EPIDERMIS SEGMENTATION TECHNIQUE.

Modules	Parameters	Values
Coarse segmentation	T_{area}	0.006M pixels
Coarse segmentation	T_{ratio}	3
Thickness measurement	τ_3	150 pixels

To determine an adaptive threshold value for T_{area} , we calculate the portion of epidermis pixels about skin tissue pixels in training images. It has been found that the portion of epidermis pixels ranged between 0.7% and 6%, and hence the

threshold T_{area} is set as $0.006M$ where M is the number of foreground pixels (i.e., skin tissue pixels) in the WSI. Similarly, we calculate the ratio r_{maj}/r_{min} for all ground truth epidermis regions in training images, and the r_{maj}/r_{min} values have been found to be in the range between 3.3 and 26.6. Therefore, the threshold T_{ratio} is set as 3.

The parameter τ_3 is determined based on experiments on training images in this work. Based on visual examination, the coarsely segmented result of training images are divided into two groups: subsets A and B. In subset A (11 WSIs), the segmented results are quite similar to ground truths, while in subset B (9 WSIs) a large number of false positive pixels in dermis area are classified as epidermis pixels. The coarsely segmented masks of subset B have markedly large thickness than masks of subset A. Table 3.3 shows the performance evaluations of subsets A, B by the area based metrics, and the corresponding average epidermis mask thickness \bar{x} . As observed in Table 3.3, the segmentation precision for the subset B is significantly low, only 38.69%. The average thickness \bar{x} for subset B is 211.60 pixels, which is much higher than 63.26 pixels for subset A. The boxplot for the epidermis thickness of subsets A and B is shown in Fig. 3.11. As observed in Fig. 3.11, the maximum thickness in subset A is smaller than 150 pixels, while the minimum thickness in subset B is larger than 150 pixels. Therefore, the threshold τ_3 is finally set as 150 pixels.

TABLE 3.3
PERFORMANCE EVALUATIONS OF EPIDERMIS COARSE SEGMENTATION IN SUBSETS OF TRAINING IMAGES.

Subsets	$\mathcal{A}_{PRE}(\%)$	$\mathcal{A}_{SEN}(\%)$	$\mathcal{A}_{SPE}(\%)$	$\bar{x}(pixels)$
A	98.11	97.04	99.96	63.26
B	38.69	99.83	93.70	211.60

In order to test how sensitivities are the parameter values to the choice of training images, we selected another set of 20 skin images randomly (from the testing images) and calculated the values of T_{area} , T_{ratio} and τ_3 following a similar process of parameters selection. Experiments show that the values of these three parameters

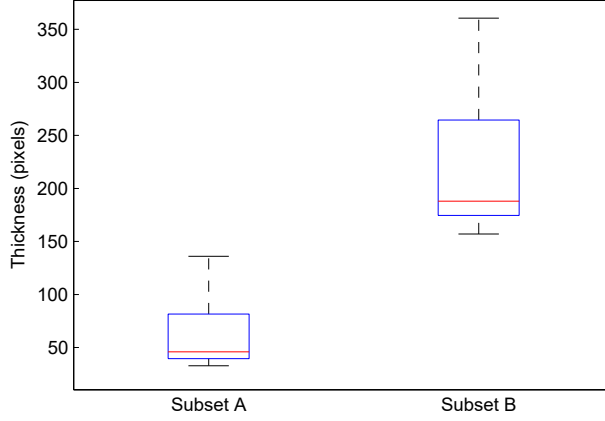


Fig. 3.11. Thickness variations of coarsely segmented epidermis masks in subsets A and B of training images. Subset A (B) includes images with correct (incorrect) segmentations after coarse segmentation module.

only have marginal variations ($T_{area} = 0.007M$, $T_{ratio} = 3$ and $\tau_3 = 155$). In other words, the parameter values do not fluctuate too much across databases.

3.4.3 Quantitative Results

To illustrate the efficacy of the proposed epidermis segmentation technique, the performance of the proposed technique is compared with the existing epidermis segmentation techniques including the GTSA [74], CET [75] and MCGT [21] techniques. The GTSA technique has two parameters T_{area} and T_{ratio} , which were set the same values as our proposed technique. The CET technique has several key parameters including the low output thresholds for contrast enhancement, the sizes of smoothing mean filter and morphological operations and the thresholds to eliminate noisy regions after thresholding. For the parameters (e.g., the size of smoothing filter) that are not used in the proposed technique, we set them following the work in [75]. While for parameters (e.g., T_{area} used to eliminate noisy regions) that are used in the proposed technique, we set them the same values as our proposed technique. The MCGT technique has only one key parameter that is the radius of the circular structuring element for the morphological closing operation. To determine an optimal radius value, we selected a set of values from 20 to 50 with a step of 5 to do experiments. 30 is finally determined as the radius of the structuring element, as

it provides the best performance of epidermis segmentation in our training images.

The average results of quantitative evaluations by Eqs. (3.12)-(3.16) on both training and testing sets are shown in Table 3.4. It is observed in Table 3.4 that the proposed technique provides an overall superior performance in epidermis segmentation than the existing techniques. Although the sensitivities of the proposed technique (90.39% and 92.78%) are marginally lower than those of the GTSA [74] technique, the proposed technique achieves much higher precisions (98.69% and 96.53%), roughly 20% higher than the GTSA technique. The k-means algorithm used by fine segmentation module of the proposed technique incorrectly classifies a small number of epidermis pixels as dermis pixels, which results in the marginal drop in sensitivity. The poor performances of the GTSA and CET techniques are mainly because a large number of dermis pixels are incorrectly classified as epidermis pixels in images where there are a large number of cell nuclei in the dermis area (see Figs. 3.13(c)(d)). The cell nuclei in the dermis area appear dark purple, and global thresholding incorrectly considers them as epidermis pixels. In addition, the CET technique [75] applies global thresholding on an equally weighted linear combination of the Y (i.e., gray) and b^* channel (i.e., b^* in $L^*a^*b^*$ color space) images, which provides a poor performance than using the red channel in the our database. The performance of the MCGT [21] technique is much poorer than that of the other techniques, as it does not work on skin WSIs which include epidermis, dermis and sebaceous areas. The MCGT technique assumes that the morphological closing operation can remove all unrelated components (typically dark appearance) in the skin dermis area, and hence the epidermis area can be segmented out by thresholding. However, the dermis areas of WSIs contain many different dark skin components such as hair follicles, sweat glands and nuclear clumps. Since the size of different skin components may vary greatly, it is difficult to define an appropriate structuring element for closing operation which can remove all unrelated skin tissues and keep the epidermis area unchanged. It is also noted from the Table 3.4 that the proposed technique has achieved relatively smaller \mathcal{D}_{HD} and \mathcal{D}_{MAD} values in both training and testing sets, and hence the proposed

technique provides a better matching between the ground truth contours and the automatically segmented contours.

TABLE 3.4
QUANTITATIVE EVALUATIONS OF EPIDERMIS SEGMENTATION BETWEEN EXISTING
TECHNIQUES AND PROPOSED TECHNIQUE.

Datasets	Techniques	$\mathcal{A}_{PRE}(\%)$	$\mathcal{A}_{SEN}(\%)$	$\mathcal{A}_{SPE}(\%)$	\mathcal{D}_{HD}	\mathcal{D}_{MAD}
Training (20 WSIs)	MCGT [21]	29.12	76.59	90.14	147.99	26.31
	CET [75]	56.53	91.44	95.14	143.45	23.75
	GTSA [74]	75.01	98.13	97.53	140.25	12.43
	Proposed	98.69	90.39	99.98	130.16	7.71
Testing (46 WSIs)	MCGT [21]	27.61	77.41	86.26	152.63	27.41
	CET [75]	49.91	91.25	93.84	139.39	24.33
	GTSA [74]	77.82	98.42	97.15	117.37	13.82
	Proposed	96.53	92.78	99.84	86.83	6.99

For further comparison of the proposed epidermis segmentation technique with existing techniques, the thickness of automatically segmented epidermis masks of different techniques was measured by the proposed thickness measurement method (see thickness measurement module), and compared with the thickness of manually labeled epidermis masks. Fig. 3.12 shows the thickness comparisons between the automatically segmented masks and ground truth masks for 46 testing images. It is observed in Fig. 3.12 that the thickness of epidermis mask obtained by the proposed technique is very close to that of manually labeled epidermis mask, whereas the segmented epidermis masks by existing techniques tend to have much larger thickness than manually labeled epidermis masks. The MCGT [21], CET [75] and GTSA [74] techniques incorrectly segment some low intensity areas (e.g., cell nuclei) in the dermis area as the epidermis area, which increases the thickness of the segmented epidermis mask.

3.4.4 Qualitative Results

Qualitative results of epidermis segmentation for a whole slide skin histopathological image is illustrated in Fig. 3.13. Note that Fig. 3.13(a) shows the WSI with the manually labeled epidermis contour, while Figs. 3.13(b)(c)(d)(e) show the

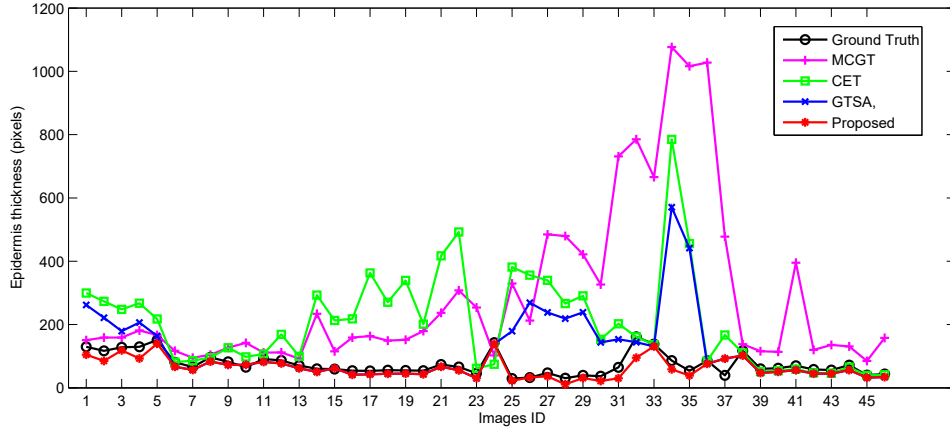


Fig. 3.12. Comparison of epidermis thickness between manually labeled epidermis masks and automatically obtained results for testing images.

corresponding automatically segmented results by the MCGT [21], CET [75], GTSA [74] and proposed technique, respectively. Figs. 3.13(a₁ – e₁) and (a₂ – e₂) show the corresponding selected parts of magnified segmentation results, respectively. Note that the magnification of selected parts are indicated by the rectangles on the WSIs. It is observed in Fig. 3.13 that the proposed technique provides more accurate segmentations than existing epidermis segmentation techniques. The MCGT [21] technique segments many false positive regions in the dermis area as epidermis area, as a simple closing operation fails to remove dark regions in the dermis area which are subsequently classified as epidermis area by thresholding. The CET [75] and GTSA [74] techniques incorrectly segment many low intensity dermis areas as epidermis areas, since these low intensity areas are segmented as binary foregrounds by thresholding but not eliminated by subsequent shape and area analysis.

3.5 Proposed Dermis Segmentation

Unlike epidermis areas which are easily distinguishable from other skin components, the dermis areas do not have clear borders from other tissues (e.g., subcutaneous tissues). In addition, pathologists are mainly interested in the DEJ areas (dermis-epidermis junction) instead of the whole dermis area when diagnosing skin melanomas. So it is not necessary to segment out the whole dermis region. In the

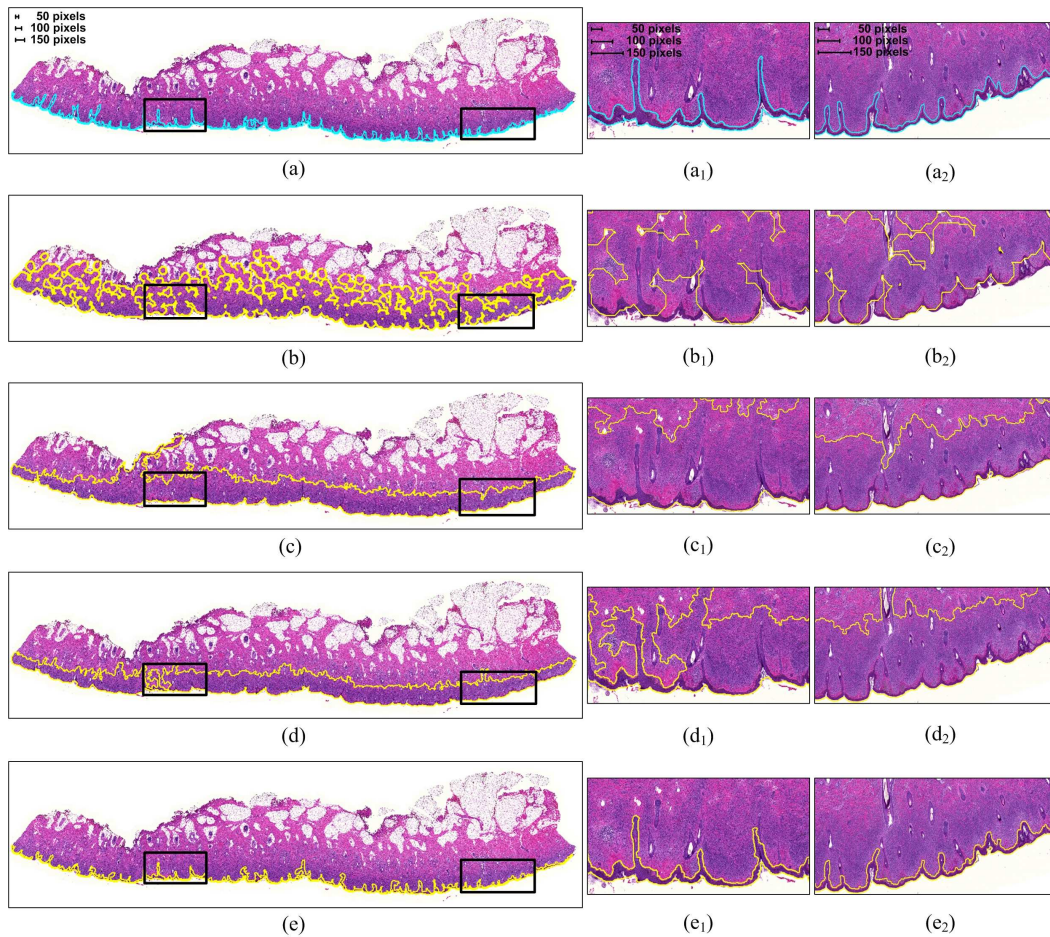


Fig. 3.13. Comparative segmentation results on a skin WSI. (a) Manually labeled epidermis contour. (a₁) & (a₂) Magnification of selected parts in (a). (b) MCGT [21]. (b₁) & (b₂) Magnification of selected parts in (b). (c) CET [75]. (c₁) & (c₂) Magnification of selected parts in (c). (d) GTSA [74]. (d₁) & (d₂) Magnification of selected parts in (d). (e) Proposed technique. (e₁) & (e₂) Magnification of selected parts in (e). Note that a large number of dermis pixels are incorrectly segmented as epidermis pixels in (b)(c)(d).

following, we present a technique to segment skin dermis based on a predefined depth of interest from epidermis outer boundary. The schematic of the proposed dermis segmentation technique is shown in Fig. 3.14. As observed in Fig. 3.14, the proposed technique has three modules. First, the epidermis area is segmented out from the skin WSI, and the epidermis outer boundary is determined. The line segments with a predefined length (i.e., depth of interest) are then generated from the epidermis outer boundary. Image morphological operations are finally applied to obtain desired dermis areas.

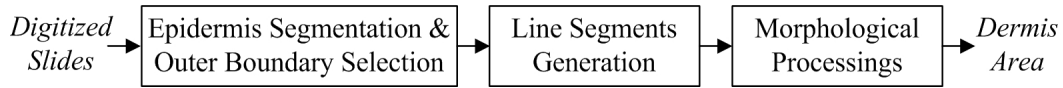


Fig. 3.14. Schematic of dermis segmentation.

3.5.1 Epidermis Segmentation & Outer Boundary Selection

To segment skin dermis areas, the first step is to locate skin epidermis areas. In this module, the coarse-to-fine epidermis segmentation technique presented in Section 3.3 is applied to segment skin epidermis areas. Given an input RGB image I_l , a binary mask b_e with epidermis areas as foreground is obtained. Fig. 3.15(a) shows a H&E stained skin WSI, while Fig. 3.15(b) shows the corresponding segmented epidermis mask b_e . On the mask b_e , the left-bottom and right-bottom points E_1 and E_2 are identified (see the solid red dots in Fig. 3.15(c)). The points E_1 and E_2 divide the boundary of epidermis mask b_e into two parts: outer boundary and inner boundary. Based on the average distance between boundary points and the geometric center of skin pixels (i.e., outer boundary on average has a longer distance), the epidermis outer boundary \mathcal{B} is determined. These two extrema points E_1 and E_2 and the epidermis outer boundary \mathcal{B} (the green contour in Fig. 3.15(c)) are to be used for generation of line segments in the next module.

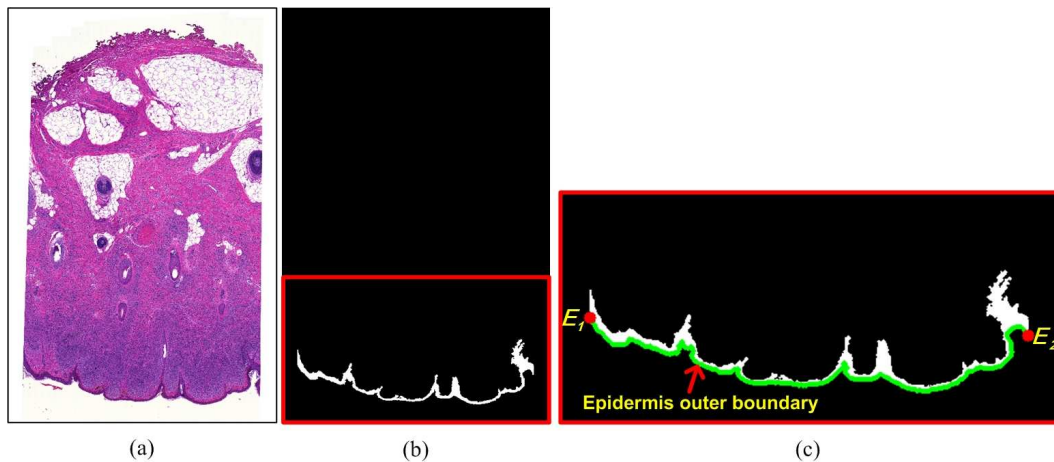


Fig. 3.15. Example of epidermis segmentation. (a) A skin WSI. (b) Epidermis mask b_e . (c) Epidermis extrema points and outer boundary. Note that (c) is cropped from (b) and zoomed up here for clear illustration.

3.5.2 Line Segments Generation

In this module, we generate a series of line segments emerging from epidermis outer boundary, and the dermis regions for segmentation will be covered by these generated line segments. The two steps of line segments generation are as follows.

(1) Down-sampling outer boundary: Given the epidermis outer boundary \mathcal{B} , we first down-sample it and select a subset of boundary points \mathcal{B}^* such that the subsequent processing is performed more efficiently. Assuming that there are z points on the boundary \mathcal{B} and the down-sampling factor is h (e.g., $h = 3$), the down-sampled boundary \mathcal{B}^* is obtained as follows:

$$\mathcal{B}^* = \{p_i \in \mathcal{B} \mid i = h, 2h, \dots, mh, m = \lfloor z/h \rfloor\} \quad (3.17)$$

For brevity, let us denote the down-sampled outer boundary \mathcal{B}^* as $\{P_i\}_{i=1\dots m}$, where m is the number of points.

(2) Generating line segments: After obtaining the boundary \mathcal{B}^* , we define a perpendicular line segment (with a length of d) for each point P_i ($2 \leq i \leq m - 1$) based on its two neighboring points. Since the outer boundary (see Fig. 3.15(c)) is not a smooth contour, it will result in noisy directions for generated line segments. In this step, we first smooth the outer boundary points $\{P_i\}_{i=1\dots m}$ using a moving average filter with a predefined length (e.g., 201 taps). A set of smoothed outer boundary points $\{P_i^*\}_{i=1\dots m}$ is obtained and used for estimating directions of line segments. Given a point P_i (see Fig. 3.16), the perpendicular line segment is generated as follows:

Step 1: Three consecutive points P_{i-1}^* , P_i^* and P_{i+1}^* are selected from the set $\{P_i^*\}_{i=1\dots m}$, and a curve \mathcal{C}_i (as shown in Fig. 3.16) is fitted based on the three consecutive points using a polynomial curve fitting method.

Step 2: A line \mathcal{T}_i , which is tangent to the fitted curve \mathcal{C}_i at point P_i^* , is calculated.

Step 3: A line \mathcal{V}_i passing through the point P_i and perpendicular to the line \mathcal{T}_i is calculated.

Step 4: A line segment P_iQ_i is generated such that it satisfies the following

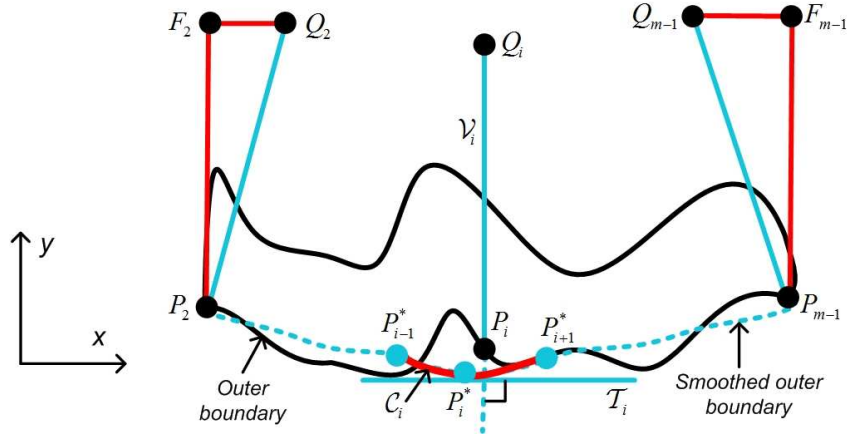


Fig. 3.16. Illustration of line segments generation.

equation:

$$d = \|P_i - Q_i\| \quad (3.18)$$

where the point Q_i is an end point of line segment P_iQ_i .

By using steps 1-4, the line segments for all outer boundary points $\{P_i\}_{i=2\dots(m-1)}$ are generated. Figs. 3.17(a)(b) separately show the mask b_e and the image I_l with pixels of line segments. As observed in Fig. 3.17(b), the line segments have covered almost all pixels of desired dermis regions (based on depth d). However, two small areas A_1 and A_2 (highlighted by the green contours) are missed. To incorporate these two regions, four more line segments P_2F_2 , F_2Q_2 , $P_{m-1}F_{m-1}$ and $F_{m-1}Q_{m-1}$ are created, where the points F_2 and F_{m-1} have coordinates (x_{P_2}, y_{Q_2}) and $(x_{P_{m-1}}, y_{Q_{m-1}})$, respectively. Note that x_{P_2} and $x_{P_{m-1}}$ are the x coordinates of the points P_2 and P_{m-1} , respectively. y_{Q_2} and $y_{Q_{m-1}}$ are the y coordinates of the points Q_2 and Q_{m-1} , respectively.

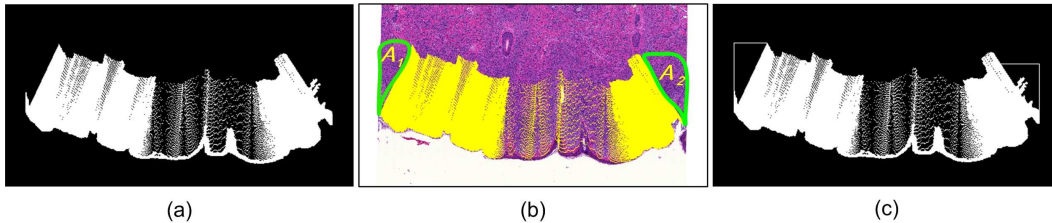


Fig. 3.17. Example of line segments generation. (a) Mask b_e with line segments. (b) Image I_l with line segments. (c) Mask b_l . Note that the depth value d is set as $650\mu m$.

After this step, a binary mask b_l with pixels of epidermis and line segments as foreground is generated. Fig. 3.17(c) shows the mask b_l with a predefined depth d of $650\mu m$. Note that the depth d is a critical parameter which determines the segmentation depth of dermis regions for analysis.

3.5.3 Morphological Processing

As observed in Fig. 3.17(c), there are gaps and holes within foreground pixels (i.e., white pixels) of the mask b_l . In order to obtain an intact dermis mask, the morphological closing operation with a disk shape structuring element (with a radius of 30 pixels) is first applied on b_l to join isolated pixels. The holes are then filled by performing morphological reconstruction operation. Let us denote the obtained mask as b_{e+d} . Fig. 3.18(a) shows the mask b_{e+d} , while Fig. 3.18(b) shows b_{e+d} in the red channel image R_l (i.e., the red channel of I_l).

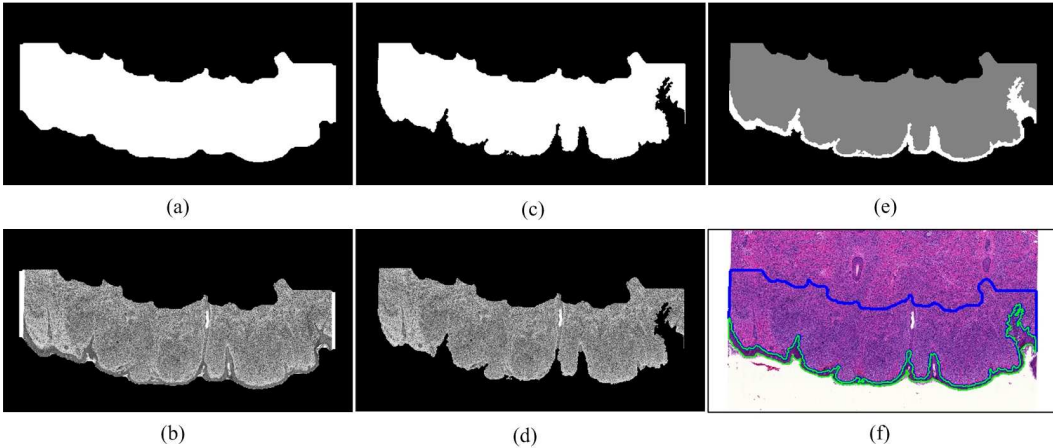


Fig. 3.18. Illustration of dermis mask generation. (a) b_{e+d} . (b) b_{e+d} in red channel R_l . (c) b_d . (d) b_d in red channel R_l . (e) Segmented epidermis (white color) and dermis (gray color) regions. (f) Image I_l overlapping with epidermis contours (green color) and dermis contours (blue color).

As observed in Fig. 3.18(b), the coarse mask b_{e+d} contains a small number of background pixels (typically white in I_l) and epidermis pixels. To generate the dermis mask, an intensity threshold τ (e.g., $\tau = 240$) is first applied (on foreground pixels in Fig. 3.18(b)) to remove white background pixels. Epidermis pixels, which have been segmented out during epidermis segmentation are then discarded. A

binary mask b_d with dermis pixels as foreground is finally generated after removing small noisy regions. Figs. 3.18(c)(d) show the dermis mask b_d and dermis pixels in the image R_l , respectively. Fig. 3.18(e) shows the segmented epidermis and dermis regions, while Fig. 3.18(f) shows the image I_l overlapping with the contours of epidermis and dermis regions.

3.6 Dermis Segmentation Evaluations

To the best of our knowledge, this is the first work to segment skin dermis from skin WSI based on a predefined depth of interest. To evaluate the proposed technique, we ran two experiments based on our dataset, where the depth value d is set as $650\mu m$ and $800\mu m$, respectively. During the experiment, the thickness of each automatically segmented epidermis and dermis mask is manually measured based on three randomly selected points using a computer program (based on Matlab). For each epidermis and dermis mask, the thickness t_i is computed as the average of three manually measured thickness, and compared with the predefined depth value d . A small difference between t_i and d indicates a good performance of the dermis segmentation.

Let the average thickness over all t_i ($1 \leq i \leq 66$) values be denoted by \bar{t} , and the standard deviation over all t_i values be denoted by σ_t . Table 3.5 shows the experimental results of \bar{t} and σ_t values. Fig. 3.19 compares the thickness t_i of automatically segmented epidermis and dermis mask with the predefined depth values d . As observed in Table 3.5 and Fig. 3.19, there is a small difference between t_i and d values, and hence the proposed technique provides a good performance in dermis segmentation.

TABLE 3.5
PERFORMANCE EVALUATIONS OF DERMIS SEGMENTATION.

Experiments	\bar{t}	σ_t
$d = 650\mu m$	$646.7\mu m$	$12.6\mu m$
$d = 800\mu m$	$797.1\mu m$	$13.6\mu m$

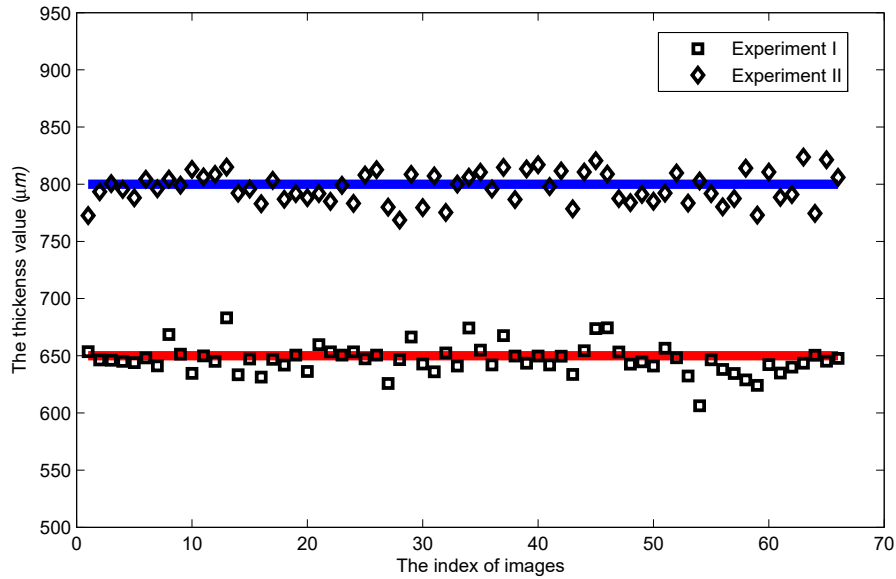


Fig. 3.19. Comparison of manually obtained dermis depth with the predefined segmentation depth. Note that the red and blue lines are the predefined segmentation depth (i.e., ground truths). The diamonds and squares indicate the manually measured segmentation depth. The manually measured values vary around ground truths mainly due to the errors caused by manual measurement.

Fig. 3.20 shows examples of epidermis and dermis segmentations, where the manually measured thickness is overlapped on the magnifications of selected image patches. Note that the green contours indicate the segmented epidermis regions, while the blue contours indicate the segmented dermis regions. In the second row, the selected image patches are zoomed in for better viewing. As observed in Fig. 3.20, the desired dermis regions have been correctly segmented by the proposed dermis segmentation technique.

3.7 Computational Complexity

All experiments of epidermis and dermis segmentations were done on a 1.80 GHz Intel Core II Duo CPU, with 16GB of RAM using MATLAB version R2013a. The proposed techniques roughly take 6.7s to segment both epidermis and dermis regions from a whole slide skin histopathological image with size of 3200x3000 pixels, where epidermis and dermis segmentations take about 4.2s and 2.5s, respectively.

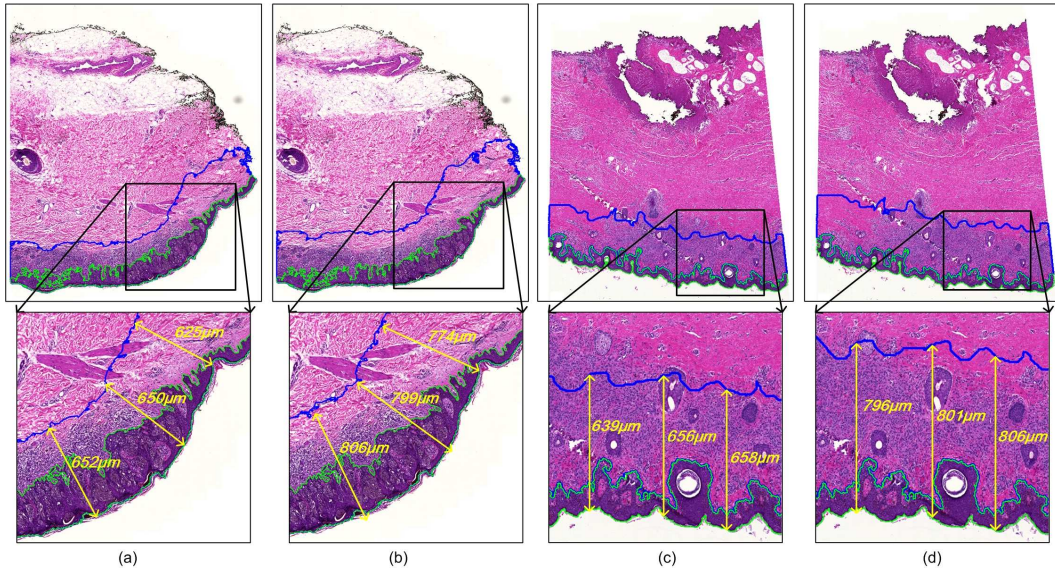


Fig. 3.20. Epidermis and dermis segmentations. (a) & (c) $d = 650\mu\text{m}$. (b) & (d) $d = 800\mu\text{m}$. Note that blue contours indicate the segmented dermis regions, while green contours indicate the segmented epidermis regions.

3.8 Summary

This chapter presents the proposed techniques for automatic segmentation of epidermis and dermis regions in skin WSIs. The proposed epidermis segmentation follows a coarse-to-fine processing procedure, where the coarse segmentation is performed using global thresholding and shape analysis, and the fine segmentation is performed using k-means algorithm. Based on epidermis segmentation, the proposed dermis segmentation is performed based on morphological operations and a predefined depth of interest. Experiments on 66 digitized skin WSIs show that the proposed techniques provide a good performance in epidermis and dermis segmentations. The proposed epidermis and dermis segmentations are important pre-requisite steps towards the automated skin histopathological image analysis, since most of diagnostic clues for skin melanoma are observed in the epidermis and DEJ junctional areas.

Chapter 4

Automated Detection of Cell Nuclei

Robust nuclei detection is an important step towards automatic pathological analysis, as the abnormal growth or loss of cell nuclei are usually indicative to a certain disease (e.g., skin cancer). In addition, automated segmentation of cell nuclei is usually dependent on accurate nuclei detection. Although there have been a number of computerized techniques proposed for nuclei detection, robust nuclei detection is still a very challenging problem due to the presence of clustered nuclei and nuclear variations in terms of color, size and shape. In this chapter, we propose two nuclei detection techniques, which are implemented based on voting analysis and gLoG kernels, respectively. Experimental results show that the proposed techniques are superior to existing techniques in nuclei detection.

4.1 Introduction

Robust nuclei detection is an important step in histopathology for disease diagnosis, as the counts of nuclei provide quantitative information when studying changes in cells, tissues and organs [38]. For example, the density of cell nuclei in histological images is an important feature for automatic breast tumor grading [24]. In addition,

Part of this chapter has been published in: Xu *et al.* “An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm”, *IEEE Journal of Biomedical and Health Informatics (JBHI)*, vol.18, no.5, pp.1-14, 2014. Xu *et al.* “Automatic nuclei detection based on generalized Laplacian of Gaussian filters”, *IEEE Journal of Biomedical and Health Informatics (JBHI)*, vol.21, no.3, pp.826-837, 2017.

nuclei detection is usually the first step for nuclei segmentation, and the accuracy of segmentation depends critically on the accuracy and reliability of the detected points (i.e., seed points) [42]. Although it is intuitive to manually identify cell nuclei in biopsy images, there still remains a huge challenge for a computer algorithm to perform automated detection. Fig. 4.1 shows examples of cell nuclei in H&E stained skin histopathological images, where cell nuclei have been stained as blue-purple and the intra-cellular matters have been stained as bright pink. Figs. 4.1(b)(d) show the magnifications of selected image patches in Figs. 4.1(a)(c), respectively, where the cell nuclei are highlighted by yellow dots. As observed in Fig. 4.1, since the image is the 2-D section of a 3-D tissue sample, the nuclear size and shape are usually different, and there exist touching and overlapping nuclei. The nuclear variations and presence of clustered nuclei are main obstacles for accurate nuclei detection.

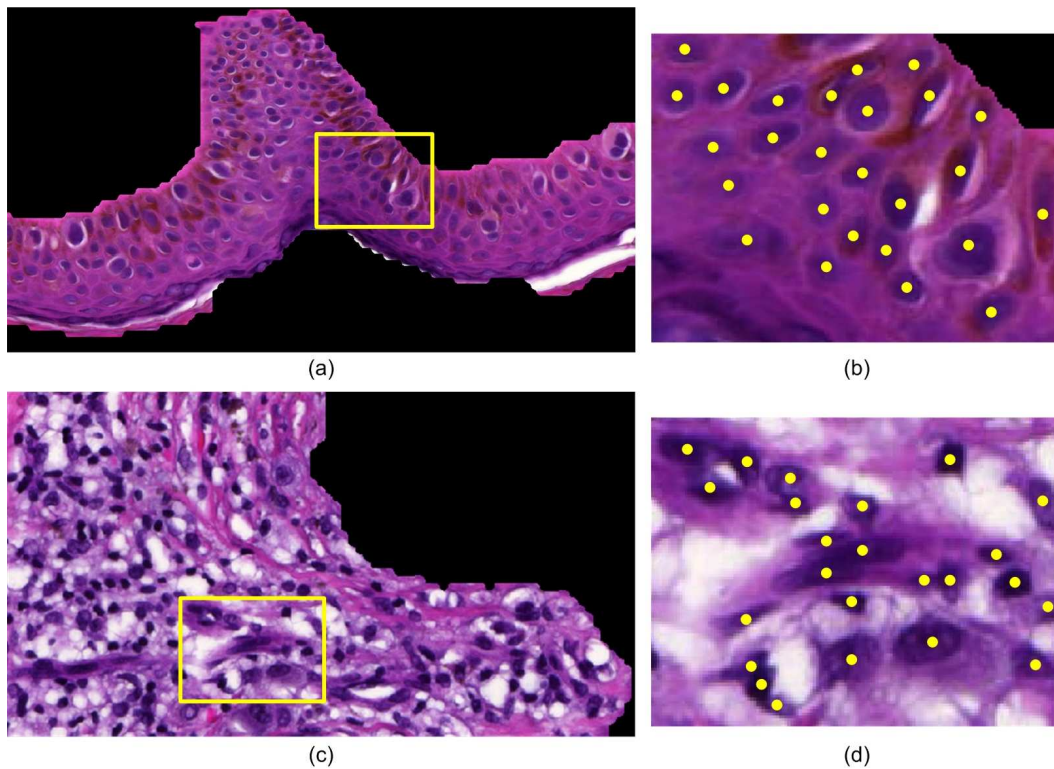


Fig. 4.1. Nuclei in skin histopathological images. (a) Epidermis. (b) Magnification of selected part in (a). (c) Dermis. (d) Magnification of selected part in (c). Note that the yellow dots in (b)(d) highlight the existence of cell nuclei.

TABLE 4.1
SUMMARY OF NUCLEI DETECTION TECHNIQUES

Authors	Key techniques	Main limitations
Parvin <i>et al.</i> [80]	Multi-pass voting	High computational complexity
Qi <i>et al.</i> [43]	Single-pass voting	Sensitive to thresholding values
Byun <i>et al.</i> [38]	Single-scale LoG filter	Poor performance on blobs with different scales
Niazi <i>et al.</i> [81]	Multi-scale DoG filter	Poor performance on elliptical blobs
Al-Kofahi <i>et al.</i> [39]	Multi-scale LoG filter	Sensitive to image binarizations
Kong <i>et al.</i> [40]	Multi-scale gLoG filter	Sensitive to image noises

Recently several nuclei detection techniques have been proposed (see the summary listed in Table 4.1). Parvin *et al.* [80] proposed a multi-pass voting technique for inferring the centers of cell nuclei. This technique initially votes along the gradient direction of high gradient pixels in a cone-shape voting area, and then iteratively votes in multiple passes with a shrunk voting area and updated voting direction (towards the highest votes in the previous pass). The points which have the votes above a predefined threshold are finally considered as nuclei seeds. To speed up the multi-pass voting in [80], Qi *et al.* [43] proposed a single-pass voting (SPV) technique for nuclei detection, which only performs one round voting followed by mean shift clustering to detect nuclei seeds. Although these techniques have been reported to provide a good performance, they are either sensitive to the predefined parameters or have high computational complexity. They tend to be infeasible for high-throughput analysis of images with a large number of cell nuclei. Byun *et al.* [38] proposed to detect nuclei seeds in retinal images by a single-scale LoG filter. This method has a low computational complexity, but it is usually difficult to determine an optimal scale for the LoG filter due to the size and shape variations of cell nuclei. Niazi *et al.* [81] proposed to detect cytotoxic T-cells by multi-scale difference of Gaussian (DoG) filters, with false seeds removed based on blobness measurement. Al-Kofahi *et al.* [39] proposed a multi-scale LoG filters (mLoG) based technique which detects nuclei seeds by mLoG filtering and adaptively selects scales for filters using inner distance transform. The LoG operators have shown a promising performance in locating the seeds of circular blobs, but they

have limitations in detecting blobs with general elliptical shapes. This is because the conventional LoG operators are generated from the uniform Gaussian kernels which are rotational symmetric in all directions. In order to identify elliptical blob structures and their scales, Kong *et al.* [40] proposed a gLoG filter based technique. The technique first generates a set of response maps by convolving gLoG kernels with the image, and then aggregates all obtained response maps together and searches blob centers (e.g., nuclei seeds) from the aggregated response map. Since the technique in [40] has not taken the advantage of gLoG kernels' directional information, it is likely to miss seeds for nuclear regions that do not have peak responses in the aggregated response map. Besides, the technique is quite sensitive to the local noise in the image background [82], as it considers all local maxima in the aggregated response map as blob seeds.

In this chapter, we propose two improved nuclei detection techniques for biopsy image analysis, which overcome many of the limitations of existing techniques. The first technique performs nuclei detection based on ellipse descriptor analysis and an improved voting algorithm, while the second technique performs nuclei detection based on directional gLoG kernels and mean-shift clustering. The details of two proposed techniques and experimental results are provided in following sections.

4.2 Proposed Voting-based Technique

The schematic of the proposed voting-based technique is shown in Fig 4.2. It is observed that the proposed technique has three modules. In the first module, an initial segmentation of cell nuclei regions is performed using an adaptive thresholding technique. In the second module, an elliptical descriptor is used to filter the isolated nuclei with elliptical shapes in the initially segmented binary image. In the third module, a voting algorithm is applied to detect seeds of nuclear clumps and the nuclei with irregular shapes.

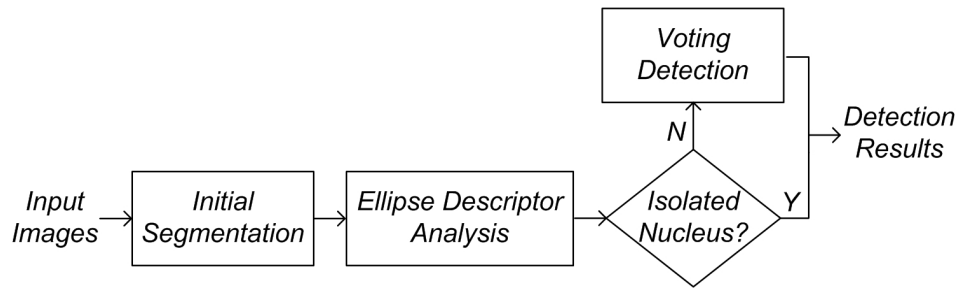


Fig. 4.2. Schematic of the voting-based technique.

