A Few Algorithms for Histopathological Images in Computational Pathology

ABHISHEK RAMNATH VAHADANE

TH-1577_11610215

A Few Algorithms for Histopathological Images in Computational Pathology

A

Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

By

ABHISHEK RAMNATH VAHADANE

under guidance of

Dr. Amit Sethi



DEPARTMENT OF ELECTRONICS AND ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI-781039, ASSAM, INDIA

April, 2017

TH-1577_11610215

To My Family

TH-1577_11610215

DECLARATION

*This is to certify that the thesis entitled "**A Few Algorithms for Histopathological Images in Computational Pathology**", submitted by me to the Indian Institute of Technology Guwahati for the award of the degree of **Doctor of Philosophy** is a bonafide work carried out by me under the supervision of **Dr. Amit Sethi**. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.*

Sign:

Abhishek Ramnath Vahadane
Department of Electronics & Electrical Engineering
Indian Institute of Technology Guwahati
Guwahati-781039, Assam, India

Date:

TH-1577_11610215

CERTIFICATE

This is to certify that the this thesis entitled "**A Few Algorithms for Histopathological Images in Computational Pathology**" submitted by **Abhishek Ramnath Vahadane (11610215)**, a research scholar in the *Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati,* for the award of degree of **Doctor of Philosophy**, is a record of an original research work carried out by him under my supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the institute and in my opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Sign:
_____

Supervisor: Dr. Amit Sethi
Assistant Professor, Department of Electronics & Electrical Engineering,
Indian Institute of Technology Guwahati,
Guwahati-781039, Assam, India.

Date:
_____

TH-1577_11610215

Acknowledgments

I feel a great privilege in expressing my deepest and most sincere gratitude to my supervisor Dr. Amit Sethi, for his excellent guidance throughout the course of this work. He was always supportive and helping hand. His kindness, dedication, friendly accessibility and attention to detail have been a great inspiration to me. My heartfelt thanks to him for the unlimited support and patience shown to me. I would particularly like to thank for all his help in patiently and carefully correcting all my manuscripts, reports and other scientific writing. I could not have imagined having a better advisor and mentor for my PhD study.

I would like to thank my doctoral committee members Prof. Prabin Bora, Prof. Rohit Sinha. Dr. Sanjukta Patra, and Dr. Manas Bhuyan for sparing their precious time to evaluate the progress of my research work. Their suggestions have been valuable. I would also like to thank other faculty members for their kind help in my academic studies. I am also grateful to technical staff of the department.

I owe my deepest gratitude to Dr. Tingying Peng, Prof. Nassir Navab and colleagues at computer aided medical procedures (CAMP) who mentored me during a yearlong PhD internship at Technical University of Munich (TUM), Germany. I also thank to the Heritage fellowship for their scholarship award to carry out this internship. I would like to thank my collaborators Dr. Katja Steiger (TUM), Dr. Christine Bayer (TUM) and Prof. Peter Gann (University of Illinois, Chicago) for their help in expert evaluation of our experimental results.

I would like to thank to all my bachelor students specially, Sanuj Sharma, Surabhi Bhargava, Dileep Munugoti, Chaitanya, and Vamsi Tallam for their crucial help in performing experiments for their Bachelor thesis and summer project. I also extend my gratitude and thanks to all my beloved friends at the IPCV laboratory. They have always been around to provide useful suggestions, companionship and created a good research environment. I thank Saroj, Ankit, Sikandar, Santosh, Gaurav and lots of other for their friendly support during my stay at IITG.

Finally, I believe this research experience will greatly benefit my career in the future.

Abhishek Vahadane

TH-1577_11610215

Abstract

Computational pathology is increasingly gaining attention due to its significant advantages over a human expert under certain circumstances. Huge stacks of histological data are made available through advances in the microscopic image acquisition. Clinicians cannot learn easily from this big data due to obvious human limitations. However, machine learning systems can learn important patterns from these huge histological data sets to improve the diagnosis and prognosis. Such automated prediction can act as a second opinion for the experts to help make their decisions. For example, in one of the breakthrough work, clinically unknown novel stromal features were discovered using machine learning to be highly predictive of mortality in breast cancer [1]. In this thesis, we propose solutions for four different problems, some of which are widely applicable to computational pathology pipelines such as include stain separation, color normalization, nuclei segmentation, and one that is highly specific, that is, HER2 score prediction. The first problem we solved is stain separation in histological images. We achieved stain separation that is better than existing methods by extending a non-negative matrix factorization (NMF) formulation to include a biologically plausible prior of sparsity of stains at each pixel as an additional constraint. We got our results evaluated by experts, and they preferred sparse NMF (SNMF) over NMF. We used SNMF stain representation in proposed color normalization which was our next goal. Color normalization is about suppressing the unwanted color variation in the histological images. Color normalization relies on the accurate stain separation representation. We proposed a color normalization method that standardizes the color appearance of a source image to that of a standard (target) image, which effectively preserves the biological structure of the source image. In the stain separation representation, we

TH-1577_11610215

estimate the color appearance matrix and structure information i.e. density maps. Since, the source image should be normalized to the expert chosen (or any) target, we estimate the SNMF stain separation representation for both source and target. To normalize, we replaced the color appearance of the source by that of target and linearly normalized the density maps of source with the target. The proposed technique was shown to outperform the compared techniques. We also validated the impact of the color normalization for the epithelial-stromal segmentation in prostate tissue. Our next goal was to detect and segment nuclei. Nuclear segmentation is still not convincingly solved because of nuclei appearance variations and overlapping nuclei. We proposed an image processing pipeline that improves upon the watershed segmentation method. We included morphological pre-processing and marking schemes for enhancement of watershed segmentation. However, it was not able to solve the problem of overlapping and enlarged nuclei which is still an open challenge. So, we devised another algorithm to apply the ability of deep convolutional neural networks (CNN) for learning three classes of pixels such that one of the classes, the ± 1 pixel boundary, helps segment touching and enlarged nuclei. A large number of hand-labeled nuclei annotations were used to train the CNN. A simple post-processing method resulted in good segmentation accuracy and overlapping nuclei were resolved as separate objects. Detection of HER2 score of whole slide image (WSI) is important for diagnosis and prognosis since HER2 overexpression is observed in 25% of the breast cancer cases. While participating in a competition [2], we developed a deep CNN to predict the HER2 score of the patches and novel criteria to score the WSI based on the patches HER2 scores. We were successfully able to detect the HER2 score of the difficult WSI in which experts gave wrong HER2 score. We also estimate percentage of complete membrane staining through a simple image processing pipeline. Proposed HER2 predictor was ranked first in HER2 scoring contest.

Keywords: Digital Pathology, Histological Images, Staining, Stain Separation, Non-Negative Matrix Factorization, Sparse Coding, Color Normalization, Fast Stain Separation, Nuclei Segmentation, Deep Convolutional Neural Network, Breast Cancer, HER2 Score Prediction

TH-1577_11610215

Contents

List of Figures iii

List of Tables ix

1 Introduction 1

1.1 Goals of thesis and organization 4

2 Stain Separation of Histological Images 9

2.1 Introduction 9

2.2 Related work 10

2.3 Proposed SNMF based stain separation 14

2.4 Experiment and Results 16

2.4.1 Ground truth generation 16

2.4.2 Comparison of stain color appearance W 17

2.4.3 Comparison of stain density maps H 18

2.4.4 Sensitivity analysis of sparse coefficient λ 24

2.4.5 Qualitative comparison by expert grading 24

2.5 Applications 25

2.5.1 Nuclei segmentation 26

2.5.2 Stain fusion for head and neck cancer 27

2.6 Discussion 28

3 Color normalization of histological images 31

3.1 Introduction 31

3.2 Background and Related Work 33

TH-1577_11610215

3.2.1 Color normalization without stain separation . . . 33
3.2.2 Color normalization methods based on stain separation . . . 34
3.3 Proposed Method . . . 35
3.3.1 Structure-preserving color normalization (SPCN) . . . 35
3.3.2 Smart patch-based acceleration scheme for WSI . . . 36
3.4 Experimental Results . . . 38
3.4.1 Quantitative results . . . 38
3.4.2 Qualitative validation . . . 39
3.4.3 Histogram interpretation of structure-preservation of SPCN . . . 44
3.4.4 Validation of acceleration scheme for WSIs . . . 44
3.4.5 Impact of color normalization . . . 45
3.5 Summary and Discussion . . . 46
4 Nuclei Segmentation 48
4.1 Introduction and related work . . . 48
4.1.1 Nuclear segmentation . . . 49
4.1.2 Nuclei overlapping resolution . . . 50
4.2 Proposed image processing pipeline . . . 51
4.2.1 Method . . . 51
4.2.2 Experimental results . . . 55
4.2.3 Summary and future directions . . . 59
4.3 Deep CNN based three class nuclear segmentation . . . 59
4.3.1 Experiments and results . . . 62
4.3.1.1 CPCTR . . . 64
4.3.1.2 TDLU . . . 64
4.4 Discussion . . . 69
5 Breast Cancer HER2 Score Prediction 70
5.1 HER2 and its clinical scoring of breast cancer . . . 71
5.1.1 Contributions . . . 73

TH-1577_11610215

5.2 HER2 predictor 73
5.2.1 Preprocessing and patch sampling of WSI 74
5.2.2 CNN patch HER2 estimator 75
5.2.3 Criteria for WSI HER2 estimation 76
5.3 Complete membrane staining estimator 77
5.3.1 Tumor detection (T) 78
5.3.2 IHC-brown stain identification (B) 80
5.3.3 Percentage of complete membrane staining (PCMS) 83
5.4 Experimental Results 83
5.5 Discussion and Conclusion 85
6 Conclusion **88**
Bibliography **92**

TH-1577_11610215

List of Figures

1.1 H&E biopsy preparation process [3]: From top left to bottom right: 1) Small cuts are made from the tissue 2-3) The smaller cuts are put into cassettes and (after processing) embedded in paraffin 4) Thin sections are made from the paraffin blocks with a microtome 5) The sections are put onto glass slides for staining 6) H&E stained slide 2

1.2 H&E stained histological image patch of (from left to right) bladder, colon, prostate, and stomach. 3

1.3 Modified computational pathology framework inspired from [4] 4

1.4 Color variation across different histological images 6

1.5 A typical gland structure as shown in (a). Outlined black are the (b) lumen area, (c) epithelial cytoplasm, and (d) epithelial nuclei [5]. 7

2.1 Breast H&E stained image patch (left) and predicted hematoxylin stain by proposed method 10

2.2 Effective stain graphical representation: cell nucleus stained with effective H, which is a mixture of a large amount of pure H and a small amount of pure E; by contrast, connective tissue stained with effective E, which is a mixture of a large amount of pure E and a small amount of pure H. H&E image copyright of G.Landini, University of Birmingham, U.K., reproduced with permission. Pure stain vectors are "built-in" vectors determined experimentally (details refer to http://www.mecourse.com/landinig/software/cdeconv/cdeconv.html). Effective stain vectors are determined by our SNMF decomposition. 13

TH-1577_11610215

2.3 Density map comparison of pathologist, NMF [6], SVD [7], and proposed SNMF against ground truth H_{gt}. 19
2.4 Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the bladder tissue input (a). 20
2.5 Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the colon tissue input (a). 21
2.6 Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the prostate tissue input (a). 22
2.7 Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the stomach tissue input (a). 23
2.8 Sensitivity of SNMF to hyper-parameter λ. Note that NMF is equivalent to $\lambda = 0$. 24
2.9 Application of NMF and SNMF based stain separation on a sample image. . . . 25
2.10 The flowchart of SNMF based approach for nuclei segmentation [8]. The original image (a) is decomposed into H stain component (b) with SNMF algorithm. The final segmentation result (d) is generated from binary image (c) after post-processing operation on the decomposed H stain component (b). 26
2.11 SNMF application to stain fusion 29
2.12 Proposed SNMF on IHC (H&DAB) stained image of head and neck tissue: (left) input image, (middle) Hematoxylin, (right) Diaminobenzidine. 30

3.1 The same tissue section taken by different microscopic slide scanners shows the unintended color variation 32
3.2 Flow diagram of the proposed structured preserving color normalization (SPCN). 35
3.3 Schematic diagram of fast WSI stain separation and global color matrix estimation. 37
3.4 Boxplot showing similarity metrics with respect to ground truth stain density map for different normalization methods: M1 (Reinhard), M2 (Macenko), M3 (Khan), M4 (NMF-CN), and M5 (SNMF-SPCN) 39

TH-1577_11610215

3.5 Color normalization results by various techniques on a sample from the data used for quantitative analysis 40
3.6 Example 1: Visual comparison of different color normalization techniques. See text for more details. 41
3.7 Example 2: Visual comparison of different color normalization techniques. SPCN outperform Khan technique since nuclei structure are well preserved comparatively. 42
3.8 3D histogram visualization of color normalization on a sample source and target image from the data used in Section 3.4.1. 43
3.9 Box plots of quality metrics between separated stain and ground truth for 27 H&E stained colon tissue WSIs for hematoxylin (H) and eosin (E) suggest consistent results between our fast patch-based scheme and the direct decomposition of WSI (slow approach). 45

4.1 Nuclei segmentation challenges: Marker shows (from left to right) blurred and partially visible nuclei, large size nucleus, rough nuclei boundaries, size and shape variations, and nuclei overlapping. 49
4.2 Flow diagram for the improved seeded watershed segmentation. Italic texts are our improvements to seeded watershed segmentation. 52
4.3 Improvement in nuclear segmentation by Gaussian smoothing (a) R channel of input color image, (b) Segmentation of (a), (c) Smoothed R channel, and (d) Segmentation of (b). Blurred nuclei missed by (b), are retained in (d), as shown by (red color) annotations in (d). 56

TH-1577_11610215

4.4 Results of the proposed algorithm (a) Input color image, (b) R channel, (c, e, g)- are the results for the background markers obtained from foreground markers scheme, and (d, f, h)- are the results for the proposed image driven background markers scheme. (c, d) - enhanced image + foreground markers (dots) + background markers (lines on the background), Some background markers in (c) crosses objects of interest unlike by proposed image driven background marking shown in (d), which leads to oversegmentation. (e, f)- marker controlled watershed segmentation output, (f) shows less false detections, and (g, h) - post-processed output, in between (g) and (h), (g) shows some nuclei missing unlike by proposed algorithm in (h). 57

4.5 Segmentation results on the Feulgen stained image: (a) Input Feulgen stained image, (b) Segmented output by proposed method. 58

4.6 Prostate H&E stained image patch (left) and corresponding three class annotations (right). 60

4.7 CNN architecture 61

4.8 Flow diagram of post-processing 63

4.9 The CNN is given an input (a) which produce output (b, c, d) corresponding to inside, boundary and outside class respectively. (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f). 65

4.10 The CNN is given an input (a) which produce output (b, c, d) corresponding to (inside, boundary and outside class). (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f). 66

4.11 The overlapped nuclei were distinctly segmented 67

TH-1577_11610215

4.12 The CNN is given an input (a) which produce output (b, c, d) corresponding to (inside, boundary and outside class). (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f). 67

4.13 The CNN is given an input (a) which produce output (b, c, d) corresponding to (inside, boundary and outside class). (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f). 68

4.14 The overlapped nuclei were segmented separately. 69

5.1 A typical whole slide image in our dataset includes patient and control tissue [2]. Control tissue was HER2 3+ for all WSI. Note that from now onward we showed down sampled version of the whole slide image for better visibility. The downsampling factor was 32. 72

5.2 Flow diagram of the HER2 predictor 74

5.3 IHC image patch that have tumor cells with complete membrane staining (in brown) 77

5.4 Flow diagram of percentage of complete membrane staining (PCMS) 78

5.5 H&E image of tumor-affected tissue. Different nuclear structures indicate a high likelihood of tumor presence at the lower bottom of the image patch. 79

5.6 H&E tumor detection 80

5.7 Incomplete (left) and complete (right) membrane staining 81

5.8 IHC image patch (left), detected brown region (middle) and detected contours (right) 81

5.9 IHC image patch (left), detected brown region (middle) and detected contours (right) 82

5.10 Test WSI (HER2 3+). Three new experts gave HER2 score as 1, 1 and 0 while our algorithm accurately predicted 3+. 84

TH-1577_11610215

5.11 Proposed patch CNN prediction on WSI. The WSI was accurately scored as HER2 3+. 85

5.12 Tumor detection in H&E (first row) and brown detection in IHC image patch (second row). 86

TH-1577_11610215

List of Tables

2.1 Comparison of rRMSE of stain color basis matrices (W) based on pathologists annotations, NMF [6], SVD [7], and the proposed SNMF with respect to ground truth W_{gt} 17

4.1 Typical Parameters Values for Proposed Segmentation 56

4.2 Performance measures on intestinal microscopic images 59

4.3 CNN architecture dimensions 62

5.1 HER2 staining score, pattern and assessment [2] 71

5.2 CNN architecture typical values 76

5.3 Guidelines for the PCMS scoring [2] 84

TH-1577_11610215

1

Introduction

Contents

1.1 Goals of thesis and organization 4

There are various reasons for computer assisted diagnosis (CAD) to become popular in many screening programs at the clinics and hospitals. Pathologists are busy human experts because of their workload which involves careful screening and grading of the disease stages. If we can remove obviously benign areas with a higher or equal confidence as produced by experts in routine cases, their precious time can be saved and efficiently utilized to diagnose other difficult cases and research to improve detection and diagnosis. There are significant inter- and intra-observer variations in the diagnosis and grading [9][10]. Also, the digital pathology data generated in the recent years is very huge and this will continue to grow at this rate [11]. This further aggravates the problem of manual screening and grading of such big data. On the other hand, CAD can benefit from such huge data to learn desired statistics and generalization. These are some of the reasons that we need computer assisted diagnosis (CAD) to support experts decision and ease their workload. In near future, researchers and experts in pathology are expected to use CAD tools in the diagnosis, prognosis and treatment planning.

TH-1577_11610215



Figure 1.1: H&E biopsy preparation process [3]: From top left to bottom right: 1) Small cuts are made from the tissue 2-3) The smaller cuts are put into cassettes and (after processing) embedded in paraffin 4) Thin sections are made from the paraffin blocks with a microtome 5) The sections are put onto glass slides for staining 6) H&E stained slide

The use of CAD tools demonstrate the effectiveness and has shown the following improvements. Beck *et al.* computationally discovered new prognosis features from stroma region in breast cancer histological images [1]. Pathologists were not looking at these features or criteria to grade the breast cancer. Treatments can vary across the patients. CAD tools allows us to precisely implement and improve the personalized treatments. We focused on processing and mining histological image data in this thesis. Histological image is one out of the many medical image types (MRI, mammograms, computed tomography) where automation is important (as discussed before in this Introduction). Histological images are taken using a brightfield microscopy or digital scanner of stained tissues and are assessed by pathologist to grade the cancer in a patient and to estimate the abnormality at the cellular and glandular level.

Tissues are translucent. Tissues are stained by chemicals to increase the contrast between the structures. For example, hematoxylin (H) and eosin (E) are used to create the contrast between epithelial (glandular) and nuclear and stromal (cytoplasm, connective tissue) structures. H&E is a common general purpose staining. A typical procedure for preparation of H&E stained biopsy in a pathology lab is depicted in the Figure 1.1 [1].

[1]Please refer electronic version for all color figures.

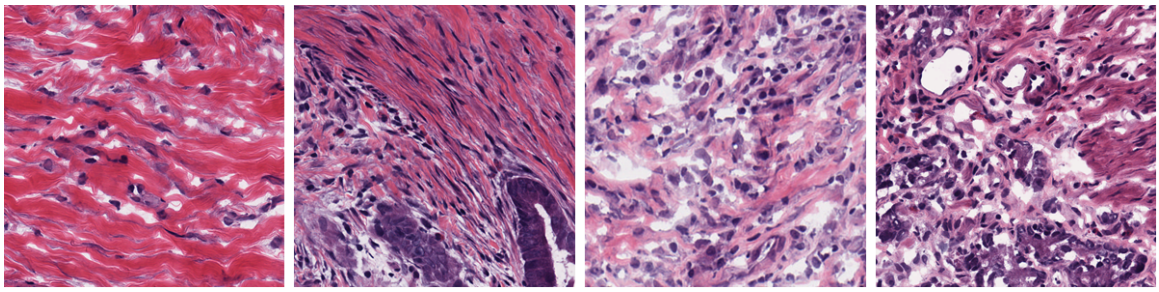


Figure 1.2: H&E stained histological image patch of (from left to right) bladder, colon, prostate, and stomach.

Digital images of H&E stained slides can be acquired through light microscopes with a camera attachment, or high throughput whole slide scanners. In this manner, histological images are available for pathologist to grade on the computer using various easy to use software such as imagescope [12] instead of observing it through microscope and grading. An example of different tissue H&E stained histological image is shown in Figure 1.2. Although, diagnosis using a digital image is not approved by medical regulatory authorities yet, it is soon expected to be approved because of its benefits.

Pathologist grade the severity of cancer through rigorous study of patient biopsy histological samples or images. The patient biopsy is prepared from the removed target tissue from the patient. The target tissue can be detected by non-invasive screening and detection techniques. The histological images are graded in the pathology labs. If the tumor is detected, it can be either malignant or benign. The amount of malignancy is measured through the grade and confined by the cancer grading gold standards such as Nottingham grading for breast cancer [13] and Gleason grading for prostate cancer [14]. The cancer severity and treatment emergency grows with the cancer grade.

The pathology problem solved by a computational resource is known as computational pathology (CP). In this thesis, we develop various image processing, machine learning and deep learning algorithms regarding computation pathology. We architect the CP framework inspired from [4]. In this, we add one of the important block called color normalization which was not included in the previous framework [4]. The modified CP framework is depicted in the

TH-1577_11610215

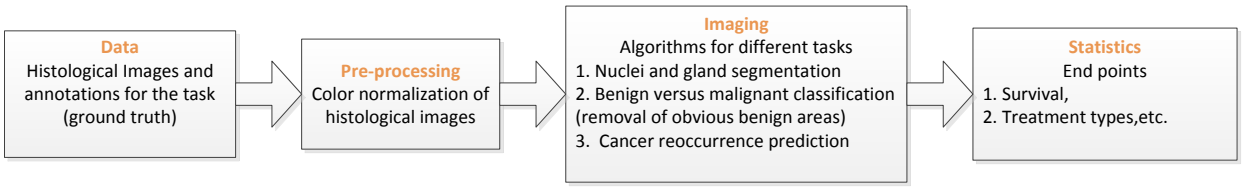


Figure 1.3: Modified computational pathology framework inspired from [4]

Figure 1.3.

A modified flow diagram of the CP framework shows four major parts: data and annotations, preprocessing, image analysis, and statistics. The data and its annotations for a particular pathology problem play significant role for the machine learning on the histological image data. Annotations are required for learning supervised or semi-supervised models, fine-tuning models learned in an unsupervised manner (e.g. stacked-autoencoders) and validation of the learned or designed algorithms. The tissue can be stained using general purpose H&E or immunohistochemistry (IHC). Given the histological images with color variations, they should be normalized to avoid any color variance while training the machine learning models or reduce the color uncertainty in developing image processing models. Hence, we introduced the color normalization as a pre-processing stage in CP framework. This stage of color normalization was shown to improve the performance of machine learning and deep learning models [15][16]. The last stage is the statistical prediction of end point of the different pathological and cancer stages like cancer grade, survival of a patient, prediction of probability of treatment type for a patient etc. This CP framework is very broad and there are various tasks under this framework. In rest of the thesis, we concentrate on few important problems regarding CP and propose a software aided solution for those.

1.1 Goals of thesis and organization

In this thesis, we developed algorithms to solve four different problems in computational pathology. The proposed algorithms were shown to outperform the state of the art.

Our first algorithm improves the stain separation in histological images [17]. We achieved

TH-1577_11610215

stain separation that is better than existing methods by extending a non-negative matrix factorization (NMF) formulation to include a biological prior of sparsity of stains at each pixel as an additional constraint. Preparation of the H&E stained biopsy slide consists of use of two chemicals – hematoxylin and eosin. Unmixing or color deconvolution or stain separation is a technique which decomposes the histological image into its constituent H and E components. It is useful to quantify the amount of each stain absorbed by the tissue specimen. This color deconvolution representation is fruitful in nuclei segmentation [8]. This is the process which decouples the color information from the intensity which can be useful color-invariant feature for automated analysis and indirectly suppress the color variation in the histology images. Additionally, various morphological features linked to diseases can be extracted from the stain separation maps. The application, stain *fusion* can use stain separation representation, where more than one physical stain can be observed in a single virtual image which can make the pathologist grading accurate and efficient as multiple stains can be available in single view. It has been demonstrated that it is easier to segment nuclei while working with a stain separation representation than with R,G,B representation [8]. We propose a stain separation technique in Chapter 2. We also used the stain separation representation to develop color normalization algorithm.

The color normalization algorithm suppresses the color variations in similar tissue compartments across histological images. Such unwanted color variation is a common problem in analysis of tissue histological images. It is caused by differences in color responses of slide scanners, raw materials and manufacturing techniques of stain vendors, and staining protocols across different pathology labs. This color variation can create difficulty in image interpretation by software trained on a particular stain appearance [16][18], and further add to already existing inter- and intra-expert variance in diagnosis and labelling among pathologists [9]. Color normalization is relied on the accurate stain separation representation. We proposed a color normalization method that standardizes the color appearance of a source image to that of a standard (target) image which effectively preserves the source structure [17]. In the stain separation representation, we predict the color appearance matrix and structure information

TH-1577_11610215

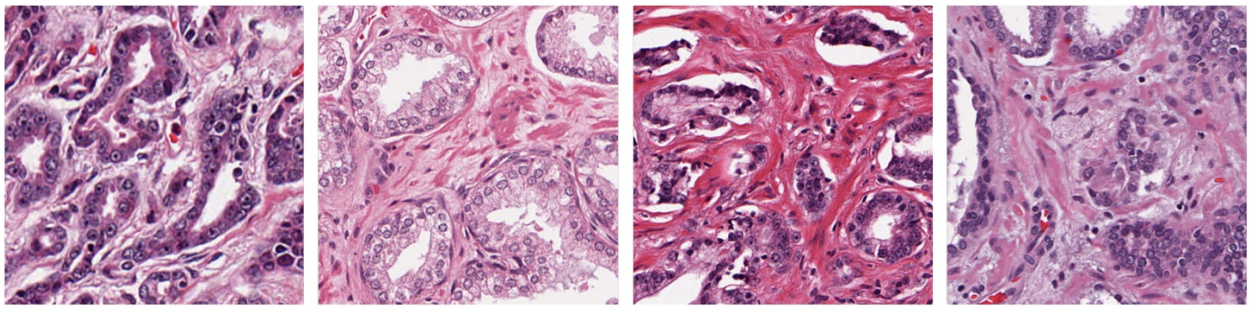


Figure 1.4: Color variation across different histological images

i.e. density maps. Since, the source image should be normalized to the expert chosen (or any) target, we predict the SNMF stain separation representation for both source and target. To normalize, we replaced the color appearance of the source by that of target and linearly normalized the density maps of source with the target. The proposed technique is shown to outperform the compared techniques. We also validated the impact of the color normalization for the epithelial-stromal segmentation in prostate tissue. Analysis suggest that color normalization improved the performance of different ways to classify epithelium (glands) from stroma (connective tissue). This establishes the color normalization as a useful pre-processing step for the histological image analysis. The color normalization contribution is included in Chapter 3.

Histopathological image consist glands (epithelial tissue) and non-glands (stromal tissue and debris) as shown in Figure 1.5. Glands are made up of epithelial cells. A typical structure of the gland consist of lumen, epithelial cytoplasm and epithelial nuclei as shown in Figure 1.5. Stroma consists of stromal nuclei and cytoplasm. Hence, nuclei are the basic unit of tissue and thus histopathological images. Nuclear morphology and morphometry are important phenotypic features that are indicative of the physiological state of the nucleus and cell. These features can help to quantify cell count, cell migration and interaction (morphodynamics). It is useful to identify cell phases (mitotic versus non-mitotic) by use of morphological features. Thus, nuclear segmentation is useful in various ways as described above. Nuclear segmentation is still not convincingly solved because of nuclei appearance variations and overlapping nuclei. We proposed an image processing pipeline that improves upon the watershed segmentation method [17]. We included morphological pre-processing and marking schemes for enhancement

TH-1577_11610215

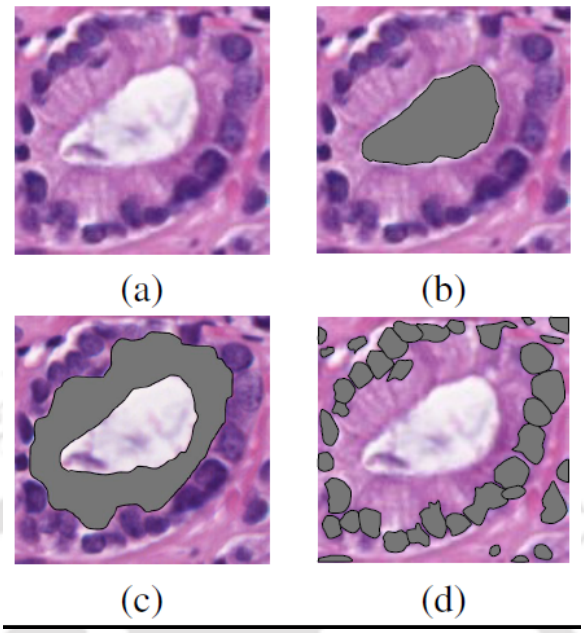


Figure 1.5: A typical gland structure as shown in (a). Outlined black are the (b) lumen area, (c) epithelial cytoplasm, and (d) epithelial nuclei [5].

of watershed segmentation. However, it was not able to solve the problem of overlapping and enlarged nuclei which is still an open challenge. So, we devised another algorithm to utilize the ability of deep convolutional neural networks (CNN) for learning three classes of pixels such that one of the classes, the ± 1 pixel boundary, helps segment touching and enlarged nuclei. A sufficient number of hand-labeled nuclei annotations were used to train the CNN. A simple post-processing method resulted in good segmentation accuracy and overlapping nuclei were resolved as separate objects. We propose these two techniques in Chapter 4.

Human epidermal growth factor receptor 2 (HER2) is a protein that can indicate the growth of malignant epithelial cells. HER2 over expression almost exists in 20 % of invasive breast cancer cases. This HER2 over expression is important in prognostic and predictive cancer model design. In clinical practice, pathologists observe and analyze morphological features to give a HER2 score from 0-3 per whole slide image. Details about the criteria for HER2 scoring will be described in the Chapter 5. We proposed a method combining deep learning and image processing based approach to achieve the state of the art HER2 score prediction as discussed in Chapter 5. Proposed HER2 predictor was ranked first in HER2 scoring contest [2]. In the

TH-1577_11610215

last Chapter, we make concluding remarks.

TH-1577_11610215

2

Stain Separation of Histological Images

Contents

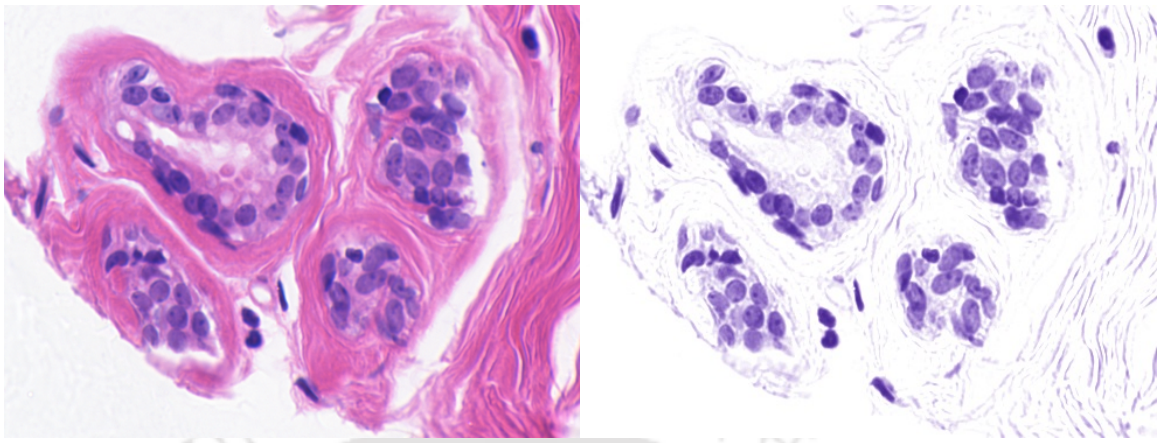
2.1 Introduction 9
2.2 Related work 10
2.3 Proposed SNMF based stain separation 14
2.4 Experiment and Results 16
2.5 Applications 25
2.6 Discussion 28

2.1 Introduction

The standard H&E stain highlights nuclei in purple and cytoplasm in pink color. This staining helps to visualize the structure of interest in the tissue by creating a perceptual contrast between structures. Different color stains are usually used to indicate the presence of different target proteins. As the hematoxylin dye stains all of the nuclei in blue color as shown in Figure 2.1, they can be easily segmented if we estimate the H-stain intensity from H&E stained image.

Immunohistochemistry (IHC) is the overarching name given to many different specialized stains that capture the expression levels associated with specific proteins. For example, CD31,

TH-1577_11610215



(a) H&E image **(b)** SNMF H stain

Figure 2.1: Breast H&E stained image patch (left) and predicted hematoxylin stain by proposed method

CD34, and CD105 expressions are found on vascular channels and some specific cell types (e.g. expression of CD34 on myofibroblasts) in human breast tumor tissue sections [8]. The degree of angiogenesis induced by a tumor in the tissue section can be evaluated from the immunostaining of these proteins[19]. Quantitative analysis of diaminobenzidine (DAB) IHC stain in the breast tissue is used to evaluate tumor growth and metastasis [20]. To assess the presence of their target proteins for their quantitative analysis, stain separation is necessary. Stain separation or color deconvolution or color unmixing are techniques that estimate individual stain components from stained histological image. Stain separation transforms the RGB channels of an image into an absorption or density map for each stain. This can also help in computational pathology, for example in quantitative analysis of cancer diagnostic factors such as count of nuclei, shape of glands [21][22] , gland differentiation, and in custom visualization of the structure information using false colors. We survey different stain separation methods in the next Section.

2.2 Related work

Since stain separation is estimation of the density map of each stain, it is instructive to understand the relation between RGB colors and stain density at each pixel. Stained tissue attenuates light in a certain spectrum depending on the type and amount of stain it has

TH-1577_11610215

absorbed. This relationship is captured in Beer-Lambert law as follows. Let $I \in R^{m \times n}$ be the matrix of RGB intensities, where $m = 3$ for RGB channels, and $n =$ number of pixels, and let I_0 be the illuminating light intensity on the sample (usually 255 for 8 bit images). Let $W \in R^{m \times r}$ be the stain color appearance matrix whose columns represent color basis of each stain such that r is the number of stains, and $H \in R^{r \times n}$ be the stain density maps, where the rows of which represent the concentration of each stain. Then, I can be written as follows[23]:

$$I = I_0 \exp(-WH) \tag{2.1}$$

Let V be the relative optical density then,

$$V = \log \frac{I_0}{I} \tag{2.2}$$

Hence, equation (2.1) becomes,

$$V = WH. \tag{2.3}$$

So, given an observation matrix V, the goal is to find stain color appearance matrix W and stain density map matrix H.

Color deconvolution [24] is one of the most widely used methods to separate light absorbing stains using the relationship between stain amount and light absorption given by Beer-Lambert Law [25]. The authors suggested measuring the stain color appearance from control slides with a single stain per slide, that is empirically estimating W (an example shown in Figure 2.2). Stain density maps H can then be computed by pre-multiplying the optical density with Moore-Penrose pseudoinverse of color appearance matrix W. This calibration-based method has a strong limitation that it is only valid for images using the same staining and imaging protocol whereas color variation in histology is almost inevitable as described in Chapter 1.

Non-negative matrix factorization (NMF) [26] was used in one of the pioneering unsupervised frameworks for stain separation proposed in [6] that derived image-specific stain color basis. The non-negative constraints on the stain density and color appearance matrix capture the important property of stains that can only absorb but can not emit light, which makes

TH-1577_11610215

both their color basis and density non-negative. It, therefore, solves the following problem:

$$\min_{W,H} \frac{1}{2}\|V - WH\|_F^2, \text{ such that } W, H \geq 0 \tag{2.4}$$

Estimation of both factors jointly is a non-convex optimization problem which may converge to a local optimum instead of the global optimum and give undesirable stain vectors [27].

Besides NMF[6], other popular unsupervised decomposition methods are based on independent component analysis (ICA) [28], singular value decomposition (SVD) [7], and blind color decomposition (BCD) [23]. As per the relative root mean square error (rRMSE) between estimated color basis and ground truth in [23], BCD outperforms NMF by 20% to 40% and ICA by 3 to 5 folds. Though BCD has a good performance, it is unclear how to fix some of its hyper-parameters. For example, the authors proposed to use a piece-wise linear decomposition instead of a linear one in case of a poor stain-cluster separation but they have not provided a quantitative threshold for fisher criteria [29] to determine 'poor' clusters. SVD [7] has a closed-form solution which can be computed efficiently, although its performance deteriorates when images contain uneven proportions of each stain [30].

We proposed a solution for stain separation which preserves biological structure information by modelling stain density maps based on the following properties [31][32]:

- Non-negativity: It is essential to assume that a stain density and its associated optical density can either be absent (zero) or present (positive) at a given location (pixel), but it cannot be negative. Negative stain density would mean emitting light.