4.2.1 Initial Segmentation

Due to the staining imperfection and variation, the pixel intensity in nuclear regions is not homogenous. Blurred nuclei boundaries and noisy pixels in the background are usually observed. The traditional low level segmentation methods (i.e., histogram based method [83]) generally fail to precisely segment nuclear regions from the background pixels. After comparing several segmentation techniques including Otsu's method [51], active contour model [84], [85], and graph cut [86], an adaptive threshold technique [30] is adopted for the initial segmentation of cell nuclei. The principal steps for initial segmentation by the adaptive threshold technique are as follows:

(1) Enhance the image using the hybrid gray-scale morphological reconstructions (HGMR) module. In this module, the R channel of the original RGB image is used as the input. Opening-by-Reconstruction and Closing-by-Reconstruction [31] are consecutively applied to reduce the impact of undesirable variations in the image and to make the nuclear region homogenous.

(2) Apply an adaptive thresholding [31] on the enhanced image. For each pixel in the enhanced image, a threshold is calculated as the mean intensity of the local neighborhood (e.g., 60x60). Since the intensity of nuclear regions is lower than the background, a pixel value below the threshold is set to the foreground value (represented by 1), and a pixel value above the threshold is set to the background value (represented by 0).

(3) Enhance the segmented result by a series of morphological operations. The

holes within the segmented regions are filled by performing the morphological reconstruction. The small isolated regions, unlikely to be the nuclei, are removed by conducting an opening operation with a disc structuring element. Fig. 4.3(a) shows an original image, and Fig. 4.3(b) shows the segmentation result.

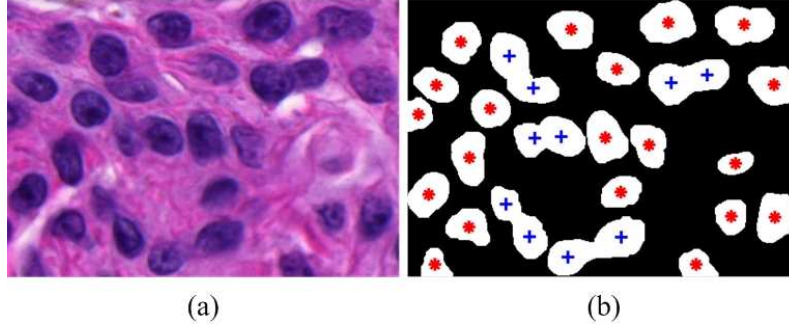


Fig. 4.3. An example of the initial segmentation. (a) The original image. (b) The segmentation result. Note that * indicate the isolated nuclei, whereas + indicate the touching nuclei.

4.2.2 Ellipse Descriptor Analysis

It is observed in Fig. 4.3(b) that initial segmentation results in both isolated nuclei and nuclei clumps. Assume that the segmented regions in Fig. 4.3(b) are denoted as $\{N_p\}_{p=1\dots Z}$, where Z is the total number of regions. In this module, we classify $\{N_p\}_{p=1\dots Z}$ as the isolated nuclei or nuclei clump. Based on the observation that isolated nuclei in skin histopathological images typically have elliptical shapes, most isolated nuclei can be detected by using an elliptical model. To build the elliptical model for each region in $\{N_p\}_{p=1\dots Z}$, we first calculate the gradient image of the initial segmentation result, and then obtain the boundary points by selecting positions with non-zero gradient magnitudes. Using the obtained boundary points, each region in $\{N_p\}_{p=1\dots Z}$ is now mathematically described by an elliptical model. Let (x, y) be a 2-D boundary point of a connected region. The ellipse is modeled by the following second-order polynomial [64]:

$$E(\bar{a}, \bar{x}) = \bar{a} \cdot \bar{x} = ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (4.1)$$

where $\bar{a} = [a \ b \ c \ d \ e \ f]^T$ denotes the ellipse coefficients, $\bar{x} = [x^2 \ xy \ y^2 \ x \ y \ 1]^T$, and $E(\bar{a}, \bar{x})$ is the algebraic distance from the point (x, y) to the conic $E(\bar{a}, \bar{x}) = 0$. The optimal ellipse fitting the nuclear shape is generated by minimizing the sum of squared algebraic distances between the ellipse and the 2-D boundary points of the connected region [64].

After generating the elliptical model for each region N_p , the ellipticity evaluation is performed by comparing the region of N_p and the corresponding ellipse. Let R_n be the set of pixels in N_p (i.e., the nucleus or the nuclear clump), and R_e be the set of pixels in the elliptical region. We use the following two parameters e_1 and e_2 to measure the ellipticity of a region:

$$e_1 = \frac{|R_n \cap R_e|}{|R_e|} \quad (4.2)$$

$$e_2 = \frac{|R_n \Delta R_e|}{|R_e|} \quad (4.3)$$

where $|\cdot|$ is the cardinality of the pixels set, \cap is the intersection operation, and Δ is the symmetric difference operation. Intuitively, the higher value of the ellipticity parameter e_1 and the lower value of the ellipticity parameter e_2 represent a closer match to an elliptical shape. Thus a candidate region with high value of e_1 and low value of e_2 is more likely to be a single nucleus. On the other hand, a region with lower value of e_1 but higher value of e_2 is expected to be a nuclei clump or a nucleus with the irregular shape.

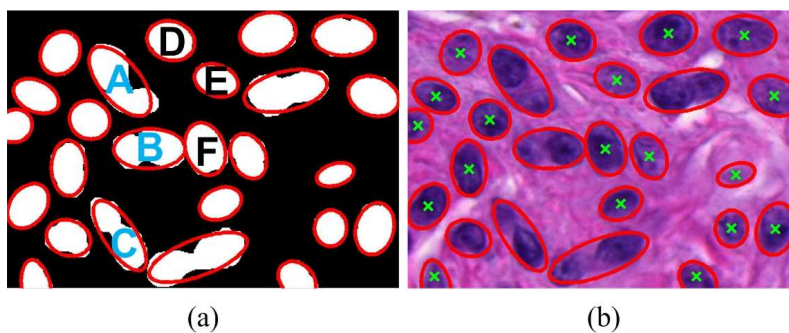


Fig. 4.4. The elliptical modeling for each connected region. (a) The ellipse model constructed from Fig. 4.3(b). (b) The ellipse model superimposed on the original image. Note that the symbols \times in (b) mark the ellipse centers for the detected isolated nuclei.

TABLE 4.2
VALUES OF ELLIPTICITY PARAMETERS FOR SEVERAL SELECTED CANDIDATE REGIONS
SHOWN IN FIG. 4.4(A).

Ellipticity Parameters	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
e_1	0.87	0.88	0.86	0.95	0.97	0.94
e_2	0.26	0.21	0.29	0.09	0.07	0.10

An example of the ellipse descriptor analysis is shown in Fig. 4.4. Fig. 4.4(a) shows the ellipses obtained by fitting the boundary points of the binary image shown in Fig. 4.3(b). The markers *A, B, C* indicate the three regions formed by the touching nuclei, while the markers *D, E, F* indicate three typically isolated cell nuclei with elliptical shapes. Fig. 4.4(b) also shows the constructed ellipses superimposed on the original image. The values of the ellipticity parameters e_1 and e_2 for the candidate regions *A – F* are shown in Table 4.2. It is observed that the regions of the single nucleus *D, E, F* have high values of e_1 (close to 1.0) and low values of e_2 (close to 0.0). In contrast, the candidate regions *A, B, C* are separately formed by two touching nuclei, and they have relatively lower values of e_1 (around $0.8 \sim 0.9$) but higher values of e_2 (around $0.2 \sim 0.3$). Therefore, a single nucleus with elliptical shape is detected by the above two defined parameters. A segmented region N_p is determined as a single nucleus if both of the following two conditions are satisfied:

$$e_1 \geq \tau_1 \tag{4.4}$$

$$e_2 \leq \tau_2 \tag{4.5}$$

where τ_1 and τ_2 are the predefined thresholds. In order to calculate the parameters τ_1 and τ_2 , 170 nuclear regions in 15 randomly selected image patches are manually labeled, and the values of e_1 and e_2 for each region are calculated. Fig. 4.5 shows the box plots of ellipticity parameters for isolated nuclei and nuclei clumps in 170 nuclear regions. It is observed that isolated nuclei on average have a high value of e_1 (0.96) and a low value of e_2 (0.07), whereas nuclei clumps have a lower value of

e_1 (0.85) and a higher value of e_2 (0.30). The isolated nuclei with elliptical shapes can be separated from nuclei clumps by thresholding the two ellipticity parameters. In our work, the values of τ_1 and τ_2 are calculated as follows:

$$\tau_1 = \frac{\bar{e}_{1N} + \bar{e}_{1C}}{2} \quad (4.6)$$

$$\tau_2 = \frac{\bar{e}_{2N} + \bar{e}_{2C}}{2} \quad (4.7)$$

where \bar{e}_{1N} and \bar{e}_{1C} are the average values of e_1 for isolated nuclei and nuclei clumps, \bar{e}_{2N} and \bar{e}_{2C} are the average values of e_2 for isolated nuclei and nuclei clumps, respectively. Note that nuclear regions (i.e., single nucleus) that satisfies the above two conditions in Fig 4.4(b) are marked with crosses at the ellipse centers.

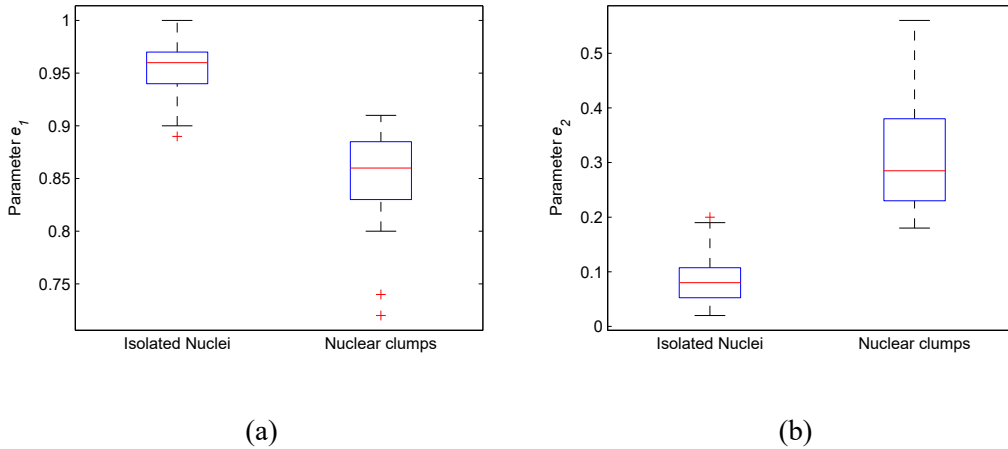


Fig. 4.5. Variations of ellipticity parameters for isolated nuclei and nuclear clumps. (a) Parameter e_1 . (b) Parameter e_2 . Note that the results are obtained based on 170 randomly selected nuclear regions.

4.2.3 Voting Detection

After the ellipse descriptor analysis, most isolated cell nuclei with elliptical shapes have been detected. To identify the number and location of the remaining cell nuclei regions, for example the regions $A - C$ shown in Fig 4.4(a), a voting algorithm is proposed in this module.

Assume that the nuclear regions, which are waiting for seed detection, are

denoted as $\{N_i\}_{i=1\dots K}$, where K is the number of connected regions. The regions $\{N_i\}_{i=1\dots K}$ are first labeled via 8-connected criterion [31], and then each region is put into a binary image which has the same size as the original image. Let $I_i(x, y)$ be a binary image which contains a region N_i . The boundary points set $T_i(x, y)$, the corresponding image gradient $\nabla T_i(x, y)$ and the gradient magnitude $\|\nabla T_i(x, y)\|$ are computed subsequently. Since the nuclear regions in the binary image $I_i(x, y)$ are the white foreground, the voting direction for a boundary point is defined as the positive gradient direction $\frac{\nabla T_i(x, y)}{\|\nabla T_i(x, y)\|} = (\cos(\theta(x, y)), \sin(\theta(x, y)))$, where θ is the angle of the gradient direction with respect to the horizontal direction of the image. Fig. 4.6 illustrates an example of voting for a boundary point (x_0, y_0) . The parameters r_{\min} , r_{\max} and Δ define a cone-shape voting area. The dashed contours represent the extracted boundary points. The potential voting region (e.g., the yellow region) is obtained by performing the erosion operation on the original nuclear region. The points (e.g., red points) in both the cone-shape voting area and the potential voting region get the votes weighed by the Gaussian kernel centered at the point (u_x, u_y) . The Gaussian kernel is defined as follows:

$$g(x, y, u_x, u_y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - u_x)^2 + (y - u_y)^2}{2\sigma^2}\right) \quad (4.8)$$

where $u_x = x_0 + r \cos \theta$, $u_y = y_0 + r \sin \theta$ and $r = \frac{(r_{\min} + r_{\max})}{2}$.

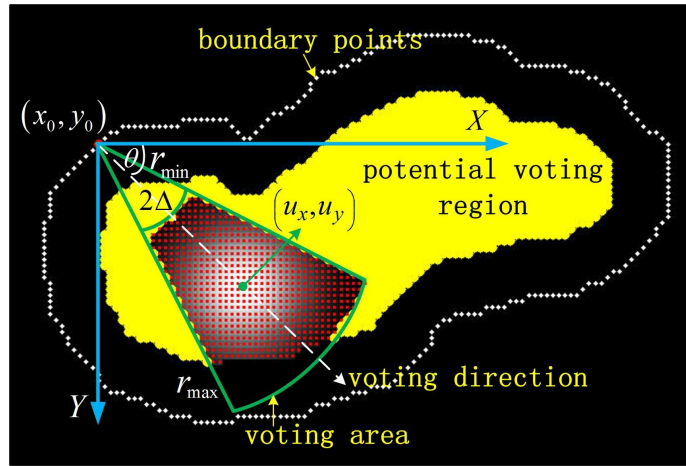


Fig. 4.6. Illustration of voting algorithm for nuclei seed detection.

Note that prior to voting towards the gradient direction of boundary points, the erosion operation is applied to shrink the nuclear region N_i . The structuring element S for the erosion operation is defined as follows:

$$S = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (4.9)$$

The disk shape structuring element with a radius of 3 pixels is empirically applied such that the erosion operation is able to retain the actual shape when reducing the size of nuclear regions. In this work, the size of each region in $\{N_i\}_{i=1\dots K}$ is reduced by at least half. But note that if the erosion operation has generated the isolated small region (i.e., ≤ 100 pixels), it should stop continually eroding the obtained small region to avoid the small nucleus disappearing. The erosion operation is performed for two reasons. First, nuclei centers are far away from nuclei boundaries, and hence there is no need to vote for pixels near nuclei edges. Second, the erosion operation contributes to reducing the time requirement of the voting algorithm by reducing the number of calculations, since the potential voting region (i.e., the original nuclei region) has been shrunk and there are fewer pixels that can get votes. For example, in Fig. 4.6 only the pixels in the eroded nuclear region (i.e., yellow region) may get votes.

Now let $V_i(x, y)$ be the vote image for the binary image I_i , and have the same size as the original image. All pixels in $V_i(x, y)$ are first initialized as zeros, and then updated as:

$$V_i(x, y) = V_i(x, y) + \sum_{(x,y) \in A(m,n)} \|\nabla T_i(m, n)\| g(x, y, u_m, u_n, \sigma) \quad (4.10)$$

where (m, n) represents a nuclear edge point, $A(m, n)$ is the voting area for the point (m, n) , $g(x, y, u_m, u_n, \sigma)$ is the Gaussian kernel centered at (u_m, u_n) . The final vote image V is the sum of all images V_i . Assume the value for a point $(x, y) \in V$ be n , we interpret that the point (x, y) is observed n times. Using the positions of the points in V as the two dimensional feature space, the nuclei seeds are computed by executing mean shift clustering algorithm [87]. The detailed process of the proposed voting algorithm is shown in Algorithm 1.

Algorithm 1 The proposed voting algorithm

Input: Candidate nuclei regions set $\{N_i\}_{i=1\dots K}$.

Initialization: $r_{\min} = 0.4d$, $r_{\max} = 1.5d$, $\Delta = \frac{\pi}{6}$, $\sigma = 7$ and $w_1 = 0.6d$, where d is the estimated average radius of the nuclei in the image, w_1 is the bandwidth of the mean shift algorithm.

for each region N_i **do**

 Compute boundary points set $T_i(x, y)$, image gradient $\nabla T_i(x, y)$, gradient magnitude $\|\nabla T_i(x, y)\|$ and gradient directions $\frac{\nabla T(x, y)}{\|\nabla T(x, y)\|}$.

 Perform erosion operation on nuclear region N_i .

 Initialize the vote image V_i , and conduct the single-path voting using Eq. (4.10).

end for

Sum all the vote images V_i and run the mean shift clustering algorithm to localize nuclei seeds.

Output: Nuclei seeds for regions $\{N_i\}_{i=1\dots K}$.

Fig. 4.7 illustrates an example of voting-based nuclei detection. Fig. 4.7(a) shows the image with nuclei boundary points (i.e., voting points) obtained after initial segmentation. Fig. 4.7(b) shows the final vote image with ellipse contours of isolated nuclei. Fig. 4.7(c) shows the original image with ellipse contours of isolated nuclei. Fig. 4.7(d) shows the nuclei detection results indicated by \times or $+$ symbols. Note that in Fig. 4.7(d) the \times symbols are obtained based on ellipse descriptor analysis, and the $+$ symbols are obtained based on the voting algorithm.

4.3 Proposed gLoG-based Technique

The schematic of the proposed gLoG-based technique is shown in Fig. 4.8. It is observed that there are three modules. In the first module, a bank of scale-normalized

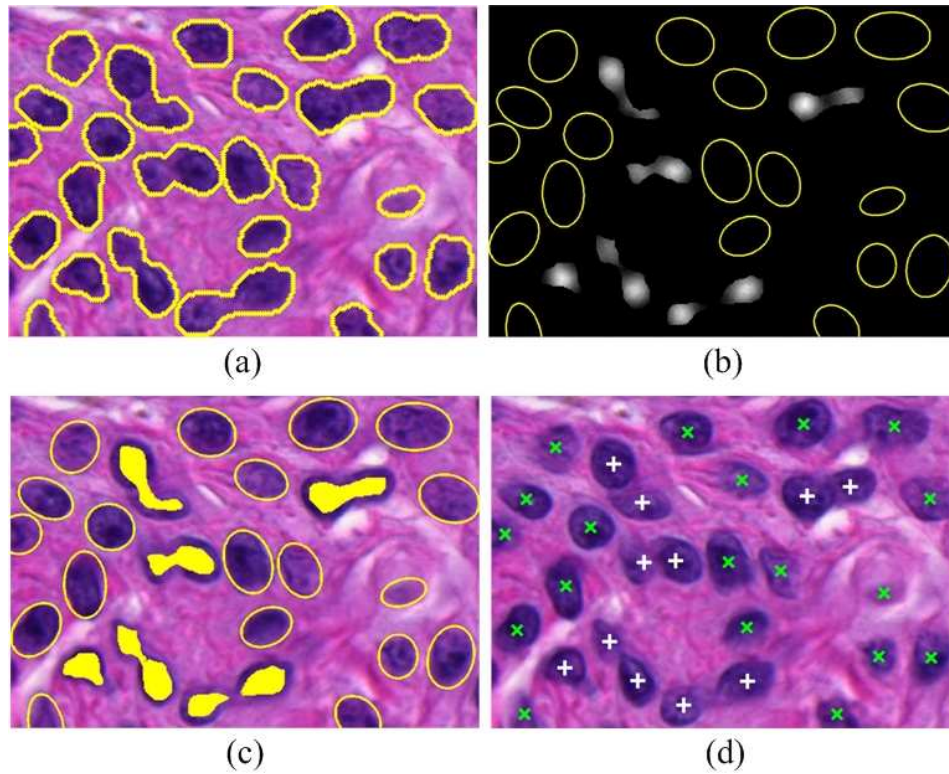


Fig. 4.7. Illustration of voting-based nuclei detection. (a) Image with extracted voting points (i.e., nuclei boundary points). (b) Final vote image with ellipse contours of isolated nuclei. (c) Image with vote points and ellipse contours of isolated nuclei. (d) Image with detected nuclei seeds. Note that green \times symbols are the seeds (i.e., geometrical centers) obtained after ellipse descriptor analysis, and white $+$ symbols are the seeds obtained by the voting algorithm.

gLoG kernels with different scales and orientations are constructed. In the second module, the constructed gLoG kernels are aggregated together according to their orientations, and the response maps are generated by convolving the aggregated gLoG kernels with the image. In the third module, nuclei seeds are detected from multiple response maps with redundant detections merged by mean-shift clustering. The details of three modules are now presented in the following.

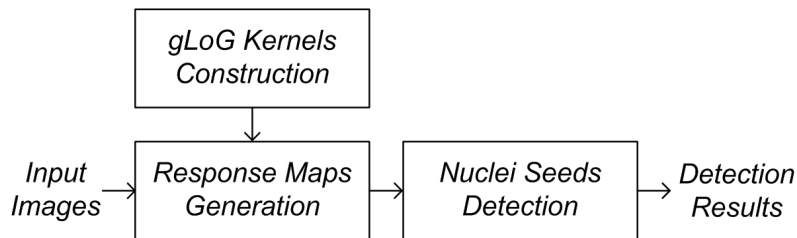


Fig. 4.8. Schematic of gLoG-based technique.

4.3.1 gLoG Kernels Construction

Because cell nuclei in digitized histopathological images typically have circular or elliptical shape, the gLoG kernels are appropriate for nuclei detection. In this module we illustrate the gLoG kernels construction and show that traditional scale-normalization for gLoG kernels performs better than log-scale normalization in [40]. The gLoG kernel $\nabla^2 G(x, y)$ is given as follows:

$$\nabla^2 G(x, y) = \frac{\partial^2 G(x, y)}{\partial x^2} + \frac{\partial^2 G(x, y)}{\partial y^2} \quad (4.11)$$

where (x, y) are the image domain coordinates, $G(x, y) = \lambda \cdot e^{-(ax^2+2bxy+cy^2)}$, and the Laplacian $\partial^2 G(x, y)/\partial x^2$, $\partial^2 G(x, y)/\partial y^2$ are given by:

$$\frac{\partial^2 G(x, y)}{\partial x^2} = [(2ax + 2by)^2 - 2a] G(x, y) \quad (4.12)$$

$$\frac{\partial^2 G(x, y)}{\partial y^2} = [(2bx + 2cy)^2 - 2c] G(x, y) \quad (4.13)$$

where λ is a normalization factor and the parameters a, b, c control the shape and orientation of the gLoG kernel, which are defined as follows:

$$a = \frac{\cos^2 \theta}{2\sigma_x^2} + \frac{\sin^2 \theta}{2\sigma_y^2} \quad (4.14)$$

$$b = -\frac{\sin 2\theta}{4\sigma_x^2} + \frac{\sin 2\theta}{4\sigma_y^2} \quad (4.15)$$

$$c = \frac{\sin^2 \theta}{2\sigma_x^2} + \frac{\cos^2 \theta}{2\sigma_y^2} \quad (4.16)$$

where (σ_x, σ_y) are the scales and θ is the orientation of the gLoG kernel. By varying the scales and orientations, a set of different gLoG kernels $\nabla^2 G(x, y)$ can be obtained. To construct a bank of gLoG kernels for nuclei identification, we first set angles that the gLoG kernel's orientation can take, which is as follows:

$$\theta = \frac{\pi}{k}(i - 1), \quad i = 1 \dots k \quad (4.17)$$

where k is the number of orientations. The range of scales σ_x and σ_y are then set as positive integers between σ_{\min} and σ_{\max} , where σ_{\min} and σ_{\max} are determined based on the estimated radius of detected nuclei [88]. Without loss of generality, we assume that $\sigma_x \geq \sigma_y$. Table 4.3 shows an example of gLoG kernels' configuration, where θ has 4 orientations and $4 \leq \sigma_y \leq \sigma_x \leq 8$.

TABLE 4.3
INDEX OF gLoG KERNELS' CONFIGURATION.

Index	σ_x, σ_y	θ (degree)	Index	σ_x, σ_y	θ (degree)
1	8,8	0	2-5	8,7	0,45,90,135
6-9	8,6	0,45,90,135	10-13	8,5	0,45,90,135
14-17	8,4	0,45,90,135	18	7,7	0
19-22	7,6	0,45,90,135	23-26	7,5	0,45,90,135
27-30	7,4	0,45,90,135	31	6,6	0
32-35	6,5	0,45,90,135	36-39	6,4	0,45,90,135
40	5,5	0	41-44	5,4	0,45,90,135
45	4,4	0			

When convolving gLoG kernels with the image containing blobs with general elliptical shapes, it produces a scale-and-orientation selective peak response at the center of each object. This can be considered as a pattern matching process, which matches the predefined gLoG kernels of different scales and orientations to the given image patterns. However, a direct application of multi-scale gLoG kernels to images of blobs would be unreliable, since the signal amplitude of the gLoG convolution at blob regions decreases monotonically as the scales increase. A straightforward way to keep the signal amplitude invariant across changes of image scale is to apply the scale normalization. Lindeberg [89] proposed a traditional normalization method for the LoG filter. This normalization method has been used by Xiao *et al.* [90] for detection of curvilinear structures using the bi-Gaussian kernel. For the scale normalization of gLoG filters in our work, the traditional normalization method in [89] can be intuitively extended and defined as:

$$\nabla^2 G_n(x, y) = \sigma_x \sigma_y \nabla^2 G(x, y) \quad (4.18)$$

where $\nabla^2 G_n(x, y)$ is the scale-normalized gLoG kernel. Kong *et al.* [40] proposed

a log-scale normalization for gLoG filter, that is:

$$\nabla^2 G_n(x, y) = (1 + \alpha \log(\sigma_x))(1 + \alpha \log(\sigma_y)) \nabla^2 G(x, y) \quad (4.19)$$

where α is a positive number (e.g., $\alpha = 0.5$) which can control the eccentricities of detected blobs. To evaluate the capacity of scale normalization by Eq. (4.18) and Eq. (4.19), we have separately applied the traditional scale normalized gLoG kernels and log-scale normalized gLoG kernels on our testing images, i.e.,

$$L_n(x, y) = \bar{I}(x, y) * \nabla^2 G_n(x, y) \quad (4.20)$$

where $I(x, y)$ is the original red channel image and $\bar{I}(x, y) = 255 - I(x, y)$. $L_n(x, y)$ is the convolution response map. Note that nuclear regions appear as dark blobs in the original red channel image, and hence the complement operation in Eq. (4.20) is performed to make nuclear regions as bright foreground.

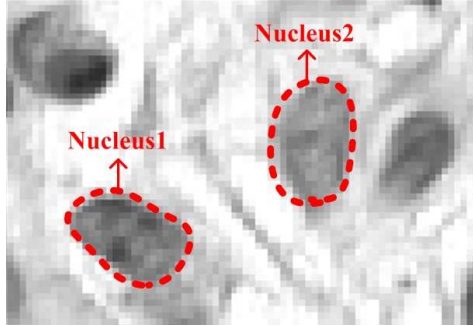


Fig. 4.9. Nuclei in a skin image. Note that the orientations for nucleus1 and nucleus2 are about 135° and 90° , respectively. The ratios between σ_x and σ_y for both nuclei are about 2.

Fig. 4.9 shows cell nuclei in a testing skin image, where the orientations of nucleus1 and nucleus2 are about 135° and 90° , respectively, and the ratios between σ_x and σ_y for both nuclei are about 2. Figs. 4.10(a)(b) show the maximum convolution responses by applying traditional scale-normalized gLoG kernels (using Eq. (4.20)) on nucleus1 and nucleus2, respectively. Figs. 4.10(c)(d) show the maximum convolution responses by applying log-scale normalized gLoG kernels ($\alpha = 1$) on nucleus1 and nucleus2, respectively. The kernel index (in horizontal axis) of Fig. 4.10 has

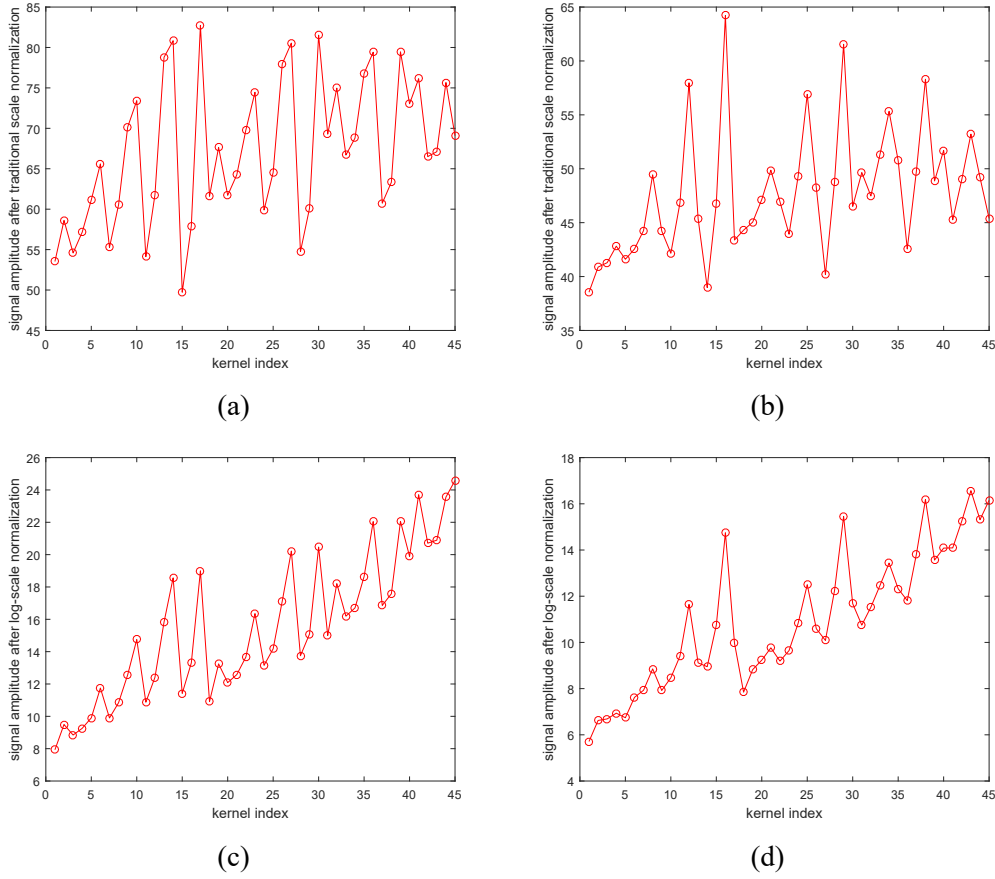


Fig. 4.10. Peak responses by traditional scale normalization and log-scale normalization. (a) & (b) Peak responses at nucleus1 and nucleus2 shown in Fig. 4.9 using traditional scale normalization. (c) & (d) Peak responses at nucleus1 and nucleus2 shown in Fig. 4.9 using log-scale normalization. Note that the kernel index in the horizontal axis has been shown in Table 4.3.

been provided in Table 4.3. As observed in Figs. 4.10(a)(b), the maximum responses are obtained when the kernel indices are 17 and 16 respectively. By contrast, in Figs. 4.10(c)(d) the maximum responses are obtained when the kernel indices are 45 and 43 respectively. Table 4.4 shows gLoG kernels' configurations which provide the maximum responses using different normalization methods. As shown in Table 4.4, it is apparent that the traditional scale normalization performs better than log-scale normalization, since the traditional normalized gLoG kernel with the maximum response has a good matching with the cell nucleus. Therefore, in our work, we utilized the traditional normalization on gLoG kernels (see Eq. (4.18)) for nuclear detection. Figs. 4.11(a)(b) show examples of traditional normalized

TABLE 4.4
INDEX OF NORMALIZED gLoG KERNELS WITH MAXIMUM RESPONSES.

Nuclei	Traditional normalization			Log-scale normalization		
	Index	σ_x, σ_y	θ	Index	σ_x, σ_y	θ
Nucleus1	17	8,4	135°	45	4,4	0°
Nucleus2	16	8,4	90°	43	5,4	90°

gLoG kernels with scales between $[8, 8]$ and $[4, 4]$ in 4 directions. As observed in Fig. 4.11(a), the gLoG kernels are rotational-symmetric when $\sigma_x = \sigma_y$, which are essentially the LoG kernels.

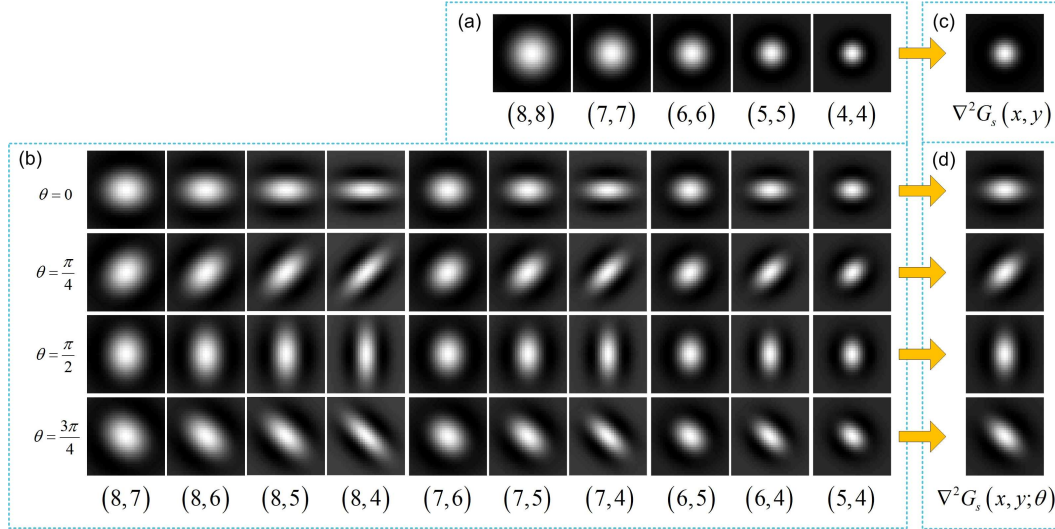


Fig. 4.11. Example of normalized gLoG kernels (with $k = 4$). (a) Rotational symmetric LoG kernels with $\sigma_x = \sigma_y \in \{8, 7, 6, 5, 4\}$. (b) gLoG kernels with $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$, $\sigma_x \in \{8, 7, 6, 5\}$, $\sigma_y \in \{7, 6, 5, 4\}$ and $\sigma_x > \sigma_y$. (c) $\nabla^2 G_s(x, y)$ (see Eq. (4.21)). (d) $\nabla^2 G_s(x, y; \theta)$ (see Eq. (4.22)). Note that in (a) and (b) (\cdot, \cdot) corresponds to (σ_x, σ_y) .

4.3.2 Response Maps Generation

After constructing a set of normalized gLoG kernels, the next is to convolve the image with gLoG kernels to obtain response maps which provide information about nuclei numbers and locations. Although it is possible to do the convolution between each gLoG kernel and the image (like the blob localization algorithm in [40]), it is time-consuming due to multiple convolutions. To speed up the algorithm, we make use of distributivity property of convolution operation and compute the

aggregation of gLoG kernels before convolution operation. Specifically, for the rotational symmetric gLoG kernels (see Fig.4.11(a)), the aggregated gLoG filter $\nabla^2 G_s(x, y)$ is given by:

$$\nabla^2 G_s(x, y) = \frac{1}{S_0} \sum_{\sigma=\sigma_{\min}}^{\sigma_{\max}} \nabla^2 G_n(x, y; \sigma) \quad (4.21)$$

where S_0 is the number of aggregated rotational symmetric gLoG kernels. For other general elliptical shapes of gLoG kernels (see Fig.4.11(b)), the aggregated gLoG filter $\nabla^2 G_s(x, y; \theta)$ is given by:

$$\nabla^2 G_s(x, y; \theta) = \frac{1}{S_k} \sum_{\sigma_x=\sigma_{\min}}^{\sigma_{\max}} \sum_{\sigma_y=\sigma_{\min}}^{\sigma_x} \nabla^2 G_n(x, y; \sigma_x, \sigma_y, \theta) \quad (4.22)$$

where S_k is the number of aggregated elliptical gLoG kernels. Note that $\nabla^2 G_s(x, y; \theta)$ is the aggregation of gLoG kernels in the orientation θ . Figs. 4.11(c)(d) show the corresponding aggregated results of gLoG kernels. With k orientations, there are $k + 1$ aggregated gLoG kernels in total. For brevity, let us denote the aggregated gLoG kernels as $\nabla^2 G_s^j(x, y)$, where $j = 0 \dots k$. The scale-space response map $L^j(x, y)$ is then obtained as:

$$L^j(x, y) = \bar{I}(x, y) * \nabla^2 G_s^j(x, y) \quad (4.23)$$

Note that the response map $L^j(x, y)$ corresponds to the pattern matching result between the image and the gLoG kernel $\nabla^2 G_s^j(x, y)$.

Using $k + 1$ normalized gLoG kernels, there are $k + 1$ response images obtained by Eq. (4.23). Fig. 4.12(a) shows two touching nuclei with elliptical shapes and (major axis) directions about 45° , while Figs. 4.12(b)-(f) show the response maps obtained by convolving Fig. 4.12(a) with gLoG kernels in Figs. 4.11(c)(d), respectively. In Figs. 4.12(b)-(f) the local maxima (usually round nuclear centers) are also searched and superimposed on the image with (red) plus symbols. As observed in Fig. 4.12, only the response map (d) provides two local maxima points, which

correctly identifies the existence of two nuclei. By contrast, nuclei are not correctly detected from response maps (b)(c)(e)(f). Fig. 4.12(d) provides a correct detection, mainly because the gLoG kernel used in (d) has a similar direction with nuclei in (a). More details about searching nuclei seeds from multiple response maps will be presented in the next section.

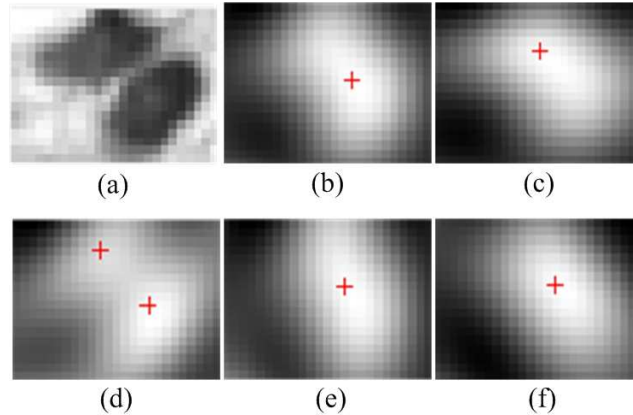


Fig. 4.12. Response maps with local maxima. (a) Two touching nuclei. (b)-(f) Response maps obtained by convolving (a) with gLoG kernels in Figs. 4.11(c)(d), respectively. Note that nuclei are correctly detected in image (d) which has two local maxima points indicating the existence of nuclei.

4.3.3 Nuclei Seeds Detection

Using the above obtained response maps $L^j(x, y)$, nuclei seeds can be detected as the points of local maxima which are usually around nuclei centers. In this module, we present the process of determining nuclei seeds from multiple response maps, which consists of four main steps as follows.

(1) Searching local maxima: We first search the points of local maxima from each response map $L^j(x, y)$. Let us denote them as $\{B_i\}_{i=1\dots M}$, where M is the number of local maxima in $k + 1$ response maps. Fig. 4.13(a) shows the image $I(x, y)$ overlapped with local maxima (red + symbols).

(2) Filtering false local maxima: To filter out false local maxima in image background, we first apply an adaptive thresholding based technique to perform image binarization. The adaptive thresholding based technique has two passes of Otsu's thresholding. In the first pass, Otsu's thresholding is applied on the whole

image $I(x, y)$, which segments nuclear regions into foreground. In the second pass, the bounding box (i.e., the smallest rectangle containing the region) for each foreground region is computed, and the pixels (retrieved from the image $I(x, y)$) within the bounding box are adaptively classified by Otsu's method. Note that two passes of Otsu's thresholding are sequentially applied which can reduce the impact of image variations on image binarization. After adaptive thresholding, a binary mask b_1 is generated with nuclei pixels as foreground. The local maxima $\{B_i\}_{i=1\dots M}$, which are in the background of mask b_1 are then eliminated, as they are generated due to the local noise in the image background. Let us denote the local maxima after this step as $\{B_i\}_{i=1\dots N}$, where N is the number of local maxima. In Fig. 4.13(a), the local maxima generated due to the local noise are highlighted by the (blue) circles. Fig. 4.13(b) shows the image $I(x, y)$ with local maxima $\{B_i\}_{i=1\dots N}$. Note that in Fig. 4.13(b) the (green) contours of nuclear regions are obtained after image binarization by adaptive thresholding.

(3) Clustering local maxima: As observed in 4.13(b), there are usually more than one point of local maxima for each nuclear region. In other words, it exists redundant detections of cell nuclei by searching local maxima from multiple response maps. But it is also noted that the local maxima corresponding to each nucleus usually cluster together with a closely geometrical distance. This is because the locations of detected nuclei seeds are mainly determined by the scales of gLoG kernels. In the previous module (i.e., Section 4.3.2), we generate the aggregated gLoG filters based on the same set of σ_x and σ_y for any specific orientation θ , and hence different orientation of responses can generate closely located local maxima. Fig. 4.14 illustrates an example of seed detection by different gLoG filters. In Fig. 4.14(a), the gLoG filters $\sigma_x = 8, \sigma_y = 7, \theta = \{45^\circ, 90^\circ\}$ are used, while in Fig. 4.14(b) the gLoG filters $\sigma_x = 8, \sigma_y = 7, \theta = 45^\circ$ and $\sigma_x = 5, \sigma_y = 4, \theta = 45^\circ$ are used. As observed in Fig. 4.14(a), the generated nuclear seeds are close with each other mainly due to the same scales of gLoG filters used. By contrast, nuclei seeds in Fig. 4.14(b) are far away from each other as the scales of gLoG filters differ greatly.

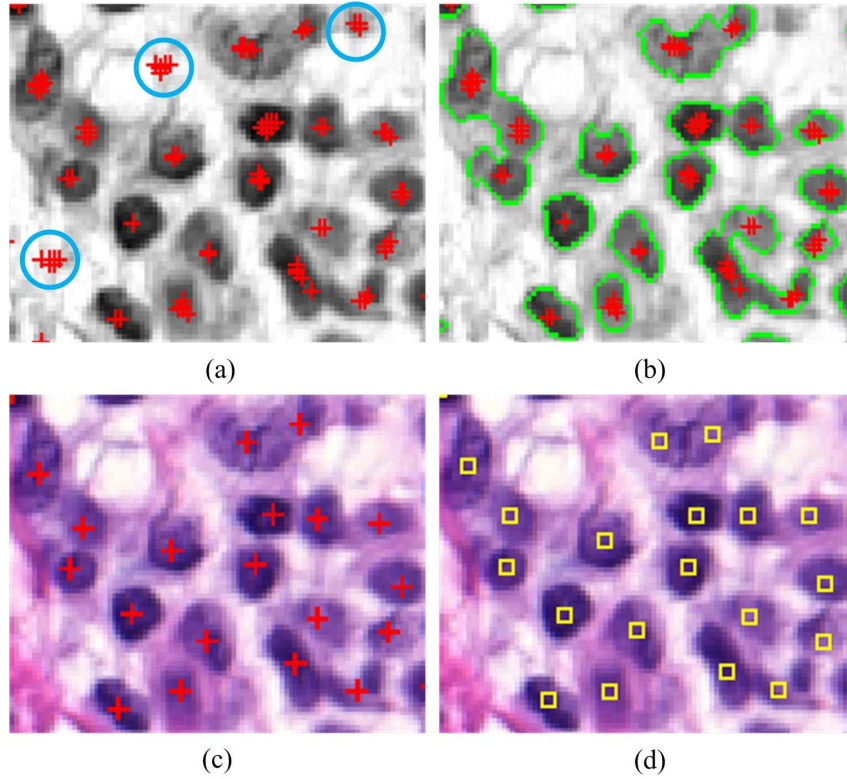


Fig. 4.13. Illustration of nuclei detection. (a) $I(x, y)$ with all local maxima (shown as +). (b) $I(x, y)$ with filtered local maxima (shown as +). (c) Original image with detected nuclei seeds. (d) Original image with ground truths (shown as squares). Note that in (a) circles indicate the local maxima generated by noise in the image background. In (b) nuclei contours are obtained by adaptive thresholding.