- Sparsity: We assume that the tissue behind most pixels represent one of a few types of biological material (such as nuclei or cytoplasm), characterized by their effective stain, allowing each pixel to be modeled as a sparse mixture of the constituent effective stains. Effective stain is illustrated in Figure 2.2. A biological structure, such as a nucleus, can bind to more than one chemical stain, such as in case of H&E staining, but we assume that their relative proportions are fixed for a given biological structure within one image. We term this mixture of stains that is uniquely present in a type of biological structure

TH-1577_11610215

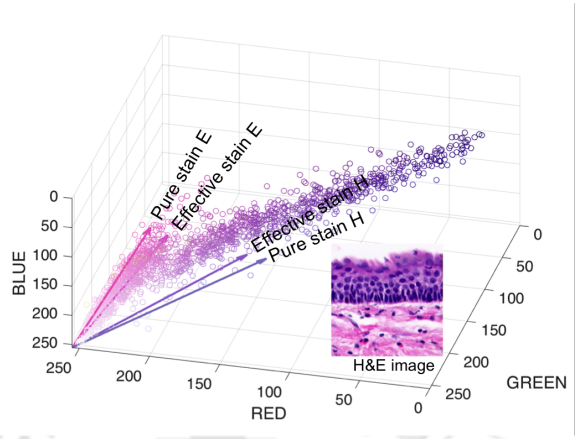


Figure 2.2: Effective stain graphical representation: cell nucleus stained with effective H, which is a mixture of a large amount of pure H and a small amount of pure E; by contrast, connective tissue stained with effective E, which is a mixture of a large amount of pure E and a small amount of pure H. H&E image copyright of G.Landini, University of Birmingham, U.K., reproduced with permission. Pure stain vectors are "built-in" vectors determined experimentally (details refer to http://www.mecourse.com/landinig/software/cdeconv/cdeconv.html). Effective stain vectors are determined by our SNMF decomposition.

as an effective stain. Thus, we estimate the effective stain (henceforth stain) basis vectors and not the pure stain basis vectors.

- Soft-classification: We can allow a small proportion of pixels on the boundaries (edges) of two biological materials to capture more than one stain, which also helps smoothen the optimization cost functions.

We do this by first casting the stain separation problem as a non-negative matrix factorization (NMF) to which we add a sparseness constraint, and refer to it as sparse non-negative matrix factorization (SNMF). An additional advantage of this formulation is that the color basis is determined in an unsupervised manner without requiring manual annotation of different areas of pure stains. We have compared our algorithm quantitatively and qualitatively with competing stain separation techniques [6][7][23]. For quantitative comparison, we generated ground truth by having pathologists annotate regions of different pure stains, and using that to compute median color basis and density maps for the stains. We validated the stain separation methods by comparing estimated color basis and stain density maps with the ground truth. We also compared the pathologists subjective grading of the NMF and SNMF. SNMF outper-

TH-1577_11610215

formed the NMF significantly, thus validating the role of sparsity in stain density estimation. In addition, we found that SNMF was robust to variations of a hyperparameter in the defined range that controls sparseness.

Addition of a sparsity constraint to NMF, as we explain in Section 2.1, reduces the solution space as well as adds another biologically-concordant principle to the model. A sparse formulation for stain separation was recently reported in [33]. However, it is only an extension of color deconvolution method [24] and still requires W to be experimentally determined, which limits its applicability over large histological image dataset with considerable color variation.

In this Chapter, we have compared the proposed SNMF-based stain separation with NMF-based [6] and SVD-based [7] methods. ICA[28] was dropped from comparison because its performance was reported to be far worse than that of NMF and BCD in [23]. BCD was dropped from comparison because neither all of the details required for its implementation nor its source code were made available in [23]. However, we indirectly compared it with SNMF based on their results relative to NMF as reported in [23].

2.3 Proposed SNMF based stain separation

In the SNMF stain separation, we add a sparsity constraint on NMF to capture the biological principle of discreteness of biological structures. That is, each biological structure, such as a nucleus, has finite and connected spatial extent, which is characterized by an absorption spectrum or effective stain [31][32]. In this Section, we provide details of implementation of SNMF. The source codes of SNMF are available for the academic use at `https://github.com/abhishekvahadane/CodeRelease_ColorNormalization.git`.

We first convert a given RGB image to optical density using equation (2.2) based on Beer-Lambert law. Then, we add a sparseness constraint to equation (2.4), i.e. we propose an improved NMF cost function for stain separation by including $l1$ sparseness regularization on stain mixing coefficients H_j, for stains indexed by $j = 1, 2, ., r$, which yields,

$$\min_{W,H} \ \frac{1}{2}\|V - WH\|_F^2 + \lambda \sum_{j=1}^{r} \|H(j,:)\|_1, \ \ W, H \geq 0, \ \ \|W(:,j)\|_2^2 = 1 \tag{2.5}$$

TH-1577_11610215

where λ is sparsity and regularization parameter. Note that the additional constraint on W in equation (2.5) is to suppress many equivalent solutions of type $(W/\alpha,\ \alpha H)$, $\alpha > 0$.

The joint non-convex optimization in equation (2.5) is solved by alternating between W and H which optimizes one set of parameters while keeping the other fixed starting with an initialization of W by random elements from the training set (V) i.e. the RGB optical density of two randomly selected pixels corresponding to two columns of W of the histological image, as follows.

$$\text{For fixed } W,\ \ \hat{H} = \min_{H} \frac{1}{2}\|V - \hat{W}H\|_F^2 + \lambda\|H\|_1,\ \ H \geq 0 \tag{2.6}$$

$$\text{For fixed } H,\ \ \hat{W} = \min_{W} \frac{1}{2}\|V - W\hat{H}\|_F^2,\ \ W \geq 0, \|W(:,j)\|_2^2 = 1 \tag{2.7}$$

This cost function is equivalent to a well-established dictionary learning objective [34], but with additional non-negative constraints on dictionary atoms W and coefficients H. These two alternating steps are called sparse coding for H and dictionary learning for W, and summarized as follows:

Sparse coding or estimation of H equation (2.6) with fixed $\hat{W}$ is an $l1-$regularized linear least-squares problem. A number of recent methods for solving this type of problems are based on coordinate descent with soft thresholding [35] and LARS-LASSO algorithm [36]. However, as pointed out in [37], when the columns of the dictionary are highly correlated, coordinate descent is found to be empirically slow. Hence, we use LARS with efficient Cholesky-based implementation which provides robust and accurate solution without the need for an arbitrary stopping criterion.

Dictionary learning or estimating W is done by using parameter free block-coordinate descent with warm restart [38] which does not require learning rate tuning and still guarantees convergence to a global optimum for convex optimization in equation (2.7). We use publicly available SPArse Modelling Software (SPAMS) [37] for sparse coding and dictionary learning. It should be noted that although SPAMS has been optimized to solve equation (2.5), the computational expense for such an iterative solver is too costly for the application to large size

TH-1577_11610215

whole slide images.

Once the solutions W and H are obtained, the separated color components H and E for H&E stained image I can be obtained as,

$$I_h = I_0 \exp(-W(:,1)H(1,:)) \tag{2.8}$$

$$I_e = I_0 \exp(-W(:,2)H(2,:)) \tag{2.9}$$

For H&E stained images, hematoxylin stain vector among the columns of the final solution matrix $\hat{W}$ was determined based on its lower relative attenuation of blue channel as compared to eosin.

2.4 Experiment and Results

The proposed SNMF-based stain separation was extensively tested qualitatively and quantitatively. All of the datasets used for testing were publicly-available benchmark data, which allow an easy and fair comparison with other competing methods. SNMF was quantitatively validated against NMF [6] and SVD [7] based stain separation on the samples from four different tissue types: stomach, prostate, colon, and bladder, available at The Cancer Genome Atlas (TCGA) data portal[1]. Blind color deconvolution (BCD) [23] does not provide source code for public access and thus we only quote their results and make a comparison relative to NMF.

2.4.1 Ground truth generation

While multiple ideas have been proposed for generating ground truth for stain density maps and their color appearance models, such as scanning slides with pure stains separately [24][6] or scanning a single tissue section using chemical de-staining [39], these methods lead to errors because of stain leakage and co-location such as base-level presence of eosin even in areas covered primarily by hematoxylin. We found the ideas proposed in [23] to be closest to capturing the actual manifestation of the stain appearances in the context of real tissue structure. Following

[1] https://tcga-data.nci.nih.gov/tcga/

TH-1577_11610215

Table 2.1: Comparison of rRMSE of stain color basis matrices (W) based on pathologists annotations, NMF [6], SVD [7], and the proposed SNMF with respect to ground truth W_{gt}

Tissue	n samples	Pathologist	NMF	SVD	SNMF
Stomach	71	0.048	0.204	0.094	0.059
Prostate	70	0.034	0.200	0.074	0.049
Colon	70	0.046	0.140	0.061	0.062
Bladder	84	0.042	0.163	0.057	0.041

this approach, experts on our team marked a subset of tissue regions that *should* be stained by hematoxylin (i.e. cell nuclei) separately from a subset that *should* be stained by eosin (i.e. stromal matrix) in n randomly selected non-overlapping patches of size 1000×1000 for each tissue type (about the size of a microscopic field of view (FOV), each FOV covering physical area of $0.25mm \times 0.25mm = 0.0625mm^2$). The median of the selected pixels of each stain from each FOV in RGB space formed the reference color bases (I_k^{median}) for that stain and FOV (median in RGB space aid to filter out non-stain related noise such as small dusts or holes) and median of all n FOVs formed the ground truth color bases (I_{gt}^{median}) for that particular tissue. The W_k and W_{gt} corresponding to I_k^{median} and I_{gt}^{median} are computed from equation (2.1) followed by normalization of each column vector to unity. The variance of W_k across FOVs can be attributed to variation of stain quality over the tissue sections as well as intra-pathologist variation. Given the color appearance matrix, the stain density can be generated using color deconvolution method [24] as:

$$H_{gt} = (W_{gt}^T W_{gt})^{-1} W_{gt}^T V \tag{2.10}$$

We also use the same comparison metrics as those in [23] to evaluate our stain separation.

2.4.2 Comparison of stain color appearance W

Relative root mean square error (rRMSE) has been used for statistical comparison of different stain separation methods on their ability to estimate stain color basis close to the ground truth basis W_{gt}. This is a measure of how close the estimated color appearance W is to the

TH-1577_11610215

ground truth W_{gt}, [23], as

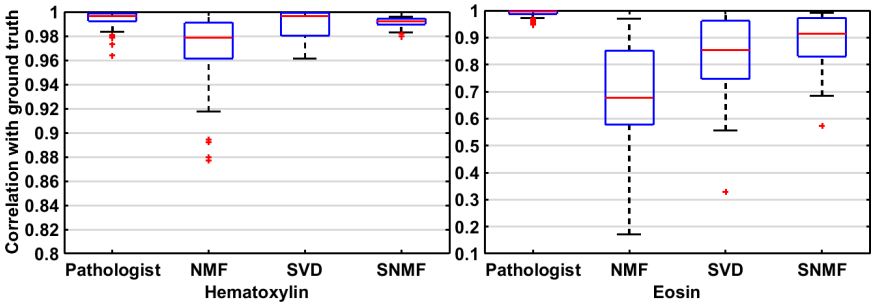
$$rRMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n} \frac{\|(W_k - W_{gt})^T(W_k - W_{gt})\|}{\sqrt{\|(W_k)^T(W_k)\| \cdot \|(W_{gt})^T(W_{gt})\|}}} \tag{2.11}$$

where $\|\cdot\|$ denotes the matrix trace. Table 2.1 summaries the results for 70 to 84 randomly selected fields of view (FOVs) for each of the four tissue types. SNMF outperformed both NMF [6] and SVD [7]. Our rRMSE improved about 3 folds as compared to that of NMF for all four tissue types and about 20-30% over SVD for three tissue types except colon where the performance of both methods was comparable. It is worth mentioning that our improvement over NMF was much larger compared to that of BCD method, since in [23] authors reported only a 20% to 40% increase over NMF. The pathologist performed better than all computational methods, as expected, because of the ground truth generation process that was based on pathologists color models to begin with.

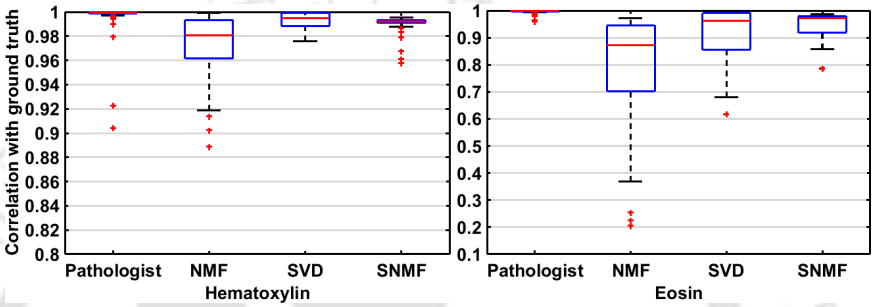
2.4.3 Comparison of stain density maps H

Although stain color appearance is vital for color normalization, comparing that alone is not sufficient to quantify the accuracy of individual stain components, because the error in hematoxylin and eosin color appearance estimation can propagate to the error in estimating density maps H. We assess the accuracy of the estimated stain density maps H by computing Pearson correlation index between them and a map obtained using ground truth. As shown in the box plots of Figure 2.3, while all of the three methods performed well in estimating hematoxylin density with average correlation >0.95 in almost all cases, SNMF improved eosin quantification significantly over the other techniques. Wilcoxon signed rank tests were statistically significant at $p < 0.01$ for all cases except colon for which the performance was approximately similar for all techniques. An example of stain separation for each tissue type using the methods compared is shown in the Figures 2.4-2.7. The results indicate the estimated stain components by SNMF are more visually similar with the ground truth stain components.

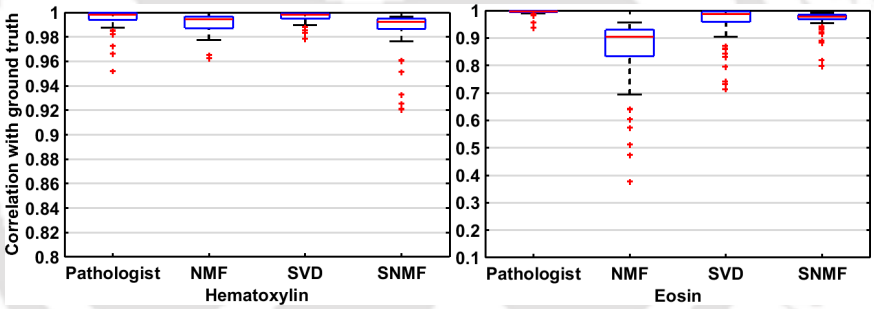
TH-1577_11610215



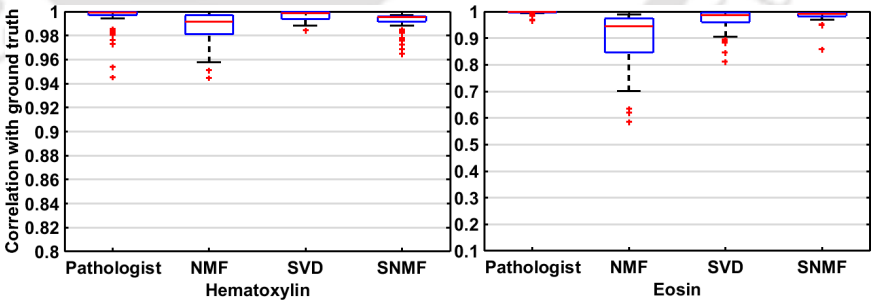
(a) Stomach



(b) Prostate



(c) Colon



(d) Bladder

Figure 2.3: Density map comparison of pathologist, NMF [6], SVD [7], and proposed SNMF against ground truth H_{gt}.

TH-1577_11610215

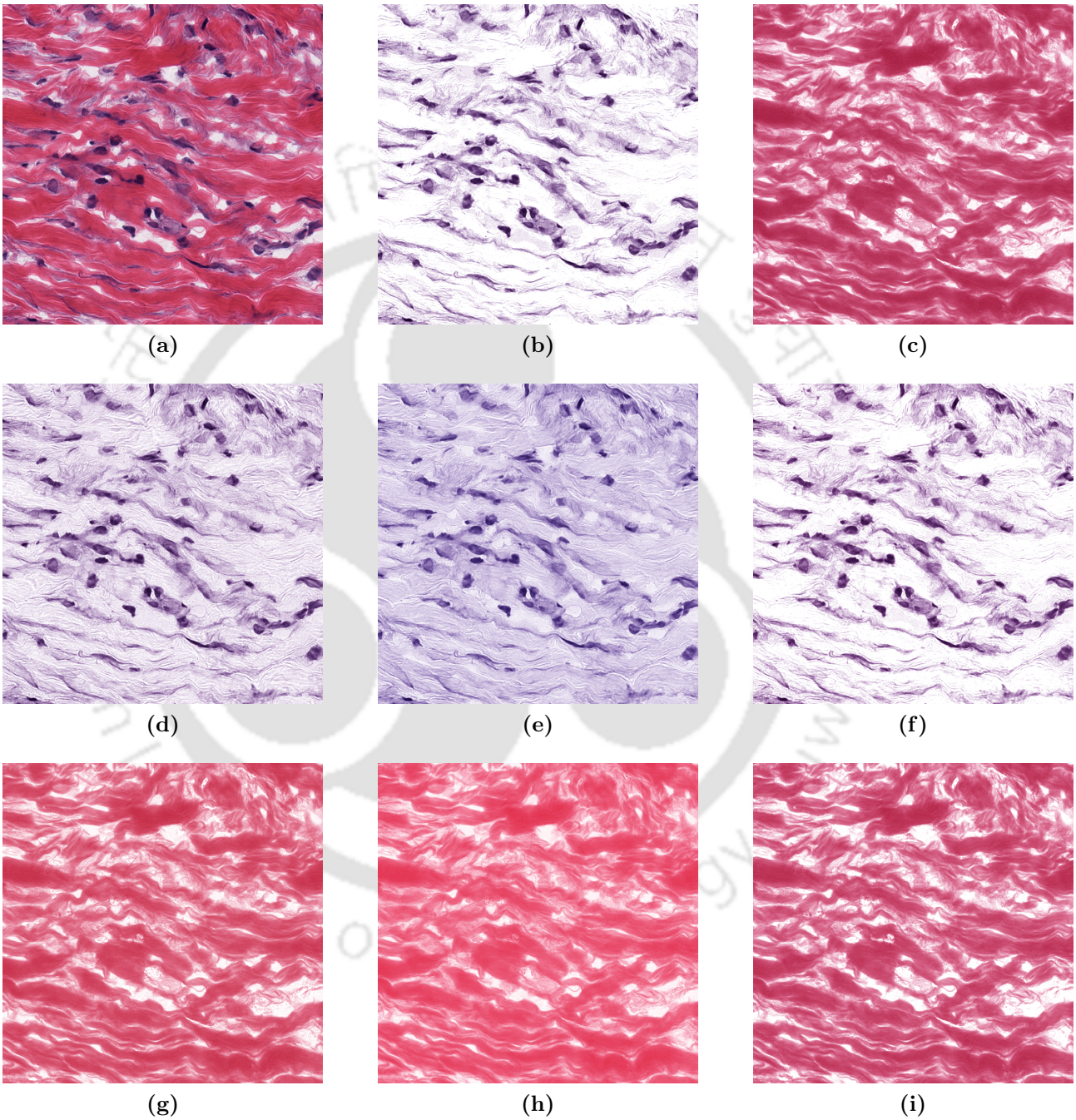


Figure 2.4: Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the bladder tissue input (a).

TH-1577_11610215

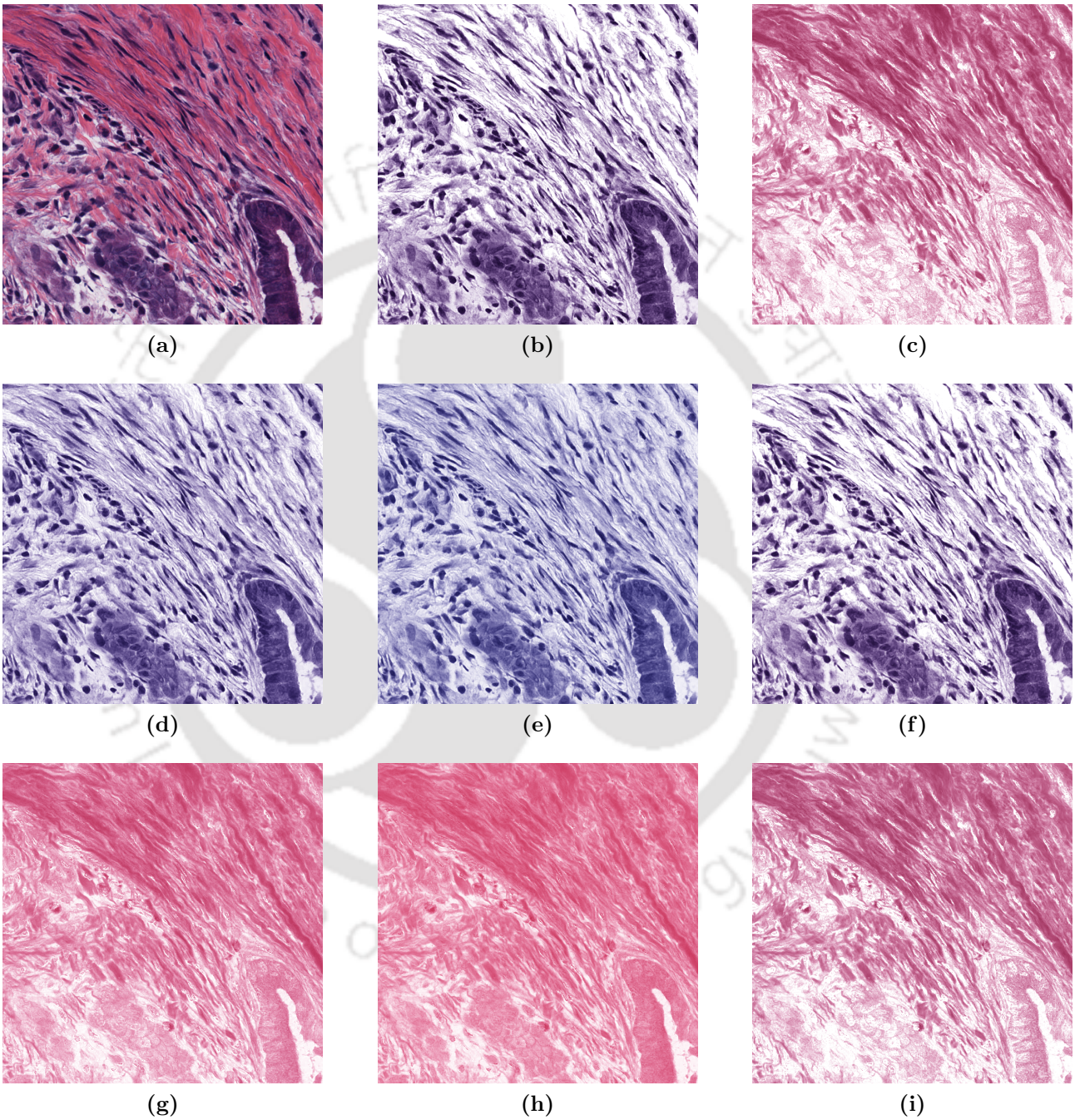


Figure 2.5: Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the colon tissue input (a).

TH-1577_11610215

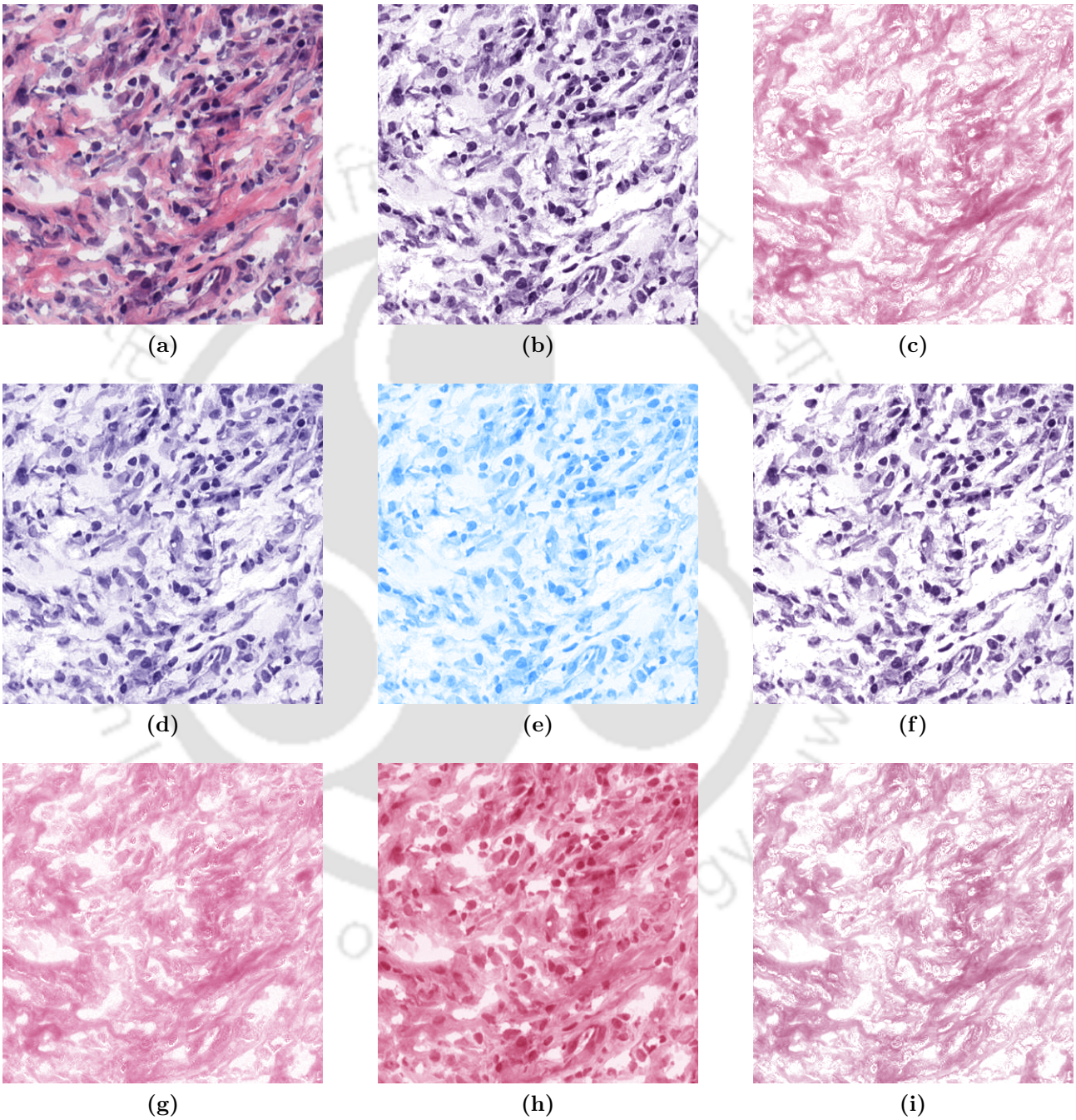


Figure 2.6: Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the prostate tissue input (a).

TH-1577_11610215

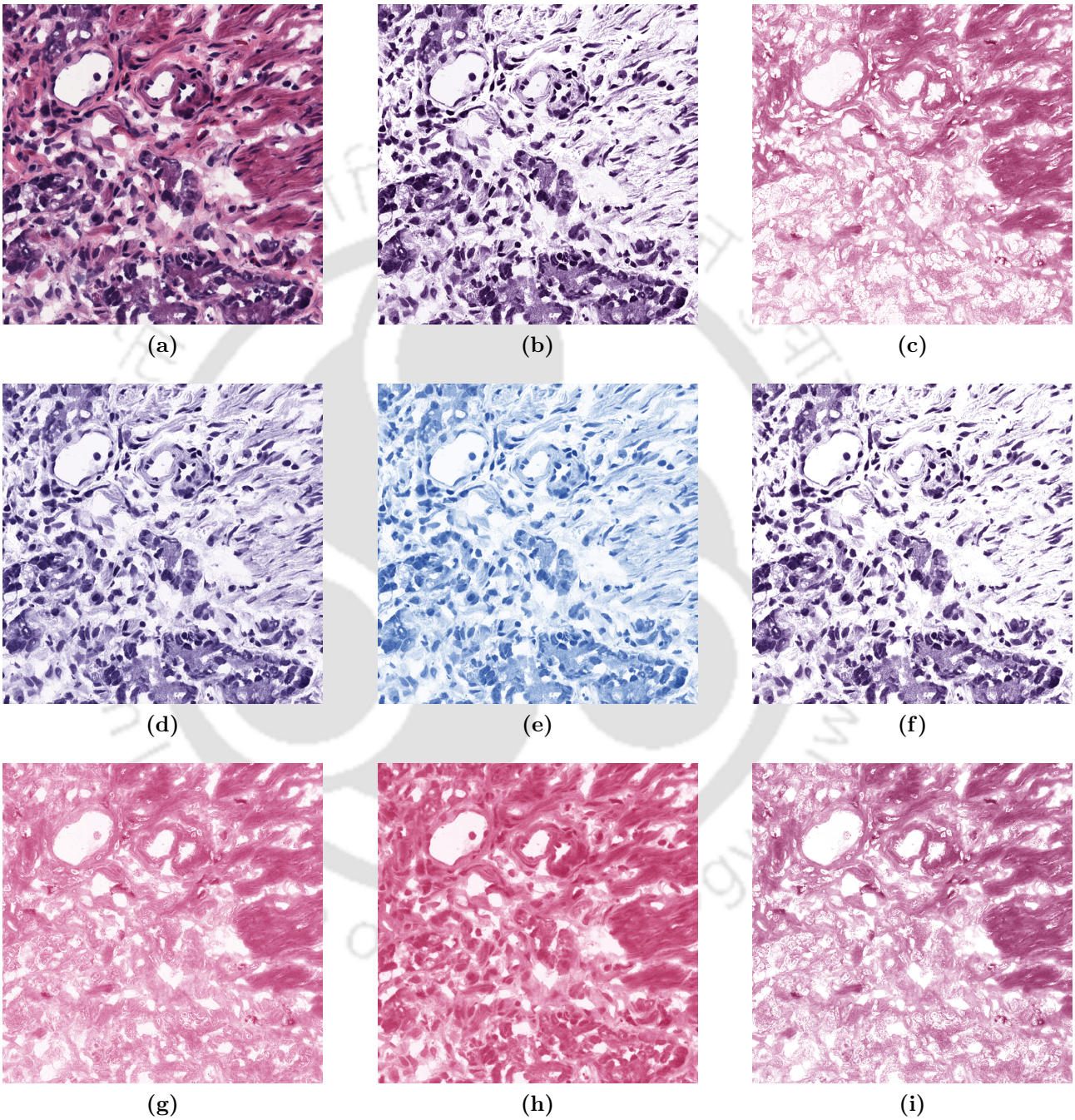
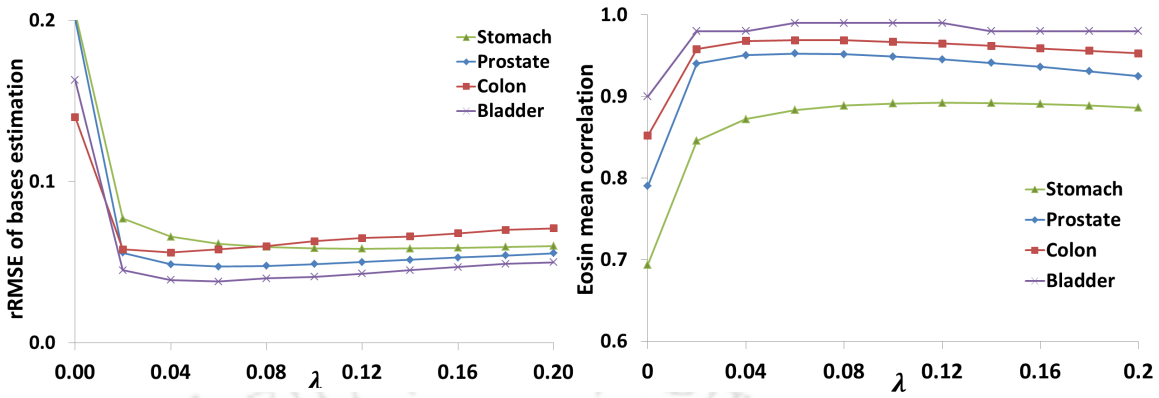


Figure 2.7: Stain separation by pathologist (b, c), NMF [6](d, g), SVD [7](e, h), and proposed SNMF (f, i) on the stomach tissue input (a).

TH-1577_11610215



(a) rRMSE of W **(b)** Pearson correlation of eosin

Figure 2.8: Sensitivity of SNMF to hyper-parameter λ. Note that NMF is equivalent to $\lambda = 0$.

2.4.4 Sensitivity analysis of sparse coefficient λ

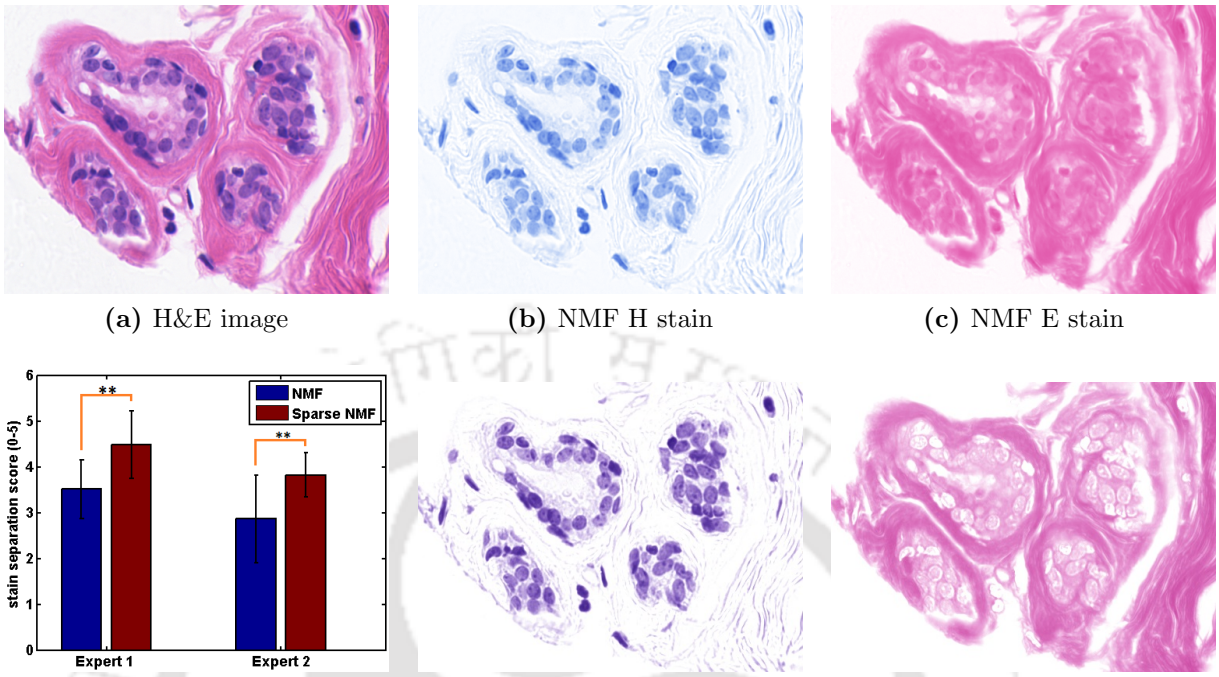
We analyzed sensitivity of SNMF performance to different values of the hyper parameter λ and found a wide range around 0.1, the value set for comparison to other techniques, in which the performance was largely the same. We varied λ from 0 to 0.2 in steps of 0.02 and computed rRMSE of the estimated stain color matrix and the Pearson correlation of the estimated stain density maps. Figure 2.8 shows the sensitivity analysis of proposed SNMF across λ. Only eosin detection results are presented here as hematoxylin detection was close to 1 regardless of the value of λ and tissue type.

Setting $\lambda = 0$ reduces SNMF to NMF. As we increased λ from 0 to 0.02, we observed a rapid reduction in rRMSE as well as a sharp increase in the correlation index, which suggests a key role of sparse regularization in improving stain separation accuracy. The performance remained largely flat around 0.1 (chosen default value). However, larger values of λ led to slight decrease in performance because increase in sparsity most likely erased local structures such as some stain co-location or higher densities.

2.4.5 Qualitative comparison by expert grading

We also evaluated SNMF-based stain separation qualitatively using expert grading on 50 H&E breast cancer image dataset made available in [40]. Two experts independently evaluated

TH-1577_11610215



(a) H&E image **(b)** NMF H stain **(c)** NMF E stain

(d) Expert scores **(e)** SNMF H stain **(f)** SNMF E stain

Figure 2.9: Application of NMF and SNMF based stain separation on a sample image.

the quality of stain separation using scores in the range of 0-5 (0 poorest, 5 perfect). Figure 2.9d shows bar plot of mean±standard deviation for the expert scores. It shows that on average SNMF was scored 4.48 and 3.82 respectively by the two pathologists, and deemed suitable for clinical use. In comparison, NMF only achieved average scores of 3.52 and 2.86 respectively by the two experts, which are significantly lower than SNMF ($p < 0.001$ using Wilcoxon signed-rank test). An example of stain separation results on sample image is shown in Figure 2.9, which illustrates that SNMF achieved a much clearer separation with less distracting background stains (e.g. less stroma in H channel and less nuclei in E channel), demonstrating the crucial role of sparse regularization.

2.5 Applications

We discuss some possible stain separation applications here. Stain density maps can be used in a variety of different tasks in computational pathology. We discuss two promising applications of stain separation – nuclei segmentation and stain fusion.