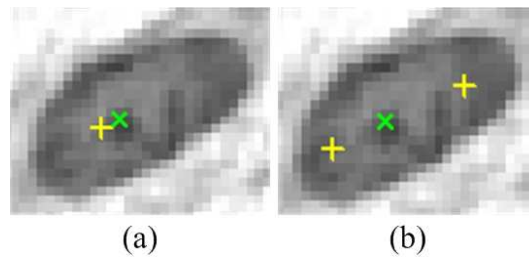


Fig. 4.14. Nuclei detection by different scales of gLoG kernels. (a) By gLoG filters $\sigma_x = 8$, $\sigma_y = 7$, $\theta = 45^\circ$ (shown as \times) and $\sigma_x = 8$, $\sigma_y = 7$, $\theta = 90^\circ$ (shown as +). (b) By gLoG filters $\sigma_x = 8$, $\sigma_y = 7$, $\theta = 45^\circ$ and $\sigma_x = 5$, $\sigma_y = 4$, $\theta = 45^\circ$ (shown as +). Note that nuclei seeds detected by gLoG filters with the same scales are close with each other in (a).

To cluster the local maxima, we apply the mean-shift algorithm [87] on N local maxima points $\{B_i\}_{i=1\dots N}$ and group them based on their geometric closeness such that each nucleus corresponds to one cluster of local maxima. The bandwidth w_2 of mean-shift algorithm is determined based on estimated radius of nuclei (e.g., 6

pixels in Fig. 4.13). Let the number of clusters be denoted by p after this step, and the p clusters be denoted by $\mathcal{T}_l, 1 \leq l \leq p$.

(4) Identifying nuclei seeds: Since each nucleus corresponds to one cluster of local maxima, it is necessary to determine one point as a nucleus seed. In this work, the point which has the maximum response value (retrieved from the corresponding response image $L^j(x, y)$ obtained in Eq. (4.23)) is selected as the nucleus seed from each cluster $\mathcal{T}_l, 1 \leq l \leq p$. Fig. 4.13(c) shows the original RGB image with the detected nuclear seeds (+ symbols) by the proposed gLoG-based technique, and Fig. 4.13(d) shows the manually labeled ground truths. As observed in Figs. 4.13(c)(d), the proposed gLoG-based technique has correctly detected most of cell nuclei, even for nuclei with light stains.

Once nuclei seeds are detected, we can easily estimate the orientation and scales of cell nuclei. Let us denote the detected nuclei seeds as $C_l, 1 \leq l \leq p$. Assume that the nucleus seed C_l corresponds to a peak response searched in the response map $L^j(x, y)$ (see Eq. (4.23)), and the orientation of the aggregated gLoG kernel ($\nabla^2 G_s^j(x, y)$) is considered as the orientation of the nucleus. The aggregated gLoG kernel $\nabla^2 G_s^j(x, y)$ is summed from the same directional gLoG filters. Let us assume that the $\nabla^2 G_s^j(x, y)$ is obtained by summing S normalized gLoG kernels $\nabla^2 G_n(x, y)$ (see Eqs (4.21), (4.22)). To estimate the scales of nuclei, we can separately convolve the nuclei with each gLoG kernel $\nabla^2 G_n(x, y)$ to obtain S response maps. The response map that has the highest response value at the nucleus seed point C_l is selected, and the corresponding scales of the gLoG kernel $\nabla^2 G_n(x, y)$ are estimated as the nuclei scales. Fig 4.15(a) illustrates an example of nuclei scales estimation. The overlapped ellipses have a semi-major axis of $\sqrt{2}\hat{\sigma}_x$, a semi-minor axis of $\sqrt{2}\hat{\sigma}_y$ and an orientation of $\hat{\theta}$, where $\hat{\sigma}_x$, $\hat{\sigma}_y$ and $\hat{\theta}$ are the estimated scales and orientations of nuclei. Using the detected seeds as markers, nuclei boundaries can also be efficiently obtained by such as marked watershed algorithm [91]. Fig. 4.15(b) shows nuclei segmentations by the marked watershed algorithm.

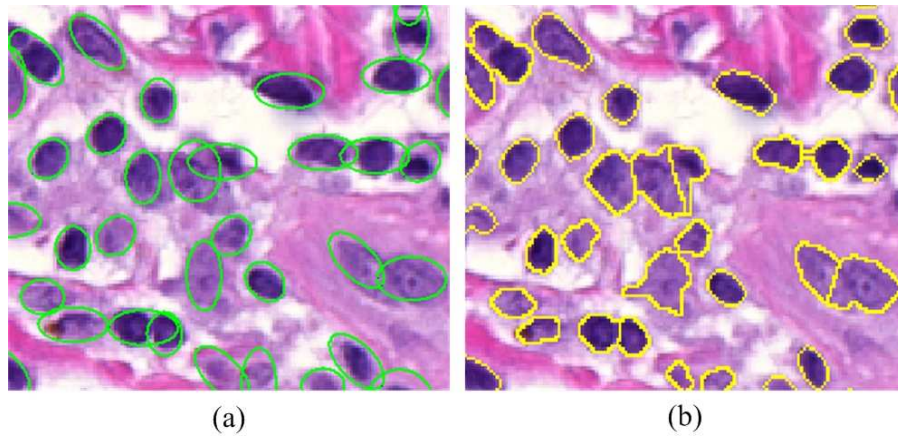


Fig. 4.15. Nuclei scales estimation and segmentation. (a) Nuclei ellipses generated based on scales and orientation estimation. (b) Marked watershed segmentation.

4.4 Performance Evaluation

To evaluate the performance of two proposed nuclei detection techniques, we carried out a series of experiments on skin and breast biopsy images. The main objective of evaluation is to determine if the automatically detected nuclei seeds are consistent with the manually labeled ground truths. In this section, we present the nuclei detection performance of two proposed techniques and compare them with several existing techniques.

4.4.1 Dataset Description

The study of nuclei detection was based on two datasets of H&E stained histopathology images. Table 4.5 lists the summary of two datasets. The first dataset (Dataset-I) includes 12 skin histopathological images, which were cropped skin DEJ regions from six WSIs (i.e., 2 images are cropped from each WSI). Note that the original skin images were captured under 40X magnification on Carl Zeiss MIRAX MI-DI scanning system. The cropped images were saved in TIFF format with 10X magnification for a good balance between image quality and size. The 12 images on average have a size of 1326×667 pixels, and have a total of 7,701 cell nuclei. The second dataset (Dataset-II) is a published online dataset [92], which consists of

18 breast histopathology images. The 18 breast images on average have a size of 896×768 pixels, and have a total of 6,048 cell nuclei. Since cell nuclei in Datasets I and II have similar size variations, we have used the same set of parameter values for two datasets during our subsequent evaluations.

TABLE 4.5
IMAGE DATASET USED FOR NUCLEI DETECTION EVALUATIONS.

Datasets	Tissue Types	Image No.	Nuclei No.
Dataset-I	Skin biopsy	12	7701
Dataset-II	Breast biopsy	18	6048

4.4.2 Parameter Configuration and Comparisons

In the proposed voting-based technique (henceforth referred to as the P-Voting technique), three parameters should be selected appropriately to obtain a good performance. These three parameters are the minimum voting radius r_{\min} , the maximum voting radius r_{\max} and the bandwidth w_1 of mean-shift clustering to find the modes of voting image. Note that all these parameters are determined by the estimated nuclei radii (see Algorithm 1). That is: $r_{\min} = 0.5d$, $r_{\max} = 1.5d$ and $w_1 = 0.6d$, where d is the estimated average nucleus radius. In order to estimate the value of d , we manually measure the radii of 170 randomly selected cell nuclei (from 15 image patches) and compute the average value as d (i.e., $d = 10$ pixels).

In the proposed gLoG-based technique (henceforth referred to as the P-gLoG technique), four parameters should be determined appropriately to obtain a good performance. These four parameters are the k (i.e., the number of orientations for gLoG kernels), scales $(\sigma_{\min}, \sigma_{\max})$ of gLoG kernels, and the bandwidth w_2 of mean-shift clustering to merge redundant seeds. The parameter k is set as 9, which corresponds to nine orientations (i.e., $\theta = m\pi/9$, $m = 0, \dots, 8$) with an angular resolution of 20° . The parameters σ_{\min} and σ_{\max} are determined based on the relationship between the radius d of detected objects and scales of gLoG kernels $\sqrt{2}\sigma_{\min} \leq d \leq \sqrt{2}\sigma_{\max}$ [88]. Here we measure the minimal and maximal

radii (d_{\min} , d_{\max}) of nuclei (in pixels) in 170 manually labeled nuclei regions, and set σ_{\min} and σ_{\max} as $d_{\min}/\sqrt{2}$ and $d_{\max}/\sqrt{2}$, respectively. The bandwidth w_2 is empirically set as the minimal radius d_{\min} of detected objects. That is: $\sigma_{\min} = 4$, $\sigma_{\max} = 8$ and $w_2 = 6$.

The proposed techniques have been compared with several existing techniques for blob (e.g., cell nuclei) detection. These techniques include the restricted randomized hough transform (RRHT) [93], DoG filters with blobness measure [81], mLoG method [39], gLoG method [40] and SPV method [43]. The parameters of these methods are mainly determined by the estimated nuclei size (i.e., nuclei radii range). For comparisons, we have experimentally set the optimal parameter values for these techniques such that the best performance is achieved with our testing images. Note that the original SPV method [43] determines the set of points for voting by empirically selecting the points with a large gradient magnitude. Since it is difficult to set an appropriate threshold to select large gradient points for different images, we use the Canny operator to compute image edges (e.g., nuclei boundaries) for voting. The original gLoG method [40] is highly sensitive to the image noise which cannot be directly applied on our testing datasets. We have determined a threshold τ based on trial and error tests to filter out false seeds in the image background. In particular, if the points of local maxima searched from the aggregated response map have a response value above τ (i.e., $\tau = 45$), they are considered as nuclei seeds. Figs. 4.16(a)(b) show nuclei detection results by the technique in [40] before and after filtering false seeds by the threshold τ . As observed in Fig. 4.16, the threshold τ can effectively remove false seeds in the image background.

4.4.3 Nuclei Detection Results

To evaluate the nuclei detection performance, the automatically detected nuclei seeds are compared with the manually labeled ground truths (provided by pathologists). An automatically detected seed s_i is counted as a true positive (TP) if and only if its location is in a detection pair (s_i, g_j) for which the corresponding (near-

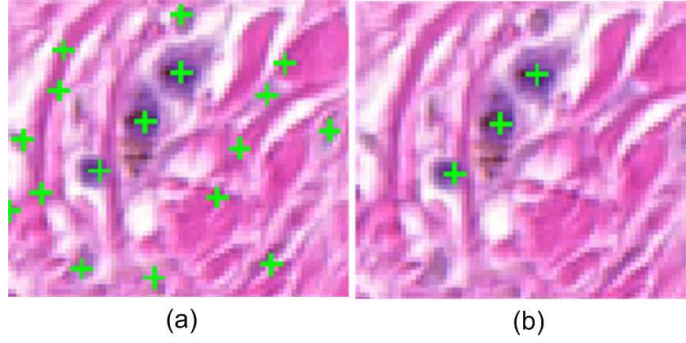


Fig. 4.16. Nuclei detection results by the technique in [40]. (a) Results before filtering false seeds. (b) Results after filtering false seeds.

est) ground truth seed g_j has not been paired [82], and also their Euclidean distance is within a certain range of t (number of pixels). Fig. 4.17 illustrates examples of TPs and false positives (FPs). Figs. 4.17(a)(b) show two simple examples of TP and FP, respectively. In Fig. 4.17(c), s_1 is counted as a TP because it has the closest distance (smaller than t) with g_1 , whereas s_2 is counted as a FP because g_1 has been used and paired with s_1 . In Fig. 4.17(d), if $t_1 < t_2 < t_3$, s_1 is paired with g_1 and counted as a TP. Meanwhile s_2 is paired with g_2 and counted as a TP. However, if $t_2 < t_1 < t_3$, s_1 is paired with g_2 and counted as a TP, but s_2 is now counted as a FP.

Let N_{GT} , N_{DS} and N_{TP} denote the numbers of ground truth seeds, detected seeds and true positives seeds, respectively. The performance is evaluated with respect to the recall (\mathcal{D}_{REC}), precision (\mathcal{D}_{PRE}) and F-measure (\mathcal{D}_F) which are defined as follows:

$$\mathcal{D}_{REC} = \frac{N_{TP}}{N_{GT}} \times 100\% \quad (4.24)$$

$$\mathcal{D}_{PRE} = \frac{N_{TP}}{N_{DS}} \times 100\% \quad (4.25)$$

$$\mathcal{D}_F = \frac{2 \times (\mathcal{D}_{REC} \times \mathcal{D}_{PRE})}{(\mathcal{D}_{REC} + \mathcal{D}_{PRE})} \quad (4.26)$$

Table 4.6 shows the quantitative evaluations of nuclei detections for the two proposed techniques (i.e., P-Voting and P-gLoG) and existing techniques, where the distance t is set as 6 pixels. Note that F-measure \mathcal{D}_F is the harmonic mean of

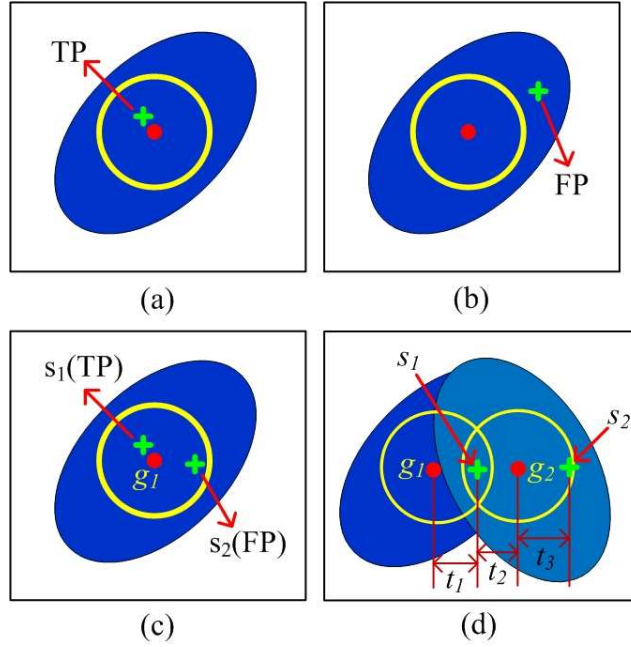


Fig. 4.17. Definition of true positives and false positives. Note that in (d) t_1 , t_2 and t_3 are the Euclidean distances corresponding to different pairs of points.

precision and recall, which provides a comprehensive measurement of the performance. It is observed in Table 4.6 that the RRHT [93] and SPV [43] techniques provide much poorer performances than other techniques, with only around 78% and 74% of \mathcal{D}_F on Datasets-I and II, respectively. The poor performance of these two techniques is mainly because they are critically relied on the accuracy of detected image edges, and hence inaccurate edge detections have significantly negative influence on nuclei detections. The DoG [81] and mLoG [39] techniques provide intermediate performances, which are mainly limited by either consideration of nuclei as circular blobs or filtering of false seeds using unstable post-pruning approaches. The gLoG method [40] and P-Voting technique achieve a similar performance. More specifically, the P-Voting technique provides a better performance (88.58% of \mathcal{D}_F) than the gLoG method [40] in Dataset-I, but a poorer performance (82.76% of \mathcal{D}_F) in Dataset-II. The relative poor performance on Dataset-II is because the P-Voting technique is highly dependent on image binarizations, while the breast images in Dataset-II tend to be more difficult to be binarized. Among all techniques, the P-gLoG technique achieves the highest F-measures for two testing datasets (on

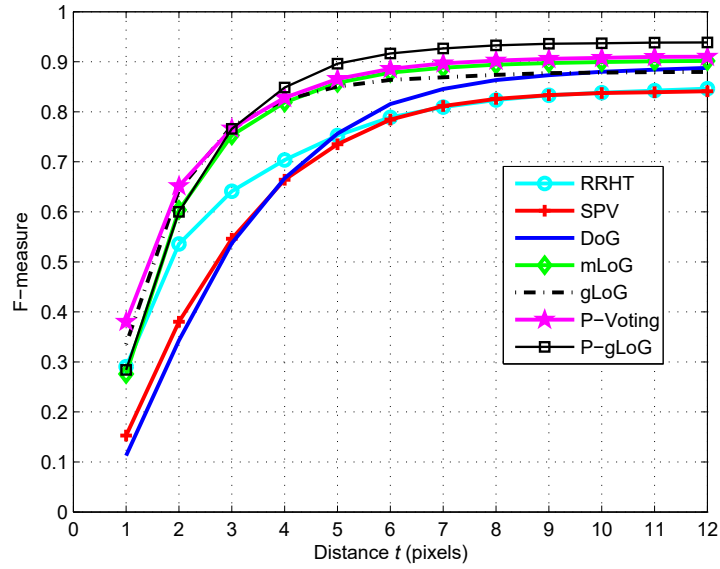
TABLE 4.6
COMPARISON OF NUCLEI DETECTION PERFORMANCE ($t = 6$).

Datasets	Techniques	N_{GT}	N_{DS}	$\mathcal{D}_{REC}(\%)$	$\mathcal{D}_{PRE}(\%)$	$\mathcal{D}_F(\%)$	$\mathcal{T}(s)$
Dataset-I (12 Images)	RRHT [93]	7701	6613	72.97	85.94	78.89	13.0
	SPV [43]	7701	6764	74.47	83.18	78.42	334.8
	DoG [81]	7701	8193	83.05	80.20	81.53	32.5
	mLoG [39]	7701	7991	88.67	87.08	87.83	7.4
	gLoG [40]	7701	7052	83.63	89.55	86.39	36.3
	P-Voting	7701	7400	86.17	91.21	88.58	173.4
	P-gLoG	7701	7424	89.45	94.20	91.72	11.3
Dataset-II (18 Images)	RRHT [93]	6048	6553	76.42	70.74	73.27	7.3
	SPV [43]	6048	5510	70.47	78.61	74.14	215.1
	DoG [81]	6048	6430	78.39	73.88	75.92	9.1
	mLoG [39]	6048	6429	85.07	79.96	82.35	3.0
	gLoG [40]	6048	5795	83.94	87.28	85.50	27.5
	P-Voting	6048	5893	81.81	83.91	82.76	92.9
	P-gLoG	6048	5963	87.44	88.42	87.87	6.7

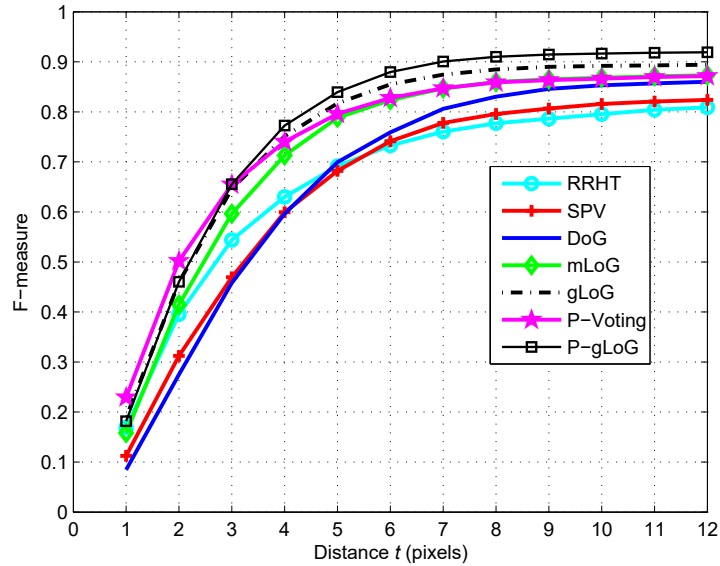
average 91.72% and 87.87%, respectively). This indicates that the gLoG kernels are superior to other methods in detecting elliptical blobs. In addition, the post-pruning algorithms based on mean-shift clustering and adaptive thresholding can efficiently remove false seeds.

Fig. 4.18 compares nuclei detection performance with respect to F-measures of different techniques, where the distance t is ranged from 1 to 12 pixels. As observed in Fig. 4.18, the \mathcal{D}_F values of different techniques go up exponentially when t values increase from 1 to 6 pixels, but then go up marginally when t values increase from 6 to 12 pixels. This indicates that most of detected true positive seeds by automated techniques are lying within a distance of 6 pixels from manually labeled ground truth seeds. The proposed techniques (especially the P-gLoG technique) provide a markedly superior performance with different t values.

For visual comparison, Fig. 4.19 shows two randomly cropped image patches with nuclei seeds detected by different techniques. In Fig. 4.19, the first and third rows show a skin image, while the second and fourth rows show a breast image. The automatically detected nuclei seeds are indicated by (green) + symbols, while the ground truths shown in Fig. 4.19(h) are indicated by yellow \square symbols. The rectangles in the images highlight missed or false seeds generated



(a)



(b)

Fig. 4.18. F-measures with different distance t values. (a) Performances on Dataset-I. (b) Performances on Dataset-II.

by computerized methods. As observed in Fig. 4.19, the proposed techniques (see Figs. 4.19(f)(g)) provide a superior performance in nuclei detections, which have less false and missed detections than existing techniques. Although a good performance is achieved by proposed techniques, there still exist failure cases when nuclei are closely clustered together. If cell nuclei are severely clustered together,

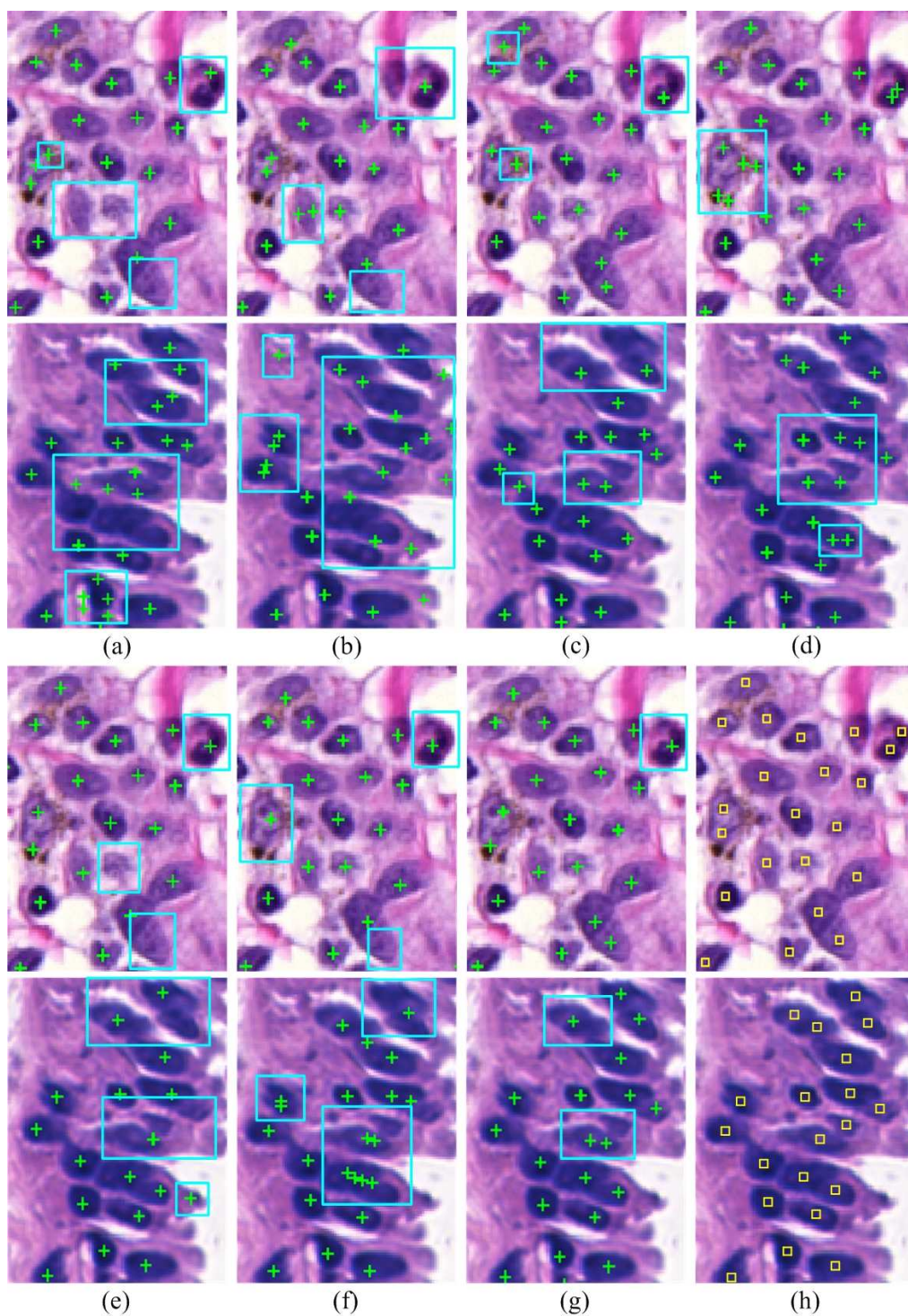


Fig. 4.19. Nuclei detection in skin (first and third rows) and breast (second and fourth rows) images. (a) RRHT [93]. (b) SPV [43]. (c) DoG [81]. (d) mLoG [39]. (e) gLoG [40]. (f) P-Voting. (g) P-gLoG. (h) Ground truths. Note that automatically detected nuclei seeds are indicated by + symbols, while ground truths in the column (h) are indicated by \square symbols. The missed or false seeds are highlighted with rectangles.

it is actually very difficult to be identified even by an experienced pathologist. Fig. 4.20 shows nuclei detections by the P-gLoG technique in a skin image, where the second row shows two zoomed-in image patches for better illustration. As seen in Fig. 4.20, the P-gLoG technique can correctly detect most of cell nuclei.

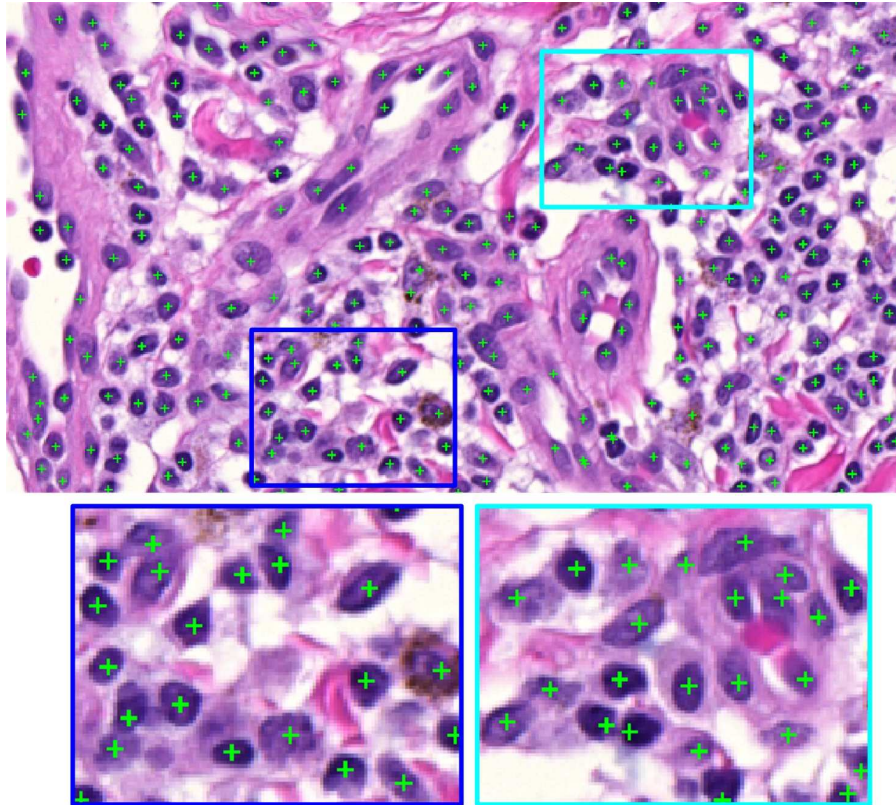


Fig. 4.20. Nuclei detection by the P-gLoG technique. Note that the second row shows zoomed-in image patches for better viewing.

4.4.4 Parameter Sensitivity Evaluations

As explained in section 4.4.2, the estimated nucleus radius d plays a key role in determining the values of other parameters in the P-Voting technique. To evaluate the sensitivity of the parameter d , its value is varied by 10% and 20% around the estimated optimal value ($d = 10$). Fig. 4.21 shows the F-measures with different values of parameter d . As observed in Fig. 4.21, a selection of smaller values for d (e.g., 8 and 9 pixels) roughly does not change F-measures in two datasets, while a selection of larger values for d (e.g., 11 and 12 pixels) has marginally decreased F-

measures about 2% for two datasets. Considering the slight changes of F-measures with different d values, it can be concluded that the P-Voting technique is not very sensitive to the parameter d , but using an appropriate value of d is still important for achieving the best performance.

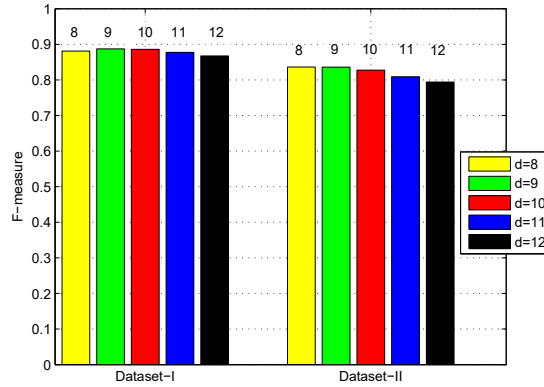


Fig. 4.21. Parameter sensitivity evaluation of the P-Voting technique.

There are four key parameters that should be predefined for the P-gLoG technique. Parameters σ_{\min} and σ_{\max} can be intuitively determined based on the average radius of detected nuclei. Parameter k determines the angular resolution of gLoG kernels, which is empirically set as 9 (corresponds to 20° angular resolution). To evaluate the sensitivity of this parameter, we test k with values $\{7, 8, 10, 11\}$ on two datasets. Fig. 4.22(a) shows the F-measures with different k values on two datasets. As seen in Fig. 4.22(a), the F-measures for the same dataset are similar with different k values. This indicates that the P-gLoG technique is generally robust to the choice of parameter k . However, note that since k determines the angular resolution of gLoG kernels, with a larger k value, more accurate estimation of nuclear orientation will be obtained. Another adjustable parameter in the P-gLoG technique is the bandwidth w_2 of mean-shift algorithm which is used for post-pruning to merge redundant seeds. The parameter w_2 is empirically set as the minimal nuclear radius (i.e., $w_2 = 6$ pixels). To evaluate the sensitivity of this parameter, we varied w_2 by 10% and 20% around 6 (or tuned it from 4.8 to 7.2 with a step of 0.6). Fig. 4.22(b) shows the F-measures with respect to different w_2 values

on two datasets. As observed in Fig 4.22(b), the F-measures only have marginal fluctuations, and hence the P-gLoG technique is not sensitive to the parameter w_2 when a reasonable value is selected.

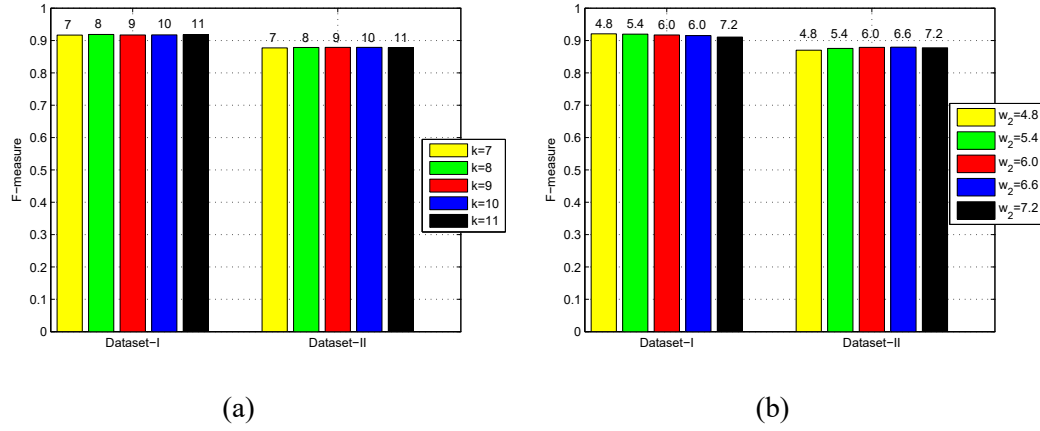


Fig. 4.22. Parameter sensitivity evaluation of the P-gLoG technique. (a) F-measures with different k values. (b) F-measures with different w_2 values.

4.4.5 Computational Complexity and Limitations

To evaluate the computational complexity, the average run time \mathcal{T} on two datasets is separately calculated. All our experiments were carried out on a 2.9-GHz CPU with 4 GB RAM using MATLAB R2013a. In Table 4.6, the eighth column shows the values of \mathcal{T} for different techniques on Datasets-I and II, respectively. It is observed in Table 4.6 that the SPV [43] and the P-Voting techniques have much higher computational complexity (e.g., the SPV [43] takes 334.8s and 215.1s on two datasets, respectively) than other techniques. These two techniques detect nuclei seeds based on a voting procedure, which requires intensive computations. However, since the potential voting regions are only the eroded nuclear regions rather than the whole image (like the SPV [43]), the P-Voting technique is much faster (about 2 times) than the SPV method. The P-gLoG technique is more efficient than voting-based techniques, which takes about 11.3s and 6.7s for Datasets-I and II, respectively. Note that the P-gLoG technique is about three times faster than the gLoG method in [40]. This is due to the accumulations of gLoG kernels with

the same orientations before performing convolution operation, which reduces the number of convolutions (convolution takes a relatively long time for a large image).

It should be mentioned that like every computerized method, the proposed techniques for nuclei detections also have limitations. Based on our experience, the P-Voting technique is heavily relied on accurate image binarizations, while the P-gLoG technique may fail to detect cell nuclei that do not have homogeneous gray level intensities. However, these weaknesses can be relieved by utilizing appropriate image binarization methods and performing image pre-processings to make nuclear region homogenous.

4.5 Summary

This chapter presents two automated techniques for nuclei detection in histopathological images. The first proposed technique, P-Voting, performs nuclei detection based on ellipse descriptor analysis and an improved voting algorithm, while the second proposed technique, P-gLoG, performs nuclei detection by using multi-scale gLoG kernels with redundant seeds merged by mean-shift clustering. The P-Voting technique provides a very good performance in nuclei detection compared to existing techniques, but it has a high computational complexity. The P-gLoG technique provides a further improvement in nuclei detection, with a significantly low computational complexity. The proposed techniques have been thoroughly evaluated on two datasets of skin and breast histopathological images. The proposed techniques can be used to evaluate nuclei densities and distributions in biopsy images, and also they are the key steps for automated nuclei segmentation.

Chapter 5

Automated Segmentation of Cell

Nuclei

Analysis of cell nuclei such as nuclear size, shape and presence or absence of nucleoli plays an important role in the histopathological examination and diagnosis. However, it is very labor-intensive and error-prone for manual analysis of cellular features, as there is usually a large amount (e.g., millions of cell nuclei) of cell nuclei in biopsy images. Although many existing computerized methods have been proposed for nuclei segmentation, robust and efficient segmentation of cell nuclei is still an open problem because of overlapping nuclei, inhomogeneous staining, and presence of noisy pixels and other tissue components. In this chapter, we propose an automated technique that segments cell nuclei using a multi-scale radial line scanning (mRLS) based method. Comparisons with several existing techniques indicate that the proposed technique can segment cell nuclei more accurately.

5.1 Introduction

Evaluation of cell nuclei plays an important role in the histopathological examination and analysis of digitized slides. The parameters like nuclear size, shape,

Part of this chapter has been published in: Xu *et al.* "Automatic nuclear segmentation using multi-scale radial line scanning with dynamic programming", *accepted by IEEE Transactions on Biomedical Engineering (TBME)*, January, 2017.

contours and presence or absence of nucleoli are generally used by pathologists for various cancer diagnosis. The abnormal cell distributions and morphological changes often indicate the malignancy of the lesion [38]. However, manual segmentation and evaluation of cell nuclei are very tedious works, and also are prone to inter- or intra-observer variability due to poor contrast and clumping of nuclei. To reduce the workload and obtain more objective results, many efforts have been devoted to develop automated nuclei segmentation techniques.

Threshold-based techniques have been widely used for nuclei segmentation in histological images. These works are related to segmentation of cell nuclei in the digitized neuroblastoma slides [27], breast histopathological images [24] and skin histopathological images [30]. A major limitation of these threshold-based techniques is the under-segmentation of clumped cell nuclei. In order to address this issue, a number of techniques for segmentation of cellular or nuclei clumps have been reported. These techniques mainly include two categories: concavity analysis based techniques [79], [94]–[99], and “seed” (or “marker”) based techniques [32]–[35], [43], [91].

For concavity analysis based techniques, the first step is to identify valid concave points for splitting. Several techniques exist for identifying concave points on the contour. For example, the concave points have been detected based on computing the cross product of vectors defined by three consecutive points [79], or minimizing a cost function to find the bottleneck positions on the contour [94]. With detected concave points, the next step is to connect them and split cellular or nuclear clumps into isolated ones. This step can be achieved iteratively by connecting the shortest path between concave points until all split regions are considered to be isolated [98]. Although concavity analysis based techniques have been reported to provide a good performance, the performance is overly sensitive to image binarizations, and the predefined thresholds (usually several thresholds) that are used for finding valid concave points [94], [97].

For “seed” (or “marker”) based techniques, the first step is to identify “seeds” for nuclear regions. These detected “seeds” are usually close to nuclei centers, and

used as the starting points for active contours or watershed segmentation. The segmentation performance depends critically on the accuracy and reliability of initial seed points. Because of the advantages of simplicity, speed and absence of adjustable parameters, several marked watershed algorithms have been developed for cellular or nuclear segmentation. For example, the markers of watershed algorithm are detected by condition erosion [32], H-minima transform [33], [34], improved voting algorithm [91] and radial symmetry transform [35]. These watershed based segmentations are sensitive to the accuracy of image binarizations as the topographic surfaces for segmentation are built from distance maps of binary images. In order to get more accurate segmentations, Al-Kofahi *et al.* [39] proposed a technique that detects nuclei seeds by mLoG filters and segments cell nuclei based on size-constrained clustering with further refinement by a graph-cuts based algorithm. Qi *et al.* [43] proposed to detect cell seeds in breast histopathological images by a SPV algorithm and delineate cell contours by a repulsive level set model [42]. Xing *et al.* [100] proposed a cell detection and segmentation technique for evaluating Ki-67 proliferation index in gauge neuroendocrine tumor images. This technique detects cell seeds by a hierarchical voting method, and delineates cell boundaries by a repulsive active contour model. The active contour based techniques tend to provide more accurate segmentations, but they (e.g., in [43], [100]) cannot delineate occluded nuclear boundaries for overlapping nuclei.

Recently a few techniques based on active shape models and deep learning (DL) have been proposed for nuclei segmentation. Plissiti *et al.* [101] proposed a technique that segments overlapping cell nuclei by driving the deformable model based on physical principals, but this technique only works automatically for the image with a known number of cell nuclei. Lu *et al.* [102] proposed a technique to perform cervical cell segmentation by joint optimization of multiple level set functions. This technique has a high computational complexity and its performance is critically relied on initial segmentations for level set functions. Janowczyk *et al.* [103] proposed a resolution adaptive deep hierarchical (RADHical) learning scheme for nuclei segmentation in histological images. The RADHical employs the

AlexNet network [104] in a multi-resolution framework which provides a similar performance with the naive DL model [105] (e.g., processing the high-resolution images directly), but it achieves about 6 times speed improvement. Although the DL methods are very powerful to learn image features, they are generally quite time-consuming to train and apply.

In this chapter, we propose a “seed” based technique for nuclei segmentation in histological images, which can resolve nuclear overlap and obtain occluded nuclei boundaries in clumped nuclei. The proposed technique detects nuclei seeds by using multi-scale gLoG kernels, and delineate nuclei boundaries by using a mRLS method. The details of the proposed technique and experimental results are provided in the following sections.

5.2 Proposed mRLS-based Technique

The schematic of the proposed mRLS-based technique for nuclei segmentation is shown in Fig. 5.1. It is observed that there are two main modules. In the first module, nuclei seeds are detected using a bank of gLoG kernels with different scales and orientations. In the second module, nuclei boundaries are delineated by using a multi-scale radial line scanning (mRLS) based method. This method consists of three sub-modules. The first module estimates candidate nuclei contours using mRLS with dynamic programming (DP) based on gradient information. The second module selects the optimal nuclei contours by integrating nuclear intensity and shape prior with gradient information. The third module resolves overlapped nuclei based on a predefined area-overlapping allowance. In the following, we explain the two modules with an emphasis on nuclei segmentation module.

5.2.1 Seeds Detection

Because cell nuclei in digitized histological images typically have circular or elliptical shapes, we apply a gLoG kernels based technique (presented in Chapter 4) to detect nuclei seeds. This technique first generates a bank of gLoG kernels with

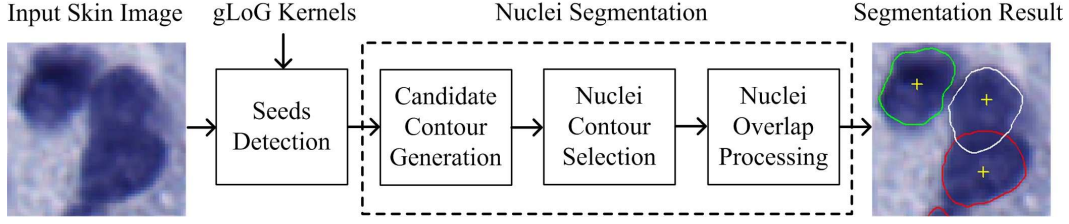


Fig. 5.1. Schematic of the proposed mRLS-based technique.

different scales and orientations. The gLoG kernel $\nabla^2 G(x, y)$ is defined as:

$$\nabla^2 G(x, y) = \frac{\partial^2 G(x, y; \sigma_x, \sigma_y, \theta)}{\partial x^2} + \frac{\partial^2 G(x, y; \sigma_x, \sigma_y, \theta)}{\partial y^2} \quad (5.1)$$

where $G(x, y; \sigma_x, \sigma_y, \theta)$ is a generalized Gaussian kernel with scales (σ_x, σ_y) and orientation θ [106]. After generating multi-scale gLoG kernels, the same directional kernels are aggregated together and convolved with the image $I(x, y)$ (the red channel of original color image) to generate response maps. The points of local maxima are then searched from multiple response maps and clustered into different groups using mean-shift algorithm [87]. Finally, the geometrical means of local maxima points in different groups are computed and considered as nuclei seeds. Fig. 5.2 shows examples of gLoG kernels with different scales and orientations. Figs. 5.3(a)(c) shows two skin histological images where nuclei are observed as blue blobs. Figs. 5.3(b)(d) shows seed detection results using gLoG kernels shown in Fig. 5.2 (e.g., 95 gLoG kernels), where the detected seeds are indicated by (yellow) crosses. Note that Figs. 5.3 (a)(b) shows skin images stained with Ki-67, while Figs. 5.3 (c)(d) shows skin images stained with H&E. As observed in Fig. 5.3, nuclei have been correctly detected by the gLoG kernels based method [106].

5.2.2 Nuclei Segmentation

After detecting nuclei seeds, the next step is to delineate nuclei boundaries. In order to capture nuclei boundaries, we propose a mRLS based technique in this module. In the following, we explain the mRLS based technique in three sub-modules: candidate contour generation, nuclei contour selection and nuclei overlap

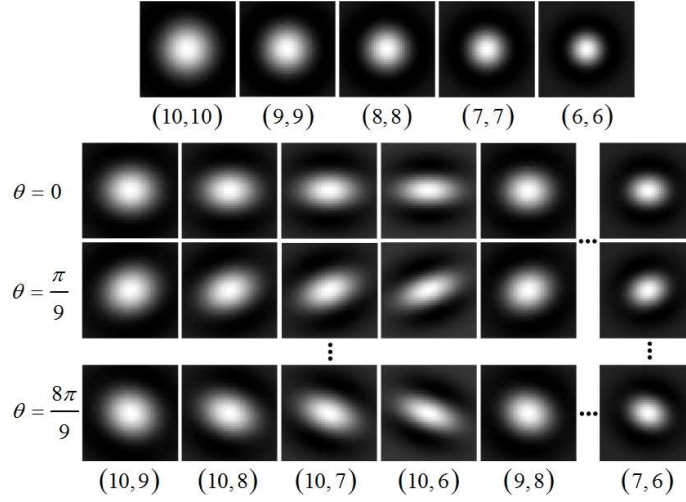


Fig. 5.2. gLoG kernels. Note that parameters (\cdot, \cdot) below a kernel correspond to (σ_x, σ_y) .

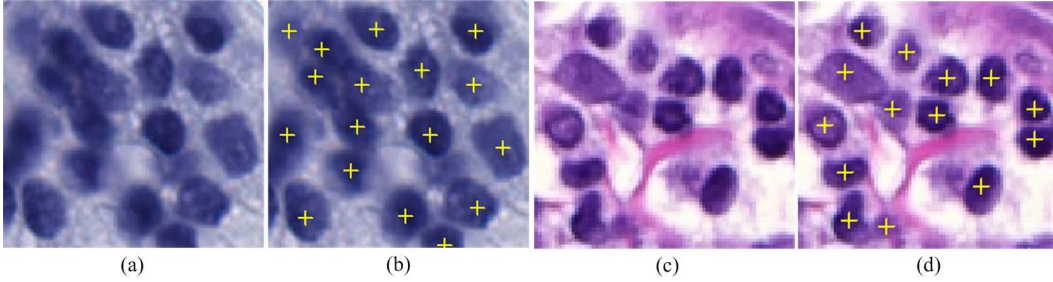


Fig. 5.3. Nuclei detections. (a) & (c) Original skin images. (b) & (d) Nuclei detection results. Symbols + in (b)(d) indicate detected nuclei seeds.

processing. Details of these three modules are as follows:

5.2.2.1 Candidate Contour Generation

The radial line scanning (RLS) method has been successfully utilized to detect melanocytes in skin epidermis areas [107]. To utilize the RLS method, the detected seeds are considered as nuclei centers. A set of radial lines emerging from the nucleus center are then computed. These radial lines have radial orientations as follows:

$$\theta_i = \frac{2\pi}{N} (i - 1), \quad i = 1 \dots N \quad (5.2)$$

where N indicates the number of radial lines. For each radial line, we denote the minimum scanning radius as R_S and maximum scanning radius as R_E . Fig. 5.4(a)

illustrates an example of scanning radii for one radial line, where $\theta_1 = 0$. Fig. 5.4(b) shows a real example with 30 radial lines (i.e., $N = 30$) emerging from the detected nucleus seed. The (yellow) contour in Fig. 5.4(b) is the reference nuclear boundary. As observed in Fig. 5.4(b), each radial line intersects the nuclear boundary at only one point. Let us assume there are M discrete points on each radial line between R_S (green circle) and R_E (blue circle). Hence, we have M^N different possible closed contours, where a contour uses only one point in each of N lines. The nuclear boundary shown in Fig. 5.4(b) is one of these possible configurations.

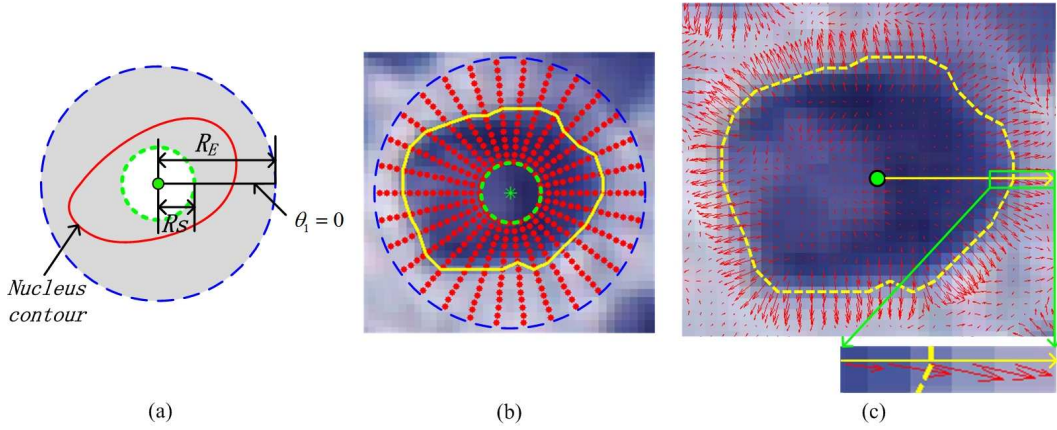


Fig. 5.4. Illustration of radial line scanning. (a) Scanning radii of one radial line. (b) 30 radial lines emerging from a detected nucleus seed. (c) Gradient map overlapped on a nuclear image. Note that in (b)(c) nuclear boundaries are indicated by yellow contours.

In order to automatically delineate a nuclear boundary (like it shown in Fig. 5.4(b)) from M^N different configurations, image gradient is utilized. Since the original image $I(x, y)$ usually includes noisy pixels, a smooth image $I_s(x, y)$ is first obtained by convolving the image with a Gaussian kernel as follows:

$$I_s(x, y) = G(x, y; \sigma) * I(x, y) \quad (5.3)$$

where $G(x, y; \sigma) = \sigma^{-1/2} e^{-(x^2+y^2)/4\sigma}$. The gradient magnitude map $\mathcal{G}_m(x, y)$ and gradient direction map $\mathcal{G}_d(x, y)$ are then generated using following equations:

$$\mathcal{G}_m(x, y) = \sqrt{\left(\frac{\partial I_s(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I_s(x, y)}{\partial y}\right)^2} \quad (5.4)$$

$$\mathcal{G}_d(x, y) = \arctan\left(\frac{\partial I_s(x, y)}{\partial y} / \frac{\partial I_s(x, y)}{\partial x}\right) \quad (5.5)$$

Fig. 5.4(c) shows a close-up view of nucleus overlapped with the gradient map, where gradients are indicated by arrows. As seen in Fig. 5.4(c), the reference nuclear boundary consists of high gradient points. In addition, if assuming that the directions of radial lines are outward (i.e., pointing from the nuclear center to the boundary), nuclear boundary points have gradient directions similar to that of corresponding radial lines. Based on these observations, we append a cost value for each point of radial lines. Let p_i denote a variable that can take any of the M discrete points on the i th radial line. The appended cost value $C(p_i)$ is computed as:

$$C(p_i) = \alpha \cos(|\theta_i - \mathcal{G}_d(p_i)|) + \beta \mathcal{G}'_m(p_i) \quad (5.6)$$

where α and β are weights of gradient directions and magnitudes, respectively. $\mathcal{G}'_m(p_i)$ is the normalized $\mathcal{G}_m(p_i)$, which is calculated as follows:

$$\mathcal{G}'_m(p_i) = \frac{\mathcal{G}_m(p_i)}{\max \mathcal{G}_m(p_i)}, \quad 1 \leq i \leq N \quad (5.7)$$

Note that in Eq. (5.6) the gradient direction is incorporated into the cost value $C(p_i)$, which can suppress undesirable influence of high gradient points from other neighboring nuclei boundaries. This is because orientations of radial lines for one nucleus, in an ideal case, are opposite to gradient directions of other neighboring nuclei boundaries.

After appending a cost value to each point of radial lines, it might appear that points of nuclei boundaries can be selected by maximizing the cost value $C(p_i)$ in each radial line. However, such simple selection of boundary points is difficult to obtain a smooth nuclear contour, as the determined boundary points are prone to be disturbed by isolated high gradient noisy pixels. In order to obtain accurate and smooth nuclear contours, we consider all radial lines together and define the

following cost function (corresponding to a nuclear *gradient*):

$$E_{gradient}(p_1, p_2, \dots, p_N) = E_1(p_1, p_2) + E_2(p_2, p_3) \\ + \dots + E_{N-1}(p_{N-1}, p_N) + E_N(p_N, p_1) \quad (5.8)$$

where each additive cost value E_i is calculated as:

$$E_i(p_i, p_{i+1}) = \begin{cases} \kappa - C(p_i) & \text{if } D(p_i, p_{i+1}) \leq \delta \\ \infty & \text{otherwise} \end{cases} \quad (5.9)$$

where $\kappa = \max C(p_i)$, $1 \leq i \leq N$. $D(p_i, p_{i+1})$ is the Euclidean distance between points p_i and p_{i+1} , and δ is a predefined distance value (e.g., $\delta = 1$). Note that the cost function in Eq. (5.8) has a typical overlapping and additive form that can be efficiently optimized by DP in $O(NM^2)$ computations [108]. The cost component in Eq. (5.9) implies that if the Euclidean distance D between two consecutive points on the contour are within a predefined distance δ , the cost value E_i equals to $\kappa - C(p_i)$. Otherwise the cost value is assigned to be positive infinity. With Eq. (5.9) as the individual cost component, the net effect of minimizing Eq. (5.8) would be to obtain a contour consisting of points with high gradient magnitudes and gradient directions similar to orientations of corresponding radial lines. Besides, any two consecutive points on the contour have the Euclidean distance smaller than δ .

Although the RLS technique with DP can generally provide smooth and accurate nuclei contours, the technique is relatively sensitive to the parameter R_E (the maximal scanning radius). This is because nuclei size and shape in histological images (even for the same type of images) usually vary from each other, and there exist touching and overlapping nuclei in 2-D images. When a small R_E value is applied, the large nuclei in the image may be truncated (see Figs. 5.6(c)(d) with arrows). However, when a large R_E value is applied, it not only increases the computational complexity but also may under-split clustered nuclei (see Figs. 5.6(e)(f) with arrows). Therefore, it is difficult to obtain accurate boundary contours for all nuclei in a large image by using a fixed R_E value. In order to overcome this

limitation, we propose to use a set of R_E values (i.e., $R_E^{\min} \leq R_E \leq R_E^{\max}$) rather than a fixed value to adapt nuclear size variations. The parameters R_E^{\min} and R_E^{\max} can be estimated based on nuclei radii in the image. Using a set of R_E values (or mRLS with DP), each nucleus has several obtained contours. For example, Fig. 5.5(a) shows a skin image with three cell nuclei. Figs. 5.5(b)-(f) show the nuclei contours obtained by the mRLS technique with DP, where $R_S = 3$ and $R_E = \{14, 15, 16, 17, 18\}$, respectively. As observed in Figs. 5.5(b)-(f), with different R_E values, each nucleus has several different obtained contours. Since there are several candidate contours for each nucleus, the optimal contour is to be determined in the next module.

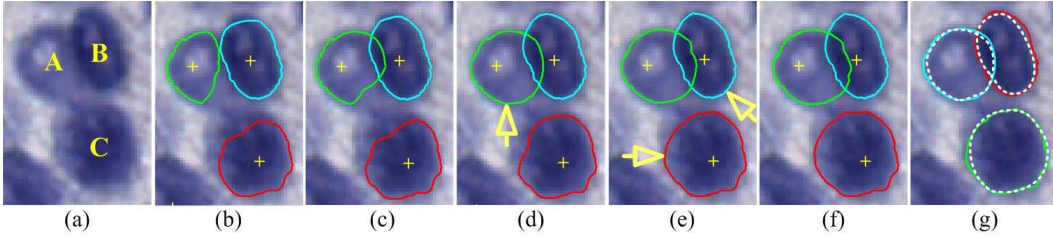


Fig. 5.5. Candidate nuclei boundaries. (a) Skin image with three cell nuclei (40X magnification). (b) $R_E = 14$. (c) $R_E = 15$. (d) $R_E = 16$. (e) $R_E = 17$. (f) $R_E = 18$. (g) Automatic segmentations compared with manual segmentations. Note that parameter settings for mRLS with DP are: $N = 100$, $R_S = 3$, $\alpha = 1.0$, and $\beta = 0.5$. In (b)-(f) symbols + are the detected nuclei seeds. In (d)(e) arrows indicate optimal nuclei contours. In (g) dotted contours indicate manual segmentations.

5.2.2.2 Nuclei Contour Selection

Several boundary contours are obtained for every nucleus by using the mRLS with DP. In this module, we select the optimal contour for each nucleus. In order to select the optimal contour, we integrate nuclear intensity variance and shape prior with the gradient information (obtained by Eq. (5.8)). Based on the observation that nuclear blobs in our dataset have relatively homogeneous intensities, the cost component corresponding to the *intensity variance* is defined as follows:

$$E_{intensity} = \iint_{(x,y) \in \Omega_N} |I(x,y) - c|^2 dx dy \quad (5.10)$$

where Ω_N denotes a segmented nucleus region, and c is the mean intensity in the region Ω_N . Note that a smaller $E_{intensity}$ value indicates a more homogeneous nucleus region.

The nuclei typically have circular or elliptical shapes, and hence the shape prior is incorporated for nuclear contour selection. Given a nuclear region Ω_N , an elliptical region Ω_E that has the same second moments as Ω_N is obtained by the ellipse-fitting algorithm [64]. Let \mathcal{A}_n and \mathcal{A}_e be the sets of pixels in Ω_N and Ω_E , respectively. The cost component E_{shape} corresponding to a nuclear *shape* is computed as follows:

$$E_{shape} = 1 - \frac{|\mathcal{A}_n \cap \mathcal{A}_e|}{|\mathcal{A}_e|} \quad (5.11)$$

where $|\cdot|$ is the cardinality of pixels set and \cap is intersection operation. E_{shape} has a dynamic range of $[0, 1)$, and a smaller value of the E_{shape} represents a closer match to an elliptical shape.

Let $\mathbf{E}_{gradient}$ be a vector of $E_{gradient}$ obtained by Eq. (5.8) with different R_E values. Similarly, let $\mathbf{E}_{intensity}$ and \mathbf{E}_{shape} be the vectors of $E_{intensity}$ (see Eq. (5.10)) and E_{shape} (see Eq. (5.11)), respectively. We integrate gradient, intensity and shape information together and define the following energy cost function:

$$\mathbf{E}_{total} = \lambda_1 \frac{\mathbf{E}_{gradient}}{\|\mathbf{E}_{gradient}\|} + \lambda_2 \frac{\mathbf{E}_{intensity}}{\|\mathbf{E}_{intensity}\|} + \lambda_3 \frac{\mathbf{E}_{shape}}{\|\mathbf{E}_{shape}\|} \quad (5.12)$$

where λ_1 , λ_2 and λ_3 are weights of each term, and $\|\cdot\|$ is the Euclidean norm. Note that \mathbf{E}_{total} is a vector of overall cost function for different R_E values. $\mathbf{E}_{gradient}$ assures that nuclei boundaries are located on high gradient points. $\mathbf{E}_{intensity}$ assures that the segmented nuclear region is homogenous. \mathbf{E}_{shape} assures that the segmented nuclear region has an elliptical shape. By selecting the minimal cost value in Eq. (5.12), it is expected to obtain a nuclear boundary lying on high gradient points and surrounding a region with an elliptical shape and homogeneous intensities. The parameters λ_1 , λ_2 and λ_3 can be adaptively adjusted according to different applications. For instance, a smaller value λ_2 can be utilized when nuclear regions in testing images are not homogeneous. In this work, we have empirically set them

TABLE 5.1
COST VALUES E_{total} FOR NUCLEI SHOWN IN FIG. 5.5.

Nuclei	E_{total}				
	$R_E=14$	$R_E=15$	$R_E=16$	$R_E=17$	$R_E=18$
A	1.66	1.57	1.47	1.52	1.50
B	1.74	1.59	1.49	1.48	1.48
C	1.62	1.64	1.60	1.52	1.54

as: $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_3 = 1$. Table 6.2 lists the costs E_{total} for nuclei A, B and C shown in Fig. 5.5(a), where $R_S = 3$, $14 \leq R_E \leq 18$. It is observed in Table 6.2 that the cost E_{total} for the nucleus A is the smallest when $R_E = 16$. In other words, the optimal nuclear contour, as indicated by the arrow in Fig. 5.5(d), is obtained when $R_E = 16$. For nuclei B and C, the optimal nuclear contours, as indicated by arrows in Fig. 5.5(e), are obtained when $R_E = 17$. Fig. 5.5(g) compares the finally obtained nuclei contours with manually labeled nuclei boundaries (shown as dotted contours). As seen in Fig. 5.5(g), the automatically obtained nuclei contours by the mRLS based technique have a good matching with manually labeled nuclei contours.

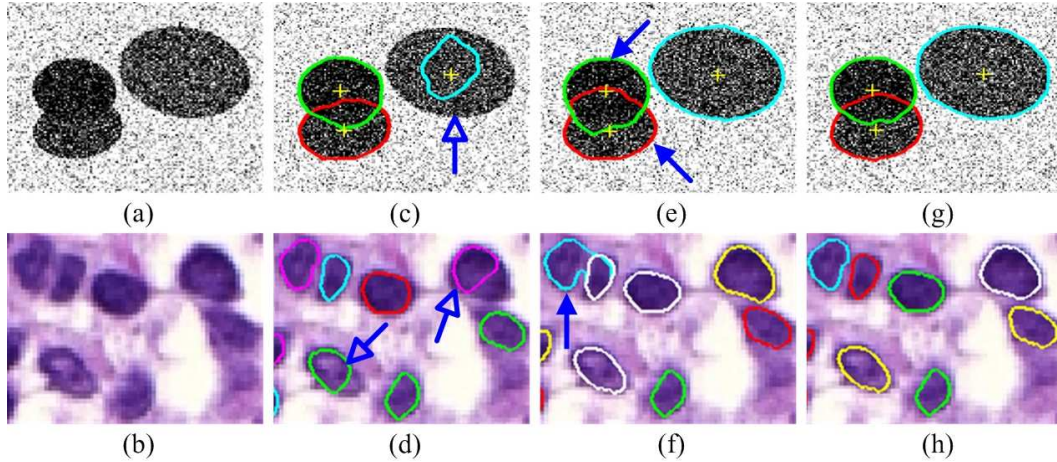


Fig. 5.6. Blob segmentations. (a) A synthetic image with three blobs. (b) A H&E stained skin image (25X magnification). (c) & (d) $R_S = 3$, $R_E = 7$. (e) & (f) $R_S = 3$, $R_E = 11$. (g) & (h) $R_S = 3$ and $7 \leq R_E \leq 11$. In (c)(d)(e)(f) arrows indicate inaccurate segmentations.

To further illustrate the advantages of the mRLS based technique, two examples of elliptical blob segmentations are shown in Fig. 5.6. The top row of Fig. 5.6 shows a synthetic image with three elliptical blobs and Gaussian white noise (with

0.1 variance), and the bottom row shows a H&E stained skin image. Figs. 5.6(c)(d) show blob segmentations with $R_S = 3$ and $R_E = 7$, where inaccurately-segmented blobs (indicated by hollow arrows) have been observed in both synthetic and real images due to a small R_E value used. Figs. 5.6(e)(f) show blob segmentations with $R_S = 3$ and $R_E = 11$. In Figs. 5.6(e)(f) although the inaccurately-segmented blobs in (c)(d) have been correctly segmented, inaccurate segmentations (indicated by solid arrows) have been observed for some other blobs because of a large R_E value used. Figs. 5.6(g)(h) show blob segmentations by the mRLS based technique, where $R_S = 3$ and $7 \leq R_E \leq 11$. It is observed that segmentation results in Figs. 5.6(g)(h) are more accurate than those in Figs. 5.6(c)(d)(e)(f), respectively.

5.2.2.3 Nuclei Overlap Processing

In previous steps, we presented a procedure to delineate each nucleus individually. The obtained nuclei contours may intersect with each other (see Fig. 5.5(g)), as cell nuclei usually overlap or touch together in 2-D images. Note that although excessive or sometimes complete overlap may also be occurred in real cases, manual evaluations usually consider excessively overlapped nuclei as single nucleus. In this module, we impose an overlap allowance limitation such that the segmented nuclear regions do not severely overlap with each other. The overlap limitation is imposed based on the Dice coefficient. Specially, for each group of overlapped nuclei, the following steps are iteratively applied until the stop criteria:

Step 1: Let s denote the number of individual candidate nucleus in the overlapped nuclei region. The Dice coefficient for two cell nuclei i and j is computed as follows:

$$D(\mathcal{A}_i, \mathcal{A}_j) = \frac{2 \times |\mathcal{A}_i \cap \mathcal{A}_j|}{|\mathcal{A}_i| + |\mathcal{A}_j|} \quad (5.13)$$

where \mathcal{A}_i and \mathcal{A}_j denote the set of pixels in the i th and j th nuclear regions, $1 \leq \{i, j\} \leq s, i \neq j$. The Dice coefficient $D(\mathcal{A}_i, \mathcal{A}_j)$ is a distance measure to evaluate the similarity of two sets [35]. The measure has a maximum value of 1 when two nuclear regions are completely overlapped with each other and a minimum value of

0 when two nuclear regions do not intersect.

Step 2: In this step, we first compute the largest Dice coefficient among $s(s-1)/2$ pairs of overlapped nuclei. Let us assume that the m th and n th nuclei have the largest Dice coefficient. If $D(\mathcal{A}_m, \mathcal{A}_n)$ satisfies the following stop condition:

$$D(\mathcal{A}_m, \mathcal{A}_n) < \tau \quad (5.14)$$

where τ is a predefined overlap allowance (e.g., $\tau = 0.6$), then the overlap processing stops. Otherwise, if $E_{total}(\mathcal{A}_m) < E_{total}(\mathcal{A}_n)$, the m th nucleus is kept and the n th nucleus is discarded, and vice versa. If this happens, the s value is reduced by 1 and the algorithm goes to the Step 3. Note that the cost values $E_{total}(\mathcal{A}_m)$ and $E_{total}(\mathcal{A}_n)$ are obtained by minimizing Eq. (5.12).

Step 3: If condition (5.14) is not satisfied and $s > 1$, we repeat Step 1 and Step 2. Otherwise, the algorithm stops.

By utilizing the above three steps, we try to add an overlap allowance for clustered nuclei such that the segmented nuclei are not severely overlapped with each other. Fig. 5.7(a) shows a cell nucleus, and Fig. 5.7(b) shows two detected nuclei seeds. Fig. 5.7(c) shows the obtained nuclei contours based on the two detected seeds, where segmented nuclear regions severely overlap with each other. Fig. 5.7(d) shows the finally obtained nuclear contour after overlap processing, where the dashed (white) nuclear contour in Fig. 5.7(c) has been discarded.

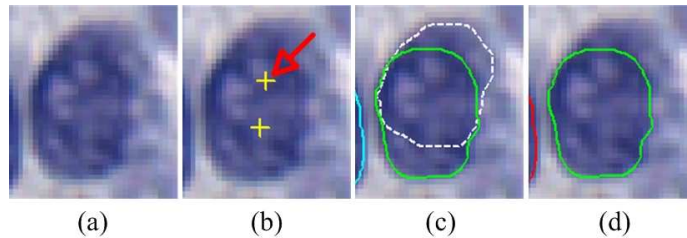


Fig. 5.7. Nuclei overlap processing. (a) An nucleus. (b) An nucleus with two seeds (where the false seed is indicated by the red arrow). (c) Nuclei boundaries obtained by the mRLS based technique. (d) Final nuclear boundary after overlap processing.

5.3 Performance Evaluations

In this section, we evaluate the performance of the proposed mRLS-based technique (henceforth referred to as the P-mRLS technique) for nuclei segmentation. First we illustrate the testing datasets used in this work. We then present the quantitative and qualitative evaluations of the P-mRLS technique and several existing techniques for nuclei segmentation.

5.3.1 Datasets Description

This study was based on two datasets of skin histopathological images which were collected from the Cross Cancer Institute, University of Alberta. Table 5.2 lists the summary of two datasets. The first dataset (Dataset-I) consists of 40 Ki-67 [109] stained skin images which were cropped from four digital biopsy slides. The biopsy slides were digitized under 40X magnification on Aperio Scanscope CS scanning system. The images have a pixel resolution of $0.248\mu m/\text{pixel}$, and include 1446 manually identified cell nuclei in total. The second dataset (Dataset-II) consists of 28 H&E stained skin images which were cropped from five biopsy slides. The biopsy slides were digitized under 25X magnification on Carl Zeiss MIRAX MIDI scanning system. These images have a pixel resolution of $0.372\mu m/\text{pixel}$, and include 714 manually identified cell nuclei in total. Note that all nuclei centroids in Datasets-I and II were first manually labeled by an imaging expert with over 3 years of experience on skin pathology. The results were then examined and fine-tuned by an experienced pathologist, and finally used as the ground truth seeds.

TABLE 5.2
IMAGE DATASET USED FOR NUCLEI SEGMENTATION EVALUATIONS.

Datasets	Tissue Stains	Image No.	Nuclei No.
Dataset-I	Ki-67	40	1446
Dataset-II	H&E	28	714

5.3.2 Quantitative Evaluations

To illustrate the efficacy of nuclei segmentation, the performance of the P-mRLS technique is compared with several existing techniques including local optimal thresholding (LOT) [30], marker optimization with watershed algorithm (MOW) [34], improved voting with watershed algorithm (IVW) [91], repulsive active contours model (RACM) [100], deep learning method (DLM) [110] and graph search based method (GSM) [111]. The LOT technique is chosen as the baseline thresholding method, and it segments cell nuclei based on two prior knowledge: nuclei size and shape. There are two key parameters namely predefined nuclei area range $[A_{\min}, A_{\max}]$ that should be adaptively selected when using the LOT technique. We have experimentally set them as [400,900] and [100,400] for our Datasets-I and II, respectively. The MOW technique segments cell nuclei based on H-minima transform and watershed algorithm. There is only one key parameter namely angular margin ω that is used to avoid over-segmentation by the MOW technique. As verified in [34], the MOW technique is not very sensitive to the parameter ω , and hence it is set as $\pi/12$ (the same as that in [34]) for comparisons. Both the IVW and RACM techniques detect nuclei seeds by a voting procedure, where the voting radial range $[r_{\min}, r_{\max}]$ and the bandwidth b_w of mean-shift clustering are the key parameters determining nuclei detection performance. The voting radial range $[r_{\min}, r_{\max}]$ are estimated based on nuclei radii, which are set as [5,16] and [2,9] for our Datasets-I and II, respectively. The bandwidth b_w is adaptively set as $(r_{\min} + r_{\max})/2$, as was done in [91]. After detecting nuclei seeds, the RACM technique delineates nuclei boundaries by the active contours model with a repulsive term to prevent evolving contours from crossing and merging with one another. There are a few weight parameters (α , β , γ , λ and η), which controls the elasticity, rigidity, pressure force, image force and repulsive force of the active contour model, respectively. We have experimentally set these parameters as: $\alpha = 0.05$, $\beta = 0$, $\gamma = 0.25$, $\lambda = 2$ and $\eta = 1$ for our two datasets. For the DLM, we implemented the VGG style CNN [110] with Batch Normalization technique [112]. Specifically, a 9

layers DL architecture which consists of 1 input layer, 6 convolution layers, 1 fully connected layer and 1 output layer is constructed. The input layer accepts an image patch of 64×64 pixels, and the output layer is a soft-max function which outputs the class probability of being nucleus or non-nucleus. During training, we randomly select 20 images from each of our Datasets-I and II as training samples, and the DL architecture is trained following the procedure proposed by [105]. During testing, a threshold of 0.3 is experimentally selected to binarize the probability map output by the DLM. The GSM is originally proposed for intestinal gland segmentation in H&E stained biopsy images. We adaptively modified it to perform nuclei segmentation for comparison with the proposed technique. In our implementation, we utilize the gradient magnitude as graph node weights instead of using image intensity information like [111], as nuclei boundaries tend to be located on high gradient points (rather than low intensity points). The graph search radii are set as 18 and 11 pixels in our Datasets-I and II, respectively. After obtaining boundary probability map by graph search, watershed segmentation is performed with false nuclear regions removed based on area and color analysis.

Automatic segmentation results by different methods have been manually inspected by an interactive software (based on Matlab). Let N_{CS} , N_{OS} , N_{US} and N_{ER} represent the numbers of correct-segmentation (CS), over-segmentation (OS), under-segmentation (US) and encroachment errors (ER). Note that OS refers to excessive splitting (i.e., one nuclear region is split into two subregions as shown in Fig. 5.8(b)). US refers to the failure to split nuclear regions into the correct number of nuclei (see Fig. 5.8(c)). ER refers to the case that automatically obtained nuclei boundaries are visually not accurate (see Fig. 5.8(d)). The performance of nuclei segmentations has been evaluated by several widely-used evaluation metrics [39], [98]: CS, OS, UR and ER rates, which are defined as follows:

$$CS = \frac{N_{CS}}{N_{GT}} \times 100\%, \quad OS = \frac{N_{OS}}{N_{GT}} \times 100\% \quad (5.15)$$

$$US = \frac{N_{US}}{N_{GT}} \times 100\%, \quad ER = \frac{N_{ER}}{N_{GT}} \times 100\% \quad (5.16)$$

where N_{GT} represents the number of manually identified (ground truth) nuclei.

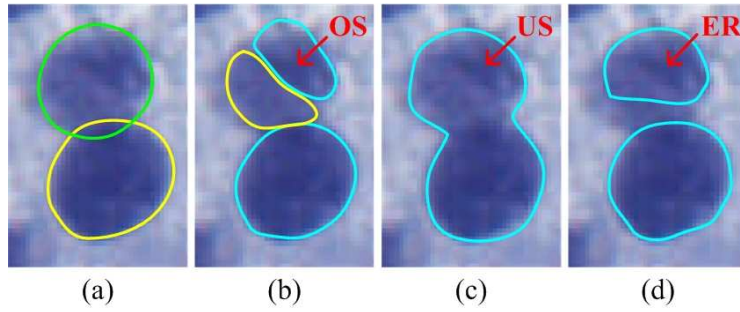


Fig. 5.8. Illustration of over-segmentation, under-segmentation and encroachment errors. (a) Two nuclei with manually labeled boundaries. (b) Over-segmentation. (c) Under-segmentation. (d) Encroachment error.

The comparison of segmentation results with different methods is shown in Table 5.3. It is observed that the LOT [30] provides only 62.03% and 67.93% of CS on our Datasets-I and II, respectively. The poor performance of the LOT technique is mainly because our testing images include a large number of nuclei clumps which cannot be efficiently split by a thresholding method. The DLM [110] only provides a slightly better performance than the LOT, with 64.72% and 68.80% of CS on our two datasets, respectively. The main reason behind this poor performance is that the DLM cannot resolve overlapping nuclei (as also mentioned in [103]), while our datasets mainly include image patches with clustered nuclei. Another reason is that the DLM requires a large number of dedicated training samples, but our training samples are limited due to the relatively small dataset. The MOW [34] provides a better performance than the LOT and DLM, which has correctly segmented 67.61% and 75.49% of cell nuclei in two datasets. The IVW [91] further improves the segmentation performance from the MOW, with around 80% of CS values on two datasets. The MOW [34] and IVW [91] perform nuclei segmentations by applying marked watershed algorithm on inner distance transform maps, both of which critically depend on the accuracy of image binarizations. In our datasets, Ki-67 stained skin images tend to be more difficult to be binarized than H&E stained skin images, and hence the performances of these two techniques [91] [34] on Ki-67 stained images are poorer than those on H&E stained images. The GSM [111]

TABLE 5.3
COMPARISON OF NUCLEI SEGMENTATIONS.

Datasets	Techniques	CS (%)	OS (%)	US (%)	ER (%)
Dataset-I (Ki-67)	LOT [30]	62.03	8.62	10.50	18.85
	DLM [110]	64.72	3.44	16.66	15.18
	MOW [34]	67.61	3.23	9.24	19.92
	IVW [91]	79.10	2.19	6.42	12.29
	GSM [111]	77.04	3.53	7.40	12.03
	RACM [100]	88.13	1.62	4.65	5.60
	P-mRLS	90.05	1.72	2.88	5.35
Dataset-II (H&E)	LOT [30]	67.93	4.76	13.59	13.73
	DLM [110]	68.80	5.66	13.39	12.15
	MOW [34]	75.49	1.40	9.24	13.87
	IVW [91]	80.81	2.38	6.44	10.36
	GSM [111]	80.95	4.90	6.58	7.56
	RACM [100]	83.89	1.96	4.62	9.52
	P-mRLS	88.10	1.12	4.90	5.88

achieves a similar performance with the IVW. This technique performs watershed segmentation on a boundary probability map obtained by graph search algorithm, which is likely to under-split overlapping nuclei or over-split heterogeneous nucleus. The RACM [100] is one of the latest nuclei segmentation methods, which provides the best segmentation performance (88.13% and 83.89%) among all tested existing methods. The P-mRLS technique achieves 90.05% and 88.10% of CS on two databases, which are slightly better than the RACM. The RACM [100] detects nuclei seeds by a voting based algorithm, and delineates nuclei boundaries by evolving active contours. As has been verified in [106], the voting algorithm is prone to be affected by image noisy pixels which will result in false positive seeds. The active contours model for nuclei segmentation not only has a high computation complexity but also is critically dependent on the accuracy of detected nuclei seeds [39]. In comparison, the P-mRLS technique detects nuclear seeds by a bank of multi-scale gLoG kernels, which is more accurate and efficient for nuclei detection. The P-mRLS technique delineates nuclear boundaries by analyzing multi-scale radial lines, which is more robust with nuclear seed detection.

We also evaluate nuclei detection performances by different techniques, where

the centroids of segmented nuclear regions are compared with manually labelled ground truth seeds. A nuclear centroid c_i (automatically obtained) is counted as a true positive if and only if its location is in a detection pair (c_i, g_j) for which the corresponding (nearest) ground truth seed g_j has been paired, and also their Euclidean distance is within a range of t (number of pixels) [106]. Let N_{DS} and N_{TP} denote the numbers of centroids for the segmented nuclei and true positive detections. The performance of nuclei detections is evaluated with respect to the recall (\mathcal{D}_{REC}), precision (\mathcal{D}_{FRE}) and F-measure (\mathcal{D}_F), as defined by Eqs. (4.24)(4.25)(4.26) in Chapter 4.

The performance of nuclei detections with different techniques are listed in Table 5.4. Note that different t values (i.e., 6 and 4 pixels) are used for evaluation as images in Datasets-I and II have different pixel resolutions. As observed in Table 5.4, the LOT [30] only provides 70.51% and 74.03% of F-measures on nuclei detections. The DLM [110] provides 74.16% and 76.27% of F-measures on nuclei detections. The MOW [34], IVW [91] and GSM [111] achieve intermediate performances with F-measures ranged between 77% and 86%. The RACM [100] provides 89.02% and 90.27% of F-measures, which are much better than other existing techniques. The P-mRLS technique provides the highest F-measure values among all techniques, with 90.77% and 93.62% on Datasets-I and II, respectively. Since the centroids of segmented nuclear regions are computed and compared with manually labeled ground truth seeds, the performance of nuclei detections is highly dependent on nuclei segmentations.

In order to further evaluate the accuracy of segmented nuclei contours, we compare the automatic segmentations with manual delineations by computing a boundary based metric, namely mean absolute distance (MAD) [91]. The MAD value estimates the disagreement averaged over two contours, and a better segmentation corresponds to a smaller MAD value (i.e., close to zero). However, since manual delineation of all cell nuclei is tedious and time consuming process, 113 and 96 nuclei are randomly chosen from Datasets-I and II, and manually delineated by using an interactive software [30]. The MAD computations are performed on these

TABLE 5.4
COMPARISON OF NUCLEI DETECTIONS.

Datasets	Techniques	N_{TP}	\mathcal{D}_{REC} (%)	\mathcal{D}_{PRE} (%)	\mathcal{D}_F (%)	MAD (px)
Dataset-I ($t = 6$)	LOT [30]	911	63.00	80.05	70.51	3.09
	DLM [110]	945	65.35	86.94	74.16	2.78
	MOW [34]	1078	74.55	79.26	76.84	1.94
	IVW [91]	1196	82.71	80.59	81.64	1.67
	GSM [111]	1214	83.96	76.54	80.08	1.52
	RACM [100]	1310	90.59	87.51	89.02	1.64
	P-mRLS	1318	91.15	90.40	90.77	1.12
Dataset-II ($t = 4$)	LOT [30]	489	68.49	80.56	74.03	1.71
	DLM [110]	497	69.61	85.10	76.27	1.53
	MOW [34]	567	79.41	81.35	80.37	1.15
	IVW [91]	640	89.64	82.69	86.02	0.88
	GSM [111]	634	88.80	79.65	83.97	0.87
	RACM [100]	654	91.60	88.98	90.27	0.92
	P-mRLS	660	92.44	94.83	93.62	0.68

manually delineated cell nuclei. In Table 5.4, the seventh column separately lists the computed MAD values for different techniques. It is observed from Table 5.4 that the P-mRLS technique has achieved the smallest MAD values (1.12 and 0.68 pixels, respectively) among all techniques, which indicates a good matching between automatically obtained nuclei boundaries and manually labeled nuclei contours.

5.3.3 Qualitative Evaluations

Fig. 5.9 shows visual examples of nuclei segmentations by the P-mRLS technique and several existing methods. In Fig. 5.9, the columns (a)(b) show the original image patches and manually labeled nuclei contours, respectively. The columns (c)-(i) separately show the automated nuclei segmentations by different techniques. For better viewing, nuclei boundaries obtained by the P-mRLS technique in the column (i) have been indicated by different color of contours (e.g., magenta, blue). As seen in Fig. 5.9, the LOT [30] and DLM [110] are likely to under-segment nuclei clumps. Although the MOW [34], IVW [91] and GSM [111] can segment nuclei clumps by watershed algorithms, the obtained nuclei boundaries are usually not very accurate. The RACM [100] can provide smooth and more accurate nuclei

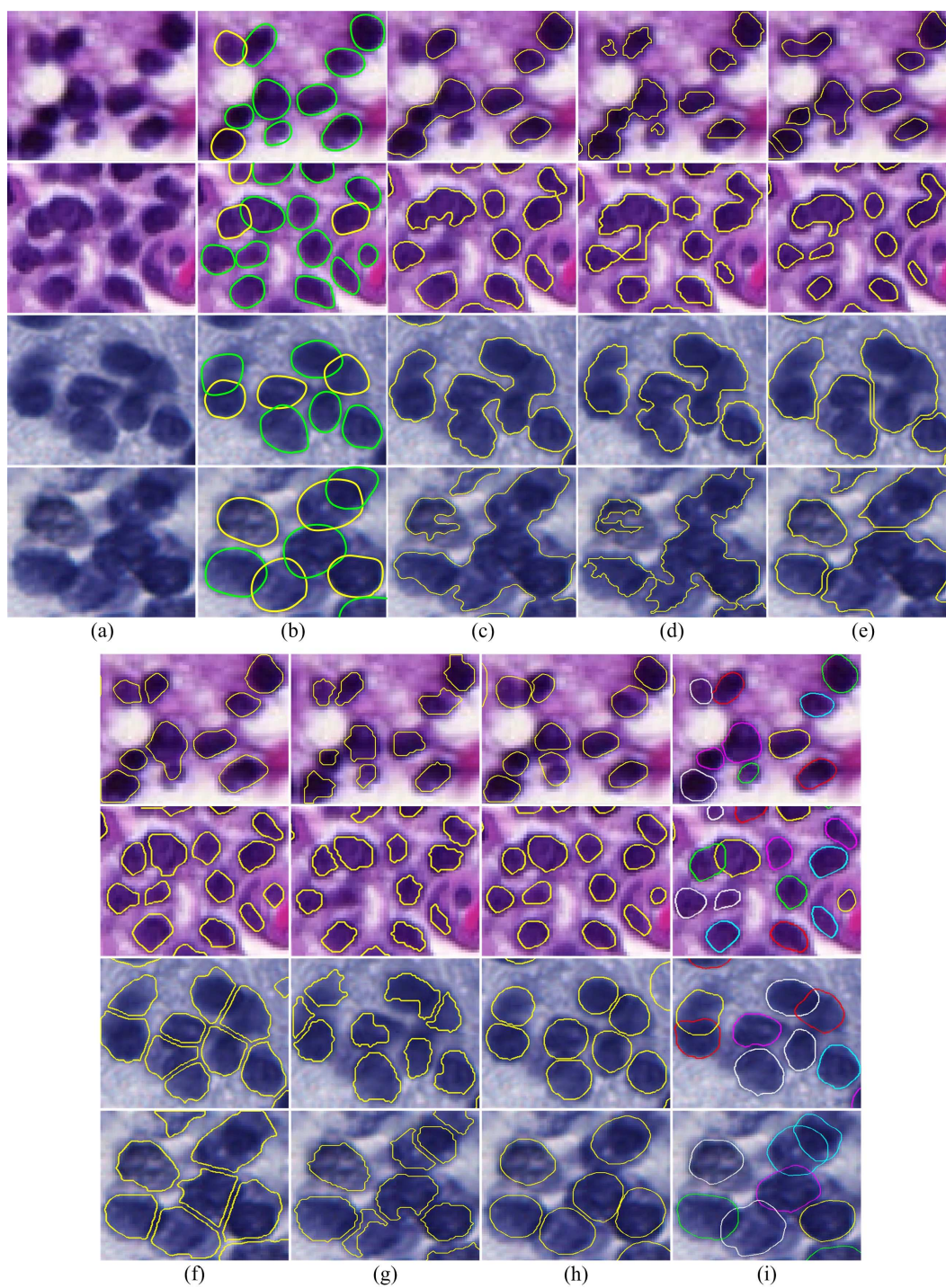


Fig. 5.9. Comparative nuclei segmentations. (a) Original image patches. (b) Ground truths. (c) LOT [30]. (d) DLM [110]. (e) MOW [34]. (f) IVW [91]. (g) GSM [111]. (h) RACM [100]. (i) P-mRLS technique.

contours, but it is limited in estimating occluded nuclei boundaries for overlapped nuclei. In comparison, the P-mRLS technique provides a more precise and reliable segmentation results compared to existing techniques. In addition, the occluded nuclei boundaries have been generated by the P-mRLS technique.

Fig. 5.10 shows qualitative evaluations of the P-mRLS technique on two large images, each of which includes hundreds of cell nuclei. Fig. 5.10(a) shows nuclei detection results by the gLoG kernels based method [106], where the detected seeds are indicated by (yellow) crosses. Fig. 5.10(b) shows nuclei segmentation results by the P-mRLS technique. Note that in Fig. 5.10 the third row shows selected zoomed-in image patches of nuclei detections and segmentations for better viewing. The superimposed circles and arrows in the third row of Fig. 5.10 highlight inaccurate nuclei detections and segmentations. It is observed in Fig. 5.10 that the P-mRLS technique have correctly detected and segmented most of the cell nuclei in testing images, although there exist a small number of false detections and segmentations.

5.3.4 Sensitivity Evaluation & Computational Complexity

In nuclei contour selection module, there are three weighting parameters $(\lambda_1, \lambda_2, \lambda_3)$ (see Eq. (5.12)) that can be adaptively adjusted for processing different biopsy images. In this work, $(\lambda_1, \lambda_2, \lambda_3)$ have been empirically set as $(1, 1, 1)$ for evaluations of our two datasets. To test the sensitivities of these three parameters, we slightly vary their values and compute the correct-segmentation (CS) rates with several different combinations of $(\lambda_1, \lambda_2, \lambda_3)$ values. This is $(\lambda_1, \lambda_2, \lambda_3) \in \{(1 \pm 0.5, 1, 1), (1, 1 \pm 0.5, 1), (1, 1, 1 \pm 0.5)\}$. Fig. 5.11 shows the correct segmentation rates of Datasets-I and II with different combinations of weighting parameters. As observed from Fig. 5.11, the correct segmentation rates vary within about 1% for both datasets when weighting parameters are slightly changed. Therefore it can be concluded that the P-mRLS technique is not sensitive to the marginal change of different $(\lambda_1, \lambda_2, \lambda_3)$ values.