TH-1577_11610215

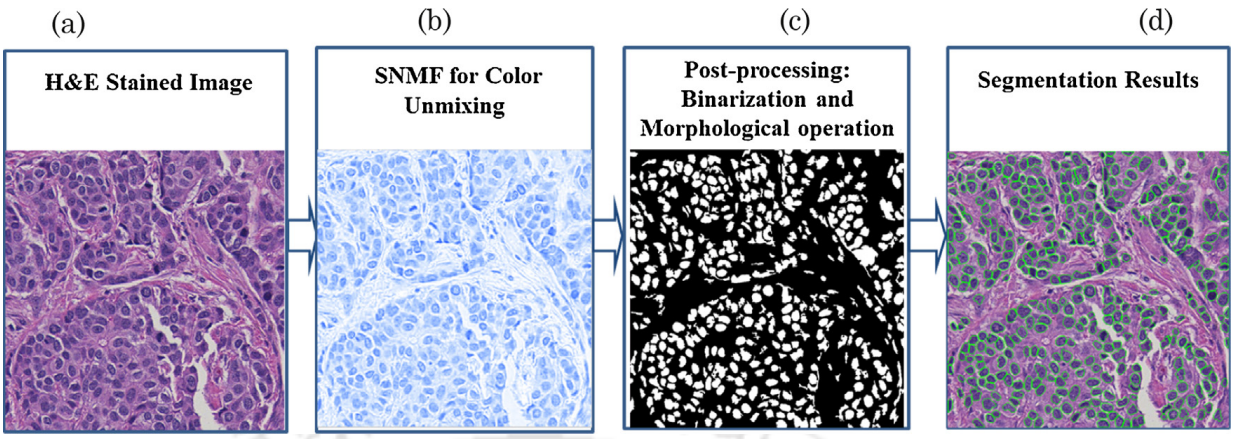


Figure 2.10: The flowchart of SNMF based approach for nuclei segmentation [8]. The original image (a) is decomposed into H stain component (b) with SNMF algorithm. The final segmentation result (d) is generated from binary image (c) after post-processing operation on the decomposed H stain component (b).

2.5.1 Nuclei segmentation

SNMF based stain separation representation can be easily used for the task of nuclei segmentation. Xu *et al.* proposed a stain separation method and showed its application on the nuclei segmentation for the immunohistochemistry (IHC) stained images. Our stain separation work [31] was developed in parallel with Xu et al.s [8] and both use sparse non-negative matrix factorization (SNMF) for stain separation, with a minor difference that the latter also penalizes l2-norm of the factors. However, there are other major differences between the two independent works. While the main objective of performing stain separation in the [31] is the color normalization (will discuss in the Chapter 3), Xu *et al.*s objective was nuclear segmentation in breast tissue images.

A simple nuclear segmentation approach was presented by Xu *et al.* with the effectiveness of stain separation-density maps. The framework is shown in the Figure 2.10 which shows that the SNMF based stain separation is followed by selection of hematoxylin component. Then, simple binarization and post-processing operations result in the fair nuclei segmentation. However, overlapping nuclei were not resolved by this approach.

TH-1577_11610215

2.5.2 Stain fusion for head and neck cancer

We also studied the application of stain separation for the stain fusion. The goal of stain fusion in histological images is to visualize the different stains (mainly, different protein markers) on a single histological image. Visualizing different stains as a single view can help pathologist diagnose the patient case effectively and accurately. Plasminogen activator inhibitor type-1 (PAI-1) and vascular endothelial growth factor (VEGF) are regulated by hypoxia and irradiation (the deciding factors for radiation therapy). A study shows that concentration of tumor-specific and host-specific VEGF during fractionated irradiation could provide considerably divergent information for the outcome of radiation therapy in the case of mice head and neck cancer [41]. PAI-1 staining was reported to be observed in the extracellular matrix, the cytoplasm of fibroblast-like stroma cells, and individual tumor cells. The co-localization of such IHC staining patterns (PAI and VEGF) provide a way to visualize both stains on the single image. Fusion and visualization techniques can help experts in taking effective and confident clinical decisions. The PAI and VEGF IHC stains are generally applied on adjacent sections of tissues so that they can be co-localized using image analysis. PAI and VEGF stains should be counter stained with hematoxylin for clinical analysis. Hence, the goal is to computationally co-localize the two adjacent sections of tissue stained with H&PAI and H&VEGF. We can do this by first registering both histological images. Then, we separate PAI and VEGF stains from hematoxylin and fuse separated PAI and VEGF to create informative fused image.

Given the H&PAI ($I1$) and H&VEGF ($I2$) images, the SNMF stain separation 2.3 can be used to get stain separation of $I1$ and $I2$. Let H_{pai} and H_{vegf} to be the stain density maps corresponding to PAI and VEGF. Let H_h^{pai} and H_h^{vegf} be the stain density maps of hematoxylin. First of all, the two hematoxylin maps are non-rigidly registered to align the structures. Since the two images are the sections of the two adjacent slices, the structures are assumed to very similar and can be registered. The registration transformation thus obtained is used to register H_{pai} and H_{vegf}. Now, all of the images under consideration are registered and ready for fusion. We use Otsu threshold on the H_{pai} and H_{vegf} to localize the PAI and VEGF (shown in Figure 2.11a-2.11d). Then, we co-localize both the stains on one of the hematoxylin components by

TH-1577_11610215

pseudo-coloring the spatial locations by different colors to indicate the PAI and VEGF presence Figure 2.11e and the set of pixels where both stains overlap can be visualized by distinct color (here, yellow). Another simple way of fusion would be to apply a max operator on the H_{pai} and H_{vegf} to retain the active components in both. Let the max fused image to be H_{fused}. We can further fuse the H_{fused} with the hematoxylin using the Beer-lambert law as follows (result shown in Figure 2.11f),

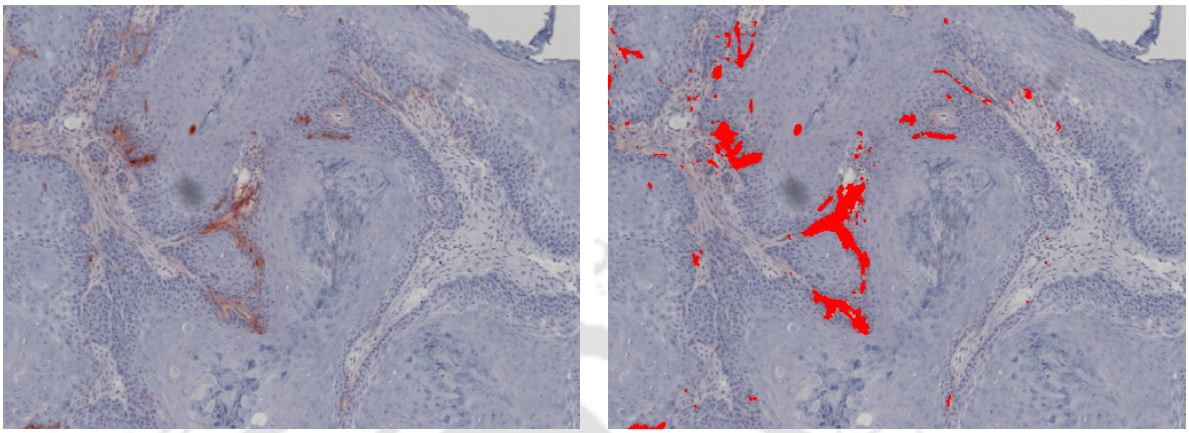
$$I = I_0 exp(-W[H_h^{pai}; H_{fused}^{vec}]) \tag{2.12}$$

where [.;.] is row concatenation operation. W can be selected to be either corresponding to H&PAI or H&VEGF for simplicity. In this fusion, we employed a max operation for simplicity. There are other standard fusion techniques which can be explored to obtain the sophisticated fusion results. Thus, stain separation can be important for such histological multi-stain fusion to create a single informative image.

2.6 Discussion

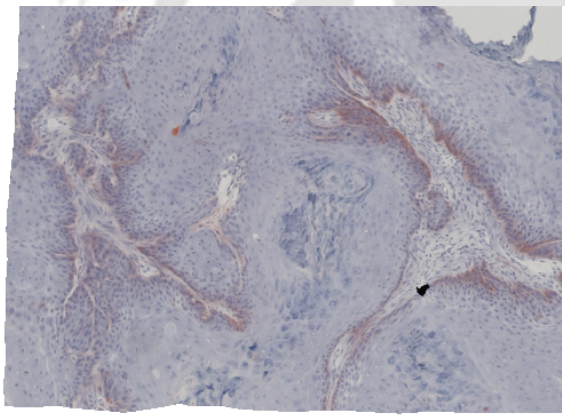
Stain leakage is a major factor for stain co-localization in histology images. It is caused by more than one target protein being present in the same area. For instance, in H&E stained specimen, hematoxylin can leak into structures that should be ideally colored by eosin only for some infinitesimal and unintended binding and vice versa. This is also a reason that the reference color of eosin-specific structure (e.g. stroma) in eosin-alone staining is different from that of in H&E combination. Our SNMF utilizes a sparse constraint to decompose an image into different tissue structures represented by effective stains, rather than the physical amount of hematoxylin and eosin like NMF (an example shown in Figure 2.9). We introduced the term effective stain to describe stain combination (usually mostly an intended and some amount of unintended stains) uniqueness to different biological structures. This clear separation of effective stains can lead to better tissue quantification as it helped nuclear segmentation due to less distraction from residual background (unintended) stains. The fusion of stains is also another potential application of this technique.

TH-1577_11610215

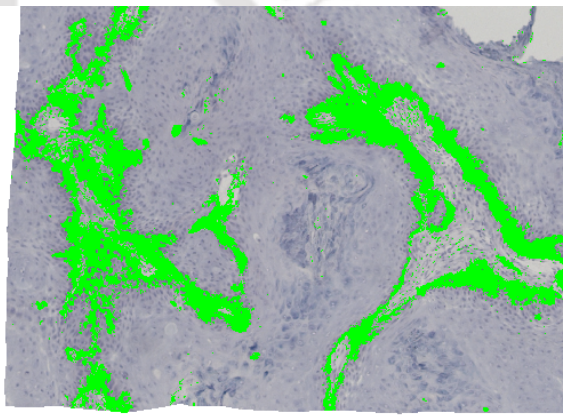


(a) H&PAI

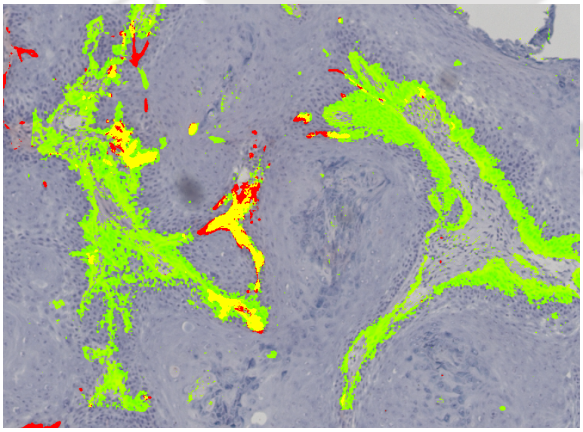
(b) Detected PAI through SNMF in red



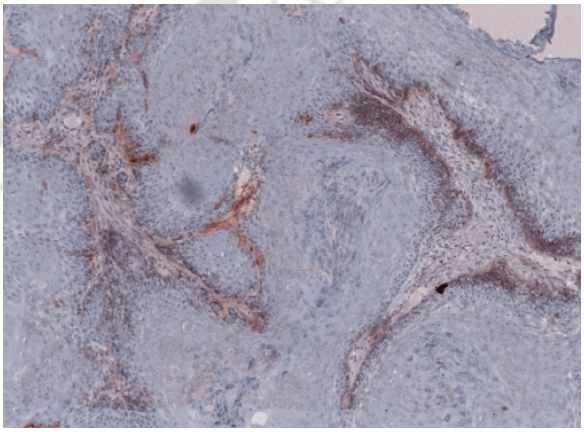
(c) H&VEGF (registered to H&PAI)



(d) Detected VEGF through SNMF in green



(e) Fused PAI and VEGF through pseudo-coloring



(f) Fused PAI and VEGF through Inverse Beer-Lambert law

Figure 2.11: SNMF application to stain fusion

TH-1577_11610215

All the data used is H&E stained histological images, which is mainly due to two reasons: 1) H&E is the most widely used histological staining in pathology labs, whose popularity is based on its comparative simplicity and ability to demonstrate on large number of different tissue structures [42], and 2) we preferred to evaluate our proposed algorithms on open source data, most of which is H&E stained, to allow a fair comparison with other competing methods. Nevertheless, we have performed some preliminary testing on the immunohistochemistry (IHC) staining such as H&DAB (DAB stands for Diaminobenzidine). Results in Figure 2.12 show that SNMF can achieve a good structure separation of H&DAB images, yet the resulted stain density map of H and DAB does not linearly correlate to the actual amount of DAB stain due to refraction. Also, the brown DAB is not a true absorber of light, but a scatterer, and has a very broad and featureless spectrum [43]. Thus, Beer-Lambert law of absorption is not valid for brown DAB. Besides, immunohistochemical DAB stains can selectively bind to cancer target structures and hence only occupy a small region of the whole image, in which case the current simple grid patch sampling scheme may miss it. One possible solution for efficient exploration of WSI for locating cancer region is the dynamic sampling approach presented in [44] where small patches were dynamically sampled using a score function and Voronoi diagram. A more thorough evaluation of stain separation on IHC stained histology needs to be done.

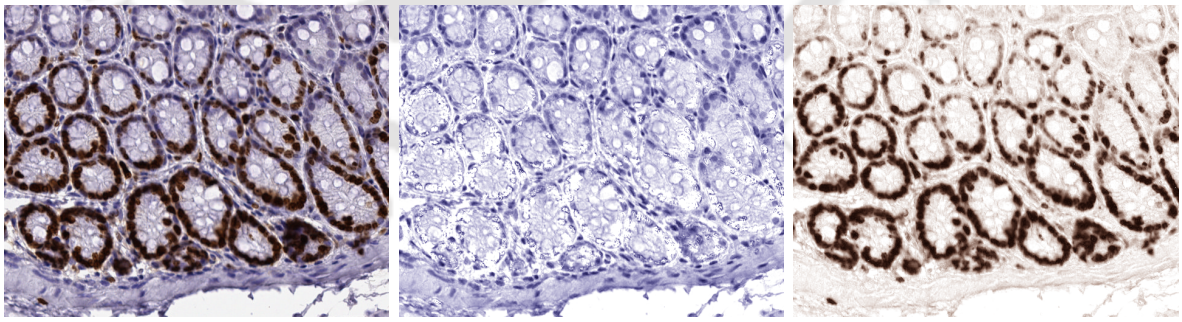


Figure 2.12: Proposed SNMF on IHC (H&DAB) stained image of head and neck tissue: (left) input image, (middle) Hematoxylin, (right) Diaminobenzidine.

TH-1577_11610215

3

Color normalization of histological images

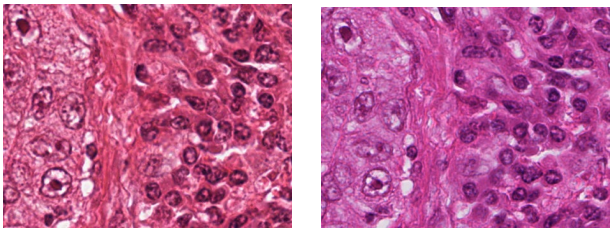
Contents

3.1 Introduction 31
3.2 Background and Related Work 33
3.3 Proposed Method 35
3.4 Experimental Results 38
3.5 Summary and Discussion 46

3.1 Introduction

A common problem in analysis of tissue samples is unintended color variation due to differences in color responses of slide scanners (as shown in Figure 3.1), raw materials and manufacturing techniques of stain vendors, and staining protocols across different pathology labs. This creates difficulty in image interpretation by software trained on a particular stain appearance [16][18], and further adds to already existing inter- and intra-expert variance in diagnosis and labeling among pathologists [9].

TH-1577_11610215



(a) Aperio scanner **(b)** Hamamatsu scanner

Figure 3.1: The same tissue section taken by different microscopic slide scanners shows the unintended color variation

Color normalization that standardizes image appearance can help both pathologists and software in comparing different tissue samples [16][31][32]. Many existing techniques for color normalization such as histogram specification [45] that were proposed for natural images may not work for histology images. Histogram normalization works best when the source image whose histogram is to be modified captures the exact same scene as the target image, except with a color change. Otherwise, the distribution of various materials and their associated true colors can be quite different between the scenes. This is often the case in histological images even for a given staining scheme, where a source image captures a different tissue slide than a target image whose appearance is preferred by a pathologist or for an algorithm. Fortunately, only a few stains are used in any given staining scheme. If the stain density maps can be estimated in the two images, then the stain color appearance in the source image can also be estimated and modified to match those of a preferred target image while preserving its stain density.

Thus, estimation of stain density maps and color appearances i.e. stain separation is central to many color normalization techniques including the proposed techniques presented in Chapter 2, which transforms the RGB channels of a histological image into an absorption or density map for each stain. Stain separation and color normalization are two sides of the same coin: an accurate stain separation requires a precise estimation of stain color appearance matrix (or color basis) whilst one key step of the structure-preserving color normalization (SPCN) proposed in [31][32] is to obtain an accurate stain density map. Other recently proposed color normalization

TH-1577_11610215

techniques have also been based on stain separation such as [16][7][46], and have shown better results than previous methods. Normalizing the appearance of each stain separately is intuitive because it takes advantage of the known properties of stained tissue images.

We propose a solution for color normalization, which preserves biological structure information by modelling stain density maps based on the SNMF stain separation presented in Chapter 2. An additional advantage of SNMF based formulation is that the color basis is determined in an unsupervised manner without requiring manual annotation of different areas of pure stains. Our color normalization is built on top of SNMF-based stain separation, and is called *structure-preserving color normalization* (SPCN). SPCN works by replacing the color basis of a source image with those of a pathologist-preferred target image, while still maintaining its original stain concentrations. The flexibility to select a preferred target appearance in different scenarios as opposed to a fixed target color appearance model is another advantage of our technique over others such as [7]. Information about most of the biological structure is captured in the stain concentration and hence it is retained in the proposed technique. Additionally, we propose a fast color prototype estimation scheme for large whole-slide images (WSI) that uses an adaptive patch sample instead of the entire image.

We have compared our algorithms quantitatively and qualitatively with competing techniques [6][7][23] and the results are in favor of our technique. We also report a 20-fold speedup on WSI using proposed patch-based color appearance estimation scheme.

3.2 Background and Related Work

In this Section, we survey color normalization methods, and how successively proposed techniques have moved closer to capturing the underlying biological structure of tissues, setting the stage for our proposed technique.

3.2.1 Color normalization without stain separation

An important way to overcome unwanted color variation in images of similar entities is to transform the color appearance of a source image into that of a target image preferred by an

TH-1577_11610215

expert. Histogram specification in RGB space [45] and matching of histogram statistics (i.e. mean and standard deviation) after transformation of an RGB image into $l\alpha\beta$ (decorrelated) color space [47] have been used to normalize color appearance in histology. Both techniques inherently assume that the proportions of tissue compartments primarily stained by a particular reagent are the same across the images being normalized, which is not a valid assumption as proportion of nuclei or whitespace varies from image to image. Color appearance of slides stained to serve as calibrated samples has also been used to model the color space of an imaging system [48][49]. Images of tissues to be examined are normalized to match the color appearance of the calibrated targets. This approach can only incorporate color variations regarding calibrated target which could not consider either differences in staining reagents and procedures or color variation due to scanners. Methods based on stain separation overcome these challenges to a large extent.

3.2.2 Color normalization methods based on stain separation

Supervised color normalization techniques such as [16] and [46] are limited to model the color variation of only the training dataset. Retraining could be required for different tissue types, staining schemes, and even same type of tissue from a different lab. Additionally, a useful application for the color normalization is to improve the color appearance or contrast of low-quality and faded histology slides. It would be very challenging for a training dataset to cover histology slides from all possible scenarios.

Among unsupervised methods, [7] used SVD to extract stain vectors followed by direction correction making it robust to inter-image variations. However, in the normalization process, authors used a built-in color appearance model to modify the color distribution of a given image, which is less flexible than having the ability to select a target image with color appearance preferred by an expert. Moreover, modeling stains as principal components does not guarantee their non-negativity or sparsity. Thus, the resultant component maps can be difficult to interpret biologically.

TH-1577_11610215

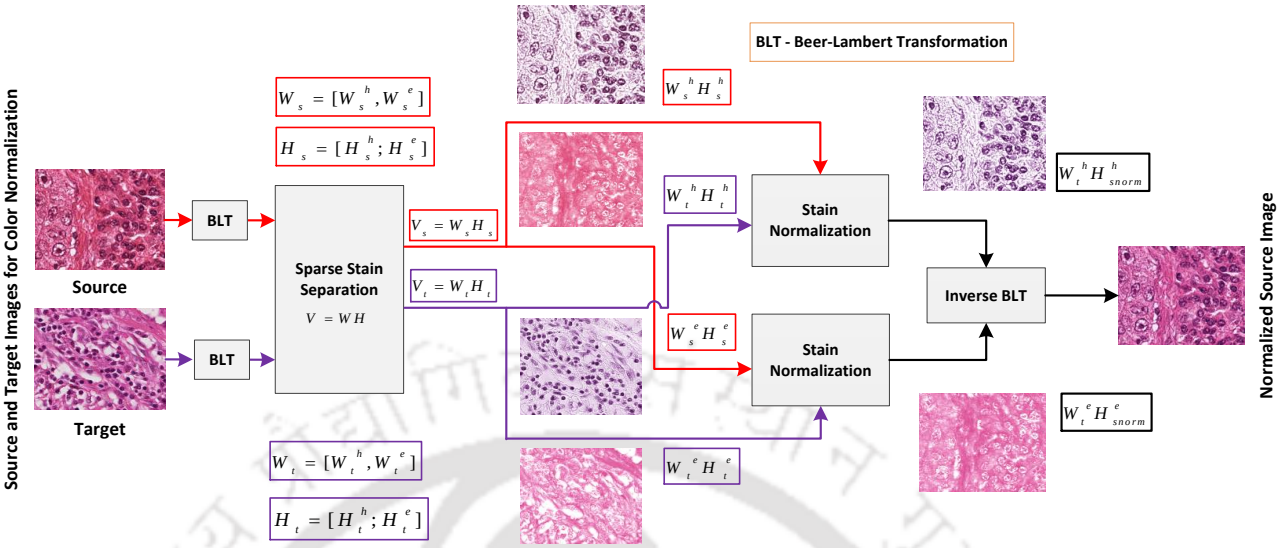


Figure 3.2: Flow diagram of the proposed structured preserving color normalization (SPCN).

3.3 Proposed Method

We used stain separation representation proposed in previous Chapter in our color normalization algorithm. Here, we provide details of implementation of SPCN and a patch-based scheme to accelerate color basis estimation. Source codes of SNMF, SPCN and acceleration scheme are available for the academic use at `https://github.com/abhishekvahadane/CodeRelease_ColorNormalization.git`.

3.3.1 Structure-preserving color normalization (SPCN)

To normalize the color appearance of a source image s to that of a target image t, first we estimate their color appearances and stain density maps by factorizing V_s into $W_s H_s$, and V_t into $W_t H_t$ using SNMF equation 3.5. Then, a scaled version of the density map of source H_s is combined with color appearance of the target W_t instead of that of the source W_s to generate the normalized source image. This preserves the structure in terms of the stain density H, and only changes the color appearance in terms of W, and can be described as follows:

$$H_s^{norm}(j,:) = \frac{H_s(j,:)}{H_s^{RM}(j,:)} H_t^{RM}(j,:), \quad j = 1, ..., r. \tag{3.1}$$

TH-1577_11610215

$$V_s^{norm} = W_t H_s^{norm} \tag{3.2}$$

$$I_s^{norm} = I_0 \exp(-V_s^{norm}) \tag{3.3}$$

where $H_i^{RM} = RM(H_i) \in R^{r\times 1}$, $i = (s,t)$ and $RM(\bullet)$ compute robust pseudo maximum of each row vector at 99%.

Unlike a non-linear mapping between statistics of H_s and H_t in [16], here we only multiply H_s by a scalar and hence keep the relative stain density maps of source image intact. This is similar to [7] where the maps of mixing coefficients of *principal* components were preserved for the source image. In this way, once accurate stain separation is done, our color normalization technique only changes the stain color appearance (basis) while preserving the source's structures. This structure-preserving property has been validated in Section 3.4.1 for its utility. As far as we know, this is the first time *structure-preservation* has been simultaneously considered in both stain separation and color normalization. Previous attempts at stain separation (for example [7]) do not start with structural properties of stains such as sparseness and non-negativity and hence cannot always guarantee the structural invariance after color normalization. Further to reduce the computation on huge WSI, we proposed a patch-based acceleration scheme.

3.3.2 Smart patch-based acceleration scheme for WSI

As explained in the previous Section, a majority of computational time of SPCN is spent in the iterative optimization of SNMF, which slows its performance on WSI, especially when a computer RAM is limited with respect to the size of the WSI. Therefore, we propose a novel acceleration scheme for estimation of global color appearance matrix W of a WSI based on smart patch sampling and patch-wise stain separation. The patches have the same resolution as the original WSI to preserve local structures, which could have been lost by using downsampling – a trivial alternative. A schematic plot of the flow diagram of this scheme is shown in Figure 3.3, and its details are described below.

We sample patches centered at corner points on grid as shown by green dots in Figure 3.3, and discard those that lie in whitespace by comparing their luminosity against a threshold (we use 0.9 for all of our testing). The luminosity is the L value in the $L*a*b$ color space. Then

TH-1577_11610215

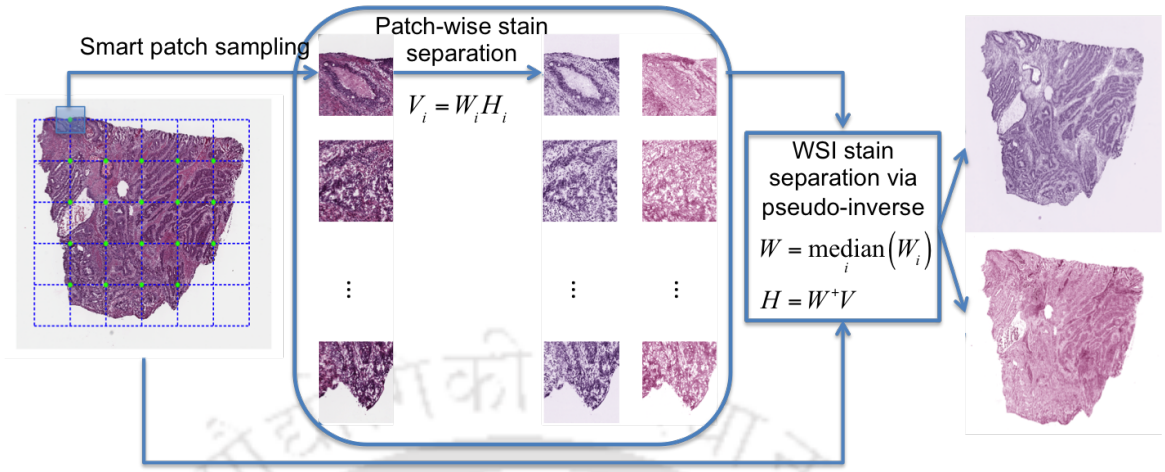


Figure 3.3: Schematic diagram of fast WSI stain separation and global color matrix estimation.

we estimate the color basis matrices W_i for each of the sampled patches indexed by i using SNMF. Stain color columns in the W_i are sorted by ranking the blue channel intensities such that first column corresponds to hematoxylin and second corresponds to eosin. Finally, we take element-wise median of these matrices to make color estimation more robust to artifacts such as folding, blurring and holes. We normalize this median matrix to have unit vector columns, and denote the final color matrix thus obtained as W. The choice of grid and patch size will be addressed in Section 3.4.

The stain separation for WSI is obtained through color deconvolution [24]:

$$H = W^+ V, \quad H \geq 0 \tag{3.4}$$

where $W^+ = (W^T W)^{-1} W^T$ is the Moore-Penrose pseudo-inverse matrix of W. Note that this operation can also be done separately for sub-images of WSI by using a single color appearance matrix W obtained for the entire image as described above i.e. obtaining H through pseudo-inverse will hold for any sub-image and hence can be parallelized.

After stain separation of source and target WSIs, $V_s = W_s H_s$ and $V_t = W_t H_t$, respectively, we change the color appearance of the source WSI to that of the target WSI while preserving original source stain concentration to obtain normalized source WSI (using Eq. (3.1)-(3.3)).

TH-1577_11610215

3.4 Experimental Results

The proposed SPCN was extensively tested qualitatively and quantitatively. All of the datasets used for testing were publicly-available benchmark data, which allow an easy and fair comparison with other competing methods.

We compared SPCN with three state-of-the-art histological color normalization techniques: Reinhard *et al.*[47], Macenko *et al.*[7] and Khan *et al.* [16]. In the original Macenko implementation, authors used built-in reference color appearance as W_t while here we compute it from the target image for a fair comparison. We start with quantitative validation first.

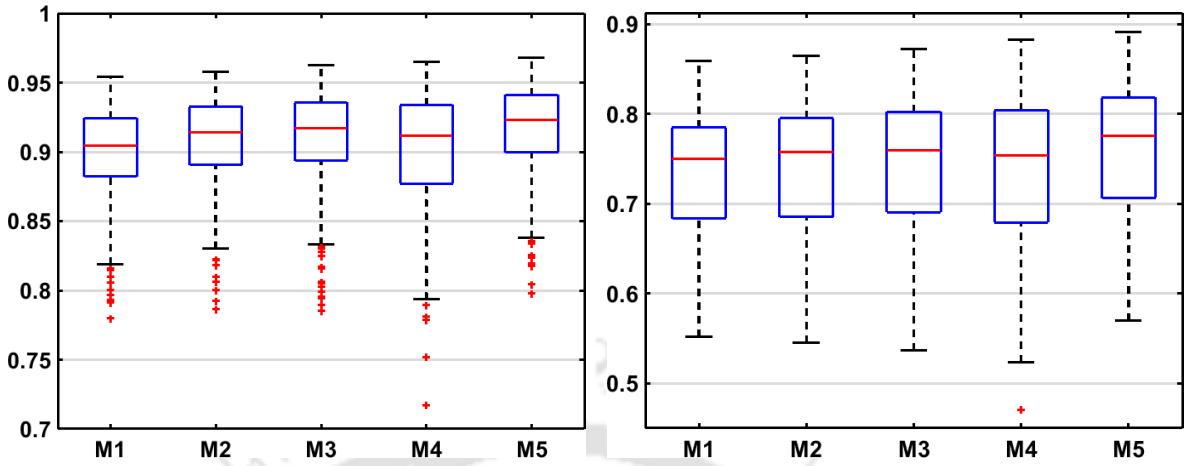
3.4.1 Quantitative results

Validation of color normalization in histology images is not an easy task, especially in a quantitative manner. The results presented in previous studies ([7][16][46]) were mostly qualitative or indirect validation (for example, improvement in the tumor segmentation performance). Here we use a direct benchmark for color variation using the ICPR2014 contest dataset[1]. In this dataset, each breast cancer section was acquired using two different scanners, which results in the color variation between them (as shown in Figure 3.1).

We randomly sampled 300 images taken using an Aperio scanner as source images, and 300 images from Hamamatsu scanner as target images. We applied different color normalization techniques to change source color appearance into target color appearance and compared normalized source to the corresponding target image of the same tissue section (ground truth). Since there is a minor misalignment between matching Aperio and Hamamatsu images due to imperfect calibration and different spatial resolutions of the scanners, we *rigidly* registered Hamamatsu image to the corresponding Aperio image to form the ground truth. We chose multichannel Pearson correlation and quaternion structural similarity index (QSSIM) [50] as similarity metrics between the normalized source and ground truth. Multichannel measure is the average of three RGB channel-wise measures. Boxplots in Figure 3.4 demonstrate SPCN superior performance over other competing methods ($p < 0.001$ by Wilcoxon signed-rank test).

[1] `http://mitos-atypia-14.grand-challenge.org/`

TH-1577_11610215



(a) Pearson Correlation **(b)** QSSIM

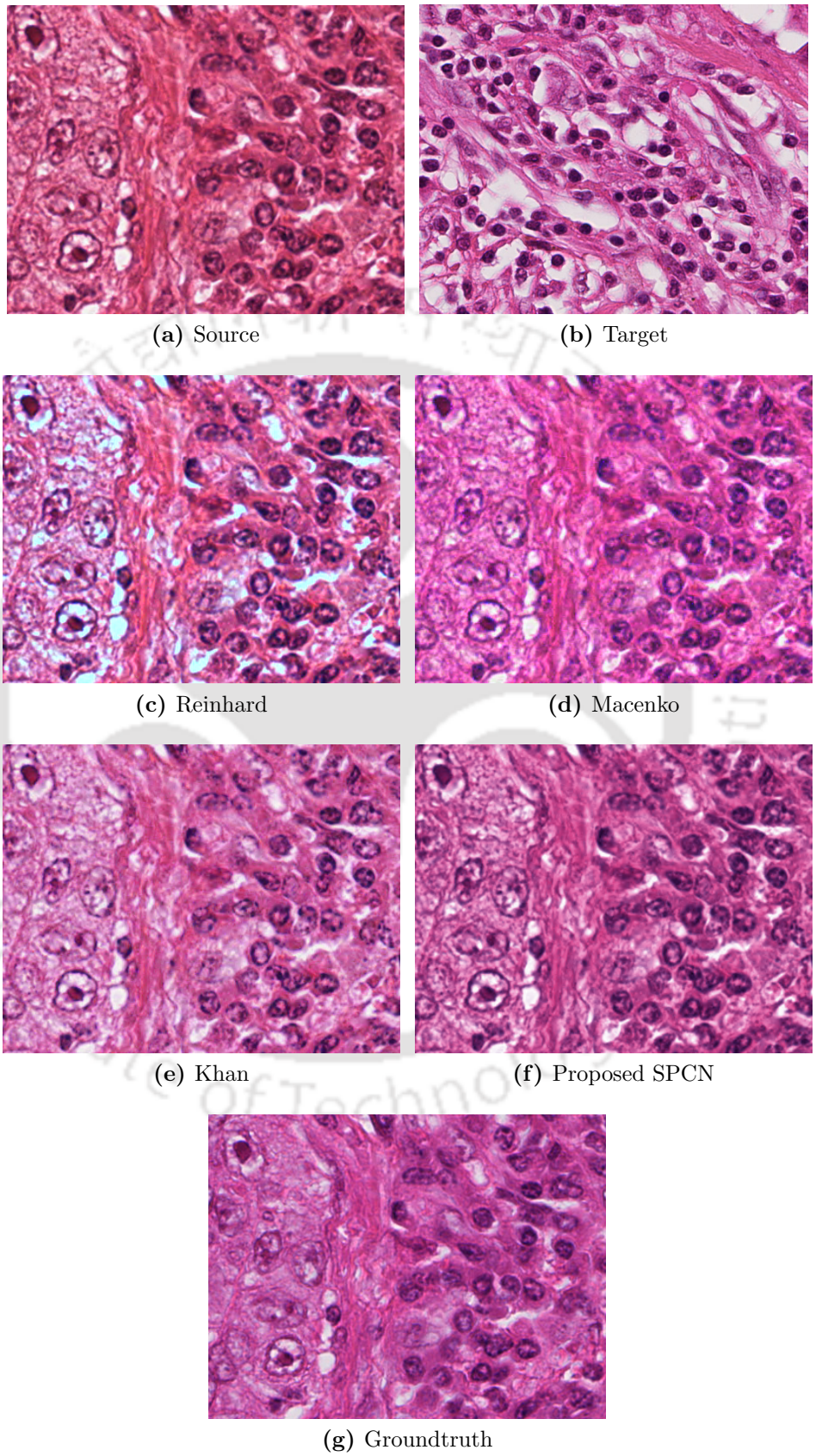
Figure 3.4: Boxplot showing similarity metrics with respect to ground truth stain density map for different normalization methods: M1 (Reinhard), M2 (Macenko), M3 (Khan), M4 (NMF-CN), and M5 (SNMF-SPCN)

A sample example of color normalization is shown in Figure 3.5.

3.4.2 Qualitative validation

The proposed color normalization method can also be used to improve contrast of low-quality histology images, such as faded stain. Figure 3.6c show that Reinhard *et al.* [47] is unable to preserve local structure and fills lumen with color when it should remain white after normalization. Figure 3.6d shows that SVD-based normalization by Macenko *et al.* [7] is largely monochrome, noisy and is also unable to preserve the lumen structure, possibly due to the unrealistic negative stain components at these regions. Normalization method by Khan *et al.* [16] in Figure 3.6e does a better job of keeping the structure intact and showing a differentiated second stain, but in some area (highlighted in green circles), it is unable to reproduce the less abundant stain (hematoxylin) faithfully. Compared to all these methods SPCN reproduces the local structure, and is able to differentially show the two stains in Figure 3.6f, even in those areas where the method by Khan *et al.* had difficulty (green circles). Another example with different source and target is shown in Figure 3.7. Here, clearly SPCN outperform Khan technique since nuclei structure are well preserved comparatively.