All experiments were carried out on a 1.8-GHZ Intel Core i7-4500U CPU with 16-GB RAM using MATLAB 8.10. The P-mRLS technique only takes around 30

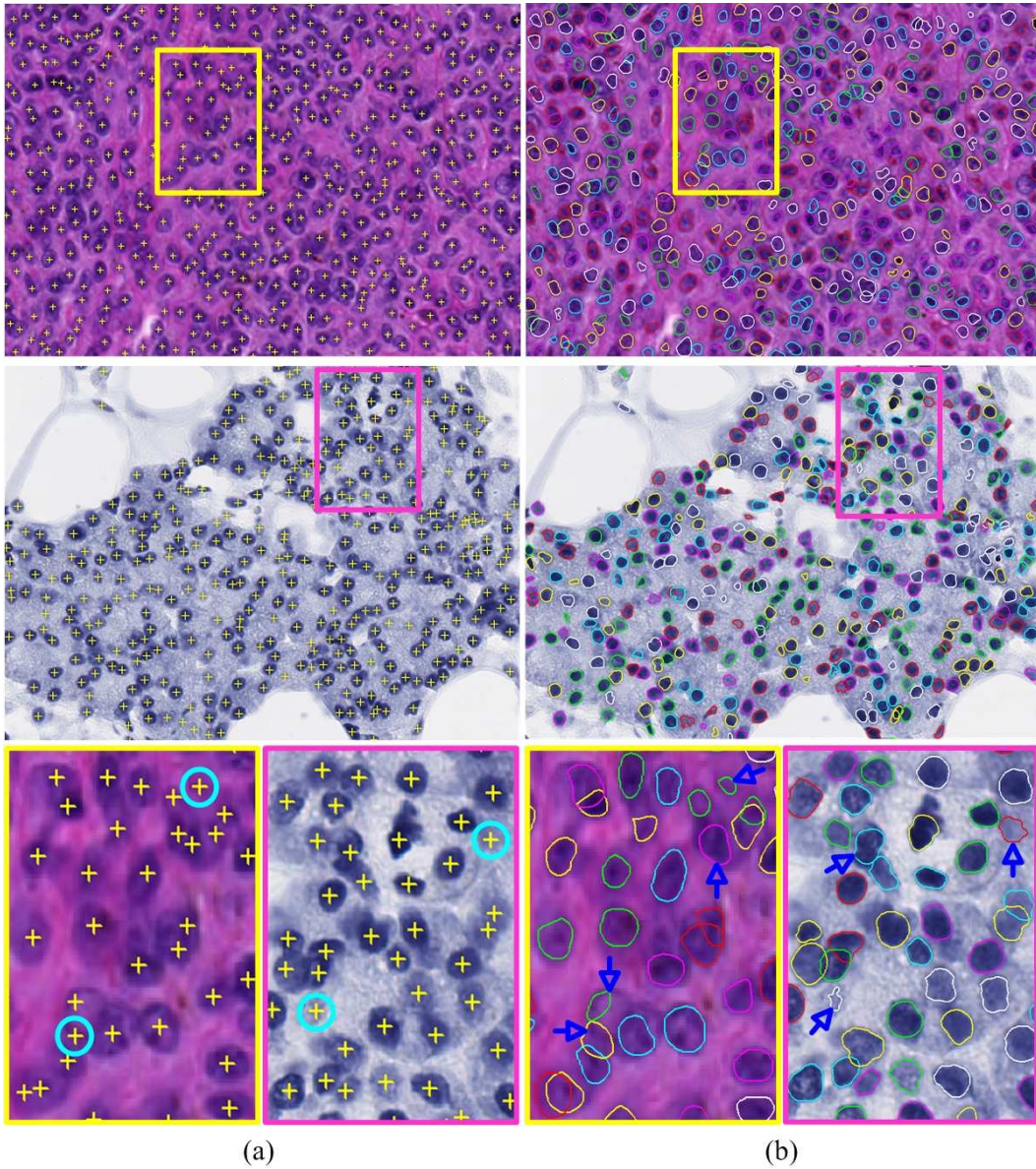


Fig. 5.10. Visual examples of nuclei detections and segmentations. (a) Nuclei detections. (b) Nuclei segmentations. Note that the first and second rows show H&E and Ki-67 stained skin images, respectively, and the third row shows the selected zoomed-in image patches. In the third row, the superimposed circles highlight false seeds detected by gLoG kernels, and the superimposed arrows indicate inaccurate segmentations by the P-mRLS technique.

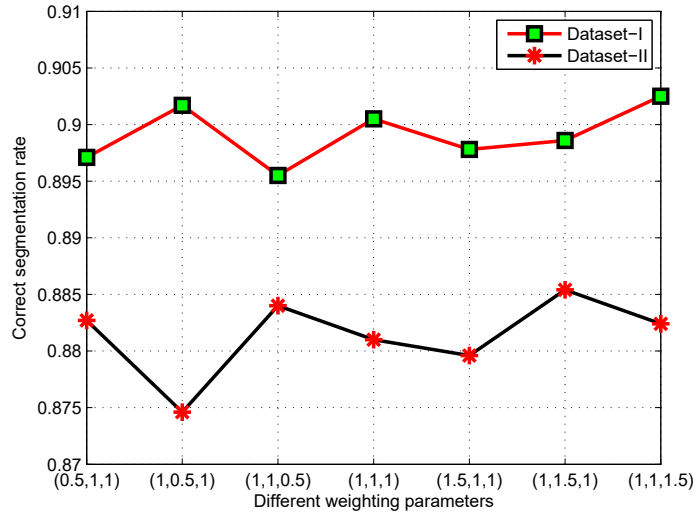


Fig. 5.11. Sensitivity evaluation of weighting parameters ($\lambda_1, \lambda_2, \lambda_3$) on nuclei segmentation.

seconds to perform nuclei detection and segmentation on a skin image like the Ki-67 stained image (800x980) shown in Fig. 5.10 with hundreds of cell nuclei, which indicates the efficiency of the proposed technique. The LOT [30], DLM [110], MOW [34], IVW [91], GSM [111] and RACM [100] separately take about 5s, 12s, 106s, 73s, 58s and 512s (or 8.5 minutes) to process the same skin image. Note that the P-mRLS technique is well suited for parallelization, and hence parallel implementations of the algorithm (e.g., to process thousands of cell nuclei simultaneously) can be many times faster than sequential implementations.

5.4 Summary

This chapter presents an automated technique for nuclei segmentation in skin biopsy images. The proposed technique relies on a bank of gLoG kernels for nuclei seed detection, and thereafter on a mRLS with DP method for nuclei contour delineation. Experiments have been thoroughly performed on two datasets of skin microscopic images with different stains, which indicates that the proposed technique is superior to several existing techniques for nuclei segmentations. In addition, the proposed technique has the advantage of recognizing occluded nuclei boundaries

for overlapped nuclei. Compared with level sets and active contours, the proposed technique is very efficient and suitable for parallel implementations, which has the potential to be used for processing the image with a large number of cell nuclei.

Chapter 6

Analysis and Classification of Skin

WSI

Histological examination of biopsy slides is the gold standard for skin melanocytic tumor diagnosis. However, visual analysis performed by pathologists is typically subjective and often leads to inter- and intra-observer variability. Besides, it is very labor-intensive to examine the whole biopsy slide due to the large volume of data to be analyzed. In this chapter, we propose a computer-aided technique for automated analysis and classification of melanocytic tumor on skin whole slide biopsy images. The proposed technique first analyzes skin epidermis and dermis regions, and then classifies skin WSI into different categories such as melanoma, nevus or normal skin using extracted textural and cytological features. Experiments on 66 skin WSIs show that the technique can achieve over 95% classification accuracies.

6.1 Introduction

A typical skin tissue slide includes three main parts [113]: epidermis, dermis and sebaceous areas (see Fig 3.1). In the diagnosis of skin melanoma by analyzing biopsy images, pathologists are mainly interested in the epidermis and DEJ areas as

Part of this chapter has been presented in: Xu *et al.* "Automated diagnosis of melanoma from skin biopsy images", *Proc. IEEE International Conference on Biomedical and Health Informatics (BHI)*, February 2017.

most of histological features are observed in these regions [17]. In the past decade, there have been a few research works that analyze skin epidermis or DEJ areas for melanocytic tumor classification (see the summary in Table 6.1). Smolle [114] proposed a tissue counter analysis (TCA) technique to recognize skin structures like epidermis and dermis. In the TCA technique, the digital images are dissected into non-overlapping subregions (elements), and the digital information such as colors and textures in each element is extracted for statistical analysis and classification. Wiltgen *et al.* [55] applied the TCA technique to classify benign nevus and malignant melanoma based on histogram and co-occurrence matrix features of skin histological images. Although a good classification accuracy (about 92%) is reported in [55], its performance is sensitive to staining variations as the TCA technique analyzes skin images by using color and texture features [115]. Miedema *et al.* [65] reported an image analysis system for melanocytic lesion classification, which distinguishes melanoma from nevus by using cytological and textural features in DEJ areas. Nielsen *et al.* [116] proposed an automated technique that differentiates melanoma and nevus based on quantification of MART1-verified Ki-67 indices in skin epidermis and dermis areas. These techniques only consider a representative region within the WSI. The selection of representative regions from WSIs requires interactive manual operation. In addition, analysis of representative regions may bring the sampling bias in the diagnostic system.

TABLE 6.1
SKIN BIOPSY IMAGE ANALYSIS AND CLASSIFICATION.

Authors	Images	Regions	Performance
Wiltgen <i>et al.</i> [55]	80 H&E biopsies	Epidermis & Dermis	~92%
Miedema <i>et al.</i> [65]	49 H&E biopsies	DE junctions	~95%
Nielsen <i>et al.</i> [116]	379 MART1-verified Ki-67 biopsies	Epidermis & Dermis	~89%
Lu <i>et al.</i> [59]	66 H&E biopsies	Epidermis	~90%

Recently Lu and Mandal [59] proposed an automated technique for analysis and diagnosis of melanoma on skin WSIs. This technique first segments epidermis areas from WSIs based on global thresholding and shape analysis [74]. The keratinocytes

and melanocytes within epidermis areas are then segmented by adaptive thresholding [30] and radial line scanning [107]. Based on the segmented regions of interest, the spatial distributions of melanocytes and morphological features of cell nuclei are computed. These computed features are finally utilized by a multi-class SVM (mSVM) method for skin melanoma diagnosis. Experimental results show that the technique achieves about 90% classification accuracies. However, the technique in [59] only analyzes skin epidermis areas and classifies skin tissues using epidermis features. During the real evaluation by pathologists, in addition to epidermis examination, dermis examination is also important because melanoma may invade into skin dermis tissue. Nuclei densities and types in dermis regions often vary from one another among different skin tissues, and hence cytological and textural features of dermis areas are important for automatic melanoma diagnosis. Fig. 6.1 shows examples of H&E stained skin histological images, where Figs. 6.1(a)(b)(c) show normal skin, benign nevus and malignant melanoma, respectively. As observed in Fig. 6.1, the epidermis presents relatively darker color than the dermis. There is a high density of cell nuclei (keratinocytes and melanocytes) in epidermis of all three skin tissues. In comparison, the dermis has a low density of cell nuclei in normal skin tissue shown in Fig. 6.1(a), but has a high density of cell nuclei in the nevus and melanoma shown in Figs. 6.1(b)(c). In Fig. 6.1(b) cell nuclei within the dermis are mainly the melanocytes which have roughly uniform color and size, while in Fig. 6.1(c) cell nuclei within the dermis includes both melanocytes (e.g., with irregular nuclei shape and size) and lymphocytes (e.g., small, round and dark nuclei).

In this chapter, we propose an automatic technique for analysis and classification of melanocytic tumor on H&E stained skin WSIs. The proposed technique is built on the work in [59] that performs skin melanoma diagnosis based on epidermis analysis. Besides skin epidermis, this work further analyzes skin dermis and computes a set of dermis features for melanocytic tumor classification. The details of the proposed technique and experimental results are provided in the following sections.

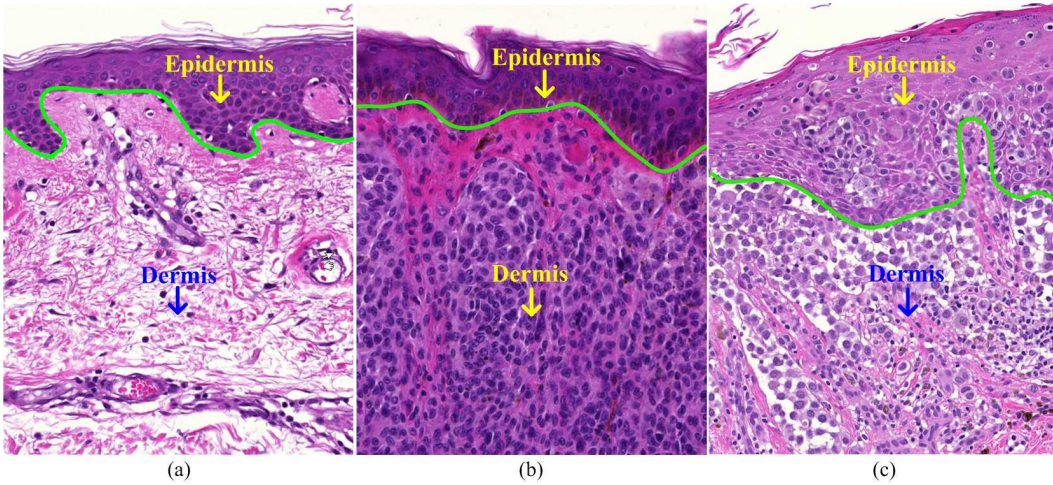


Fig. 6.1. Examples of H&E stained skin images. (a) Normal skin. (b) Benign nevus. (c) Malignant melanoma. Note that in (a)(b)(c) green contours indicate the borders of epidermis and dermis regions. Cell nuclei in both epidermis and dermis regions are observed as blue blobs.

6.2 Proposed Technique

The schematic of the proposed technique is shown in Fig. 6.2. It is observed that the proposed technique consists of four main modules. First epidermis and dermis regions are segmented, and a number of epidermis and dermis tiles are generated for subsequent analysis. Next, epidermis and dermis analysis are performed in a parallel manner. In the epidermis analysis, we first segment cell nuclei and detect melanocytes, and then construct a set of morphological and spatial features from epidermis regions. In the dermis analysis, we first perform color normalization and nuclei segmentation, and then compute a set of textural and cytological features from dermis regions. Finally, the computed epidermis and dermis features are utilized by a mSVM method that classifies the input skin WSI into different categories. Details of the four modules are presented in the following.

6.2.1 Epidermis & Dermis Segmentations

The epidermis and DEJ areas are the most important observation regions for skin melanoma diagnosis. The cytological features in these regions are very critical for pathological diagnosis. In this module, we segment skin epidermis and dermis

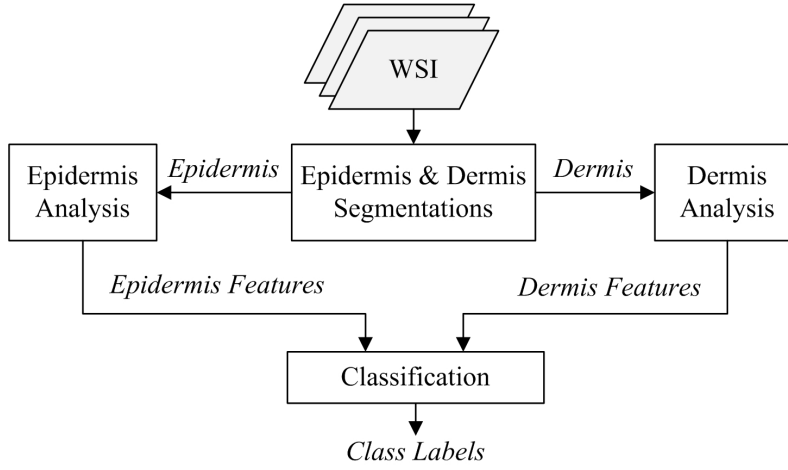


Fig. 6.2. Schematic of the proposed technique.

from the WSI such that the subsequent image analysis is focused on these regions. The schematic of epidermis and dermis segmentations is shown in Fig. 6.3. In the following, we explain the four steps shown in Fig. 6.3.

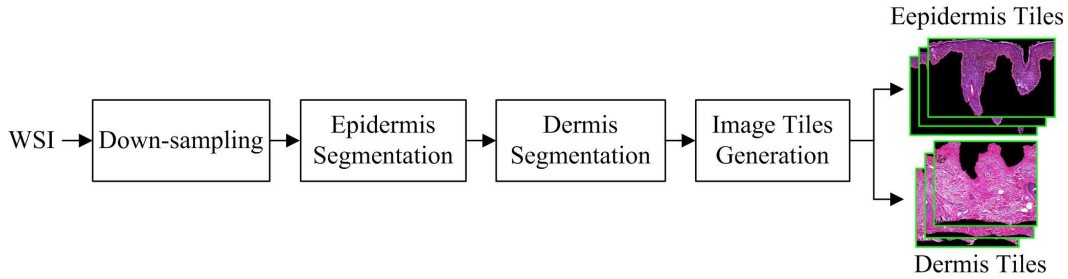


Fig. 6.3. Schematic of epidermis and dermis segmentations.

6.2.1.1 Down-sampling

Since the skin WSI has a large size, we employ a multi-resolution framework [74] to segment skin epidermis and dermis such that the segmentation is performed more efficiently and accurately. Given a high resolution color skin WSI I_h with a size of $m \times n \times 3$, it is down-sampled by a factor of k to produce a low resolution image I_l . The image I_l has a size of $\lfloor m/k \rfloor \times \lfloor n/k \rfloor \times 3$. In this work, k is empirically set as 4, which reduces the size of high resolution image by 16 times after down-sampling.

6.2.1.2 Epidermis Segmentation

After down-sampling the high resolution skin WSI, a coarse-to-fine segmentation method [113] is utilized to segment skin epidermis from the image I_l . This method first performs a coarse epidermis segmentation using global thresholding and shape analysis on the red channel of image I_l . The epidermis thickness is then measured by a series of line segments perpendicular to the main axis of the initially segmented epidermis mask. If the segmented epidermis mask has a thickness greater than a predefined threshold (i.e., the expected largest epidermis thickness), a second-pass of fine segmentation using k-means algorithm is applied to enhance the coarsely-segmented result. Finally, morphological operations such as opening and closing are utilized to fill holes and smooth boundaries of the epidermis region. Let the obtained epidermis mask be denoted by b_e . Fig. 6.4(a) shows a skin image I_l , and Fig. 6.4(b) shows the mask b_e where the binary foreground corresponds to segmented epidermis areas.

6.2.1.3 Dermis Segmentation

In this step, we segment skin dermis based on a predefined interest of depth value d [117]. First epidermis outer boundary is determined from the mask b_e . A series of line segments, which are perpendicular to epidermis outer boundary and have a length of d , are then generated. Next, morphological closing operation is applied to join line segments together, and a mask b_{e+d} including both epidermis and dermis regions (see Fig. 6.4(c) where $d = 600\mu m$) is obtained. Finally, the dermis mask b_d is obtained by removing epidermis pixels from the mask b_{e+d} . Fig. 6.4(d) shows the dermis mask b_d , and Fig. 6.4(e) shows the epidermis and dermis segmentations which are labeled by the cyan and blue contours.

6.2.1.4 Image Tiles Generation

After segmenting epidermis and dermis regions, we generate high resolution epidermis and dermis tiles for further analysis. The steps of epidermis and dermis tiles

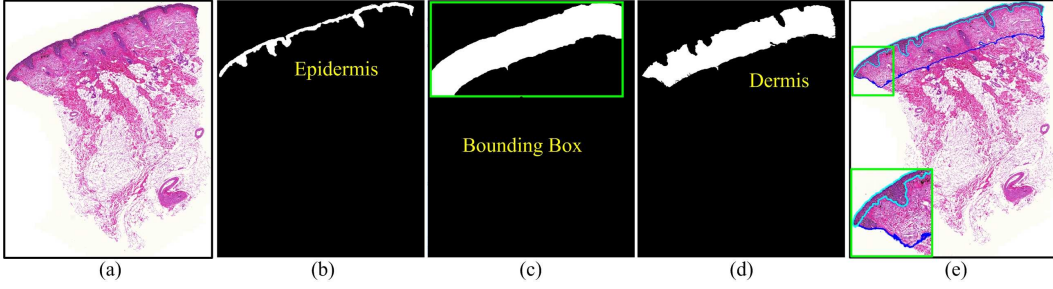


Fig. 6.4. Epidermis and dermis segmentations. (a) Skin WSI I_l . (b) Epidermis mask b_e . (c) Epidermis and dermis mask b_{e+d} . (d) Dermis mask b_d . (e) Segmentation results. Note that in (e) segmented epidermis and dermis regions are labelled with cyan and blue contours, respectively. The bottom-left corner in (e) shows a small zoomed-in image patch.

generation are as follows:

Step 1: The binary masks b_e , b_{e+d} and b_d are mapped to the original high resolution field before down-sampling.

Step 2: The bounding box of the foreground region in mask b_{e+d} (see Fig. 6.4(c)) is computed, and a mask b_{e+d}^c (i.e., cropped b_{e+d}) which contains only the regions within the bounding box is cropped from the mask b_{e+d} . We also crop the same regions as surrounded by the bounding box in Fig. 6.4(c) from the images I_h , b_e and b_d , respectively, to obtain the images I_h^c , b_e^c and b_d^c .

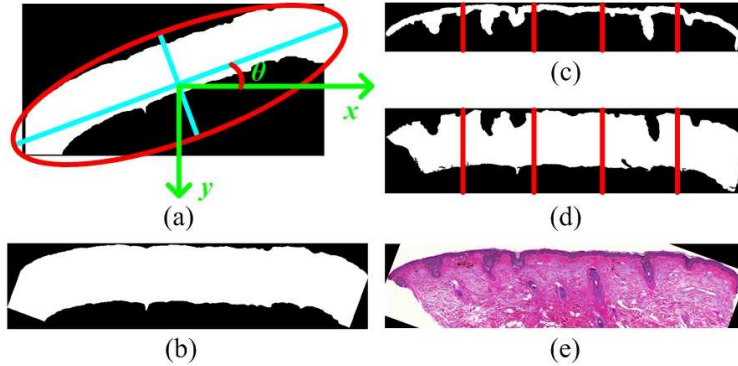


Fig. 6.5. Image rotation and division. (a) b_{e+d}^c with the best fitted ellipse (red contour). (b) b_{e+d}^r . (c) b_e^r . (d) b_d^r . (e) I_h^r . Note that in (c)(d) the red vertical lines illustrate the division of b_e^r and b_d^r into several image tiles.

Step 3: The best fitted ellipse [64] that has the same second-moments as the foreground region in b_{e+d}^c is computed. Let the angle between the x-axis and the major axis of the fitted ellipse be denoted by θ (see Fig. 6.5(a)). We rotate the masks b_{e+d}^c , b_e^c and b_d^c by $-\theta$ (e.g., clockwise rotation of θ in Fig. 6.5) to obtain the rotated

masks b_{e+d}^r , b_e^r , and b_d^r . Similarly, we rotate the image I_h^c to obtain the rotated image I_h^r . Figs. 6.5(b)(c)(d)(e) separately show the b_{e+d}^r , b_e^r , b_d^r and I_h^r , respectively.

Step 4: For easy of processing, the masks b_e^r and b_d^r are equally divided into several image tiles (e.g., 5 tiles in Figs. 6.5(c)(d)). The foreground regions in image tiles are mapped to skin tissue pixels in I_h^r . Fig. 6.6 shows an example of generated high resolution epidermis and dermis tiles. Let the obtained epidermis and dermis tiles be denoted by I_e^n and I_d^n , respectively, where $1 \leq n \leq T$ and N is the number of image tiles.

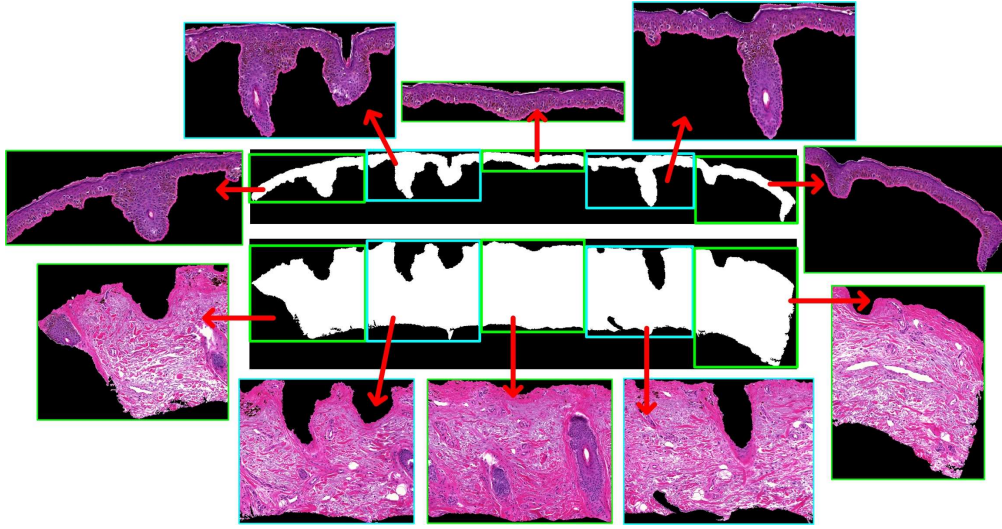


Fig. 6.6. Epidermis and dermis tiles.

6.2.2 Epidermis Analysis

After obtaining epidermis tiles I_e^n , $1 \leq n \leq N$, epidermis analysis is now performed. The schematic of epidermis analysis is shown in Fig. 6.7. As observed in Fig. 6.7, there are three modules for epidermis analysis. First, we segment cell nuclei (including keratinocytes and melanocytes) in the epidermis area. The melanocytes are then detected from the pre-segmented cell nuclei. Finally, epidermis features reflecting nuclear morphologies and spatial distributions of melanocytes are constructed. The steps of epidermis analysis are explained below.

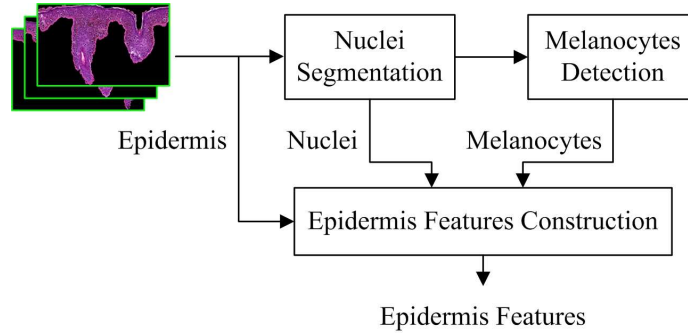


Fig. 6.7. Schematic of epidermis analysis.

6.2.2.1 Nuclei Segmentation

In H&E stained skin images, cell nuclei are mainly observed as dark blue blobs. In order to efficiently segment cell nuclei in epidermis regions, an adaptive thresholding technique [30] is applied. This technique first applies a series of gray-scale morphological reconstructions [31] (on the red channel of I_e^n) to enhance the image and make nuclear regions homogeneous. An adaptive threshold selection method is then used to achieve nuclei segmentation. Fig. 6.8(a) shows the red channel of I_e^n after gray-scale morphological reconstructions. Figs. 6.8(b)(c) show nuclei segmentations in the image tile. Figs. 6.8 (d) shows two close-up views of image patches in Fig. 6.8(c), where segmented nuclei in Fig. 6.8 are labeled with cyan or yellow contours.

6.2.2.2 Melanocytes Detection

In skin tissues, cell nuclei in epidermis are mainly keratinocytes and melanocytes. In normal skin tissue, the number of melanocytes accounts for 5% to 10% of all nuclei in epidermis [59]. However, in skin melanoma or nevus, there will be an increasing proportion of melanocytes in epidermis. In addition, melanocytes are not only found in the basal layer (i.e., epidermis close to dermis) but also found in other epidermis layers (e.g., spinous layer) in melanomas [17]. Therefore, identification of melanocytes in epidermis is important for melanoma diagnosis.

Figs. 6.9(a)(b) illustrate the typical appearance of melanocytes in epidermis,

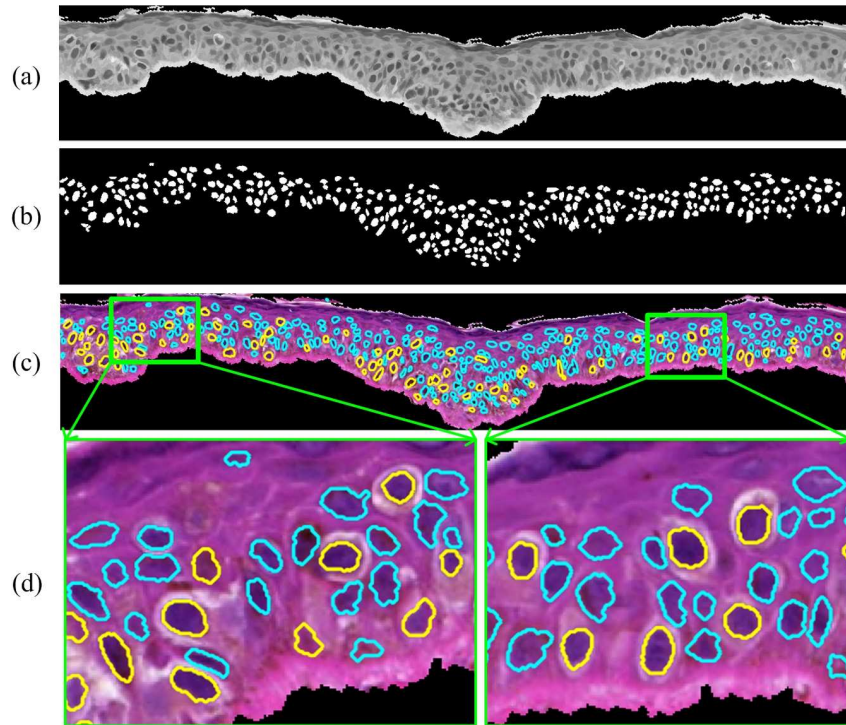


Fig. 6.8. Nuclei segmentation and melanocytes detection. (a) Red channel image after morphological reconstruction operations. (b) Nuclei segmentation. (c) Nuclei boundaries superimposed on the color image. (d) Zoomed-in image patches. Note that in (c)(d) yellow contours indicate detected melanocytes, and cyan contours indicate segmented keratinocytes.

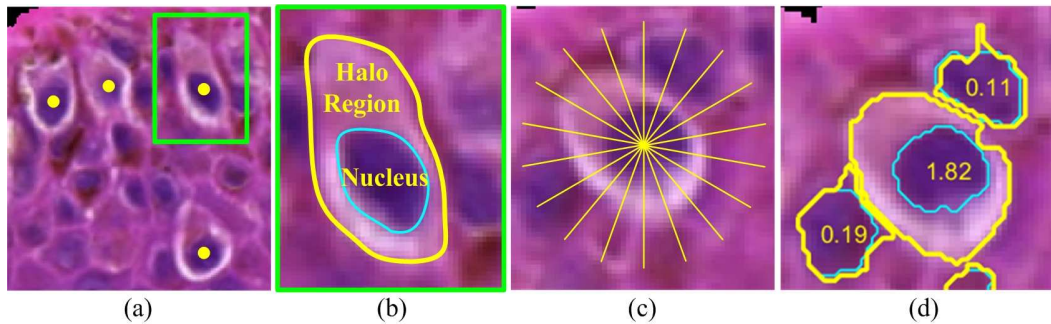


Fig. 6.9. Illustration of melanocytes detection. (a) Epidermis image patch with melanocytes (highlighted by yellow dots). (b) Close-up view of a melanocyte. (c) Radial line scanning. (d) Melanocytes detection overlapped with R_{HN} values.

where melanocytes are highlighted by yellow dots in Fig. 6.9(a). As observed in Figs. 6.9(a)(b), melanocytes have brighter halo-like surrounding regions compared with other keratinocytes (i.e., blue blobs) [107]. Based on this observed pattern, a radial line scanning (RLS) method [107] is employed to detect melanocytes. First, a set of radial lines emerging from nuclear centers (i.e., obtained based on nuclei

segmentation) are generated (see Fig. 6.9(c)). The outer boundary points of the halo region are then estimated based on the gradient information on each radial line. Intuitively, the outer boundary of the halo region consists of points that have a large gradient magnitude and gradient direction pointing to the nuclear center. Finally, melanocytes are detected using a ratio parameter R_{HN} which is calculated as follows:

$$R_{HN} = \frac{A_{HR}}{A_{NR}} \quad (6.1)$$

where A_{HR} and A_{NR} are the areas of the halo region and corresponding nuclear region, respectively. If a nucleus has a ratio R_{HN} larger than the pre-defined threshold τ (e.g., $\tau = 0.8$), the nucleus is determined as a melanocyte. In Fig. 6.8 the detected melanocytes are highlighted by yellow contours. Fig. 6.9 (d) shows the obtained boundaries (yellow contours) of halo regions, and the values of R_{HN} overlapped inside nuclei.

6.2.2.3 Epidermis Features Construction

Since spatial distributions of melanocytes are important factors considered by pathologists for melanoma diagnosis, we follow the work in [59] to construct a set of spatial features reflecting melanocyte distributions. In order to construct spatial features, we first select a set of sampling points from epidermis outer boundary, and then measure the local thickness from the selected outer boundary points to inner boundary points (red dots in Fig. 6.10). Based on the measured thickness, the epidermis is evenly divided into three sub-layers: outer layer, middle layer and inner layer. Fig. 6.10 illustrates an example of dividing epidermis into three layers. The ratio between the number of melanocytes and the total number of cell nuclei within these sub-layers reflects spatial distributions of melanocytes. Thus the spatial features R_{mn}^i , R_{mn}^m and R_{mn}^o are calculated as follows:

$$R_{mn}^i = \mathcal{N}_{mel}^i / \mathcal{N}_{nuc}^i \quad (6.2)$$

$$R_{mn}^m = \mathcal{N}_{mel}^m / \mathcal{N}_{nuc}^m \quad (6.3)$$

$$R_{mn}^o = \mathcal{N}_{mel}^o / \mathcal{N}_{nuc}^o \quad (6.4)$$

where \mathcal{N}_{mel}^i , \mathcal{N}_{mel}^m and \mathcal{N}_{mel}^o denote the numbers of melanocytes in the inner layer, middle layer and outer layer, respectively, and \mathcal{N}_{nuc}^i , \mathcal{N}_{nuc}^m and \mathcal{N}_{nuc}^o denote the total number of cell nuclei in three corresponding layers.

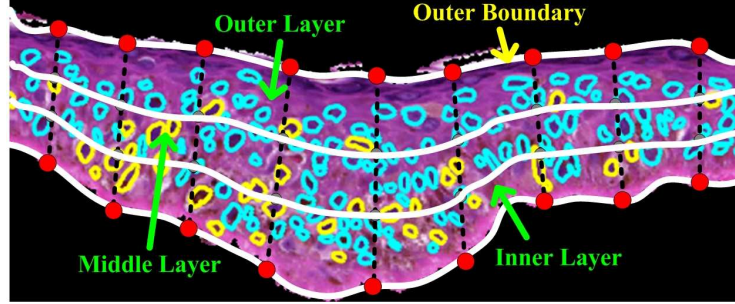


Fig. 6.10. Division of epidermis into three sub-layers: inner layer, middle layer and outer layer.

Besides spatial features, morphological features of cell nuclei are also important for melanoma diagnosis. To capture nuclear morphologies, we compute five widely-used morphological features from segmented keratinocytes and melanocytes for skin tissue classification. The five computed morphological features of cell nuclei include area, perimeter, eccentricity, equivalent diameter (i.e., the diameter of a circle that has the same area as the nuclear region) and ellipticity [59]. For each morphological feature, both the mean and standard deviation (SD) values are computed, which results in 10 morphological features in total. Table 6.2 lists 13 constructed epidermis features and their corresponding labels.

TABLE 6.2
SUMMARY OF EPIDERMIS FEATURES.

Feature Type	Feature Name	Label
Spatial	Ratios R_{mn}^i , R_{mn}^m and R_{mn}^o for the inner, middle and outer layer, respectively	F1-F3
Morphological	Area, Perimeter, Eccentricity, Equivalent diameter and Ellipticity (mean and SD values)	Mean: F4-F8 SD: F9-F13

6.2.3 Dermis Analysis

In parallel with epidermis analysis, dermis analysis is also performed on image tiles $I_d^n, 1 \leq n \leq N$. Fig. 6.11 shows the schematic of dermis analysis. It is observed in Fig. 6.11 that there are three modules for dermis analysis. First color normalization is performed to reduce the effect of staining variations on dermis textural analysis. Next, cell nuclei (mainly melanocytes and lymphocytes) in skin dermis are detected and segmented. Finally a set of textural and cytological features are computed as dermis features. Details of these steps are described below.

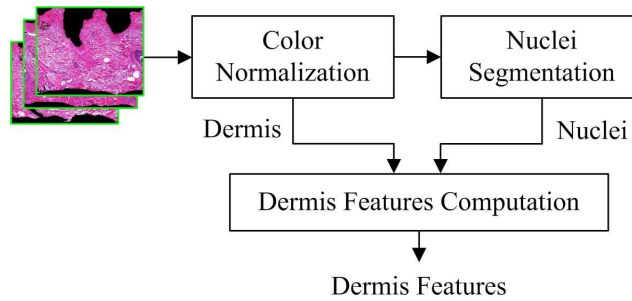


Fig. 6.11. Schematic of dermis analysis.

6.2.3.1 Color Normalization

Because of different storage time, procedures and stains of biopsy slides, the digitized histological images generally have different color appearances. Since textural features will be computed from dermis regions for melanocytic tumor classification, in this module, we perform color normalization on dermis pixels to suppress the influence of staining variations. There have been several existing color normalization methods on histopathology images [118], [119]. In this work, a color deconvolution based method [120] is applied to normalize dermis pixels. For each image, we first adaptively determine the H&E stain vectors by using a singular value decomposition method [120]. The color deconvolution [121] is then utilized to normalize the image into standard color appearance based on pre-determined H&E stain vectors. Fig. 6.12(a) shows two dermis image patches with different color appearances, and

Fig. 6.12(b) shows the normalized image patches. As observed in Fig. 6.12, the color appearance of dermis images look similar after color normalization.

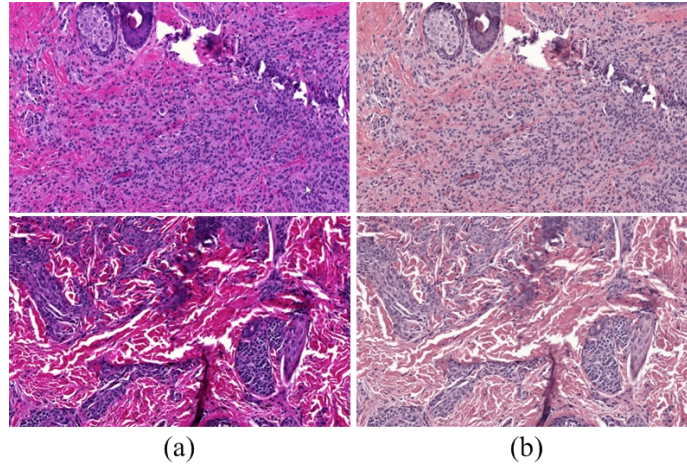


Fig. 6.12. Example of color normalization. (a) Before color normalization. (b) After color normalization.

6.2.3.2 Nuclei Segmentation

After color normalization, we segment cell nuclei (most of them are melanocytes in melanoma and nevus images) in dermis regions. Since cell nuclei usually touch or overlap together in 2-D dermis images, the adaptive thresholding method [30] used during epidermis analysis usually cannot provide a good performance. In this module, we utilize a multi-scale radial line scanning (mRLS) based method (presented in the Chapter 5) to detect and segment cell nuclei. The technique first detects nuclei seeds by using a bank of gLoG kernels with different scales and orientations [106]. Based on the detected nuclei seeds, the mRLS method [122] is then applied to delineate nuclei boundaries. Since nuclear size usually varies from one another and there exist nuclei clumps in dermis regions, the mRLS method delineates the nuclear boundary by adaptively adjusting the maximal scanning radius R_E . For each nucleus, a set of candidate nuclear boundaries (with different R_E values) are first generated by analyzing image gradient on radial lines. The gradient, intensity and shape information are then integrated together to determine

the optimal boundary for the nucleus, i.e.,

$$E(R_E) = \lambda_1 E_g(R_E) + \lambda_2 E_i(R_E) + \lambda_3 E_s(R_E) \quad (6.5)$$

where $R_E^{\min} \leq R_E \leq R_E^{\max}$, and parameters R_E^{\min} and R_E^{\max} are estimated based on nuclei radii in the image. $E_g(\cdot)$, $E_i(\cdot)$ and $E_s(\cdot)$ represent the gradient, intensity and shape terms, respectively, and λ_1 , λ_2 , λ_3 are the corresponding weights. By selecting the minimal cost in Eq. (6.5), it is expected to obtain a nuclear boundary lying on high gradient points, and surrounding a region with an elliptical shape and homogeneous intensities.

Fig. 6.13(a) shows a dermis image patch with detected nuclei seeds (indicated by + symbols). Fig. 6.13(b) shows the image with candidate nuclei boundaries, where $7 \leq R_E \leq 11$. Note that in Fig. 6.13(b) the candidate nuclei boundaries are represented by five different colors: green, magenta, cyan, red and yellow. Fig. 6.13(c) shows the final nuclei boundaries with $\lambda_1 = \lambda_2 = \lambda_3 = 1$. Fig. 6.14 shows examples of nuclei detections (yellow + symbols) and segmentations, where the first, second and third rows are the normal, nevus and melanoma images, respectively.

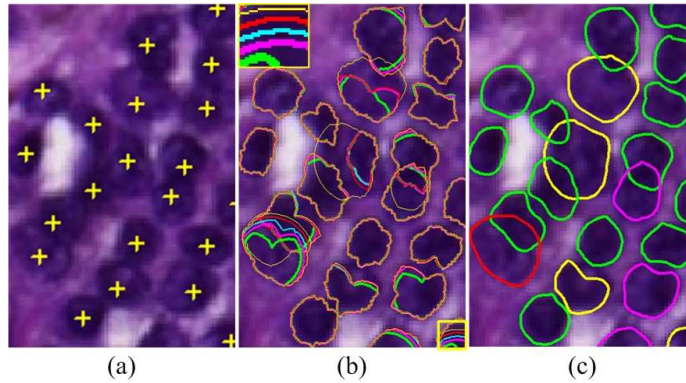


Fig. 6.13. Illustration of nuclei segmentation. (a) Nuclei seeds + detected by gLoG kernels. (b) Nuclei candidate boundaries obtained by the mRLS method. (c) Finally obtained nuclei boundaries. Note that in (b) it contains five different color of contours each of which corresponds to different R_E values. That is green: $R_E = 7$, magenta: $R_E = 8$, cyan: $R_E = 9$, red: $R_E = 10$ and yellow: $R_E = 11$. In (b) the top-left corner shows a zoomed-in image patch in the bottom-right corner for illustrating different colors.