TH-1577_11610215



(a) Source **(b)** Target

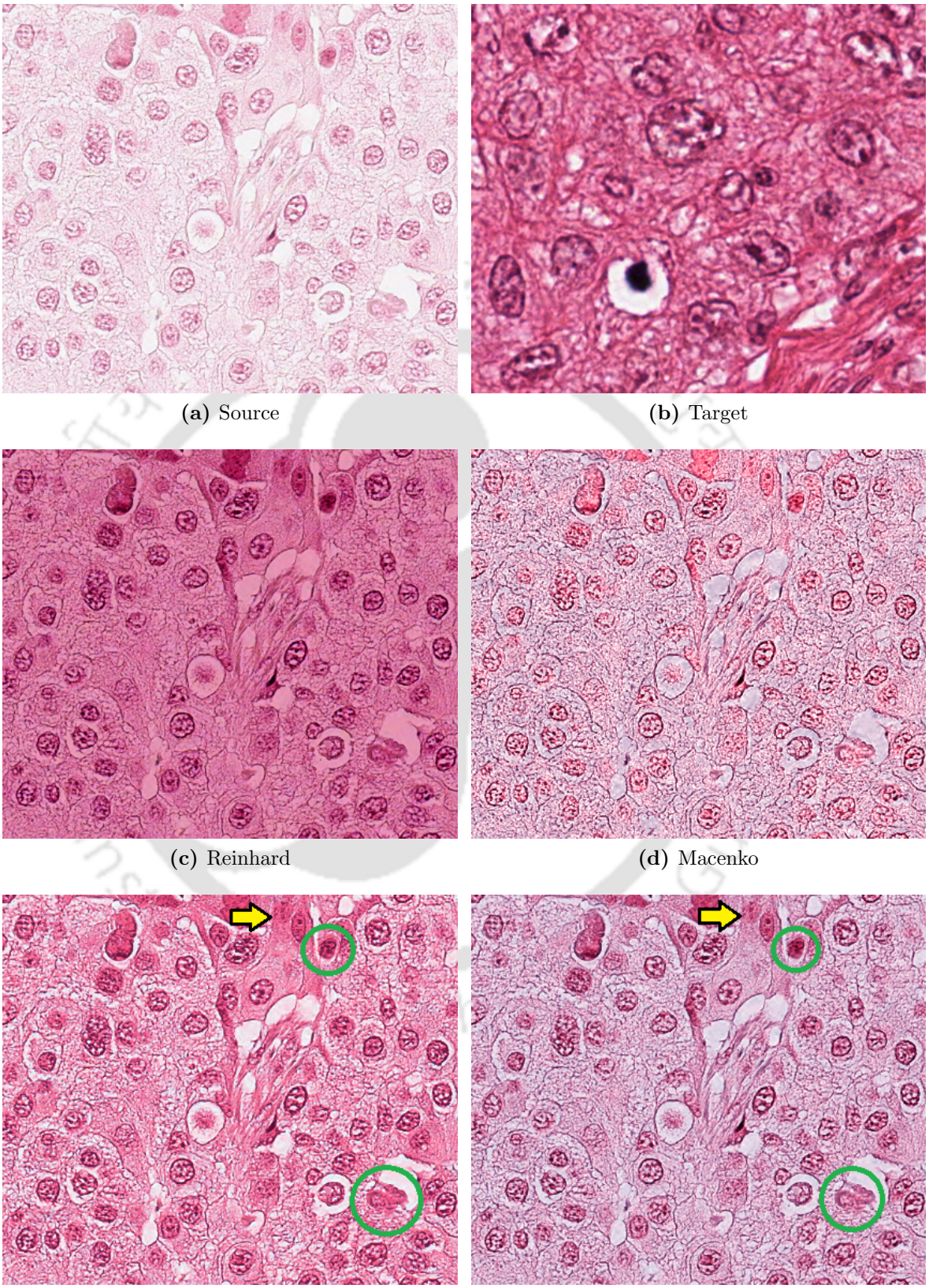
(c) Reinhard **(d)** Macenko

(e) Khan **(f)** Proposed SPCN

(g) Groundtruth

Figure 3.5: Color normalization results by various techniques on a sample from the data used for quantitative analysis

TH-1577_11610215



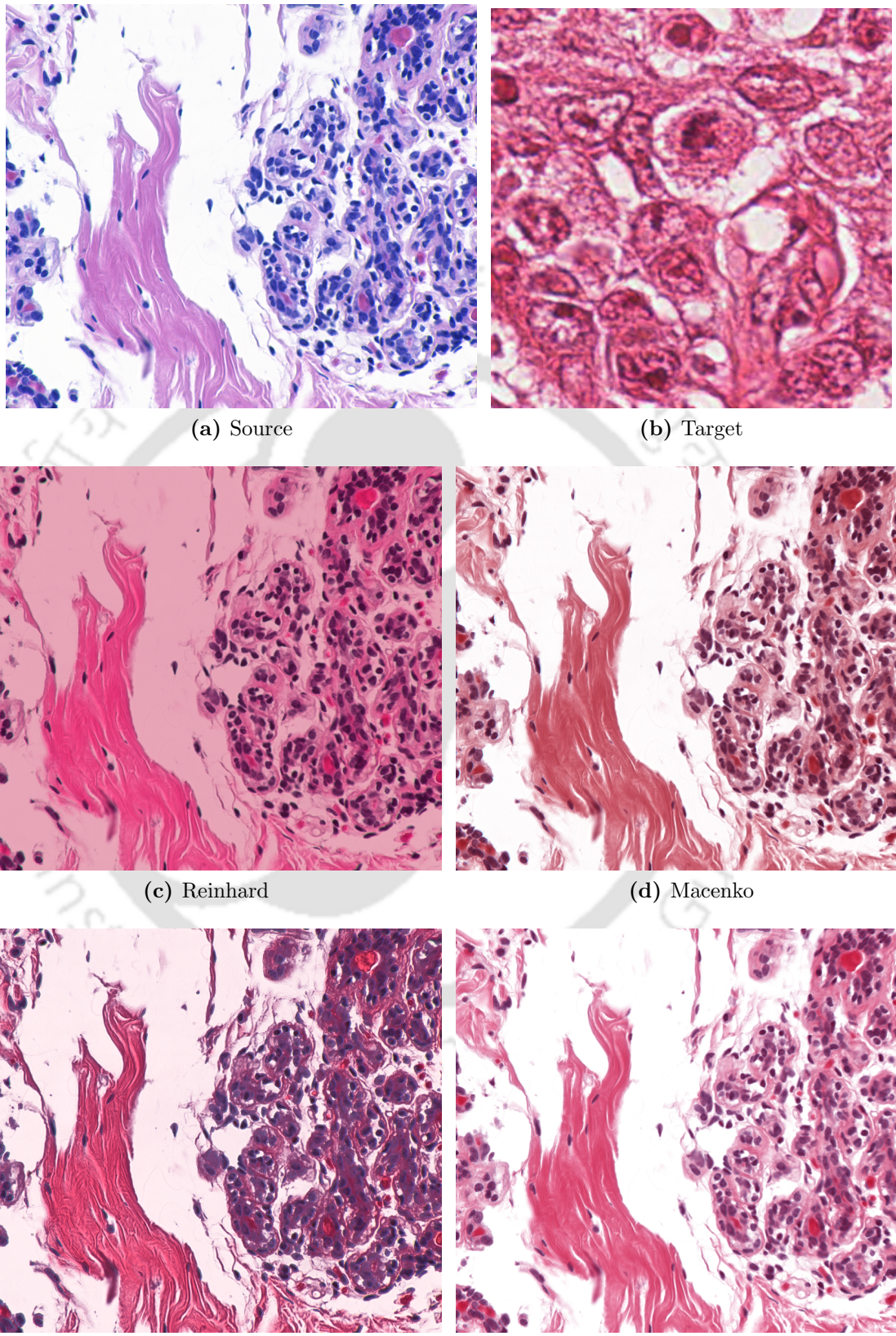
(a) Source **(b)** Target

(c) Reinhard **(d)** Macenko

(e) Khan **(f)** Proposed SPCN

Figure 3.6: Example 1: Visual comparison of different color normalization techniques. See text for more details.

TH-1577_11610215



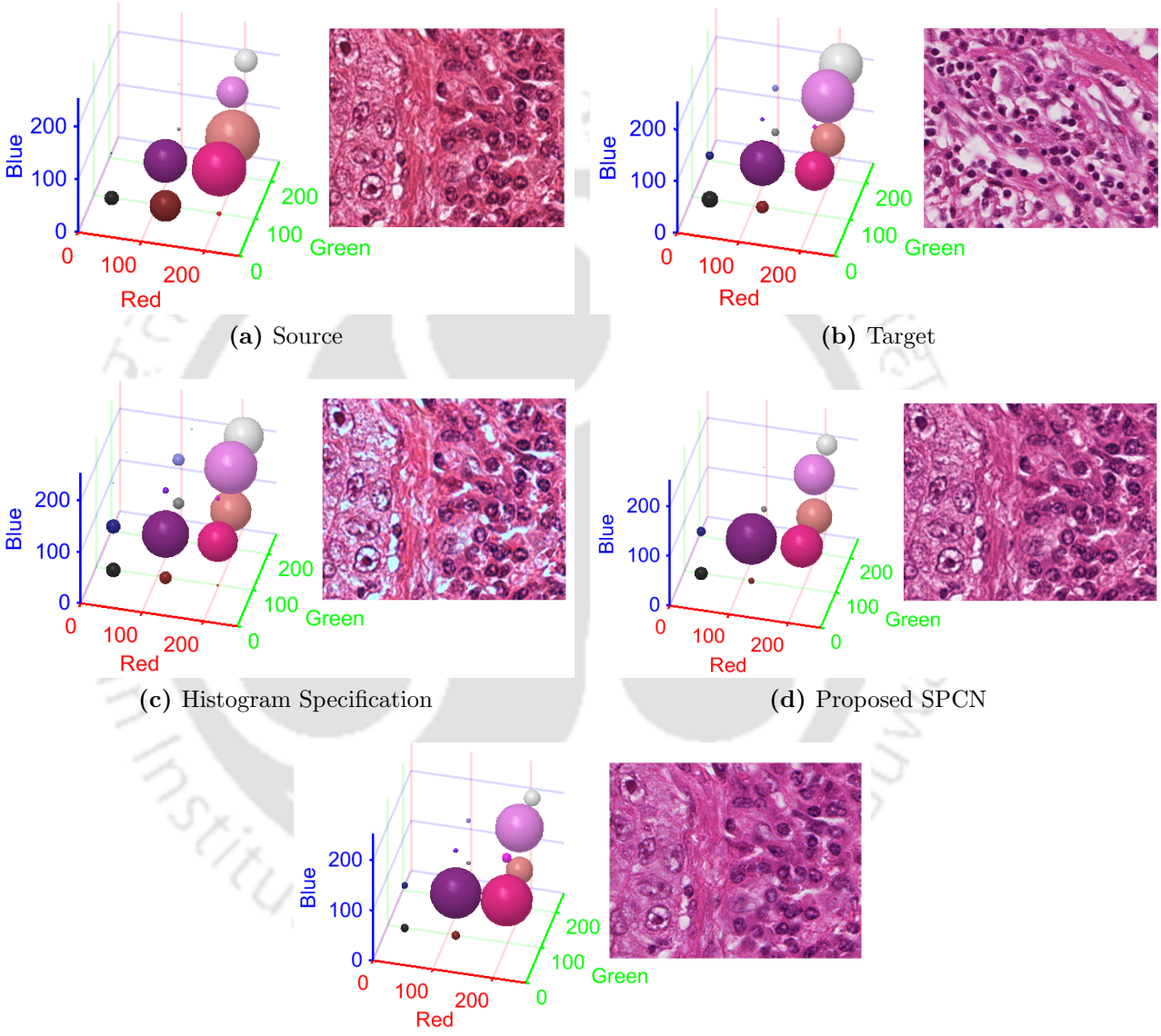
(a) Source **(b)** Target

(c) Reinhard **(d)** Macenko

(e) Khan **(f)** Proposed SPCN

Figure 3.7: Example 2: Visual comparison of different color normalization techniques. SPCN outperform Khan technique since nuclei structure are well preserved comparatively.

TH-1577_11610215



(a) Source **(b)** Target

(c) Histogram Specification **(d)** Proposed SPCN

(e) Ground Truth

Figure 3.8: 3D histogram visualization of color normalization on a sample source and target image from the data used in Section 3.4.1.

TH-1577_11610215

3.4.3 Histogram interpretation of structure-preservation of SPCN

We plotted 3D histogram of RGB image to visualize color normalization process. As shown in Figure 3.8a and 3.8b, the source and target images have different color distribution (the volume of the ball indicates histogram density). Unlike histogram specification (Figure 3.8c), SPCN does not seek an identical match to the histogram of target (Figure 3.8d). For example, SPCN retains equal amount of white lumens as that in the source image. This is reflected by the comparable size of white balls in the two corresponding histograms. By contrast, histogram specification tries to achieve an equal amount of white lumens as in the target image and therefore creates artificial structures.

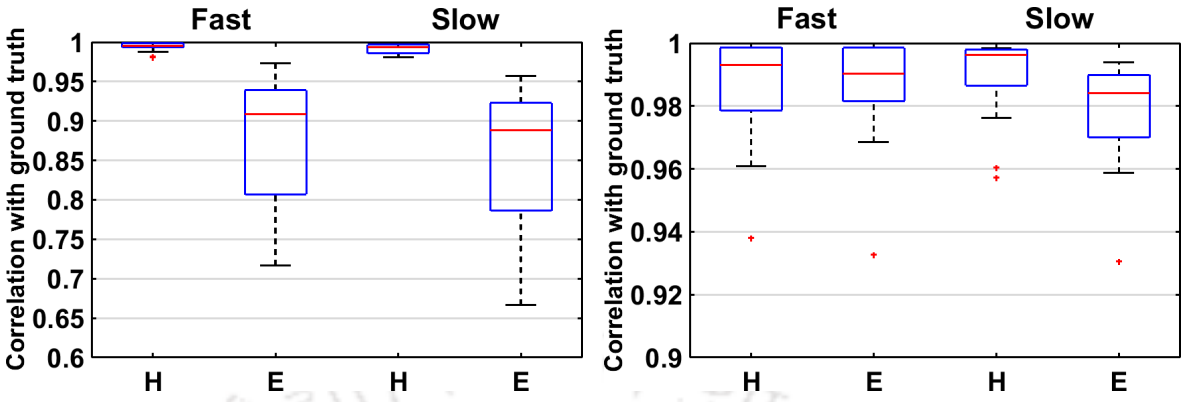
3.4.4 Validation of acceleration scheme for WSIs

We validated the proposed acceleration scheme on 27 H&E stained colon tissue WSIs of size 20000×20000 covering a physical area of $5mm \times 5mm = 25mm^2$ available at TCGA data portal[2]. An example sub-image is shown in schematic plot of Figure 3.3. Ground truth stain density was obtained through expert stain separation as described in Section 2.4.1. We performed the stain separation using the proposed patch-based scheme and compared it to a direct stain separation on WSI.

We use 6x6 grid in which around 20 patches were sampled for patch wise decompositiom. Each patch was of size 1000x1000. Figure 3.9 shows box plots of Pearson correlation between ground truth and methods (NMF and SNMF). Similar to small-size images, SNMF outperformed NMF in WSI stain separation, particularly for the estimation of eosin component ($p < 0.001$ by Wilcoxon signed-rank test). The rRMSE measures of W estimation for slow and fast scheme were 0.066 and 0.068 for SNMF; 0.115 and 0.097 for NMF respectively. The proposed acceleration scheme achieved approximately 20-fold acceleration over a direct decomposition on WSI for both NMF and SNMF methods while maintaining a consistent or improved performance. It took an average of 82 seconds on a four core processor for patch-based scheme, compared to an average 25 minutes for the direct stain separation for each of

[2] http://tcga-data.nic.nih.gov/tcga/

TH-1577_11610215



(a) NMF **(b)** SNMF

Figure 3.9: Box plots of quality metrics between separated stain and ground truth for 27 H&E stained colon tissue WSIs for hematoxylin (H) and eosin (E) suggest consistent results between our fast patch-based scheme and the direct decomposition of WSI (slow approach).

the 27 H&E stained breast WSIs. There are several factors that contribute to this acceleration: (1) whitespaces were excluded based on the proposed sampling scheme while computing SNMF or NMF step, (2) Several patches can be processed in parallel on different CPU cores for a theoretical speed up of (sizeofWSI×numofCPUcores) / (numofpatches×sizeofpatch), and (3) Inputs and temporary variables corresponding to a small patch can be accommodated in the RAM, thus reducing time-consuming read-write calls to the hard disk.

3.4.5 Impact of color normalization

The applicability and effectiveness of proposed color normalization technique was analyzed for the goal of epithelial-stromal segmentation in H&E stained prostate tissue [15]. Authors compared two normalization techniques for their utility in epithelial-stromal classification task and also compared with the original images (without normalization). One of the two techniques was Khan *et al.* method [16] and other method was proposed above in this Chapter i.e. SPCN. Thirty diverse images of H&E stained prostate microarray cores were selected for this task. Three training and testing sets were formed using Khan normalization, SPCN and third set was original i.e. without color normalization. To robustly assess the contribution of each color normalization technique, three classifiers were trained, viz. wndchrm, deep convolutional

TH-1577_11610215

neural network (CNN), and logistic regression-based super-pixel classification based on multi resolution segmentation (MRS). Wndchrm is an open source utility for biological image analysis in which we can train a supervised model on a large and fixed set of image features based on color and texture that are already hard-coded in its software [51]. CNN learns its own feature maps in a supervised manner. For MRS, images were over-segmented into super-pixels such that each super-pixel contained mostly epithelial or mostly stromal pixels, and then about a hundred hand-crafted features were extracted using Definiens Developer XD®. Penalized logistic regression was used on the super-pixel features to classify each pixel of the super-pixel with the same label as the super-pixel. These features were representing the color, texture, shape, size, relation to neighbors, and relations to sub-objects of the super-pixel objects. Test area under curve (AUC) were compared for the two normalized and original images using all three classifiers. When visually analyzed, Khan method reduced color saturation while SPCN reduced hue variance. Super pixel test-AUC for the MRS was higher for the normalized sets as compared to original sets. This improvement suggest the crucial role of the color normalization for the any histological image analysis. Similar trend were reported for the CNN and wndchrm. This demonstrated the ability of color normalization to improve the performance of most of the machine learning and image analysis methods on histological images.

3.5 Summary and Discussion

Color normalization is an essential step to remove undesirable color variation in histopathological images and is useful for analyzing disease and its progression on large datasets from different pathology labs. A potential danger of such a technique that concerns pathologists is that tissue structures presented in the original image could be altered after normalization. In this Chapter, we proposed a novel structure-preserving color normalization (SPCN) scheme, which changes color of one image (source) to match that of another (target) while reliably keeping source structural information intact. A key step in our normalization scheme is an accurate stain separation of both source and target images based on sparse regularized NMF. Extensive qualitative and quantitative validation demonstrated superior performance of the

TH-1577_11610215

SPCN techniques. Additionally, we proposed a smart patch-based scheme to accelerate computation on WSI. It is worth mentioning that the proposed estimation of global color appearance based on smart patch sampling is a general strategy, which can be easily combined with other stain separation and color normalization techniques to accelerate their implementation on WSI. Also, the proposed algorithms were insensitive to the different random initialization of W when initialized by the elements from the training set.

SPCN is an improvement over pioneering method of Macenko *et al.* [7]. We assumed certain structural properties of histopathological images that are biologically more plausible. First, we assumed non-negativity. That is, a biological material is either absent (zero light absorption) or present (positive light absorption), and thus optical density cannot be negative. Second, we assumed sparsity. That is, a given location (pixel) is likely to be occupied by one or the other type of biological material, but not a mix of the two. Based on the two assumptions, we modeled stain or structure density maps as sparse and non-negative. By contrast, [7] computes singular value decomposition, which can lead to dense and even unrealistic negative stain amount and cannot guarantee the preservation of underlying biological structures (as shown in Figure 3.6d).

In the presented color normalization evaluation, SNMF-SPCN attained a higher accuracy than NMF-CN. A likely biological interpretation is that the level of stain leakage is dependent on tissue preparation and staining procedure and hence is image-specific, which also contributes to inter-image color variation. NMF does not take this factor into account during the normalization process. On the other hand, SNMF-based color normalization, by contrast, is based on a clear separation of the structures, and therefore overcomes this problem and acts like a repainting of the source structures.