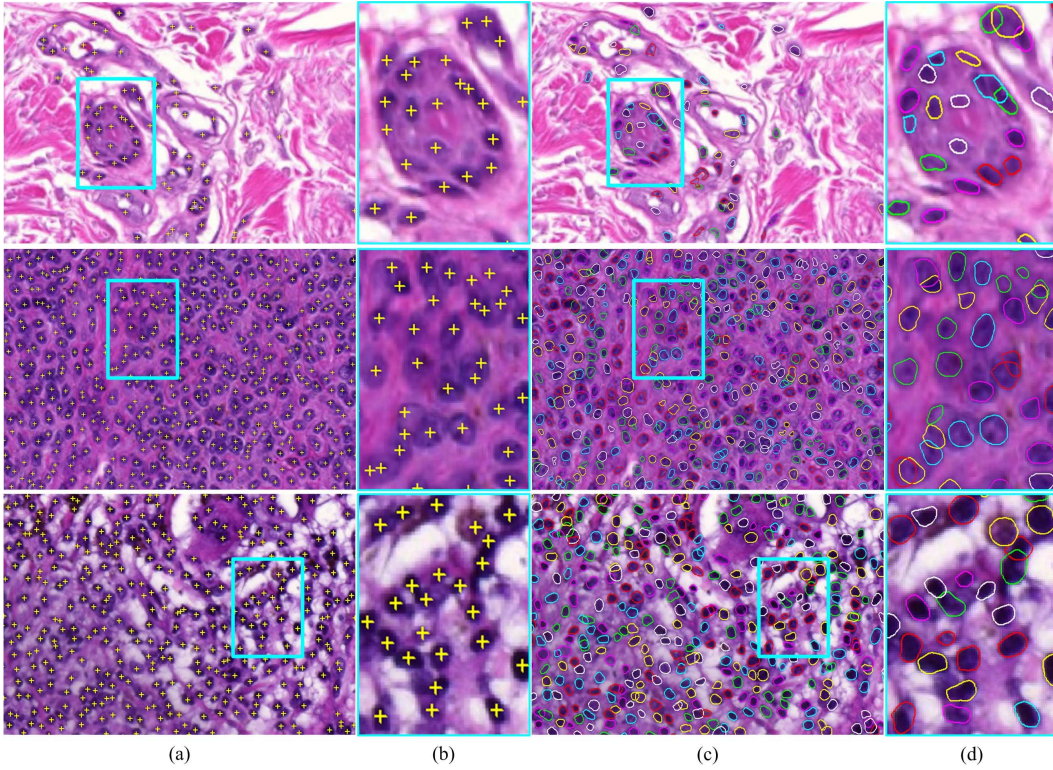


Fig. 6.14. Nuclei detection and segmentation in dermis. (a) Nuclei detection. (b) Close-up view of image patches in (a). (c) Nuclei segmentation. (d) Close-up view of image patches in (c). Note that the first row is an image from normal skin, while the second and third rows are the images from skin nevus and melanoma, respectively. In (a)(b) detected seeds are indicated by yellow + symbols. In (c)(d) segmented nuclei are indicated by different color of contours.

6.2.3.3 Dermis Features Computation

Since textural appearance of dermis regions usually varies from one another for different types of skin tissues, textural features are important for melanocytic tumor classification. In this work, we compute a set of textural features from the histogram and gray-level co-occurrence matrix (GLCM) of dermis pixels. Note that although the dermis tiles I_d^n , $1 \leq n \leq N$, are color images, only red channel images are utilized for computing textural features as the red channel of H&E stained skin images has the most information compared with green and blue channels [74]. Given a red channel image of dermis tile I_d^n , 6 histogram features including mean value, variance, skewness, kurtosis, energy and entropy of the grey level distribution [55] are computed. The GLCM records the joint frequency of intensity levels for a pair of pixels. Studies [26], [55], [58] suggest that Haralick features derived from

the GLCM can be used to distinguish different tissue types, especially contrast, correlation, energy, entropy, homogeneity, and dissimilarity. Therefore, these six Haralick features are selected and computed. For calculation of the GLCM, we use a distance of one pixel in four directions (i.e., 0° , 45° , 90° and 135°), and hence it gives 24 Haralick features in total.

Besides textural features, we also compute a set of cytological features based on pre-segmented nuclear regions. These cytological features reflect nuclear morphologies, intensity statistics and spatial densities. To capture nuclear morphologies, 5 morphological features including area, perimeter, eccentricity, equivalent diameter and ellipticity are computed from nuclear regions. To capture nuclear intensity statistics, 6 statistical moments are computed (from the red channel image) to estimate pixel properties within nuclear regions. These statistical moments [26] include average intensity, average contrast, smoothness, third moment, uniformity and entropy. To capture spatial densities of cell nuclei, 4 architectural features are computed from Voronoi diagrams and Delaunay triangulations. The Voronoi diagram and Delaunay triangulation [60] are constructed by considering the detected nuclei seeds (by gLoG kernels) as nuclei centroids. The area and perimeter of Voronoi diagrams, and the area and perimeter of Delaunay triangles are computed as architectural features to reflect nuclear densities. Intuitively, the smaller area and perimeter of polygons indicate a larger density of cell nuclei in dermis. Note that unlike existing works such as [67] which computes geometry features based on all obtained polygons, this work only considers the polygons within dermis regions, which is intuitively more accurate for nuclear density estimation. For example, Figs. 6.15(a)(b) show a dermis tile with constructed Voronoi diagrams, and Figs. 6.15(c)(d) show the image with constructed Delaunay triangles. In Fig. 6.15 only the blue polygons within dermis regions are used for computing architectural features, while the yellow polygons in the (black) background are discarded. For each cytological feature, both the mean and SD values are computed, and hence there are 30 cytological features in total. Table 6.3 summarizes 60 computed dermis features and their corresponding labels.

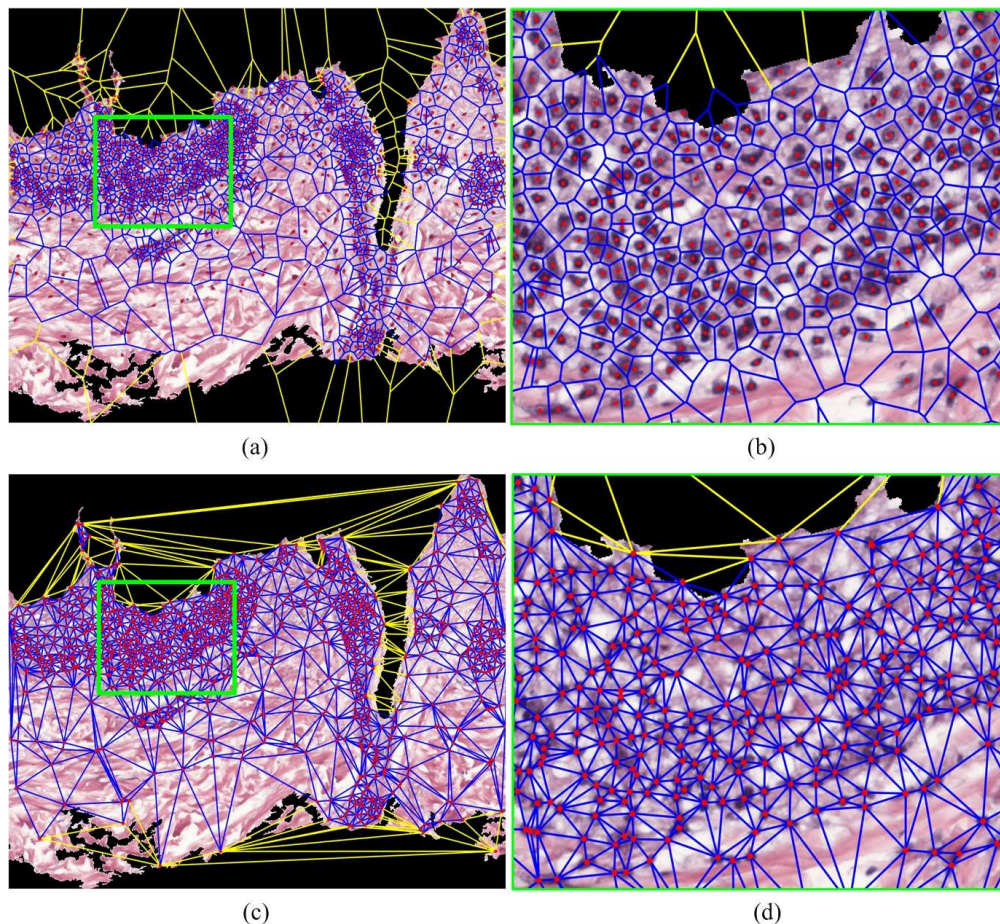


Fig. 6.15. Architectural constructions based on detected nuclei seeds. (a) Voronoi diagrams. (b) Close-up view of (a). (c) Delaunay triangles. (d) Close-up view of (c). Note that only blue polygons within dermis regions are used for architectural features computation. The red dots are the automatically detected nuclei seeds.

6.2.4 Classification

After epidermis and dermis analysis, 73 meaningful features (13 epidermis and 60 dermis features) are extracted. These extracted features are now used by classification techniques to classify the skin WSI into different categories. In this work, the “one-against-one” multi-class support vector machine (mSVM) method [123] is utilized to perform melanocytic tumor classification. Let k denote the number of skin tissue classes. In “one-against-one” mSVM method, we construct $k(k - 1) / 2$ SVM classifiers, and each classifier is trained on data from two classes during the training phase. Given training data from the i th and j th classes, the hyperplane of

TABLE 6.3
SUMMARY OF DERMIS FEATURES.

Feature type		Feature name	Label
Textural	Histogram	Mean, Variance, Skewness, Kurtosis, Energy and Entropy	F14-F19
	Haralick	Contrast, Correlation, Energy, Entropy, Homogeneity and Dissimilarity (in 4 directions: 0°, 45°, 90° and 135°)	0°: F20-F25 45°: F26-F31 90°: F32-F37 135°: F38-F43
Cytological	Morphology	Area, Perimeter, Eccentricity, Equivalent diameter and Ellipticity (mean and SD values)	Mean: F44-F48 SD: F49-F53
	Statistics	Average intensity, Average contrast, Smoothness, Third moment, Uniformity and Entropy (mean and SD values)	Mean: F54-F59 SD: F60-F65
	Architecture	Area and Perimeter of Voronoi diagrams, Area and Perimeter of Delaunay triangles (mean and SD values)	Mean: F66-F69 SD: F70-F73

SVM classifier to separate these two classes is obtained by optimizing the following function:

$$\min_{\omega^{ij}, b^{ij}, \xi^{ij}} \frac{1}{2} \|\omega^{ij}\|^2 + C \sum_t^n \xi_t^{ij} \quad (6.6)$$

with the following constraints:

$$(\omega^{ij})^T \phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij}, \quad \text{if } y_t = i \quad (6.7)$$

$$(\omega^{ij})^T \phi(x_t) + b^{ij} \leq \xi_t^{ij} - 1, \quad \text{if } y_t = j \quad (6.8)$$

where ω^{ij} , b^{ij} are the coefficients of the hyperplane, ξ^{ij} is the slack variable, C is a predefined regularization parameter, n represents the number of training data and $\phi(\cdot)$ is the kernel function. Note that x and y represent the extracted features set and manually identified class label, respectively. In this work, the mSVM classifier with a linear kernel is utilized for classification.

After obtaining $k(k-1)/2$ SVM classifiers, the test image will be labeled by all of them in the testing phase. Specially, if $\text{sign}\left((\omega^{ij})^T \phi(x_t) + b^{ij}\right)$ indicates

that x_t belongs to the i th class, then the vote for the i th class is increased by 1. Otherwise, the vote for the j th class is increased by 1. After evaluation by all $k(k-1)/2$ SVM classifiers, x_t is predicted to be in the class with the largest vote.

6.3 Performance Evaluations

In this section, we evaluate the performance of the proposed technique for skin melanocytic tumor diagnosis. First we illustrate the dataset and test strategies used in this work. We then present classification results and experiments with feature selection. Finally comparisons and computational complexity of the proposed technique are provided.

6.3.1 Image Dataset

This study was based on a dataset of 66 H&E stained skin WSIs which were collected from the Cross Cancer Institute, University of Alberta. The skin biopsies used for digitizing images include 17 normal skin tissues, 17 nevi and 32 melanomas. Based on pathologists' manual analysis, all nevus biopsies are categorized into the same type: compound nevus. Melanoma tissues are categorized into three sub-types: lentiginous, superficial spreading, and nodular melanoma. The original skin WSIs were captured under 40X magnification on Carl Zeiss MIRAX MIDI scanning system. For higher efficiency and accuracy, these skin images were saved in TIFF format with 25X magnification ($0.372\mu\text{m}/\text{pixel}$) by MIRAX viewer software. Each testing image has a size between 1 and 2 gigabytes, with tens of thousands of cell nuclei. Table 6.4 lists the distribution of different classes in the dataset.

6.3.2 Experimental Design

Considering the class distributions of our dataset, four test strategies were designed to evaluate the performance of the proposed technique, which are described as follows:

TABLE 6.4
DESCRIPTION OF THE IMAGE DATASET.

Tissue Classes		No.	Pct.
Normal		17	25.7%
Nevus (compound)		17	25.7%
Melanoma	lentiginous	9	13.7%
	superficial spreading	18	27.3%
	nodular	5	7.6%
Total		66	100.0%

6.3.2.1 Test Strategy I

The whole dataset is tested with 2 classes: 32 melanomas and 34 non-melanomas (including 17 normal tissues and 17 nevi). The objective of this testing is to evaluate if the proposed technique can efficiently diagnose melanomas from a set of skin biopsy tissues.

6.3.2.2 Test Strategy II

The whole dataset is tested with 3 classes: 17 normal skins, 17 nevi and 32 melanomas. The objective of this testing is to evaluate if the proposed technique can efficiently distinguish melanomas, nevi and normal skin tissues.

6.3.2.3 Test Strategy III

The melanoma WSIs are tested with 3 sub-types: 9 lentiginous melanomas, 18 superficial spreading melanomas and 5 nodular melanomas. The objective of this testing is to evaluate if the proposed technique can efficiently distinguish different melanoma sub-types.

6.3.2.4 Test Strategy IV

The whole dataset is tested with 5 classes: 17 normal skins, 17 nevi, 9 lentiginous melanomas, 18 superficial spreading melanomas and 5 nodular melanomas. The objective of this testing is to evaluate the overall capability of the proposed technique on classifying different skin WSIs.

For each test strategy, if the automatically classified result is the same as the expert class label, it is considered as a correct classification. The performance of the proposed technique is evaluated by the classification accuracy \mathcal{A}_{cc} , which is defined as follows:

$$\mathcal{A}_{cc} = \frac{\sum_{i=1}^k \sum_{j=1}^{N^k} 1(B_j^i = i)}{\sum_{i=1}^k N^k} \times 100\% \quad (6.9)$$

where N^k is the number of images belonging to the class k , B_j^i represents the automatically classified result, and $1(\cdot)$ is an indicator function, which is defined as follows:

$$1(x) = \begin{cases} 1 & \text{if } x \text{ is true} \\ 0 & \text{if } x \text{ is false} \end{cases} \quad (6.10)$$

For test strategy I, the skin WSI is classified as either melanoma or non-melanoma, which is a typical two-class classification problem. Besides the overall classification accuracy \mathcal{A}_{cc} , we further evaluate the test strategy I using the following widely-used evaluation metrics: sensitivity (SEN), specificity (SPE) and precision (PRE), which are defined as follows:

$$SEN = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\% \quad (6.11)$$

$$SPE = \frac{N_{TN}}{N_{FP} + N_{TN}} \times 100\% \quad (6.12)$$

$$PRE = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \quad (6.13)$$

where N_{TP} , N_{FP} , N_{TN} and N_{FN} denote the number of true positives, false positives, true negatives and false negatives, respectively.

By using the above evaluation criteria, we perform the 10-fold cross validation [59] on each test strategy. In the 10-fold cross validation, the whole dataset is first divided into 10 subsets. 9 subsets are used as the training set, and the remaining 1 subset is used as the testing set. The evaluations are then performed on the testing

set. By using different training and testing sets, the evaluations are performed by 10 times. Since the division of the dataset may slightly impact the performance of the technique, we repeat the 10-fold cross validation by 100 times, and the average values of different evaluation criteria are used as the final results.

6.3.3 Classification Results

Note that the regularization parameter C in Eq. (6.6) is a key parameter for classification by SVM with a linear kernel. A large C value generally achieves a good performance on the training set but a poor performance on the testing set due to overfitting. On the other hand, a small C value may ignore constraints and produce sub-optimal classification results [124]. In order to select an appropriate C value for our model, we test it with several empirically-selected values (i.e., 0.01, 0.1, and 1.0). Table 6.5 lists the classification accuracies of four test strategies using different C values, and Table 6.6 lists the evaluations of melanoma diagnosis in test strategy I using different C values. It is observed from Table 6.5 and Table 6.6 that the proposed technique achieves the overall best performance with $C = 0.1$. In particular, when $C = 0.1$ the proposed technique with all features achieves 97.90%, 95.78%, 91.32% and 91.98% classification accuracies (see Table 6.5) for four test strategies, respectively. When $C = 0.1$ the proposed technique with all features achieves 99.39%, 96.98% and 97.51% of SEN , SPE and PRE values (see Table 6.6). Therefore, 0.1 is a suitable value for parameter C in our SVM model.

Table 6.5 and Table 6.6 also compare the performance of the proposed technique by using epidermis, dermis and all features, respectively. It is observed from Table 6.5 that the proposed technique with all extracted features provides better performances (i.e., higher classification accuracies) than those by only using dermis features or epidermis features in all four test strategies. The proposed technique with dermis features provides overall better performances than the technique with epidermis features in test strategies I, II and IV, but poorer performances in test strategy III. This indicates that the spatial and morphological features extracted

TABLE 6.5
CLASSIFICATION RESULTS OF FOUR TEST STRATEGIES.

Tests	Classification Accuracies \mathcal{A}_{cc} (%)			
	Strategies	$C=0.01$	$C=0.1$	$C=1.0$
Proposed technique with epidermis features	I	84.74	92.88	92.98
	II	82.33	94.43	91.61
	III	84.38	88.04	89.60
	IV	81.23	85.75	88.80
	Strategies	$C=0.01$	$C=0.1$	$C=1.0$
Proposed technique with dermis features	I	82.88	97.68	97.31
	II	91.39	94.38	94.13
	III	84.93	86.92	87.45
	IV	82.63	90.16	90.66
	Strategies	$C=0.01$	$C=0.1$	$C=1.0$
Proposed technique with all features	I	92.66	97.90	96.71
	II	95.14	95.78	95.75
	III	87.84	91.32	90.82
	IV	88.45	91.98	92.15
	Strategies	$C=0.01$	$C=0.1$	$C=1.0$

from epidermis regions contribute more for the discrimination of melanoma subtypes in our dataset. It is observed from Table 6.6 that the proposed technique with all features provides an overall better performance than the technique with only epidermis features or dermis features in melanoma diagnosis (for test strategy I). This indicates that the proposed technique with all features can more accurately diagnose melanomas.

6.3.4 Experiments with Feature Selection

As in many other tissue classification applications, some of the extracted features might be redundant, without contributing much to the discrimination power. In fact, if more than the necessary number of features are used, it can degrade the overall performance due to the “peaking phenomenon” [72]. Since the proposed technique achieves the overall best performance when $C = 0.1$, in this section, we evaluate the proposed technique with feature selection using $C = 0.1$.

Let X be the original set of features with m (i.e., $m = 73$ as mentioned in Section 6.2) dimensions. The problem of feature selection is to select a subset X_s ,

TABLE 6.6
PERFORMANCE EVALUATION OF MELANOMA DIAGNOSIS.

Tests	Evaluations on Strategy I			
	C	SEN (%)	SPE (%)	PRE (%)
Proposed technique with epidermis features	0.01	73.96	94.94	93.70
	0.1	93.49	92.00	93.12
	1.0	93.19	93.08	94.21
Proposed technique with dermis features	C	SEN (%)	SPE (%)	PRE (%)
	0.01	71.65	93.11	91.81
	0.1	98.82	96.70	97.14
	1.0	97.85	96.92	97.40
Proposed technique with all features	C	SEN (%)	SPE (%)	PRE (%)
	0.01	93.83	91.81	93.13
	0.1	99.39	96.98	97.51
	1.0	97.63	95.01	95.92

where $X_s \subseteq X$, such that the classification accuracy over all subsets of X to be maximized, i.e.,

$$X_s = \arg \max_{X_s \subseteq X} \mathcal{A}_{cc}(X_s) \quad (6.14)$$

where $\mathcal{A}_{cc}(X_s)$ is the predictive accuracy over the subset X_s . Although the optimal subset X_s can be obtained by an exhaustive search (i.e., to evaluate all possible candidate subsets of X), it is usually impractical, even for moderate values of m , as the number of possible candidate subsets of X increases exponentially with its dimensions m . Several algorithms have been proposed for finding the optimal subset of features without searching the entire solution space [125]. For efficiency and accuracy, a sequential forward selection (SFS) algorithm [69] is employed to select the optimal subset of features in this work. This algorithm starts from an empty feature set, and then sequentially adds features until $\mathcal{A}_{cc}(X_s)$ is maximized. During the feature selection, we utilize the mSVM classification method to perform the 10-fold cross validation over our dataset. Fig. 6.16 illustrates feature selection by the SFS algorithm. As observed in Fig. 6.16, in the test strategy I, the classification accuracy reaches the maximum value when 29 features are used, and this classification accuracy stays unchanged over the range from 29 features to 73 features. In the test strategy II, the classification accuracy has the maximum

TABLE 6.7
SELECTED EPIDERMIS AND DERMIS FEATURES.

Tests	Epidermis features	Dermis features	Dim.
I	F1, F2, F3, F4, F7, F9, F10, F11, F12, F13	F14, F15, F16, F18, F19, F46, F47, F48, F49, F53, F54, F55, F57, F59, F63, F68, F69, F71, F73	29
II	F4, F9, F6, F11	F14, F18, F45, F49, F50, F52, F54, F59, F60	13
III	F1-F13	F14, F15, F16, F17, F20, F26, F32, F60	21
IV	F2-F13	F14, F15, F20, F43, F60, F69	18

value over the range from 13 features to 63 features. In the test strategy III, the classification accuracy has the maximum value over the range from 21 features to 65 features. In the test strategy IV, the classification accuracy has the maximum value over the range from 18 features to 47 features.

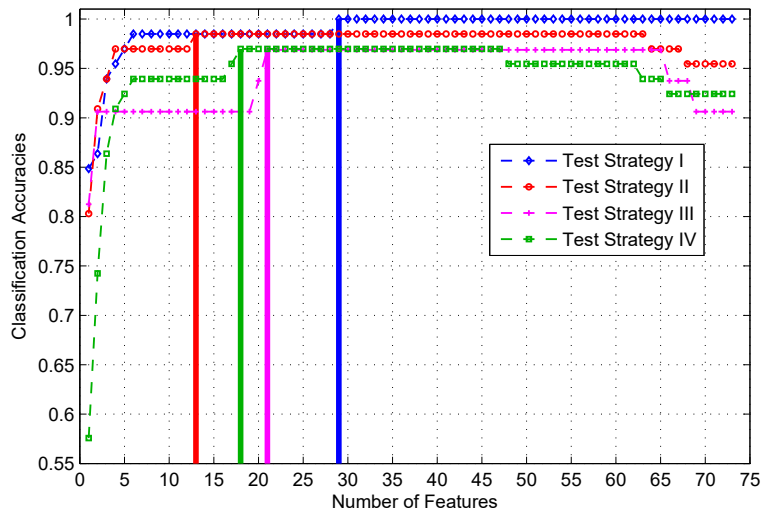


Fig. 6.16. Feature selection by 10-fold cross validation. Note that the classification accuracy reaches the maximum when 29, 13, 21 and 18 features are used for strategies I, II, III and IV, respectively.

Although the classification accuracy achieves the maximum value over a range of different number of features, for simplicity, we test proposed technique with the minimum number (i.e., 29, 13, 21 and 18) of selected features for each test strategy. Table 6.7 lists the labels of selected features in four test strategies I, II, III and IV, respectively. As observed in Table 6.7, in test strategies III and IV, most of epidermis

TABLE 6.8
COMPARISONS OF SKIN WSI CLASSIFICATION.

Techniques	Average classification accuracies \mathcal{A}_{cc} (%)			
	I	II	III	IV
LM technique [59]	88.86±1.71	89.07±2.72	90.17±2.23	90.24±1.51
Proposed technique	97.80±1.27	98.08±0.80	95.81±1.89	95.73±1.16

TABLE 6.9
COMPARISONS OF SKIN MELANOMA DIAGNOSIS.

Techniques	Test strategy I		
	$SEN(\%)$	$SPE(\%)$	$PRE(\%)$
LM technique [59]	90.35±3.22	87.97±3.37	87.42±3.83
Proposed technique	99.81±0.66	95.63±1.41	96.39±1.06

features are selected as they contribute more for classifying melanoma sub-types, which is consistent with our results in Table 6.5. To evaluate the proposed technique with selected features, we run 10-fold cross validation by 100 times for each test strategy, and compute the average values of different evaluation criteria as the final results. In the second rows of Table 6.8 and Table 6.9, it lists the evaluations of the proposed technique with selected features. It is observed from Tables 6.8, 6.9 that the proposed technique with selected features provides an overall superior performance to the proposed technique with all features (see Tables 6.5, 6.6). Specifically, the classification accuracies of four test strategies are all more than 95%. For test strategy I, the proposed technique also provides more than 95% of sensitivity, specificity and precision values. Therefore, the proposed technique with selected features can more accurately diagnose melanomas and classify skin WSIs.

6.3.5 Comparisons

To our best knowledge, there is only one existing technique [59], henceforth referred to as LM technique, which has been proposed for skin WSI classification in the literature. The LM technique analyzes skin epidermis areas and performs skin WSI classification by utilizing epidermis features. The skin WSI is classified into three categories: melanoma, nevus or normal tissue. For comparison, we

implemented the LM technique and carried out more experiments based on our experimental design such that the proposed technique is statistically compared with the LM technique. In the first rows of Table 6.8 and Table 6.9, the experimental results obtained by the LM technique are tabulated. It is observed that the proposed technique achieves a superior performance compared to LM technique [59]. There are two main reasons for the superior performance. First, instead of just analyzing epidermis areas (like the LM technique), the proposed technique computes both epidermis and dermis features, and thus more meaningful features are used for skin tissue classification. Second, the SFS algorithm [69] is utilized to select the discriminating features, which improves the capability of the classifier.

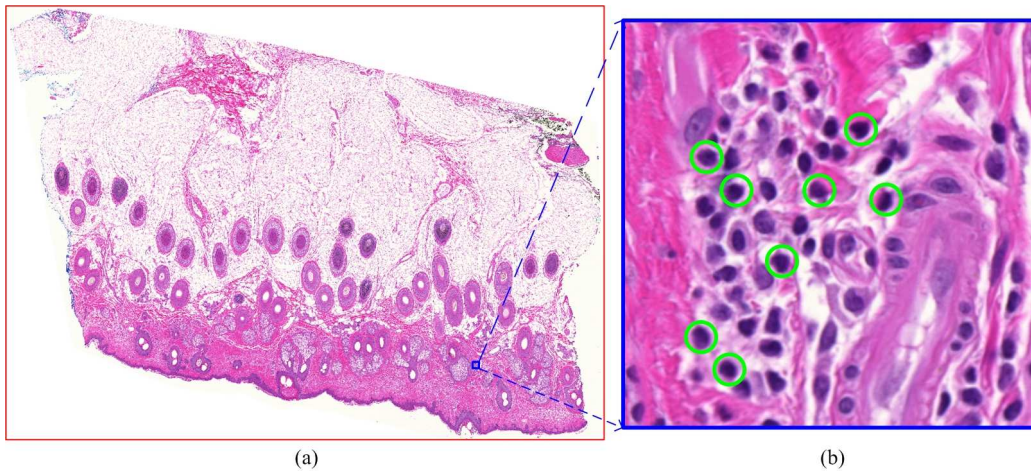


Fig. 6.17. An example of false diagnosis. (a) Skin melanoma WSI. (b) Zoomed-in image patch in (a). Note that the superimposed green circles in (b) highlight some lymphocytes which are related to lymphocytic infiltration.

Although a good performance for melanoma diagnosis is achieved by the proposed technique, there still exist cases of missed detection. An example of missed detection by the proposed technique is shown in Fig. 6.17. The skin tissue in Fig. 6.17 belongs to the melanoma class (i.e., nodular melanoma), but the automated technique incorrectly classifies it as a non-melanoma tissue. The missed detection occurs because this skin melanoma is in-transit metastasis, where cytological and textural features in the skin WSI (e.g., skin epidermis and dermis) are similar with those in the normal skin tissue. To correctly classify this melanoma tissue, it would be necessary to analyze much deeper skin dermis regions and incorporate other

image features (e.g., by evaluating lymphocytic infiltration).

6.3.6 Computational Complexity

All experiments were carried out on a 1.8-GHZ Intel Core i7-4500U CPU with 16-GB RAM using MATLAB 8.10. Because of different image size and nuclear densities, the processing time for different skin WSIs varies from each other. For processing a $13786 \times 10320 \times 3$ skin WSI with 2761 cell nuclei in epidermis and 4130 cell nuclei in dermis, the proposed technique takes about 38.5 minutes in total. The epidermis and dermis segmentation module (see Fig. 6.2) takes 2.47% of the whole processing time. The epidermis analysis module takes 77.67% of the whole processing time, and the dermis analysis module takes 19.86% of the whole processing time. Compared with the first three modules, the processing time (less than 0.1%) for feature classification module is negligible. Among the four modules, the epidermis analysis takes the largest proportion of the whole processing time. This is due to the high computational complexity of melanocytes detection in epidermis regions, which takes about 90.62% of epidermis analysis time. It should be noted that our current implementation has not used any parallelization. In the proposed technique, epidermis and dermis analysis can be performed in a parallel manner. In addition, melanocytes detection in epidermis analysis module and nuclei segmentation in dermis analysis module are also suitable for parallelization. Parallel implementation can speed up the proposed technique significantly.

6.4 Summary

This chapter presents a computer-aided technique for skin WSIs analysis and classification, especially for melanocytic tumor classification. The proposed technique first analyzes skin epidermis and dermis regions to extract a set of 73 meaningful features, and then classifies skin WSIs into different categories by using a mSVM classification method. Experiments on 66 H&E stained skin WSIs indicate that the proposed technique provides a good performance in skin tissue classification.

In particular, it provides more than 95% classification accuracies by utilizing a selected feature set. The proposed technique can potentially be used by pathologists to obtain a second opinion during their melanoma diagnosis.

Chapter 7

Automatic Measurement of Melanoma Depth of Invasion

Measurement of melanoma depth of invasion (DoI) in skin tissues is of great significance in grading the severity of skin disease and planning patient's treatment. However, accurate and automatic measurement of melanocytic tumor depth is a challenging problem mainly due to the difficulty in skin granular identification and melanoma detection. In this chapter, we propose an automated technique for measuring melanoma DoI in MART-1 stained skin histopathological images. The technique detects skin granular layer by a Bayesian classification based method, and measures melanoma DoI by a Hausdorff distance measure. Experimental results show that the proposed technique provides a superior performance in measuring the melanoma DoI than two closely related techniques.

7.1 Introduction

The digitized skin histopathological images used in this study have been obtained using the MART-1 stain (a type of immunohistochemical stain specific to melanoma). Fig. 7.1 shows a typical skin image in our database, which mainly includes image

Part of this chapter has been published in: Xu *et al.* "Automatic measurement of melanoma depth of invasion in skin histopathological images", *Micron*, vol.97, pp.56-67, 2017.

background, skin epidermis and dermis regions. As shown in Fig. 7.1, the epidermis can be divided into three layers: cornified layer (L1), granular layer (L2) and malpighian layer (L3). The dermis is located below the DEJ (indicated by the cyan curve). In the skin dermis, the regions with dark brown colors represent MART-1 verified skin melanoma areas. In clinical setting, the melanoma DoI is measured as the maximum distance between the malignant cells and the top of skin granular layer [19]. In Fig. 7.1, the manually identified skin granular layer is indicated by the red curve and the melanoma invasion depth is measured as $0.46mm$. Note that since skin granular layer is a middle (thin) layer of skin epidermis and looks similar to other skin epidermis layers, the accurate detection of skin granular layer is critical in achieving accurate melanoma DoI measurement.

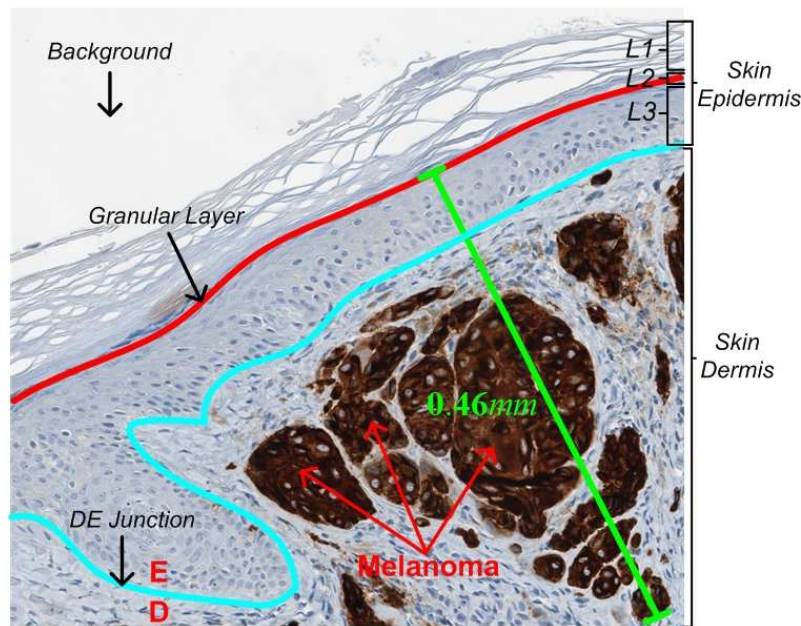


Fig. 7.1. Melanoma DoI measurement in a MART-1 stained skin image. L1: Cornified layer, L2: Granular layer, L3: Malpighian (squamous and basal) layer.

Recently, there have been two closely related works in the literature, where the melanocytic tumor DoI in H&E stained skin images is measured. Mokhtari *et al.* [21] proposed a technique to measure the melanoma DoI in skin microscopic images. The technique first segments skin epidermis based on morphological closing and thresholding. The melanocytes containing melanin are then detected

based on color features by the SVM classifier. Finally, the epidermis orientation is estimated based on Hough transform. The distance between the outermost pixel of epidermis and the farthest melanocyte along a direction perpendicular to epidermis orientation is calculated as the invasion depth. This technique has a few problems which may result in an incorrect DoI measurement. For example, the skin cornified layer is usually included in the segmented epidermis, which tends to increase the measured invasion depth. Noroozi *et al.* [19] proposed an improved technique that segments the skin cornified layer from epidermis based on entropy analysis. The cell nuclei in skin dermis are segmented by using color features and active contour model [84]. The segmented melanocytes are distinguished from lymphocytes based on shape analysis. The maximum distance between the detected melanocyte and the outer boundary of skin epidermis is finally measured as the invasion depth. Since both cornified and granular layers in skin biopsy images have ridged shapes, the technique in [19] may incorrectly segment granular layer into the cornified layer. In addition, the technique assumes that there are only melanocytes and lymphocytes present in skin dermis, which is usually not true. Besides melanocytes and lymphocytes, the skin dermis may include many other types of cell nuclei such as red blood cells, endothelial cells, and nevus cells. Some cell nuclei (e.g., endothelial cells) look similar to melanocytes in H&E stained skin images, and thus can be incorrectly detected as melanocytes.

In this chapter, we propose a robust technique for measuring the melanoma DoI in MART-1 stained skin microscopic images, which has overcome the limitations of existing techniques for tumor invasion measurement. The proposed technique identifies skin granular layer by a Bayesian classification based method, and measures melanoma DoI by a multi-resolution approach with Hausdorff distance measure. Because of accurate detection of melanoma areas and skin granular layer, the proposed technique provides a superior performance compared to existing techniques for measuring melanoma DoI. The details of the proposed technique and experimental results are provided in the following sections.

7.2 Dataset Description

In this study, all skin histopathological images are collected from the Cross Cancer Institute, University of Alberta in accordance with the protocol for the examination of specimens with skin melanoma. The histological sections of skin tissues are about $4\mu m$ thick each and are stained with anti-MART-1 clone A103, 1:25 dilution using an avidin-biotin procedure with 3,3'-Diaminobenzidine (DAB) as the chromogen. The skin images were captured under 40X magnification on Aperio Scanscope CS scanning system ($0.25\mu m/\text{pixel}$). The dataset used in this work consists of 29 MART-1 stained skin microscopic images. Among 29 images, one representative image is selected as the training image, and the other 28 images are used for testing. The ground truths of tumor invasion depth for all test images were manually assigned by an expert pathologist using the Image-Scope slide viewing software. Since the skin granular layer identification is of great importance for measurement of melanoma DoI, the ground truths of skin granular layers for all testing images are also provided by the expert pathologist. Fig. 7.1 shows the manually labeled skin granular layer (red curve) and measured melanoma invasion depth ($0.46mm$) in a skin image.

7.3 Proposed Technique

The schematic of the proposed technique for measurement of melanoma invasion depth is shown in Fig. 7.2. It is observed that there are four modules. In the first module, the skin melanoma areas are detected based on RGB color features. In the second module, the skin epidermis is segmented by a multi-thresholding method. In the third module, the skin granular layer is identified from the segmented skin epidermis using a supervised classification technique. Finally, the tumor DoI is measured as the distance between skin granular layer and the deepest pixel of melanoma areas. The details of four modules are presented in the following.

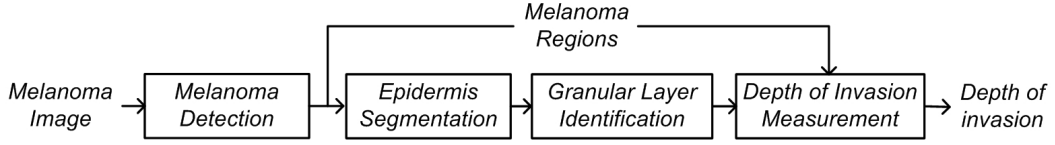


Fig. 7.2. Schematic of the proposed technique.

7.3.1 Melanoma Detection

In this module, we segment melanoma areas from MART-1 stained skin histopathological images. As observed in Fig. 7.1, melanoma areas in our MART-1 stained skin images present dark brown colors. Based on this observed feature, skin melanoma areas can be effectively detected using RGB color features. To segment melanoma areas based on color features, we manually select a set of melanoma pixels $\hat{x}_k, k = 1, \dots, K$, from the training image I_t . Fig. 7.3 shows the distribution of selected melanoma pixels in RGB color space, where $K = 10454$. As observed in Fig. 7.3, the 3D shape of point cloud for melanoma pixels can be roughly described by a solid 3-D elliptical model. A candidate image pixel within or on the surface of the 3-D elliptical model can be considered as a melanoma pixel. Let \hat{m} be a 3×1 vector corresponding to the mean RGB values of K selected melanoma pixels, and C be the 3×3 covariance matrix for the K selected melanoma pixels, i.e.,

$$C = \frac{1}{K-1} \sum_{k=1}^K (\hat{x}_k - \hat{m})(\hat{x}_k - \hat{m})^T \quad (7.1)$$

If \hat{z} denotes a column vector containing RGB values for an arbitrary pixel in a test image I , the pixel classification is performed based on the following distance measure:

$$DIS(\hat{z}, \hat{m}) = \left[(\hat{z} - \hat{m})^T C^{-1} (\hat{z} - \hat{m}) \right]^{1/2} \quad (7.2)$$

where $DIS(\hat{z}, \hat{m})$ is the Mahalanobis distance between two vectors [31]. After computing the Mahalanobis distance between the testing pixel and the training data center, the pixel \hat{z} is determined as a melanoma pixel if it satisfies the following

inequality:

$$DIS(\hat{z}, \hat{m}) \leq \tau_1 \quad (7.3)$$

where τ_1 is a specified threshold that determines if the pixel is contained within, or on the surface of the 3-D elliptical model. In this work, the threshold τ_1 is empirically determined as follows:

$$\tau_1 = \min \left\{ \sqrt{C_{11}}, \sqrt{C_{22}}, \sqrt{C_{33}} \right\} \quad (7.4)$$

where C_{11} , C_{22} and C_{33} are the main diagonal values in the matrix C . For our dataset, the obtained covariance matrix C is as follows:

$$C = \begin{bmatrix} 977 & 682 & 425 \\ 682 & 527 & 358 \\ 425 & 358 & 271 \end{bmatrix} \quad (7.5)$$

where the main diagonal elements (977, 527, 271) represent the variances of the R, G, and B components, respectively. Following Eq. (7.4), the threshold τ_1 was selected as $\sqrt{271}$ or 16.5.

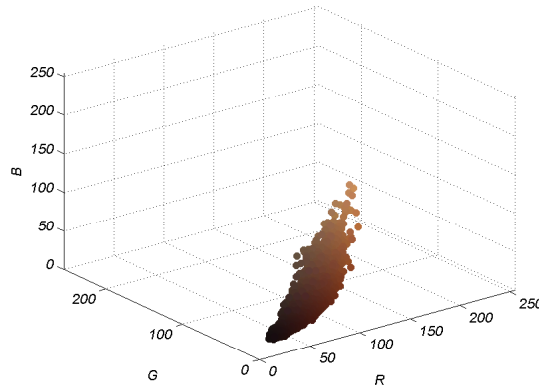


Fig. 7.3. RGB values for manually selected melanoma pixels.

After classifying all skin pixels using (7.3) in image I , a binary mask b_1 with melanoma pixels as foreground is first generated. A set of morphological operations are then applied on the mask b_1 to remove noisy pixels and fill the holes

within melanoma regions. Fig. 7.4 shows a MART-1 stained skin image, where the contours of foreground regions in mask b_1 (after morphological post-processings) are shown in blue. It is observed that all visually discernable melanoma areas have been correctly detected. Let us denote the boundaries of melanoma regions as $\mathcal{M} = \{m_i | i \in (1, 2, \dots, p)\}$, where p is the number of melanoma boundary points.

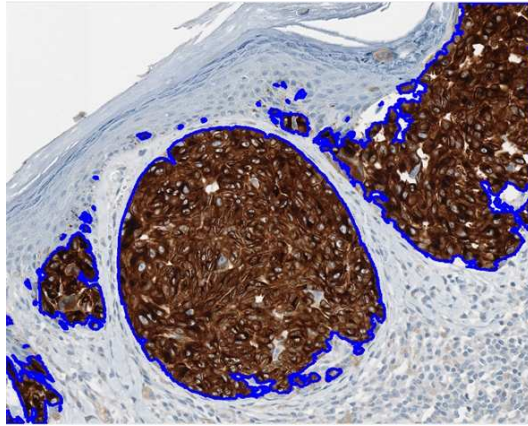


Fig. 7.4. Melanoma detection results (indicated by blue contours).

7.3.2 Epidermis Segmentation

After melanoma detection, the next step is to segment the skin epidermis. A few epidermis segmentation techniques in H&E stained skin histopathological images have been proposed in previous works [74], [113]. Note that the skin images used in this work are stained using MART-1, and the contrast between epidermis and dermis is not as strong as that in H&E stained images. Therefore, a similar technique with appropriate modifications for skin epidermis segmentation in MART-1 stained images is proposed. The principal steps for epidermis segmentations are as follows:

7.3.2.1 Pre-processing

In this step, the red channel R of the original RGB image I is used as input (just like in H&E stained images in [74], [113]). To suppress the unwanted components (e.g., noisy regions), the morphological closing operation with a disk-shape structuring

element (with a radius r_d of 15 pixels) is first applied on the image R . The histogram equalization [31] is then performed to increase the contrast between epidermis regions and other skin tissues.

7.3.2.2 Multi-thresholding

Except segmented melanoma areas (M), the image R mainly includes skin epidermis (E), dermis (D) and background (B) regions (see Fig. 7.5(a)). It is observed that background (white), dermis (gray) and epidermis (relatively darker) have different intensity levels. Based on this observation, we use two thresholds, $0 < t_1 < t_2 < 255$, for separating three classes: epidermis, dermis and background pixels. The thresholds t_1 and t_2 are automatically obtained using Otsu's method [51], which computes the optimal global thresholds by maximizing the inter-class variances. Since the epidermis area has the darkest color, a binary mask b_2 with epidermis as foreground is generated as follows:

$$b_2(u, v) = \begin{cases} 1 & \text{if } R(u, v) < t_1 \\ 0 & \text{otherwise} \end{cases} \quad (7.6)$$

where $R(u, v)$ indicates the pixel intensity at (u, v) .

7.3.2.3 Post-processing

The morphological reconstruction operation is first performed to fill the holes in the mask b_2 . We then label each connected region in the mask b_2 , and compute the area of each region. The largest foreground region in b_2 is finally selected as the epidermis region, as other small regions are generated from noisy regions in the dermis area. In Fig. 7.5(b) the segmented epidermis region has been shown with green pixels, and the segmented background and dermis regions have been shown with red and blue pixels, respectively. Fig. 7.5(c) shows the image R with segmented epidermis (indicated by red contours).

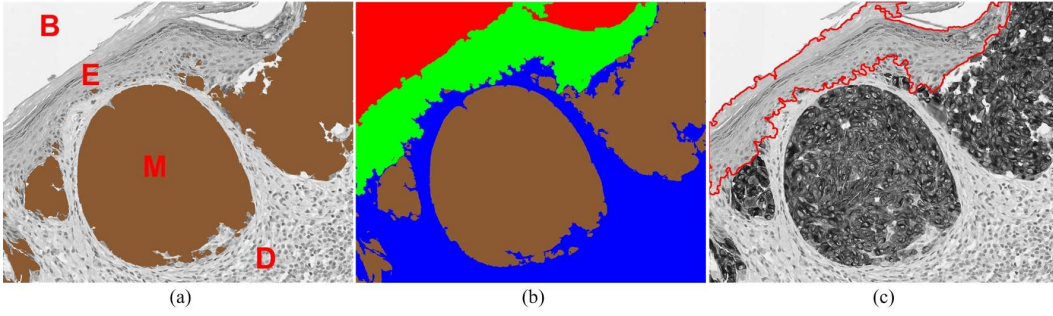


Fig. 7.5. Melanoma and epidermis segmentations. (a) R with segmented melanoma areas (brown regions). (b) Segmented epidermis (green), dermis (blue) and background (red). (c) R with segmented skin epidermis (red contours).

7.3.3 Granular Layer Identification

As observed in Fig. 7.5(c), the segmented skin epidermis includes skin cornified layer, granular layer and malpighian layer. Since the melanoma DoI is measured as the maximum distance between skin granular layer and the deepest tumor pixel in dermis region, in this module, we propose a supervised technique to detect the skin granular layer. The technique first classifies epidermis pixels into two groups: cornified layer pixels, and granular-malpighian (GM) layer pixels. The skin granular layer is then determined as the outer boundary of the skin GM layer. The details are listed as follows.

7.3.3.1 Histogram specification

Since there is a large number of cell nuclei (e.g., keratinocytes) in skin GM layer but no nuclei in skin cornified layer, the skin cornified layer has relatively lighter color (i.e., brighter) than the GM layer. Therefore, the gray-level intensity feature can be used to distinguish the cornified layer and GM layer. However, because there usually exist staining inconsistencies for different skin images, the performance of the supervised classification technique based on color features is highly sensitive to the training image. To address this problem, we perform histogram specification on the testing image R to match it the histogram of epidermis pixels in the training image R_t (the red channel of the training image I_t). Let the histograms of epidermis pixels in the training and testing images be denoted by h_{tr} and h_{te} , respectively.

The histogram specification is performed in three steps [31]. First, the cumulative distribution function (CDF) $S(\cdot)$ is computed from the testing image R as follows:

$$S(j) = \frac{1}{M} \sum_{i=0}^j h_{te}(j) \quad j = 0, 1, \dots, 255 \quad (7.7)$$

where M is the total number of epidermis pixels in image R . Next, the CDF $V(\cdot)$ is computed from the training image R_t as follows:

$$V(j) = \frac{1}{N} \sum_{i=0}^j h_{tr}(j) \quad j = 0, 1, \dots, 255 \quad (7.8)$$

where N is the total number of epidermis pixels in image R_t . Finally, for each gray level $x \in [0, 255]$, it is mapped to the gray level \tilde{x} after histogram specification, which is determined as follows [31]:

$$\tilde{x} = \arg \min_k |S(x) - V(k)| \quad k = 0, 1, \dots, 255 \quad (7.9)$$

Fig. 7.6(a) shows the manually segmented epidermis in the training image R_t , and Fig. 7.6(d) shows the histogram h_{tr} . Figs. 7.6 (b)(c) separately show the automatically segmented epidermis in R and \tilde{R} (after histogram specification), while Figs. 7.6(e)(f) separately show the corresponding histograms h_{te} before and after histogram specification. As observed from Fig. 7.6, the epidermis area in the image R (see Figs. 7.6 (b)(e)) presents brighter color than that in the image R_t (see Figs. 7.6 (a)(d)). However, after histogram specification, the grey-level appearance of epidermis area in \tilde{R} (see Figs. 7.6 (c)(f)) has looked like that in R_t . The histogram specification is applied to normalize the gray level intensities of epidermis pixels in the testing image to the same range of that in the training image, and hence the subsequent Bayesian classification based on gray level intensities is not sensitive to the R_t .

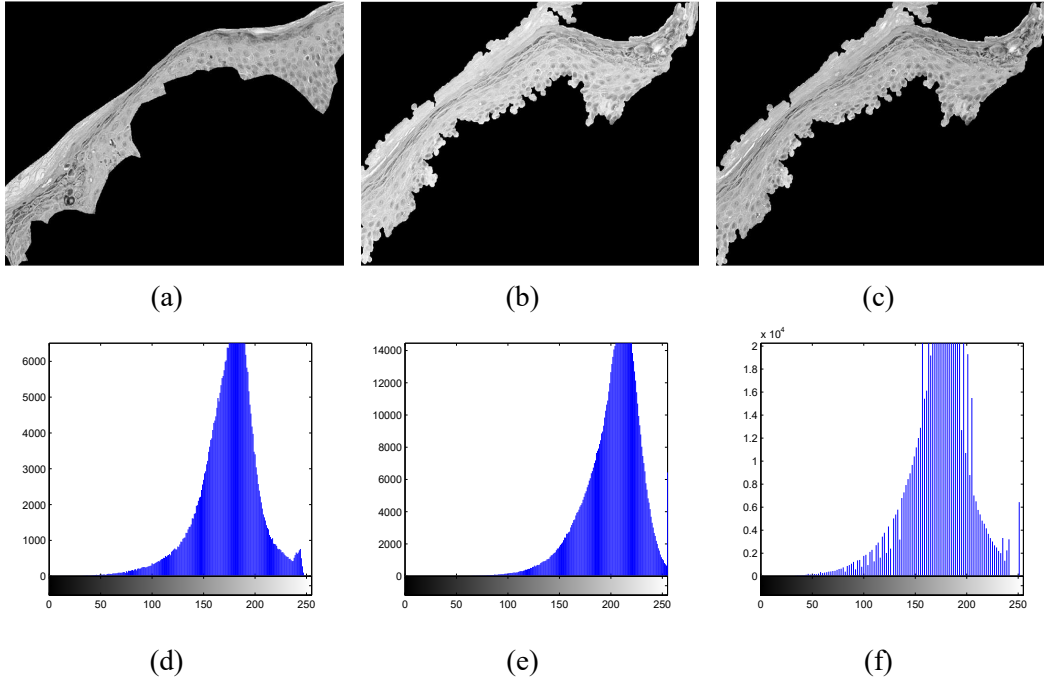


Fig. 7.6. Histogram specification for skin epidermis. (a) Manually segmented epidermis in R_t . (b) Originally segmented skin epidermis. (c) Skin epidermis after histogram specification. (d) Histogram of epidermis pixels shown in (a). (e) Histogram of epidermis pixels shown in (b). (f) Histogram of epidermis pixels shown in (c).

7.3.3.2 Low pass filtering

Although the skin cornified layer presents lighter appearance compared to other skin epidermis layers, there are noises and heterogeneities in the image. Therefore, it is not reliable to directly utilize the gray level intensities as features. For robustness, we compute the mean intensity within small neighborhoods centered at every image pixel as the feature to distinguish skin cornified layer and the GM layer. The mean intensity around the epidermis pixel $\tilde{R}(u, v)$ is computed as:

$$\bar{R}(u, v) = \frac{1}{(2z + 1)^2} \sum_{i=u-z}^{u+z} \sum_{j=v-z}^{v+z} \tilde{R}(i, j) \quad (7.10)$$

where $\bar{R}(u, v)$ is the mean intensity value in a $(2z + 1) \times (2z + 1)$ sliding window, and the parameter z (e.g., $z = 8$) determines the window size.

7.3.3.3 Bayesian classification

Let w_j , $j \in \{1, 2\}$, indicate the class label of epidermis pixels. Class w_1 corresponds to the cornified layer pixels (Class I), while class w_2 corresponds to the GM layer pixels (Class II). Using the mean intensity $\bar{R}(u, v)$ as the feature, the classification of epidermis pixels in \bar{R} is performed using the Bayesian theorem:

$$p(w_j | \bar{R}(u, v)) = \frac{p(\bar{R}(u, v) | w_j) p(u, v; w_j)}{\sum_{i=1}^2 p(\bar{R}(u, v) | w_i) p(u, v; w_i)} \quad (7.11)$$

where $p(w_j | \bar{R}(u, v))$ is the posterior probability of $\bar{R}(u, v)$ belonging to class w_j , $p(\bar{R}(u, v) | w_j)$ is the likelihood of observing feature value $\bar{R}(u, v)$ in class w_j , and $p(u, v; w_j)$ is the prior probability of coordinate (u, v) belonging to class w_j . Note that unlike existing works such as [50], [61] which use the same prior probability for the same class of objects by Bayesian classification, in this work, the prior probabilities adaptively change based on pixel coordinates.

To estimate the prior probability $p(u, v; w_j)$, we make use of a prior knowledge that the granular layer is the middle layer of the epidermis. In other words, if a pixel is closer to the image background, it is more likely to be in the cornified layer or malpighian layer. By utilizing this prior knowledge, we compute the $p(u, v; w_j)$ for different classes based on the inner distance transform [126] on mask b_2 (with segmented epidermis as foreground). Let $D(u, v)$ be the Euclidean distance between the foreground point (u, v) and the nearest background point (binary 0) in mask b_2 . The pixels belonging to class I generally have smaller $D(u, v)$ values, compared to the pixels belonging to class II. Thus the prior probabilities are computed as follows:

$$p(u, v; w_2) = D(u, v) / D_{\max} \quad (7.12)$$

$$p(u, v; w_1) = 1 - p(u, v; w_2) \quad (7.13)$$

where D_{\max} is the normalization term which is the maximum $D(u, v)$ value for all

(u, v) pixels in the epidermis.

To estimate the likelihood term $p(\bar{R}(u, v) | w_j)$, we manually segment cornified layer pixels (20,166 pixels) and GM layer pixels (76,767 pixels) from the training image R_t (shown in Fig. 7.6(a)). The mean intensity (e.g., $\bar{R}_t(u, v)$) within each training pixel's neighborhood is computed by Eq. (7.10). The probability density functions (PDFs) for the Class I and Class II pixels are separately shown in Figs. 7.7(a)(b). In order to reduce the local irregularities and create a smooth, continuous distribution, the PDFs can be modeled by many different types of continuous distributions [127]. To select the best fitted model for our training pixels, we have tested with 10 commonly-used parametric probability distributions such as Normal, Gamma, and Generalized extreme value (Gev) distributions. Table 7.1 lists the Bayesian information criterion (BIC) [128] values for 10 fitted models. The BIC is a criterion for models selection, and the model with the lowest BIC value is preferred. In Figs. 7.7(a)(b), the overlapped continuous curves separately indicate the top four preferred models for class I and II pixels. As observed from Table 7.1 and Figs. 7.7(a)(b), the Gev model achieves the smallest BIC values for two classes, and hence the Gev model is applied to estimate the likelihood term, which is as follows:

$$p(x | w_j) \approx \frac{1}{\sigma} \exp\left(-(\Delta)^{-\frac{1}{k}}\right) (\Delta)^{-1-\frac{1}{k}} \quad (7.14)$$

where $\Delta = 1 + k \frac{(x-\mu)}{\sigma}$, μ , σ and k are the location, scale and shape parameters of the Gev model. In Fig. 7.7(c), the two dashed curves indicate the PDFs of training pixels for two classes, while the two solid curves indicate the estimated likelihood terms for two classes.

With the estimated likelihoods and prior probabilities, the epidermis pixel $\bar{R}(u, v)$ is classified using Bayesian theorem as follows:

$$\bar{R}(u, v) = \begin{cases} I & \text{if } p(w_1 | \bar{R}(u, v)) > p(w_2 | \bar{R}(u, v)) \\ II & \text{otherwise} \end{cases} \quad (7.15)$$

TABLE 7.1
BAYESIAN INFORMATION CRITERION (BIC) VALUES FOR DIFFERENT MODELS.

Class I	Normal	Logistic	Nakagami	Gamma	Gev
	159860	160550	159840	159850	159800
Class II	Rician	Weibull	Extreme value	Rayleigh	t Location-Scale
	159860	162030	163110	226170	159870
Class II	Normal	Logistic	Nakagami	Gamma	Gev
	628350	627890	630360	632510	621510
Class II	Rician	Weibull	Extreme value	Rayleigh	t Location-Scale
	628370	622060	622880	827620	627310

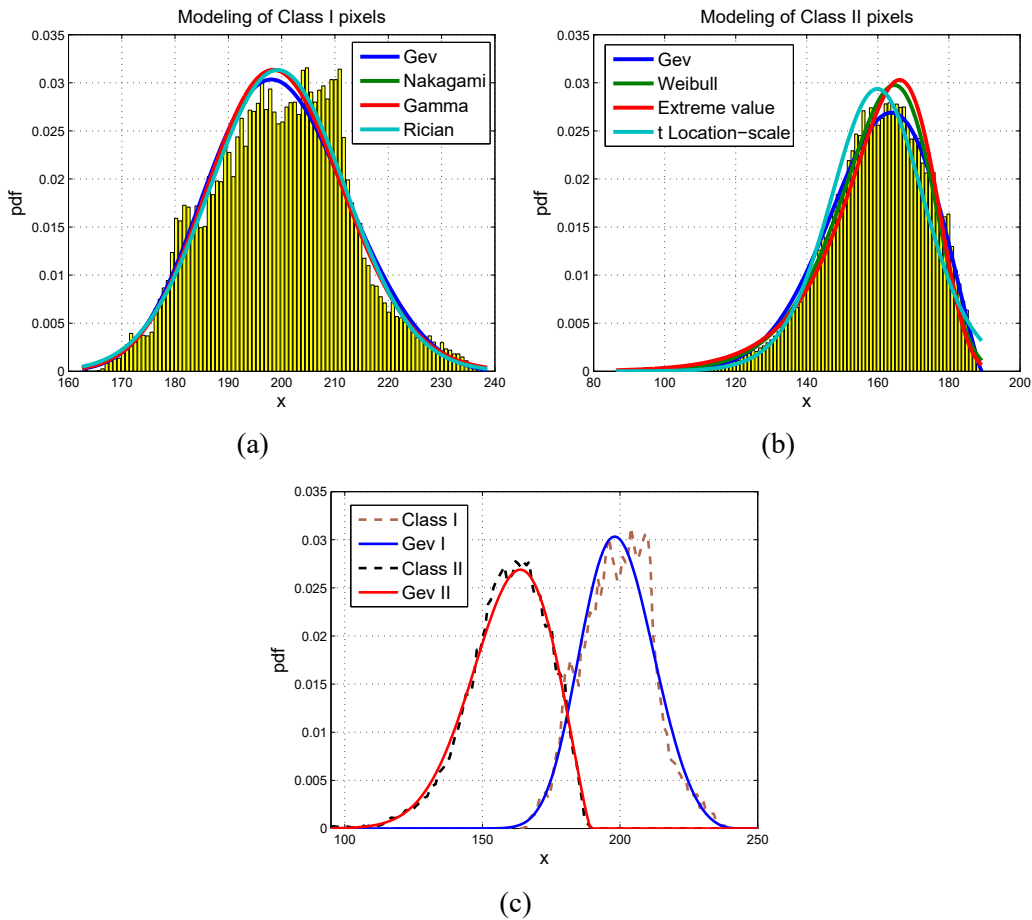


Fig. 7.7. PDFs estimated from empirical training data. (a) Class I (cornified layer pixels). (b) Class II (GM layer pixels). (c) Gev fitted curves. Note that Class I and Class II plots in (c) correspond to the original pdfs, and the parameters for fitted models Gev I and Gev II are $(\mu = 194.57, \sigma = 12.55, k = -0.25)$ and $(\mu = 155.44, \sigma = 15.53, k = -0.46)$, respectively.

After classification of all epidermis pixels, we generate a binary mask b_3 with GM layer pixels as foreground. The small noisy regions (if any) in mask b_3 are removed

by morphological opening operation, and the holes within the foreground are filled by morphological reconstruction operation. Fig. 7.8(a) shows the image R , where the contours of foreground regions in mask b_3 are shown in red. By comparing Fig. 7.5(c) and Fig. 7.8(a), it can be seen that most skin cornified layer pixels have been correctly removed from the GM layer.

7.3.3.4 Skin granular layer detection

In order to find the outer boundary of the granular layer from the mask b_3 , we first detect the extrema points [117] of the foreground region in b_3 . In Fig. 7.8(b), the detected extrema points are indicated by (blue and green) squares. The extrema points that are close to dermis pixels are then removed. Note that the dermis pixels have intensity values between the thresholds t_1 and t_2 which have been determined during epidermis segmentation. In Fig. 7.8(b), blue squares are removed as they are close to dermis pixels. The boundary pixels between the two green squares in Fig. 7.8(b) are now selected as the outer boundary of granular layer. In order to remove the irregularities, a moving average filter [113] with a predefined length s_t (e.g., $s_t=101$ pixels) is finally applied to generate a smooth boundary of the granular layer. Fig. 7.8(c) shows the identified granular layer with two extrema points (green squares). Let us denote the detected granular layer as $\mathcal{G} = \{g_j | j \in (1, 2, \dots, q)\}$, where q is the number of granular layer points.

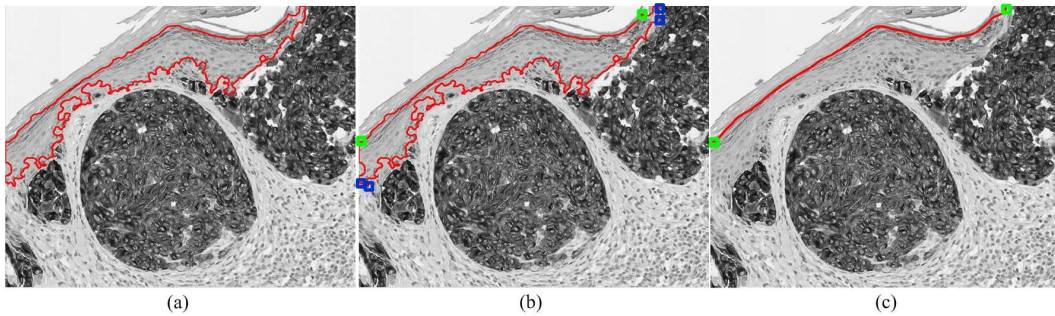


Fig. 7.8. Skin granular layer identification. (a) Skin GM layer (area enclosed by red contour). (b) Extrema points (indicated by blue and green squares). (c) Skin granular layer with two extrema points.

7.3.4 Depth of Invasion Measurement

The melanoma boundary points $\mathcal{M} = \{m_i | i \in (1, 2, \dots, p)\}$ were obtained in melanoma detection module and the granular layer points $\mathcal{G} = \{g_j | j \in (1, 2, \dots, q)\}$ were obtained in granular layer identification module. Based on these two point sets, the melanoma DoI can now be measured. Although it is possible to evaluate the invasion depth for each melanoma boundary point, it is usually computationally intensive due to the large number of melanoma boundary points. To speed up the computation, we apply a multi-resolution approach where we first determine a coarse set of \mathcal{M} to evaluate the invasion depth and then measure the DoI with higher precision. The steps for invasion depth measurement are as follows:

- 1) The points set \mathcal{M} is down-sampled with a factor d_f , that is:

$$\mathcal{M}_d = \{m_k \in \mathcal{M} | k = d_f \times i, i = (1, 2, \dots, \lfloor p/d_f \rfloor)\} \quad (7.16)$$

where the d_f is empirically set as 10 in this work, and \mathcal{M}_d is the point set after down-sampling.

- 2) The point index \tilde{k} corresponding to the deepest tumor pixel in the set \mathcal{M}_d is determined based on the Hausdorff distance measure as follows:

$$\tilde{k} = \arg \max_k \left[\min_j \|m_k - g_j\| \right], m_k \in \mathcal{M}_d \quad (7.17)$$

where $\|\cdot\|$ is the 2D Euclidean distance between two point coordinates. Note that \tilde{k} denotes the index of the maximal invasion depth for the decimated melanoma boundary pixels.

- 3) To obtain a more precise invasion measurement, we now calculate the DoI for all pixels in the neighborhood of index \tilde{k} in \mathcal{M} , that is:

$$DoI = \max_i \left[\min_j \|m_i - g_j\| \right], \tilde{k} - d_f < i < \tilde{k} + d_f \quad (7.18)$$

Our melanoma invasion measurement follows a coarse-to-fine procedure, which

greatly reduces the computational complexity. For example, with $d_f = 10$, the computational complexity reduces from $O(pq)$ to $O(pq/10 + 20q)$. In other words, the multi-resolution approach reduces the computational complexity by a factor of d_f (assuming $pq \gg 1$). Fig. 7.9 shows the automatic measurement of melanoma DoI, which is 1661.2 pixels (around $0.415mm$).

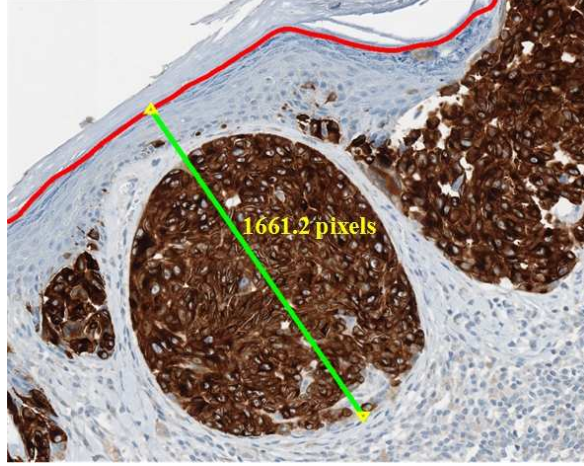


Fig. 7.9. Automatic measurement of melanoma DoI. In this example, the melanoma DoI is measured as 1661.2 pixels (around $0.415mm$).

7.4 Performance Evaluation

To evaluate the performance of the proposed technique, we carried out a series of experiments on our database. The main objectives of the evaluation are: 1) to evaluate if the automatically detected skin granular layer is consistent with the ground truth, 2) to evaluate if the automatically measured melanoma DoI is consistent with the ground truth. In this section, we present the evaluation results of the proposed technique and compare it with two existing techniques.

7.4.1 Parameters Configuration

There are mainly four adjustable parameters in the proposed technique. Table 7.2 lists the adjustable parameters and their corresponding values used in this study. Note that our testing images were captured under 40X magnification and have a

pixel resolution of $0.25\mu m/\text{pixel}$. These parameter values should adaptively change with image resolutions.

TABLE 7.2
SUMMARY OF PARAMETERS CONFIGURATION IN THE PROPOSED TECHNIQUE.

Modules	Parameter names	Parameter values
Epidermis Segmentation	Radius r_d	15 pixels
Granular Layer Detection	Window size z	8 pixels
Granular Layer Detection	Smoothing length s_t	101 pixels
DoI Measurement	Down-sampling factor d_f	10

To the best of our knowledge, there are only two existing works including [21] and [19] that try to measure the tumor DoI. In [21], the epidermis is segmented based on morphological closing operation and Otsu’s thresholding. The malignant melanocytes are detected using statistical color features, and the melanoma invasion is measured based on the estimation of epidermis orientation by Hough transform. The radius of the structuring element for closing operation is the key parameter for this technique, which is set as 15 pixels like the proposed technique. In [19], the skin cornified layer is detected based on the observation that skin cornified layers are typically ridge shaped regions. After segmenting the skin cornified layer, the boundaries of remaining epidermis regions, which are not close to skin dermis, are considered as skin granular layer. For this technique, three parameters for skin cornified layer detection should be carefully determined, which includes the block size b_s for partitioning the image, the threshold T_λ to determine if a pixel is on the ridge, and the threshold T_{ent} for pixel orientation entropy. After experiments on our database, we set them as: $b_s = 32$, $T_\lambda = 0.2$ and $T_{ent} = 4.0$.

7.4.2 Granular Layer Detection Performance

Let us denote the manually labeled skin granular layer as $\mathcal{L} = \{l_k | k \in (1, 2, \dots, r)\}$, where r is the number of manually labeled boundary points. To evaluate the granular layer detection performance, we calculate the distance of every point in \mathcal{G} (automatically obtained boundary points) from all points in \mathcal{L} . The evaluation

metrics are defined as follows:

$$\mathcal{D}_{MAD} = \frac{1}{q} \sum_{j=1}^q \left[\min_k \|l_k - g_j\| \right] \quad (7.19)$$

$$\mathcal{D}_{HD} = \max_j \left[\min_k \|l_k - g_j\| \right] \quad (7.20)$$

where the mean absolute distance \mathcal{D}_{MAD} measures the average disagreement between the two boundaries, and the Hausdorff distance \mathcal{D}_{HD} measures the worst possible disagreement between the two boundaries.

Table 7.3 shows the average evaluation results over 28 test images for different techniques. The \mathcal{D}_{MAD} and \mathcal{D}_{HD} have been expressed in microns by multiplying the values obtained by Eq. (7.19) and Eq. (7.20) with pixel resolution ($0.25\mu m/\text{pixel}$). Note that the technique in [21] does not explicitly detect the skin granular layer, and hence the outer boundary of segmented skin epidermis has been considered as skin granular layer for comparison. As observed in Table 7.3, the proposed technique achieves the smallest \mathcal{D}_{MAD} and \mathcal{D}_{HD} values ($2.95\mu m$ and $12.24\mu m$, respectively) among three techniques. The technique in [21] provides the poorest performance, with $9.59\mu m$ of \mathcal{D}_{MAD} and $33.65\mu m$ of \mathcal{D}_{HD} . This is mainly because the skin cornified layer is generally included in the segmented skin epidermis, and thus the outer boundary of skin cornified layer is incorrectly considered as skin granular layer. The technique in [19] provides a better performance than that in [21], with $7.22\mu m$ of \mathcal{D}_{MAD} and $27.03\mu m$ of \mathcal{D}_{HD} . This technique tries to segment the skin cornified layer before the skin granular layer detection. However, the skin cornified layer is segmented based on the assumption that cornified layer is the ridged shape region. In fact the skin granular layer is also the ridged shape region, which is likely to be segmented out with skin cornified layer in [21].

Fig. 7.10 shows the visual comparison of skin granular detection by different techniques. Figs. 7.10(a)(b) separately show the detected skin granular layers by [21] and [19], while Fig. 7.10(c) shows the detected skin granular layer by the proposed technique. In Fig. 7.10 the green curves indicate the manually labeled ground truths, and the red curves indicate the automatically detected results. As

TABLE 7.3
COMPARISON OF GRANULAR LAYER DETECTION BY DIFFERENT TECHNIQUES.

Techniques	\mathcal{D}_{MAD} (microns)	\mathcal{D}_{HD} (microns)
Mokhtari <i>et al.</i> [21]	9.59	33.65
Noroozi <i>et al.</i> [19]	7.22	27.03
Proposed	2.95	12.24

seen in Fig. 7.10, the skin granular layer obtained by the proposed technique has a good matching with the ground truth. By contrast, the existing techniques tend to produce unsmooth and inaccurate skin granular layers.

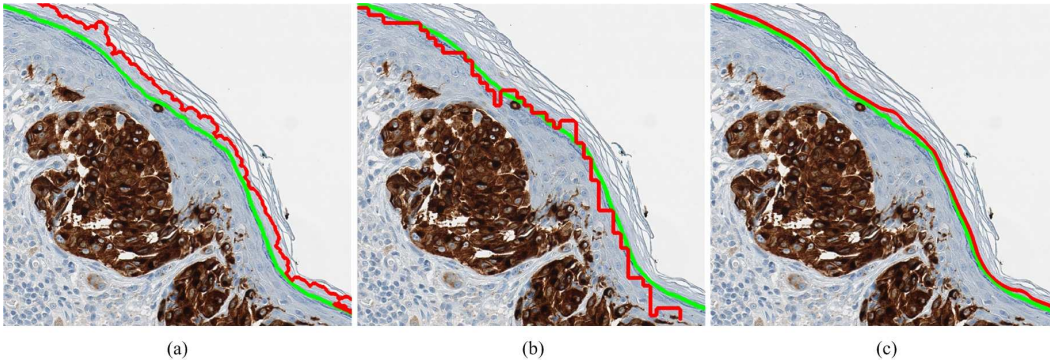


Fig. 7.10. Skin granular layer detection. (a) Mokhtari *et al.* [21]. (b) Noroozi *et al.* [19]. (c) Proposed technique. Note that green curves indicate the manually labeled ground truths, while red curves indicate the automatically obtained results.

7.4.3 Melanoma Depth Measurement Performance

Let d_i and h_i denote the automatically measured invasion depth and manually measured ground truth, respectively, where $i \in (1, 2, \dots, s)$ and s ($s=28$ in this work) is the number of testing images. The performance of melanoma invasion measurement is evaluated with respect to the average error (AE), standard deviation (SD) and average percentage of error (APE) which are defined as follows:

$$AE = \frac{1}{s} \sum_{i=1}^s |d_i - h_i| \quad (7.21)$$

$$SD = \sqrt{\frac{1}{s} \sum_{i=1}^s (|d_i - h_i| - AE)^2} \quad (7.22)$$

$$APE = \left(\frac{1}{s} \sum_{i=1}^s \frac{|d_i - h_i|}{h_i} \right) \times 100\% \quad (7.23)$$

where $|\cdot|$ is the absolute distance between two values.

The performance of the proposed technique has been compared with that of [21] and [19]. Table 7.4 shows the evaluation results of different techniques in terms of AE , SD and APE . It is observed in Table 7.4 that the proposed technique provides the best performance in melanoma DoI measurement, with $10.95\mu m$ average error, $17.49\mu m$ standard deviation and 3.53% relative error compared to ground truths. The technique in [19] has relatively poorer performance than our technique, which has $20.54\mu m$ average error and 6.81% relative error. The technique in [21] has the poorest performance with $28.03\mu m$ average error and 10.66% relative error. The inaccuracies of melanoma DoI measurement for existing techniques are mainly caused by inaccurate skin granular detections. Fig. 7.11 compares the automatically measured melanoma DoI by different techniques with ground truths in 28 testing images. Considering the visual assessment of samples as the reference, the percentage of error achieved by different computerized techniques on 28 testing images is shown in Fig. 7.12. As observed in Fig. 7.11 and Fig. 7.12, the technique in [21] tends to produce positive errors (i.e., $d_i > h_i$), as the melanoma DoI measurement is usually performed from the skin cornified layer instead of skin granular layer. The technique in [19] is likely to produce negative errors (i.e., $d_i < h_i$). This is because the skin granular layer is usually segmented out with the skin cornified layer, and the melanoma DoI measurement is usually from the outer boundary of skin malpighian layer. Overall, the proposed technique provides a more reliable melanoma invasion measurement due to accurate detections of melanoma regions and skin granular layer.

For visual comparison, Fig. 7.13 shows a testing image with measured melanoma DoI by different automatic techniques and visual assessment. Fig. 7.13(a) shows the tumor invasion measurement by [21]. Note that the blue straight line in Fig. 7.13(a) indicates the estimated epidermis orientation by Hough transform. The tumor in-

TABLE 7.4
COMPARISON OF DOI MEASUREMENT BY DIFFERENT TECHNIQUES.

Techniques	AE (microns)	SD (microns)	APE (%)
Mokhtari <i>et al.</i> [21]	28.03	29.70	10.66
Noroozi <i>et al.</i> [19]	20.54	17.21	6.81
Proposed	10.95	17.49	3.53

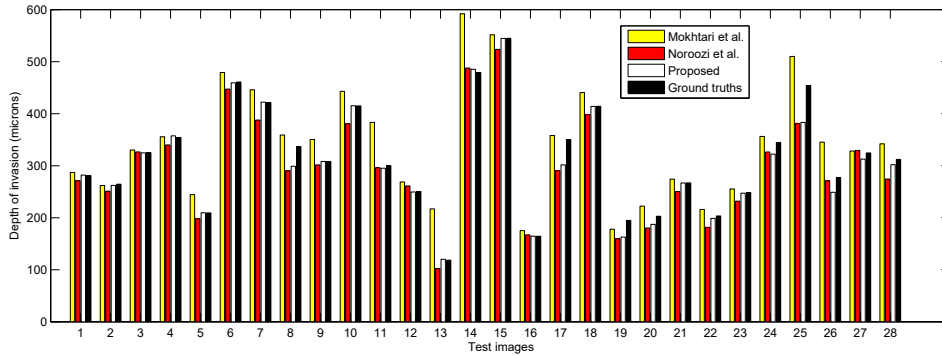


Fig. 7.11. Comparison of automatically measured melanoma DoI and ground truths.

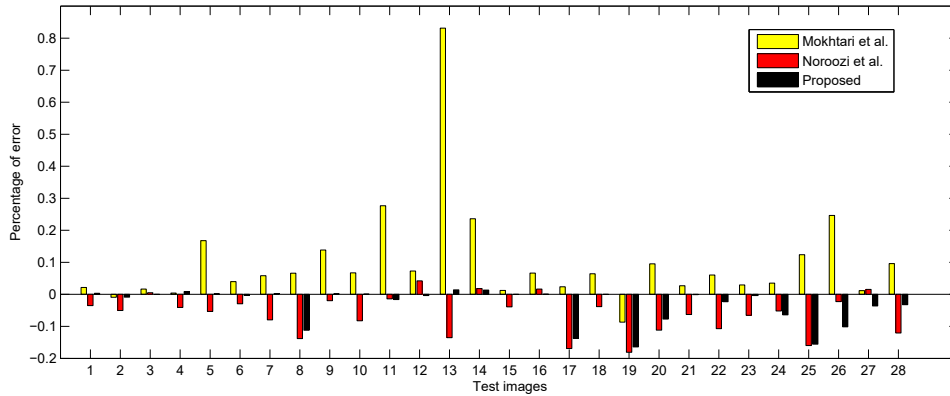


Fig. 7.12. Normalized errors of computerized methods with ground truths as the gold standard.

vasion measurement is performed along the direction that is perpendicular to the epidermis orientation (e.g. blue line). As has been verified in [19], Hough transform may fail to detect epidermis orientation when the epidermis is not a roughly straight structure, and hence it produce an inaccurate tumor invasion measurement. Fig. 7.13(b) shows the tumor invasion measurement by [19]. This technique has a main limitation that skin granular layer is usually segmented as skin cornified

layer, as both layers usually have ridged shapes in skin images. The tumor invasion measurement from the outer boundary of skin malpighian layer generally produces a depth less than the ground truth. Figs. 7.13(c)(d) show the melanoma invasion measurement by the proposed technique and visual assessment, respectively. As seen in Fig. 7.13, the proposed technique provides the smallest error in measuring the melanoma DoI.

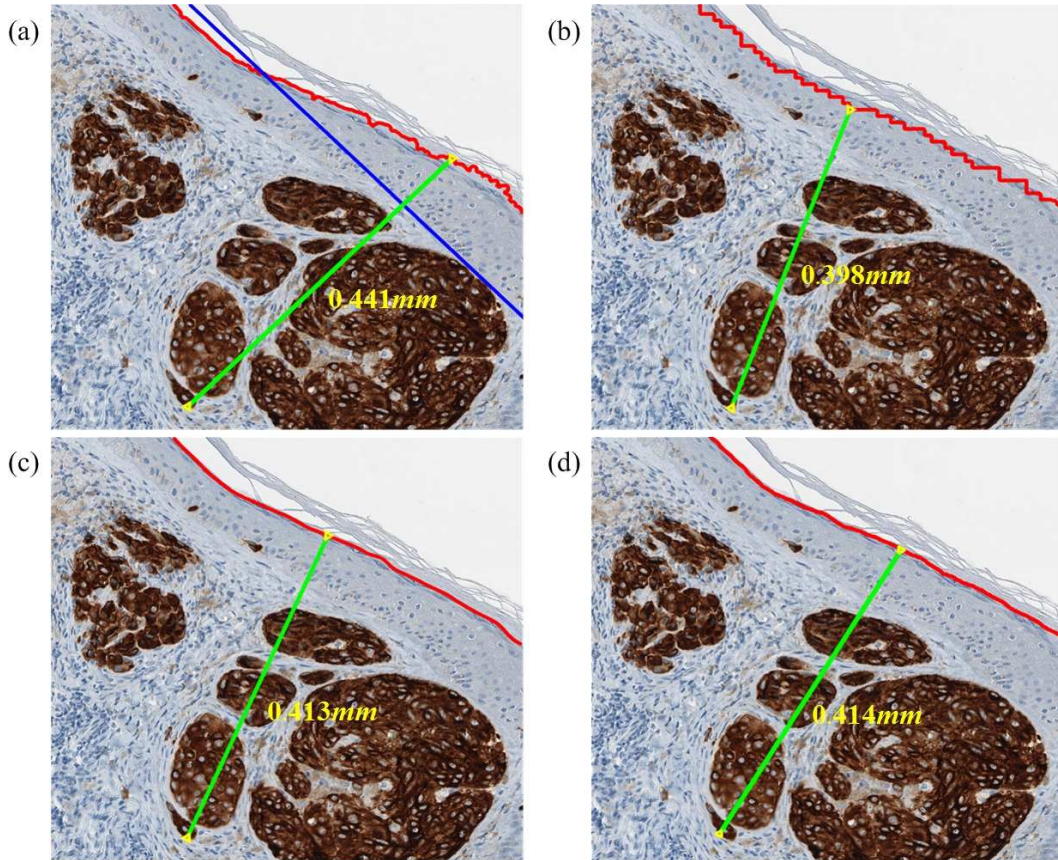


Fig. 7.13. Visual comparison of melanoma invasion measurement. (a) Mokhtari *et al.* [21]. (b) Noroozi *et al.* [19]. (c) Proposed technique. (d) Ground truth. Note that the blue straight line in (a) indicates the estimated epidermis orientation by Hough transform.

7.4.4 Parameter Sensitivity Evaluation

As listed in Table 7.2, the proposed technique has four adjustable parameters. Among these parameters, the parameters r_d , s_t and d_f can be intuitively estimated according to the resolution of testing images. The parameter z determines the

window size to compute the average intensity for Bayesian classification, which is empirically set as 8 pixels for our 40X magnification skin images. To evaluate the sensitivity of this parameter, we also test z with 7 and 9 pixels, which correspond to the window size of 15×15 and 19×19 , respectively.

TABLE 7.5
PARAMETER SENSITIVITY EVALUATIONS OF THE PROPOSED TECHNIQUE.

Parameters	Granular layer detection		Melanoma invasion measurement		
	$\mathcal{D}_{MAD} (\mu m)$	$\mathcal{D}_{HD}(\mu m)$	$AE(\mu m)$	$SD(\mu m)$	$APE(\%)$
$z=7$	3.02	12.72	10.93	17.48	3.53
$z=8$	2.95	12.24	10.95	17.49	3.53
$z=9$	2.94	12.29	10.94	17.48	3.53

Table 7.5 shows the experimental results with different z values in terms of skin granular detection and melanoma invasion measurement. It is observed in Table 7.5 that the performance of the proposed technique is not very sensitive to the minor changes of z values. In particular, the average percentage of error (APE) for three z values are the same (3.53%). Therefore, it can be concluded that the proposed technique is robust to the parameter z when an appropriate value is selected.

7.5 Summary

This chapter presents an automatic technique for measuring the melanoma DoI in MART-1 stained skin histopathological images. The proposed technique detects melanoma regions using color features with Mahalanobis distance measure, and identifies skin granular layer based on a Bayesian classification method. The melanoma DoI is computed by using the Hausdorff distance measure. Experimental results show that the proposed technique can more robustly detect skin granular layer and measure melanoma DoI than two closely related existing techniques.

Chapter 8

Conclusions and Future Works

Over the last few years, computerized image analysis and diagnosis have been playing an increasingly important role in assisting pathologists for their clinical routines. However, because of great variation in different tissue biopsies, much more efforts are still required to be devoted to develop robust computer-aided techniques for biopsy image analysis and cancer diagnosis. This thesis has developed several key techniques for quantitative analysis of skin biopsy images. In this chapter, we conclude the works presented in this thesis and discuss some future research directions.

8.1 Conclusions

In Chapter 1, we mainly introduced histopathology imaging and histological features used by pathologists for skin melanoma diagnosis and grading. In Chapter 2, we briefly reviewed the widely-used image processing techniques in histopathology and existing WSIs analysis systems for different cancers. The main contributions of this thesis were presented in Chapters 3 to 7. In the following, we summarize the contributions of this thesis.

Chapter 3 presented the proposed techniques for automated epidermis and dermis segmentations in H&E stained skin WSIs. The epidermis segmentation follows a coarse-to-fine processing procedure, where the segmentation is performed by

thresholding and k-means clustering and the segmentation quality is evaluated by thickness measurement. The dermis segmentation is performed based on diffusing line segments from epidermis outer boundaries and morphological operations. Experimental results show that the proposed techniques achieve a good performance (e.g., with over 90% sensitivities and precisions) in epidermis and dermis segmentations.

Chapter 4 presented two proposed techniques for nuclei detection in skin biopsy images. In the first technique (i.e., P-Voting), the isolated nuclei are detected by ellipse descriptor analysis on binarized images, while the clumped nuclei are detected by a voting algorithm with votes from nuclei boundaries. In the second technique (i.e., P-gLoG), the nuclei are detected by a bank of multi-scale gLoG kernels with redundant seeds merged by mean-shift clustering. Experimental results on H&E stained biopsy images show that the proposed techniques achieve higher precision and sensitivity than existing techniques for nuclei detection. Between the two proposed techniques, the P-gLoG technique provides a superior performance, and also has relatively lower complexity.

Chapter 5 presented the proposed technique for nuclei segmentation in skin biopsy images. The proposed technique first generates several candidate nuclei contours by analyzing gradient information on multi-scale radial lines. The intensity variance and shape prior are then incorporated with gradient information to determine the best contour for each nucleus in the image. Severely overlapped nuclei are finally merged together based on the Dice coefficient analysis. Experiments on Ki-67 and H&E stained skin images indicate that the proposed technique is efficient and accurate in delineating nuclei boundaries. In particular, the proposed technique has the advantage of estimating occluded nuclei boundaries in the clustered nuclei.

Chapter 6 presented an automated melanoma diagnosis technique that diagnoses melanoma by analyzing H&E stained skin WSIs. The technique first segments skin epidermis and dermis regions by a multi-resolution framework, and generates high resolution image tiles. Epidermis and dermis analysis are then performed in a parallel manner, where a set of 73 cytological, textural and architectural features are

computed from segmented epidermis and dermis regions. The computed features are finally utilized by a SVM classification method that categorizes the test skin WSI into different categories such as melanoma, nevus or normal tissue. Experiments on 66 H&E stained skin WSIs show that the proposed technique with feature selection can provide over 95% classification accuracies, which has the potential to be used for assisting pathologists in melanoma diagnosis.

Chapter 7 presented the proposed technique for measuring melanoma DoI in MART-1 stained skin biopsy images. The proposed technique detects melanoma regions based on RGB color features with a Mahalanobis distance measure. The skin granular layer is identified by a Bayesian classification based method, and the melanoma DoI is measured by using a Hausdorff distance measure. Experiments show that the proposed technique can more accurately identify skin granular layer and measure melanoma DoI than existing techniques.

8.2 Future Works

Although this thesis has developed several key techniques for computerized analysis of skin biopsy images, there are still many research works that should be done to assist pathologists in real clinical settings. A few selected future research directions are presented in the following.

Speedup & Statistic Testing

Chapter 6 presents the proposed technique for diagnosing skin melanomas from H&E stained skin WSIs. The technique makes the judgement (e.g., melanoma or non-melanoma) by analyzing both skin epidermis and dermis features, and it has been designed and tested with sequential implementations using Matlab software. Due to the large number of cell nuclei in skin WSIs, the technique has a high computational complexity which may take one to two hours (depending on biopsy image size and computer hardwares) to process a skin WSI. However, epidermis and dermis analysis can be performed in a parallel manner in the proposed framework,

and nuclei segmentations are also suitable for parallel computations. Therefore, parallelized design and implementation by using such as Graphic Processing Unit (GPU) [43] can be done for faster processing and diagnosis.

Another open problem is to test the technique (presented in Chapter 6) on more skin WSIs, especially for identifying melanoma subtypes. In this study, the proposed technique has been evaluated on 66 different skin WSIs which includes 32 melanoma biopsy images. This evaluation sample size is still quite small, and hence more evaluations with statistic tests (e.g., Wilcoxon test [21]) will be necessary to verify the robustness of the technique on skin melanoma diagnosis. In addition, since the technique is currently only implemented in Matlab for testing, it would be useful to design an user-friendly interface or integrate the technique into some medical image processing softwares such that it can be tested or used in a more practical manner by pathologists.

Lymph Node Analysis

If melanoma is found on skin and also has invaded into skin dermis with a large thickness (e.g., more than $1mm$), it would be necessary to exam the biopsy of nearby (sentinel) lymph nodes to determine whether melanoma cells have spread through the body's lymphatic system. The lymph node evaluation for melanoma diagnosis is mainly performed manually by pathologists. Computerized algorithms are desired to analyze lymph node biopsy images and help doctors in grading the severity of the disease. One efficient way to determine if melanoma has metastasized to the local lymph nodes is to analyze the MART-1 stained biopsy images. Since MART-1 stain is specific to melanoma cells and produces distinct colors, it would be easy to detect melanoma regions based on image color features using computer algorithms.

Proliferation Index Evaluation

Once the melanoma is detected in lymph nodes, it would be critical to evaluate the proliferation index of melanoma cells within lymph nodes. High proliferative index (e.g., 20% or more of tumor cells showing nuclear immunoreactivity) is

significantly correlated with increased patient mortality [11]. In practice pathologists mainly evaluate proliferation index by analyzing Ki-67 stained biopsy images, where tumor cell nuclei with positive responses are stained with dark brown colors. Since there is a large number of tumor cells to be analyzed, manual evaluations of proliferation index are very subjective and often prone to inter- and intra-observer variations. In order to obtain more objective and reliable proliferation index evaluation, it is desirable to develop robust computer algorithms which could automatically detect melanoma regions and compute the number of tumor cells with positive responses.

References

- [1] A. Kokolakis, G. Zacharakis, K. Krasagakis, K. Lasithiotakis, R. Favicchio, G. Spiliopoulos, E. Giannikaki, J. Ripoll, and A. Tosca, “Prehistological evaluation of benign and malignant pigmented skin lesions with optical computed tomography,” *Journal of Biomedical Optics*, vol. 17, no. 6, pp. 0 660 041–0 660 049, 2012.
- [2] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, “A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI,” *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2007–2016, 2011.
- [3] I. Maglogiannis and C. Doukas, “Overview of advanced computer vision systems for skin lesions characterization,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 721–733, 2009.
- [4] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2015,” *CA: A Cancer Journal for Clinicians*, vol. 65, no. 1, pp. 5–29, 2015.
- [5] N. V. Orlov, A. T. Weeraratna, S. M. Hewitt, C. E. Coletta, J. D. Delaney, D. Mark Eckley, L. Shamir, and I. G. Goldberg, “Automatic detection of melanoma progression by histological analysis of secondary sites,” *Cytometry Part A*, vol. 81, no. 5, pp. 364–373, 2012.
- [6] R. Rubin, D. S. Strayer, E. Rubin, and M. J. M., *Rubin’s pathology: clinicopathologic foundations of medicine*. Lippincott Williams & Wilkins, 2008.

- [7] F. Dick, S. VanLier, P. Banks, G. Frizzera, G. Witrak, R. Gibson, G. Everett, L. Schuman, P. Isacson, G. O’Conor, C. Ken, B. William, and B. Aaron, “Use of the working formulation for non-hodgkin’s lymphoma in epidemiologic studies: agreement between reported diagnoses and a panel of experienced pathologists,” *Journal of the National Cancer Institute*, vol. 78, no. 6, pp. 1137–1144, 1987.
- [8] I. Wikimedia Foundation, “Wikipedia, the free encyclopedia,” Website, June 2015. [Online]. Available: <https://en.wikipedia.org/?title=Histology>
- [9] D. E. Chandler and R. W. Roberson, *Bioimaging: current concepts in light and electron microscopy*. Jones & Bartlett Publishers, 2009.
- [10] I. Wikimedia Foundation, “Wikipedia, the free encyclopedia,” Website, June 2015. [Online]. Available: <https://en.wikipedia.org/wiki/Histopathology>
- [11] C. Hazan, K. Melzer, K. S. Panageas, E. Li, H. Kamino, A. Kopf, C. Cordon-Cardo, I. Osman, and D. Polsky, “Evaluation of the proliferation marker mib-1 in the prognosis of cutaneous malignant melanoma,” *Cancer*, vol. 95, no. 3, pp. 634–640, 2002.
- [12] R. S. Weinstein, A. R. Graham, L. C. Richter, G. P. Barker, E. A. Krupinski, A. M. Lopez, K. A. Erps, A. K. Bhattacharyya, Y. Yagi, and J. R. Gilbertson, “Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future,” *Human pathology*, vol. 40, no. 8, p. 1057, 2009.
- [13] R. Rocha, J. Vassallo, F. Soares, K. Miller, and H. Gobbi, “Digital slides: present status of a tool for consultation, teaching, and quality control in pathology,” *Pathology-Research and Practice*, vol. 205, no. 11, pp. 735–741, 2009.
- [14] A. H. Olson, “Image analysis using the aperio scanscope,” *Technical manual*. Aperio Technologies Inc, 2006.

- [15] L. U. Chicago, "Introduction to the skin," Website, 1996. [Online]. Available: <http://www.meddean.luc.edu/lumen/MedEd/medicine/>
- [16] N. V. Online, "Skin anatomy," Website, June 2015. [Online]. Available: <http://www.cancer.gov>
- [17] B. R. Smoller, "Histologic criteria for diagnosing primary cutaneous malignant melanoma," *Modern Pathology*, vol. 19, pp. S34–40, 2006.
- [18] G. Massi and P. E. LeBoit, *Histological Diagnosis of Nevi and Melanoma*. Springer Science & Business Media, 2013.
- [19] N. Noroozi and A. Zakerolhosseini, "Computerized measurement of melanocytic tumor depth in skin histopathological images," *Micron*, vol. 77, pp. 44–56, 2015.
- [20] D. Roberts, A. Anstey, R. Barlow, N. Cox, J. Bishop, P. Corrie, J. Evans, M. Gore, P. Hall, and N. Kirkham, "Uk guidelines for the management of cutaneous melanoma," *British Journal of Dermatology*, vol. 146, no. 1, pp. 7–17, 2002.
- [21] M. Mokhtari, M. Rezaeian, S. Gharibzadeh, and V. Malekian, "Computer aided measurement of melanoma depth of invasion in microscopic images," *Micron*, vol. 61, pp. 40–48, 2014.
- [22] C. Lu, M. Mahmood, N. Jha, and M. Mandal, "Automated segmentation of the melanocytes in skin histopathological images," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 2, pp. 284–296, 2013.
- [23] S. M. Ismail, A. B. Colclough, J. S. Dinnen, D. Eakins, D. Evans, E. Gradwell, J. P. O'sullivan, J. M. Summerell, and R. G. Newcombe, "Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia." *British Medical Journal*, vol. 298, no. 6675, p. 707, 1989.

- [24] S. Petushi, F. U. Garcia, M. M. Haber, C. Katsinis, and A. Tozeren, "Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer." *BMC Medical Imaging*, vol. 6, p. 14, 2006.
- [25] L. Brochez, E. Verhaeghe, E. Grosshans, E. Haneke, G. Piérard, D. Ruiters, and J.-M. Naeyaert, "Inter-observer variation in the histopathological diagnosis of clinically suspicious pigmented skin lesions," *The Journal of Pathology*, vol. 196, no. 4, pp. 459–466, 2002.
- [26] Y. Wang, D. Crookes, O. S. Eldin, S. Wang, P. Hamilton, and J. Diamond, "Assisted diagnosis of cervical intraepithelial neoplasia (CIN)," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 112–121, 2009.
- [27] M. N. Gurcan, T. Pan, H. Shimada, and J. Saltz, "Image analysis for neuroblastoma classification: Segmentation of cell nuclei," in *Proceedings of IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2006, pp. 4844–4847.
- [28] V. Korde, H. Bartels, J. Barton, and J. Ranger-Moore, "Automatic segmentation of cell nuclei in bladder and skin tissue for karyometric analysis," *Analytical and Quantitative Cytology and Histology*, vol. 31, no. 2, p. 83, 2009.
- [29] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, p. 146, 2004.
- [30] C. Lu, M. Mahmood, N. Jha, and M. Mandal, "A robust automatic nuclei segmentation technique for quantitative histopathological image analysis," *Analytical and Quantitative Cytology and Histology*, vol. 34, no. 6, pp. 296–308, 2012.

- [31] R. Gonzalez and R. Woods, *Digital image processing*, 3rd ed. USA: Prentice Hall, 2008.
- [32] X. Yang, H. Li, and X. Zhou, "Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and kalman filter in time-lapse microscopy," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 11, pp. 2405–2414, 2006.
- [33] J. Cheng and J. Rajapakse, "Segmentation of clustered nuclei with shape markers and marking function," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 741–748, 2009.
- [34] C. Jung and C. Kim, "Segmenting clustered nuclei using h-minima transform-based marker extraction and contour parameterization," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2600–2604, 2010.
- [35] M. Veta, P. J. van Diest, R. Kornegoor, A. Huisman, M. A. Viergever, and J. P. Pluim, "Automatic nuclei segmentation in H&E stained breast cancer histopathology images," *PloS one*, vol. 8, no. 7, p. e70221, 2013.
- [36] H. Masmoudi, S. M. Hewitt, N. Petrick, K. J. Myers, and M. A. Gavrielides, "Automated quantitative assessment of HER-2/neu immunohistochemical expression in breast cancer," *IEEE Transactions on Medical Imaging*, vol. 28, no. 6, pp. 916–925, 2009.
- [37] Y. M. George, B. M. Bagoury, H. H. Zayed, and M. I. Roushdy, "Automated cell nuclei segmentation for breast fine needle aspiration cytology," *Signal Processing*, vol. 93, no. 10, pp. 2804–2816, 2013.
- [38] J. Byun, M. R. Verardo, B. Sumengen, G. P. Lewis, B. Manjunath, and S. K. Fisher, "Automated tool for the detection of cell nuclei in digital microscopic images: application to retinal images," *Molecular Vision*, vol. 12, pp. 949–960, 2006.

- [39] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 841–852, 2010.
- [40] H. Kong, H. C. Akakin, and S. E. Sarma, "A generalized laplacian of gaussian filter for blob detection and its applications," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1719–1733, 2013.
- [41] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [42] P. Yan, X. Zhou, M. Shah, and S. T. Wong, "Automatic segmentation of high-throughput rna fluorescent cellular images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 109–117, 2008.
- [43] X. Qi, F. Xing, D. J. Foran, and L. Yang, "Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 3, pp. 754–765, 2012.
- [44] S. Ali and A. Madabhushi, "An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery," *IEEE Transactions on Medical Imaging*, vol. 31, no. 7, pp. 1448–1460, 2012.
- [45] H. Chang, Q. Wen, and B. Parvin, "Coupled segmentation of nuclear and membrane-bound macromolecules through voting and multiphase level set," *Pattern recognition*, vol. 48, no. 3, pp. 882–893, 2015.
- [46] O. Sertel, J. Kong, U. V. Catalyurek, G. Lozanski, J. H. Saltz, and M. N. Gurcan, "Histopathological image analysis using model-based intermediate representations and color texture: Follicular lymphoma grading," *Journal of Signal Processing Systems for Signal Image and Video Technology*, vol. 55, no. 1-3, pp. 169–183, 2009.

- [47] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi, “Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information,” in *Proceedings of International Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB)*, 2007, pp. 1–8.
- [48] O. Sertel, U. Catalyurek, H. Shimada, and M. Guican, “Computer-aided prognosis of neuroblastoma: Detection of mitosis and karyorrhexis cells in digitized histological images,” in *Proceedings of International Conference on Engineering in Medicine and Biology Society (EMBC)*, 2009, pp. 1433–1436.
- [49] A. N. Basavanhally, S. Ganesan, S. Agner, J. P. Monaco, M. D. Feldman, J. E. Tomaszewski, G. Bhanot, and A. Madabhushi, “Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, pp. 642–653, 2010.
- [50] C. Lu and M. Mandal, “Toward automatic mitotic cell detection and segmentation in multispectral histopathological images,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 2, pp. 594–605, 2014.
- [51] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [52] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 281-297, 1967, p. 14.
- [53] T. Xu, I. Cheng, and M. Mandal, “An improved fluid vector flow for cavity segmentation in chest radiographs,” in *2010 20th International Conference on Pattern Recognition (ICPR)*. IEEE, 2010, pp. 3376–3379.

- [54] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Journal of Reviews on Biomedical Engineering*, vol. 2, pp. 147–171, 2009.
- [55] M. Wiltgen, A. Gerger, and J. Smolle, "Tissue counter analysis of benign common nevi and malignant melanoma," *International Journal of Medical Informatics*, vol. 69, no. 1, pp. 17–28, 2003.
- [56] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [57] L.-K. Soh and C. Tsatsoulis, "Texture analysis of sar sea ice imagery using gray level co-occurrence matrices," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 780–795, 1999.
- [58] O. Sertel, J. Kong, H. Shimada, U. V. Catalyurek, J. H. Saltz, and M. N. Gurcan, "Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development," *Pattern Recognition*, vol. 42, no. 6, pp. 1093–1103, 2009.
- [59] C. Lu and M. Mandal, "Automated analysis and diagnosis of skin melanoma on whole slide histopathological images," *Pattern Recognition*, vol. 48, no. 8, pp. 2738–2750, 2015.
- [60] S. Doyle, M. Hwang, K. Shah, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated grading of prostate cancer using architectural and textural image features," in *Proceedings of International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2007, pp. 1284–1287.
- [61] S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi, "A boosted bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1205–1218, 2012.

- [70] A. Tabesh, M. Teverovskiy, H.-Y. Pang, V. P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, “Multifeature prostate cancer diagnosis and gleason grading of histological images,” *IEEE transactions on medical imaging*, vol. 26, no. 10, pp. 1366–1378, 2007.
- [71] M. Mete, X. Xu, C.-Y. Fan, and G. Shafirstein, “Automatic delineation of malignancy in histopathological head and neck slides,” *BMC Bioinformatics*, vol. 8, no. Suppl 7, p. S17, 2007.
- [72] J. Kong, O. Sertel, H. Shimada, K. L. Boyer, J. H. Saltz, and M. N. Gurcan, “Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation,” *Pattern Recognition*, vol. 42, no. 6, pp. 1080–1092, 2009.
- [73] S. Ali, J. Lewis, and A. Madabhushi, “Spatially aware cell cluster (spaccl) graphs: Predicting outcome in oropharyngeal p16+ tumors,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2013, pp. 412–419.
- [74] C. Lu and M. Mandal, “Automated segmentation and analysis of the epidermis area in skin histopathological images,” in *Proceedings of International Conference of on Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 5355–5359.
- [75] J. M. Haggerty, X. N. Wang, A. Dickinson, C. J. O’SMalley, and E. B. Martin, “Segmentation of epidermal tissue with histopathological damage in images of haematoxylin and eosin stained human skin,” *BMC medical imaging*, vol. 14, no. 1, p. 1, 2014.
- [76] S. Kusuma, R. K. Vuthoori, M. Piliang, and J. E. Zins, “Skin anatomy and physiology,” in *Plastic and Reconstructive Surgery*. Springer, 2010, pp. 161–171.

- [77] Z. Guo and R. W. Hall, “Parallel thinning with two-subiteration algorithms,” *Communications of the ACM*, vol. 32, no. 3, pp. 359–373, 1989.
- [78] P. Soille, *Morphological image analysis: principles and applications*. New York: Springer-Verlag, 2003.
- [79] H. Fatakdawala, J. Xu, A. Basavanhally, G. Bhanot, S. Ganesan, M. Feldman, J. E. Tomaszewski, and A. Madabhushi, “Expectation–maximization-driven geodesic active contour with overlap resolution (emagacor): Application to lymphocyte segmentation on breast cancer histopathology,” *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1676–1689, 2010.
- [80] B. Parvin, Q. Yang, J. Han, H. Chang, B. Rydberg, and M. H. Barcellos-Hoff, “Iterative voting for inference of structural saliency and characterization of subcellular events,” *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 615–623, 2007.
- [81] M. K. K. Niazi, A. A. Satoskar, and M. N. Gurcan, “An automated method for counting cytotoxic T-cells from CD8 stained images of renal biopsies,” in *SPIE Medical Imaging*, 2013, pp. 867 606–867 606.
- [82] M. Zhang, T. Wu, and K. M. Bennett, “Small blob identification in medical images using regional features from optimum scale,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 4, pp. 1051–1062, 2015.
- [83] R. Guo and S. Pandit, “Automatic threshold selection based on histogram modes and a discriminant criterion,” *Machine Vision and Applications*, vol. 10, no. 5-6, pp. 331–338, 1998.
- [84] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

- [85] T. F. Chan, B. Y. Sandberg, and L. A. Vese, “Active contours without edges for vector-valued images,” *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 130–141, 2000.
- [86] S. Acton and N. Ray, “Biomedical image analysis: segmentation,” *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 4, no. 1, pp. 1–108, 2009.
- [87] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [88] J. Stegmaier, J. C. Otte, A. Kobitski, A. Bartschat, A. Garcia, G. U. Nienhaus, U. Strähle, and R. Mikut, “Fast segmentation of stained nuclei in terabyte-scale, time resolved 3d microscopy image stacks,” *PloS one*, vol. 9, no. 2, p. e90036, 2014.
- [89] T. Lindeberg, “Feature detection with automatic scale selection,” *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.
- [90] C. Xiao, M. Staring, Y. Wang, D. P. Shamonin, and B. C. Stoel, “Multiscale bi-gaussian filter for adjacent curvilinear structures detection with application to vasculature images,” *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 174–188, 2013.
- [91] H. Xu, C. Lu, and M. Mandal, “An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 5, pp. 1729–1741, 2014.
- [92] C. for Bio-Image Informatics, “Bio-segmetnation,” Website, November 2015. [Online]. Available: <http://bioimage.ucsb.edu/>
- [93] Z. Cheng and Y. Liu, “Efficient technique for ellipse detection using restricted randomized hough transform,” in *Proceedings of International*

Conference on Information Technology: Coding and Computing (ITCC), vol. 2, 2004, pp. 714–718.

- [94] H. Wang, H. Zhang, and N. Ray, “Clump splitting via bottleneck detection and shape classification,” *Pattern Recognition*, vol. 45, no. 7, pp. 2780–2787, 2012.
- [95] Q. Wen, H. Chang, and B. Parvin, “A delaunay triangulation approach for segmenting clumps of nuclei,” in *Proceedings of International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2009, pp. 9–12.
- [96] O. Schmitt and S. Reetz, “On the decomposition of cell clusters,” *Journal of Mathematical Imaging and Vision*, vol. 33, no. 1, pp. 85–103, 2009.
- [97] X. Bai, C. Sun, and F. Zhou, “Splitting touching cells based on concave points and ellipse fitting,” *Pattern recognition*, vol. 42, no. 11, pp. 2434–2446, 2009.
- [98] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, “Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting,” *IEEE Transactions on Medical Imaging*, vol. 30, no. 9, pp. 1661–1677, 2011.
- [99] L. Zhang, H. Kong, C. T. Chin, S. Liu, Z. Chen, T. Wang, and S. Chen, “Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts,” *Computerized Medical Imaging and Graphics*, vol. 38, no. 5, pp. 369–380, 2014.
- [100] F. Xing, H. Su, J. Neltner, and L. Yang, “Automatic ki-67 counting using robust cell detection and online dictionary learning,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 859–870, 2014.
- [101] M. E. Plissiti and C. Nikou, “Overlapping cell nuclei segmentation using a spatially adaptive active physical model,” *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4568–4580, 2012.

- [102] Z. Lu, G. Carneiro, and A. P. Bradley, "An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells," *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1261–1272, 2015.
- [103] A. Janowczyk, S. Doyle, H. Gilmore, and A. Madabhushi, "A resolution adaptive deep hierarchical (RADHicaL) learning scheme applied to nuclear segmentation of digital pathology images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp. 1–7, 2016.
- [104] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems (NIPS)*, 2012, pp. 1097–1105.
- [105] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *Journal of Pathology Informatics*, vol. 7, 2016.
- [106] H. Xu, C. Lu, R. Berendt, N. Jha, and M. Mandal, "Automatic nuclei detection based on generalized laplacian of gaussian filters," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 826–837, 2017.
- [107] C. Lu, M. Mahmood, N. Jha, and M. Mandal, "Detection of melanocytes in skin histopathological images using radial line scanning," *Pattern Recognition*, vol. 46, no. 2, pp. 509–518, 2013.
- [108] S. T. Acton and N. Ray, "Biomedical image analysis: tracking," *Synthesis Lectures on Image, Video, and Multimedia Processing*, vol. 2, no. 1, pp. 1–152, 2006.
- [109] P. S. Nielsen, R. Riber-Hansen, J. Raundahl, and T. Steiniche, "Automated quantification of MART1-verified Ki67 indices by digital image analysis in

- melanocytic lesions,” *Archives of pathology & laboratory medicine*, vol. 136, no. 6, pp. 627–634, 2012.
- [110] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *CoRR*, vol. abs/1409.1556, 2014.
- [111] Y. Zhang, L. Yang, J. D. MacKenzie, R. Ramachandran, and D. Z. Chen, “A seeding-searching-ensemble method for gland segmentation in h&e-stained images,” *BMC Medical Informatics and Decision Making*, vol. 16, no. 2, p. 80, 2016.
- [112] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *Proceedings of International Conference on Machine Learning (ICML)*, 2015, pp. 448–456.
- [113] H. Xu and M. Mandal, “Epidermis segmentation in skin histopathological images based on thickness measurement and k-means algorithm,” *EURASIP Journal on Image and Video Processing*, vol. 2015, no. 1, pp. 1–14, 2015.
- [114] J. Smolle, “Computer recognition of skin structures using discriminant and cluster analysis,” *Skin Research and Technology*, vol. 6, no. 2, pp. 58–63, 2000.
- [115] J. Smolle, A. Gerger, W. Weger, H. Kutzner, and M. Tronnier, “Tissue counter analysis of histologic sections of melanoma: Influence of mask size and shape, feature selection, statistical methods and tissue preparation,” *Analytical Cellular Pathology*, vol. 24, no. 2-3, pp. 59–67, 2002.
- [116] P. S. Nielsen, E. Spaun, R. Riber-Hansen, and T. Steiniche, “Automated quantification of mart1-verified ki-67 indices: useful diagnostic aid in melanocytic lesions,” *Human pathology*, vol. 45, no. 6, pp. 1153–1161, 2014.
- [117] H. Xu, C. Lu, and M. Mandal, “Automated segmentation of regions of interest in whole slide skin histopathological images,” in *Proceedings of*

International Conference on Engineering in Medicine and Biology Society (EMBC), 2015, pp. 3869–3872.

- [118] J. Vink, M. Van Leeuwen, C. Van Deurzen, and G. De Haan, “Efficient nucleus detector in histopathology images,” *Journal of microscopy*, vol. 249, no. 2, pp. 124–135, 2013.
- [119] B. E. Bejnordi, G. Litjens, N. Timofeeva, I. Otte-Höller, A. Homeyer, N. Karssemeijer, and J. A. van der Laak, “Stain specific standardization of whole-slide histopathological images,” *IEEE transactions on medical imaging*, vol. 35, no. 2, pp. 404–415, 2016.
- [120] M. Macenko, M. Niethammer, J. Marron, D. Borland, J. Woosley, X. Guan, C. Schmitt, and N. Thomas, “A method for normalizing histology slides for quantitative analysis,” in *Proceedings of International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, 2009, pp. 1107–1110.
- [121] A. C. Ruifrok, D. A. Johnston *et al.*, “Quantification of histochemical staining by color deconvolution,” *Analytical and Quantitative Cytology and Histology*, vol. 23, no. 4, pp. 291–299, 2001.
- [122] H. Xu, C. Lu, R. Berendt, N. Jha, and M. Mandal, “Automatic nuclear segmentation using multi-scale radial line scanning with dynamic programming,” *accepted for publication by IEEE Transactions on Biomedical Engineering*, 2017.
- [123] C. Hsu and C. Lin, “A comparison of methods for multiclass support vector machines,” *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.
- [124] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, p. 27, 2011.

- [125] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, no. 2, pp. 153–158, 1997.
- [126] C. R. Maurer Jr, R. Qi, and V. Raghavan, "A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 265–270, 2003.
- [127] G. J. Hahn and S. S. Shapiro, *Statistical models in engineering*. Wiley-Interscience, 1994.
- [128] A. W. Bowman and A. Azzalini, *Applied smoothing techniques for data analysis*. Clarendon Press, 2004.
- [129] P. A. Fetsch, F. M. Marincola, A. Filie, Y. M. Hijazi, D. E. Kleiner, and A. Abati, "Melanoma-associated antigen recognized by t cells (MART-1)," *Cancer Cytopathology*, vol. 87, no. 1, pp. 37–42, 1999.
- [130] A. Biernacka, K. D. Linos, P. A. DeLong, A. A. Suriawinata, V. Padmanabhan, X. Liu *et al.*, "A case of s-100 negative melanoma: A diagnostic pitfall in the workup of a poorly differentiated metastatic tumor of unknown origin," *CytoJournal*, vol. 13, no. 1, p. 21, 2016.

Appendix A

Histology Images and Features

A.1 Histology images with different stains

In order to diagnose and grade skin cancers such as melanomas, pathologists may examine and analyze a set of of biopsy images with different stains, as biopsies with different histochemical stains can provide more information for making a judgment. The typical stains used in skin biopsies consists of the following six categories:

- **H&E:** Hematoxylin stains cell nuclei blue, while Eosin stains cytoplasm and connective tissue pink (see Fig. A.1(a)). Pathologists generally diagnose skin cancers by observing cytological features in H&E stained skin biopsy images.
- **MART-1:** MART-1 is a melanocyte specific transmembrane protein that is present in normal melanocytes and widely expressed in malignant melanomas. The MART-1 antibody has been considered as a superior immunohistochemical marker for the diagnosis and grading of melanomas [129]. Fig. A.1(b) shows a MART-1 stained skin biopsy image, where melanoma regions have been stained with brown hues.
- **Ki-67 (MIB-1):** Ki-67 is a nuclear protein expressed exclusively during the active cell cycle phases (i.e., interphase and cell division), but it is absent from resting cells (i.e., resting phase). Ki-67 antibody has been used to

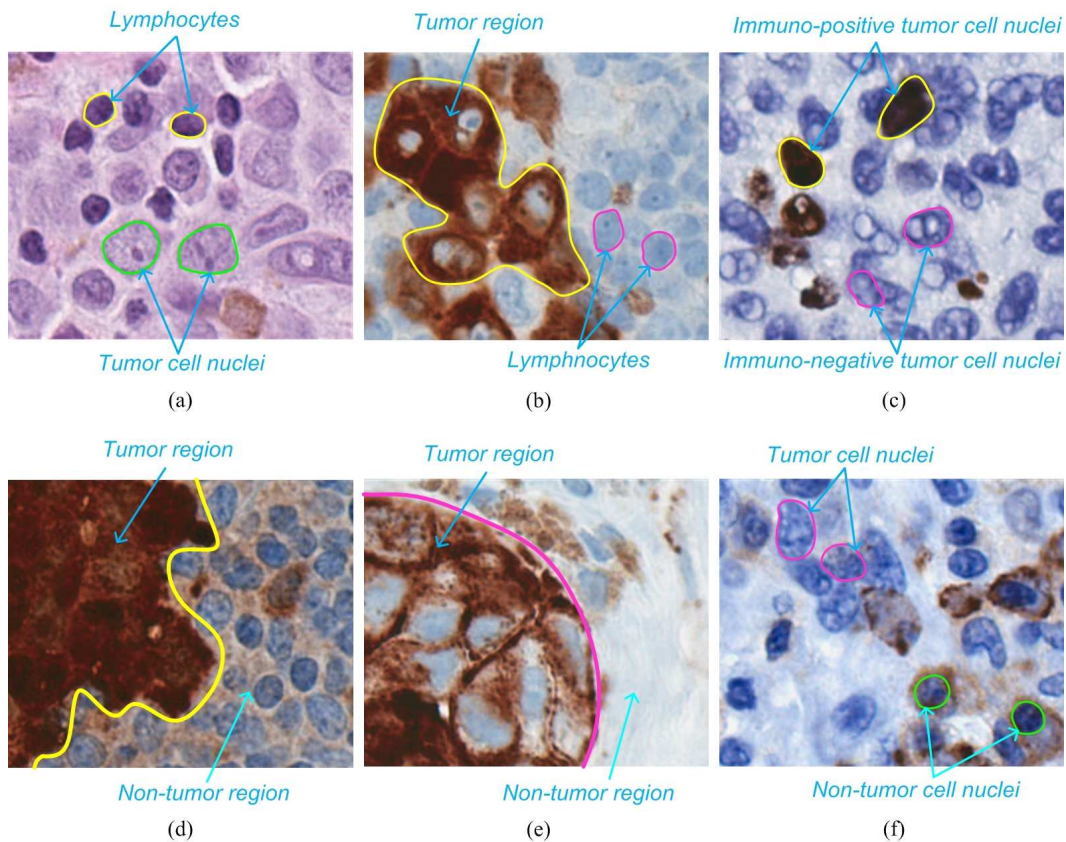


Fig. A.1. Skin histopathological images with different stains. (a) An H&E stained image. (b) A MART-1 stained image. (c) A Ki-67 stained image. (d) A S-100 stained image. (e) A HMB-45 stained image. (f) A CD-45 stained image. Note that the images are viewed at 40X.

evaluate the tumor cell proliferation, which is an important factor to grade the severity of the tumor. Fig. A.1(c) shows a Ki-67 stained skin biopsy image, where immuno-positive and immuno-negative tumor cell nuclei are observed as brown and blue colors, respectively.

- **S-100:** S-100 protein is expressed in all melanocytic cells and in most of malignant melanomas. Although very sensitive, S-100 is not specific to melanoma as it also stains most peripheral nerve sheath tumors and some other sarcomas, as well as a minority of Langerhans cell histiocytosis [130]. Because of its low specificity, S-100 is of diagnostic value for melanoma only in selected cases [129]. Fig. A.1(d) shows a S-100 stained skin biopsy image, where melanoma regions have been stained with dark brown colors.

- **HMB-45:** HMB-45 (Human Melanoma Black) antibody is similar with MART-1 antibody, which is often diffusely positive in melanomas. Pathologists usually diagnose and grade malignant melanomas by using HMB-45 stained biopsy images, especially before MART-1 antibodies become commercially available. Fig. A.1(e) shows a HMB-45 stained skin biopsy image, where melanoma regions have been stained with brown hues.
- **CD-45:** CD-45 is known to have high sensitivity and specificity for lymphoid tumors. CD-45 immunoreactivity has been considered exquisitely specific for hematopoietic cells. Melanoma tissues in CD-45 stained biopsy images have immuno-negative responses. Fig. A.1(f) shows a CD-45 stained skin biopsy image. Note that since melanoma cells in Fig. A.1(f) have immuno-negative responses, the tumor cell nuclei are observed with blue hues.

A.2 Features of histology images

In this section, definitions and explanations of some texture features are provided.

A.2.1 Textural features

A.2.1.1 Statistical features from intensity histogram

A frequently used approach for texture analysis is based on statistical properties of the intensity histogram. Given a segmented region R with Λ pixels. Let us assume z_i is a random variable indicating pixel intensity in R . $p(z)$ is the histogram of intensity levels in R , and L is the number of possible intensity levels. The descriptions of statistical features including mean, standard deviation, smoothness, third moment, uniformity and entropy are listed in Table A.1.

A.2.1.2 Haralick features

Haralick features including energy, contrast, correlation, homogeneity, dissimilarity and entropy are widely-used in histological image analysis, which are defined in

TABLE A.1
DESCRIPTION OF STATISTICAL TEXTURE FEATURES.

Name	Expression	Measure of Texture
Mean	$m = \sum_{i=0}^{L-1} z_i p(z_i)$	a measure of average intensity
Standard deviation	$\sigma = \sqrt{\frac{\sum_{i=0}^{L-1} (z_i^2 p(z_i)) - \Lambda m^2}{\Lambda - 1}}$	a measure of average contrast
Smoothness	$R = 1 - 1/(1 + \sigma^2)$	a measure of relative smoothness
Third moment	$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$	a measure of skewness of the histogram
Uniformity	$U = \sum_{i=0}^{L-1} p^2(z_i)$	a measure of uniformity
Entropy	$e = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$	a measure of randomness

Table A.2.

TABLE A.2
DESCRIPTION OF HARALICK FEATURES.

Name	Expression	Measure of Texture
Energy	$\sum_i \sum_j p(i, j)^2$	a measure of uniformity.
Contrast	$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}$, where $ i - j = n$	a measure of variation between intensity levels
Correlation	$\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y / \sigma_x \sigma_y$	a measure of correlation between a pixel and its neighborhood
Homogeneity	$\sum_i \sum_j \frac{p(i, j)}{1 + (i - j)^2}$	a measure of homogeneity
Dissimilarity	$\sum_i \sum_j i - j p(i, j)$	a measure of dissimilarity in intensity values
Entropy	$-\sum_i \sum_j p(i, j) \log(p(i, j))$	a measure of randomness

Note that in Table A.2, μ_x , μ_y , σ_x and σ_y are the means and standard deviations of the partial probability density functions, which are given by:

$$\mu_x = \sum_i \sum_j i \cdot p(i, j) \quad (\text{A-1})$$

$$\mu_y = \sum_i \sum_j j \cdot p(i, j) \quad (\text{A-2})$$

$$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 \cdot p(i, j) \quad (\text{A-3})$$

$$\sigma_y = \sum_i \sum_j (j - \mu_y)^2 \cdot p(i, j) \quad (\text{A-4})$$

A.3 Examples of skin WSIs

Fig. A.2 shows examples of H&E stained normal skin biopsies, where the first row shows three skin WSIs and the second row shows selected skin epidermis and dermis image patches. Note that most cell nuclei in epidermis layer of normal skin tissue are the keratinocytes, and only a very small number of melanocytes are present in the skin DEJ regions. In the skin dermis layer of normal tissue, cell nuclei mainly includes fibroblasts, macrophages and adipocytes.

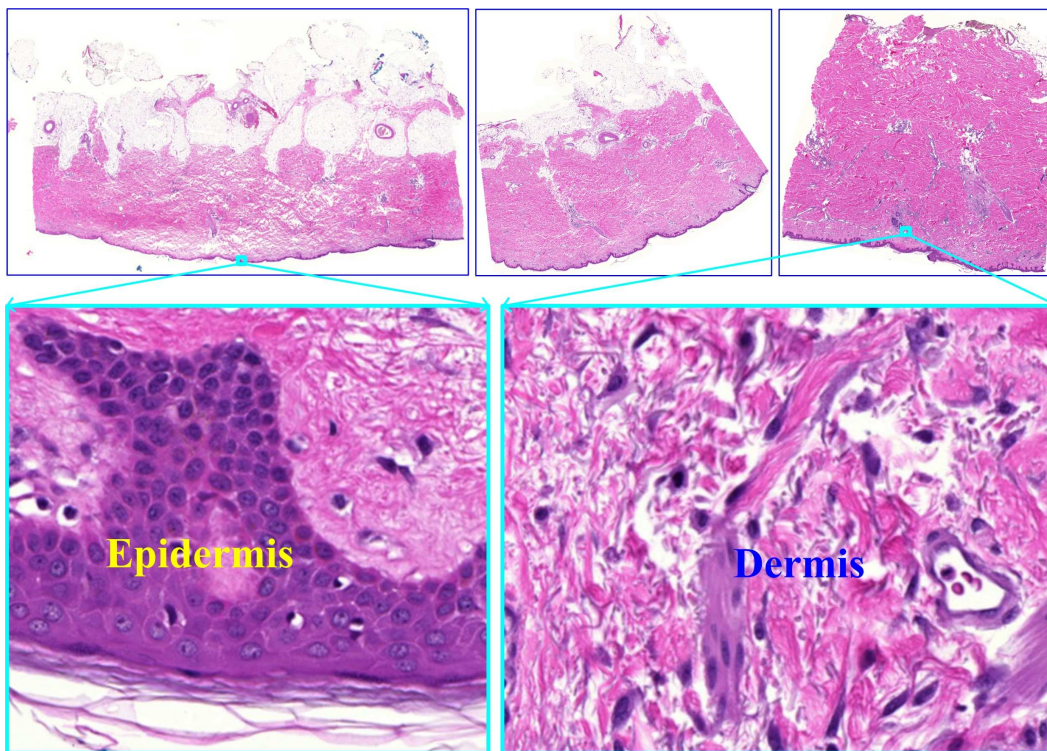


Fig. A.2. Examples of H&E stained normal skin biopsies.

Fig. A.3 shows examples of H&E stained skin nevus biopsies, where the first row shows two skin WSIs and the second row shows selected skin epidermis and dermis image patches. Note that in the epidermis layer of nevus biopsies, there is an

increasing number of melanocytes distributed along the skin basal layer. In the skin dermis layer of nevus biopsies, there is a large number of melanocytes, but these melanocytes have low degree of nuclei pleomorphism (i.e., shape, size and color do not vary greatly from each other).

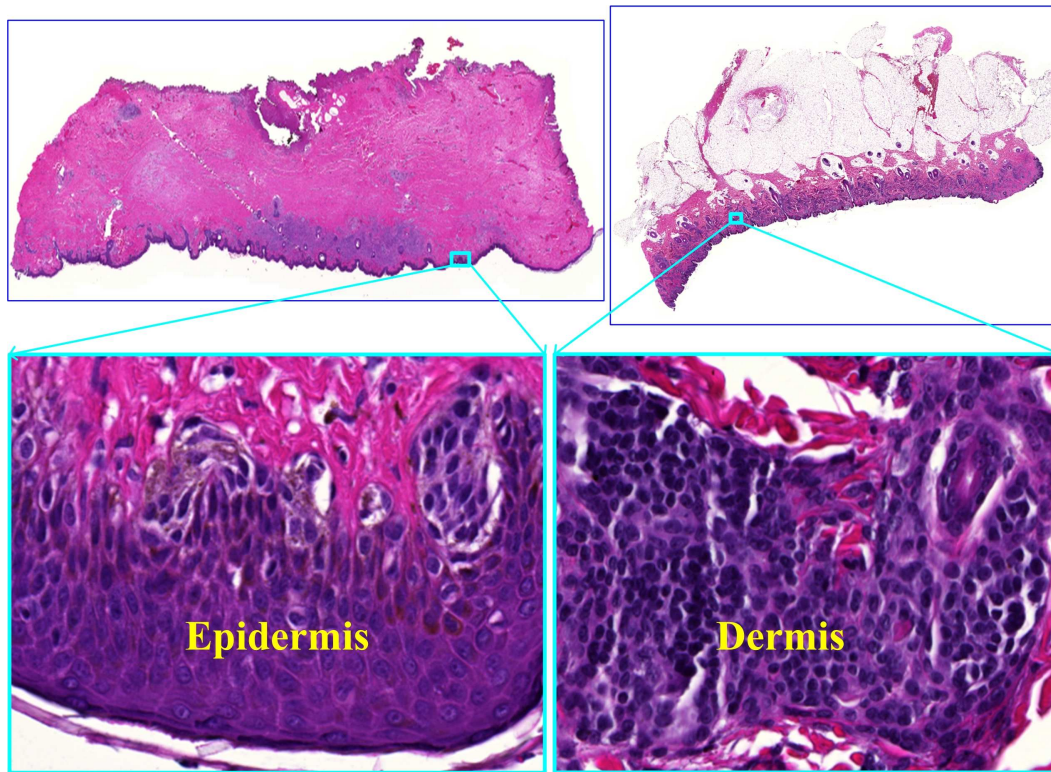


Fig. A.3. Examples of H&E stained skin nevus biopsies.

Fig. A.4 shows examples of H&E stained skin biopsies with melanomas. In the first and third rows of Fig. A.4, it shows five skin WSIs, while in the second and fourth rows it shows selected skin epidermis and dermis image patches. Note that in the epidermis layer of melanoma biopsies, there is an increasing number of melanocytes which are not only found in the basal layer but also in the spinous layer. In the skin dermis layer of melanoma biopsies, there is a large number of melanocytes and lymphocytes. The melanocytes in the skin dermis layer have high degree of nuclei pleomorphism in term of shape, color and size.

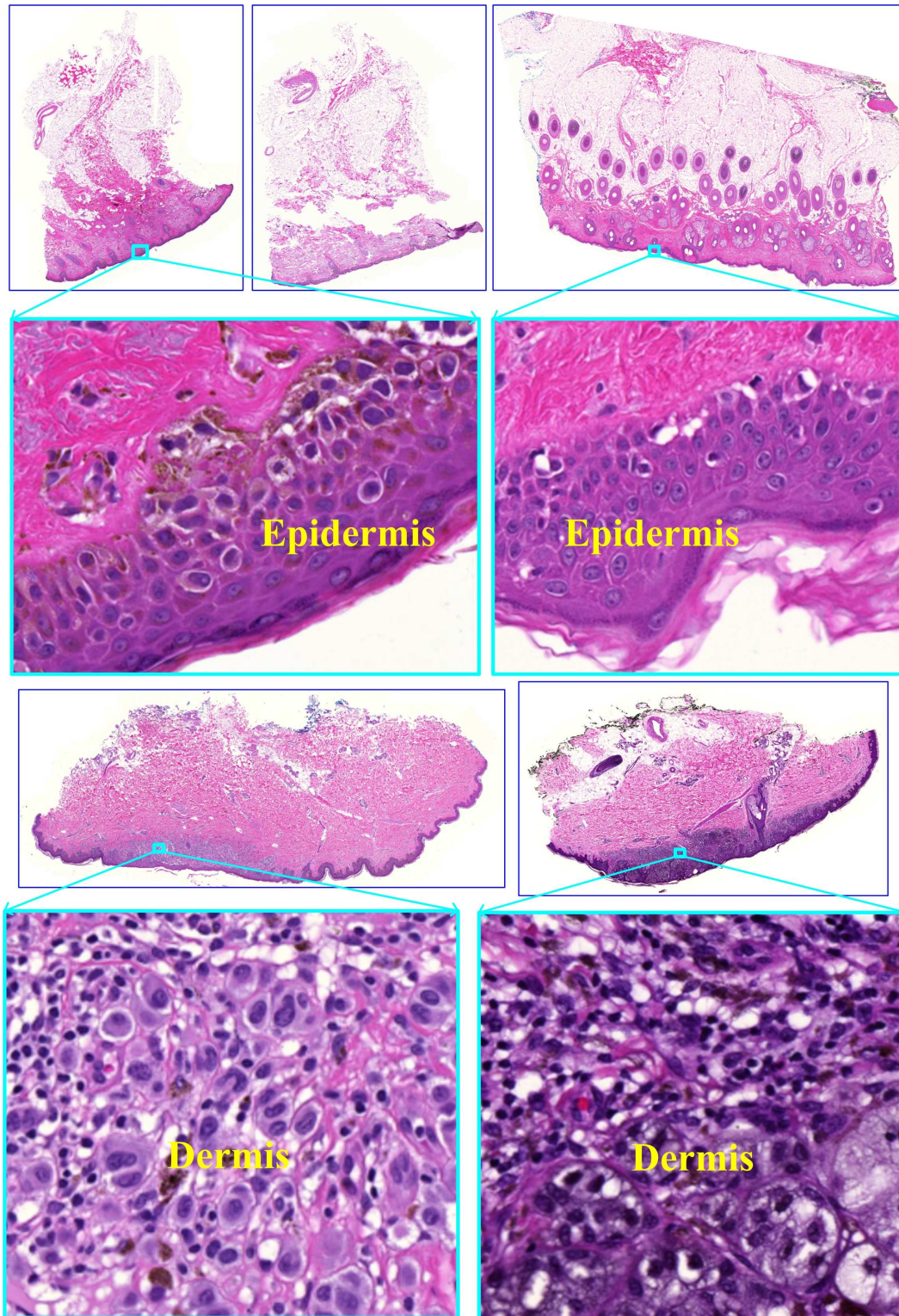


Fig. A.4. Examples of H&E stained skin melanoma biopsies.