TH-1577_11610215

4

Nuclei Segmentation

Contents

4.1 Introduction and related work 48
4.2 Proposed image processing pipeline 51
4.3 Deep CNN based three class nuclear segmentation 59
4.4 Discussion 69

4.1 Introduction and related work

Nuclei density, nucleus-to-cytoplasm ratio and average nuclear size are the quantitative parameters which can be helpful for the assessment of cancer grades since regions of tumor growth have an increased cell nuclei density. Various diagnostic features like mitotic count, nucleopleomorphism etc. can be extracted from segmented nuclei. Nuclei are parts of cells that form the relatively macro-structures like glands. Cancer grading involves differentiating between different gland architectures and shape of nuclei. Thus, segmentation of nuclei from the histological image is an important and first step for further computational interpretation of prognosis and predictive models. Segmentation of nuclei is complicated by variations in their background (epithelial and stromal cytoplasm, connective tissue, debris etc.), as well as

TH-1577_11610215

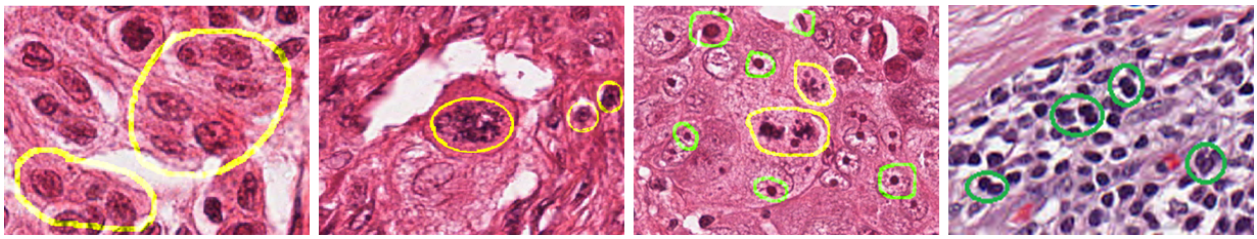


Figure 4.1: Nuclei segmentation challenges: Marker shows (from left to right) blurred and partially visible nuclei, large size nucleus, rough nuclei boundaries, size and shape variations, and nuclei overlapping.

variations of nuclei size and shape, complex heterogeneity within nucleus and in stroma, color variation, imaging artifacts, variation of nuclei density within and across images, and nuclei overlapping [52]. A visual about the typical segmentation challenges is depicted in the Figure 4.1. Let us look at the different nuclei segmentation approaches in the literature.

4.1.1 Nuclear segmentation

The nuclear segmentation methods can be based on image processing, machine learning and deep learning. Nuclear segmentation were approached by using classical image segmentation methods such as thresholding, region based approaches (like region growing), energy minimization techniques (like snake model, active contour level sets), machine and deep learning based segmentation [53].

Most of the approaches were tailored to tissue or staining type and are not generalizable across tissue type and nuclei variations [54]. A machine learning based framework was used by A. Veillard *et al.* [52] to find the probability map of pixels belonging to the nuclei. Linear discriminant analysis (LDA) was performed on Laws texture measure (180 dimensional feature space) for optimizing inter- to intra-scatter ratio and it was followed by an application of softmax function [55] to map it to the probability. This method could not be generalizable across different histological image types as they only trained on breast tissue histological image set. Also, color variation may pose a problem for this type of framework.

E. Bengtsson *et al.* proposed a comparatively robust and an upgraded seeded watershed model [53] where the seeds are generated by H-maxima transform of the distance transform

TH-1577_11610215

of the estimated background map. However, the parameters are manually tuned and hence not generalizable. Authors reported that use of intensity, gradient, connectivity, and shape information in the watershed model leads to robust segmentation methods. M.Veta *et al.* proposed an unsupervised segmentation in which marker controlled watershed transform was seeded by a foreground marking scheme based on fast radial symmetry transform (FRST) [21]. Authors reported improved results than regional minima marking scheme that was popular in watershed segmentation. The post-processing of solidity based rejection, small and big sized region removal was included to further improve the segmentation accuracy. The same framework is reflected in the work of Huang and Lai [56] with a different marking scheme. It uses a regional minima based foreground markers. This marking scheme works well with a simple tissue structure but creates problems with more complex tissue appearance and at the densely packed nuclei. To avoid over-segmentation due more than one foreground marker per nucleus, transformation of more than one markers to one per nucleus by not allowing the markers to be present within the some (threshold) Euclidean distance between the markers. It is not a robust approach for the marker filtering which is prone to errors by removing wrong markers.

4.1.2 Nuclei overlapping resolution

Overlapping nuclei is the other challenging problem in nuclei segmentation. A deformable model was proposed to resolve overlapping by Plissiti *et al.*, which combines local characteristics of the nuclei boundary and a priori knowledge about the expected shape of the nuclei learned from the training data [57]. While testing on the new image, the spatial location of initialization of model should be accurate otherwise it leads to localized wrong boundary. H-minima transform on the original image was used for initialization and it could lead to error due to tissue folding, stain spill over etc.

Thus, most of literature on nuclear segmentation fits the model to a specific application. Hence, there is a scope to develop a powerful generalizable segmentation algorithm. A generalizable algorithm is such that it is applicable for a wide range of applications (different

TH-1577_11610215

types of microscopic/histological stained images), while achieving high sensitivity and specificity. Although watershed segmentation is quite popular, it is marred by the problem of over-segmentation. We proposed an image processing pipeline to enhance the watershed based nuclei segmentation [17] and deep CNN based framework for nuclei detection and segmentation with overlapping resolution.

4.2 Proposed image processing pipeline

We incorporate the following enhancements to the watershed segmentation algorithm as shown in the flow diagram Figure 4.2: 1) Gaussian smoothing for high frequency noise removal and blurred nuclei segmentation. It enhances the image by smoothing foreground and background; 2) Proposed background markers based on image information to reduce over-segmentation; 3) Tentative foreground map for removing false detections.

4.2.1 Method

The proposed framework reported an improved performance with respect to over-segmentation, maximal nuclei segmentation, smooth boundary preservation, and false detections. This improved performance was due to inclusion of following stages: Gaussian smoothing for enhancement with its implicit advantages, an image driven background markers, and Otsu thresholding based tentative foreground estimation.

Channel selection and pre-processing : The R channel of an RGB image was selected for further ease of processing because it shows significant contrast between nuclei and background. The preprocessing was employed for the high frequency noise removal using Gaussian smoothing. It enhances the segmentation of diverse and poorly stained nuclei. The parameter of Gaussian kernel was kept low so that we will not miss the small sized nuclei. Since hematoxylin binds to the nuclei, we can use a hemtoxylin channel as an alternative to the R channel here [21]. Hematoxylin stain density maps can be obtained through the process of stain separation which we have covered in Chapter 2.

Image enhancement : A gray-scale morphological filtering (opening followed closing by

TH-1577_11610215

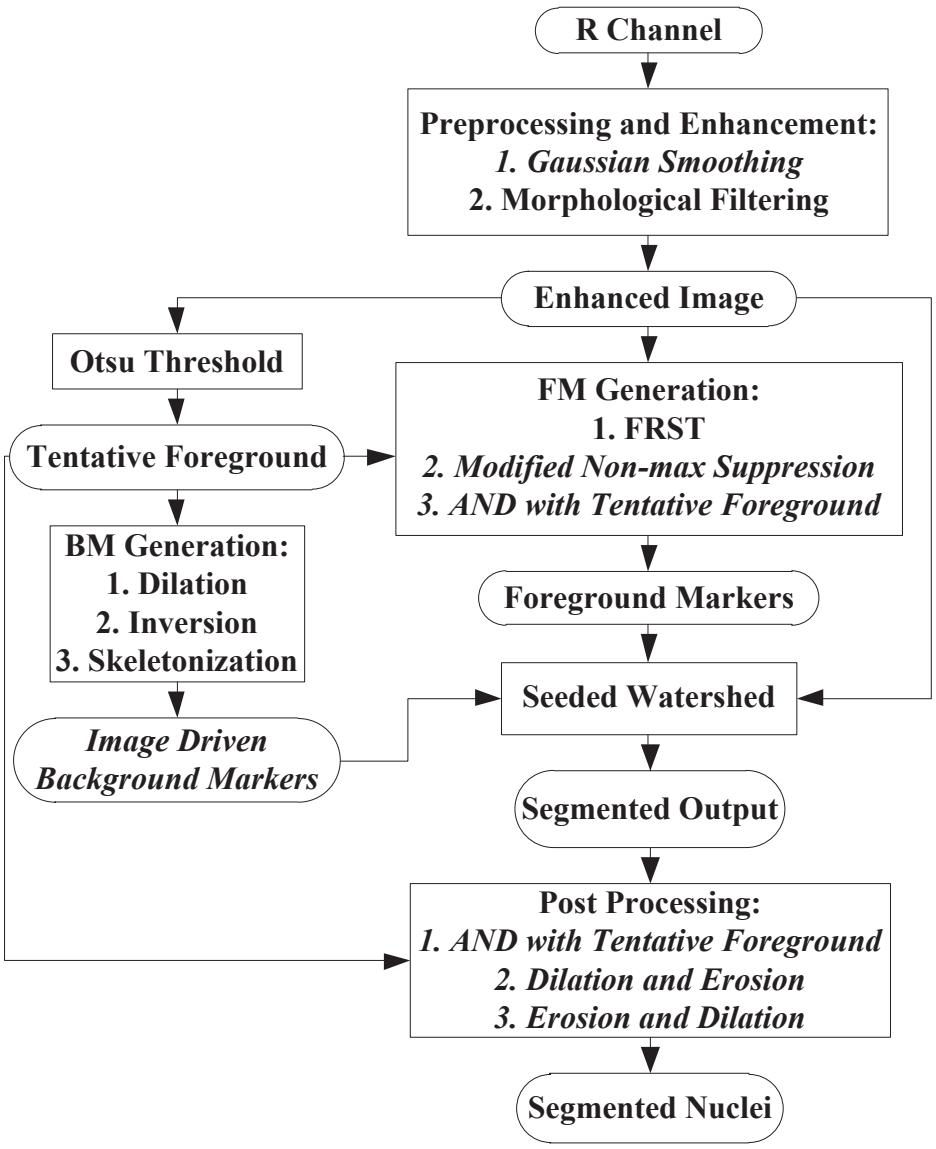


Figure 4.2: Flow diagram for the improved seeded watershed segmentation. Italic texts are our improvements to seeded watershed segmentation.

TH-1577_11610215

reconstruction) was introduced as an image enhancement step to further smoothen the high intensity background (mainly, cytoplasm) and the low intensity foreground (mainly, nuclei), while preserving the structural or edge information. Gray scale morphological operations are the extension of basic operations of dilation, erosion, opening, and closing to gray-scale images.

Gray-scale morphological reconstruction is defined identically as for binary images [58]. Let f and g be the marker and mask images and assume that both are gray-scale images of the same size and that $f \leq g$. The geodesic dilation of size 1 of f with respect to g is defined as,

$$D_g^1(f) = (f \oplus b) \wedge g \tag{4.1}$$

where $\wedge$ denotes the point-wise minimum operator. Thus, the geodesic dilation of size 1 is obtained by computing the dilation of f by b and then selecting the minimum between the result and g at every point (x, y). The dilation expression depends on the type of structuring element (SE) - flat or non-flat. The geodesic dilation of size n of f with respect to g is defined as,

$$D_g^n(f) = D_g^1(f)[D_g^{n-1}(f)] \tag{4.2}$$

where $D_g^0(f) = f$. Similarly, the geodesic erosion of size n of f with respect to g is defined as,

$$E_g^n(f) = E_g^1(f)[E_g^{n-1}(f)] \tag{4.3}$$

where $E_g^0(f) = f$. The morphological reconstruction by dilation of a gray-scale image g by a gray-scale marker image f is defined as the geodesic dilation of f with respect to g repeated (iterated) until stability is reached,

$$R_g^D(f) = D_g^k(f) \tag{4.4}$$

where k such that $D_g^{k+1}(f) = D_g^k(f)$. Identically, morphological reconstruction by erosion of a gray-scale image g by a gray-scale marker image f is defined as,

TH-1577_11610215

$$R_g^E(f) = E_g^k(f) \tag{4.5}$$

where k such that $E_g^{k+1}(f) = E_g^k(f)$. The opening by reconstruction of a gray-scale images first erodes the input image and uses it as a marker. The opening by reconstruction of size n of an image f is defined as the reconstruction by dilation of f from the erosion of size n of f:

$$O_R^n(f) = R_f^D[f \ominus nb] \tag{4.6}$$

where $(f \ominus nb)$ denotes n erosions of f by b. The objective of opening by reconstruction is to preserve the shape of the image components that remain after erosion. Similarly, the closing by reconstruction of size n of an image f is defined as the reconstruction by erosion of f from the dilation of size n of f:

$$C_R^n(f) = C_f^E[f \oplus nb] \tag{4.7}$$

where $(f \oplus nb)$ denotes n dilations of f by b. Because of duality, the closing by reconstruction of an image can be obtained by complementing the image, obtaining the opening by reconstruction, and complementing the result.

In our case, g is the R channel. The f is obtained by a simple gray-scale erosion of g by a disk structuring element (SE) of size 10. The same setting was kept for the closing by reconstruction operation to smooth the non-nuclei background effectively. This is followed by opening by reconstruction operation to smoothen the nuclei by a SE of small size 5.

As the watershed segmentation is a function of the gradient magnitude of the image, the enhancement parameters affect the performance of the segmentation. The best enhancement parameters are numerically computed and listed in Table 4.1.

Foreground and background markers : Foreground and background markers should be provided to the watershed segmentation model to avoid the over-segmentation. Assuming radial symmetry of the nuclei, we employed fast radial symmetry transform (FRST) foreground markers on the enhanced image using a discrete set of radii $(5, 7, 9, 11, 13, 15)$. The maximum

TH-1577_11610215

discrete radius ($r = 15$) was manually selected to be the average size of the nucleus. Spurious multiple maxima were removed using non-maximum suppression parametrized by a square window of size $2 \times r$. To remove multiple maxima from asymmetrical and large size nuclei, we leverage iterative non-maximal suppression with square mask of size $(2 \times r + 10)$. At convergence, there is no change in the current and previous marker image, which acts as a stopping criteria. The output of iterative non-max suppression acts as foreground markers (FMs). One can easily generate background markers (BMs) from foreground markers (FMs) by presuming an average nuclei size. However, nuclei in histological images are neither truly circular in shape nor ideal symmetrical. Even after repeated non-max suppression, some nuclei may have more than one markers. Thus, if we use those FMs to estimate BMs, there is possibility of under- or over-segmentation. We added an image driven background marking scheme to avoid the above mentioned artifacts. This scheme employ estimated foreground map to compute BMs. The foreground map was a binary image obtained after Otsu threshold based segmentation of the enhanced image. The morphological skeleton of dilated and inverted foreground map results in the true BMs. Here, dilation is done to avoid BMs on nuclei.

Segmentation and post-processing : Marker-controlled watershed segmentation was performed on the enhanced image. The FMs and BMs acts as local minima for initialization. At convergence, obtained watershed lines represented boundaries of the nuclei. In post-processing, false detected regions in the output of watershed model were removed by use of tentative foreground map. The regions, separated by less than two pixels were connected by use of morphological dilation-erosion by SE of size 2. Also, small segmented regions (less than disk SE 3 pixels) were removed.

4.2.2 Experimental results

The enhanced seeded watershed model was evaluated on the two tissue type - H&E stained breast and Feulgen stained prostate. The H&E stained breast histological images were available through ICPR2012 MITOS contest with the motive of mitosis detection in histological images. The Feulgen stained images were available through our clinical collaborator from University of

TH-1577_11610215

Table 4.1: Typical Parameters Values for Proposed Segmentation

Processing stage	Parameters	Value in pixels
Gaussian smoothing	Variance, σ	3
	Window size	10
Morphological filtering	Opening by reconstruction disk SE size	30
	Closing by reconstruction disk SE size	5
FRST	Set of radii	5,7,9,11,13,15
	Strictness of symmetry, α	2.0 (Unit less)
Non-max suppression	Window size	30
	Looping window size	40
BM from FM	Dilation disk size	20
Image driven BM	Dilation disk size	7

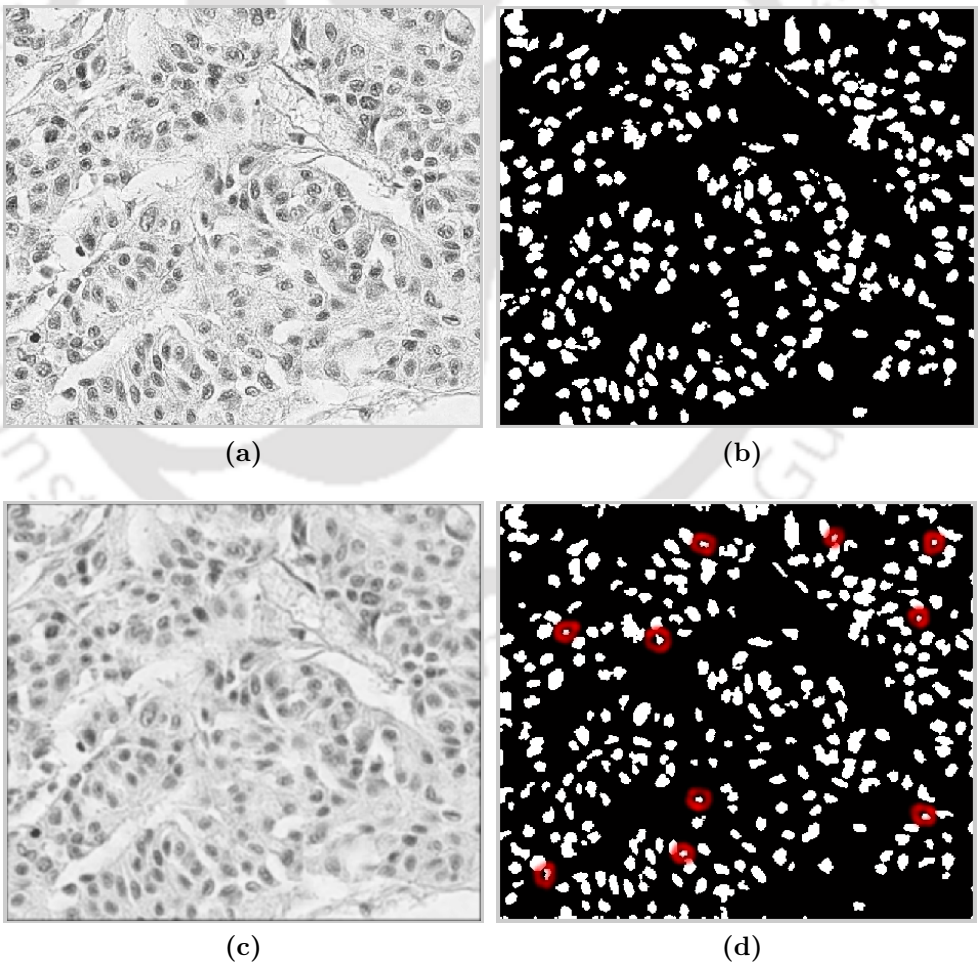


Figure 4.3: Improvement in nuclear segmentation by Gaussian smoothing (a) R channel of input color image, (b) Segmentation of (a), (c) Smoothed R channel, and (d) Segmentation of (b). Blurred nuclei missed by (b), are retained in (d), as shown by (red color) annotations in (d).

TH-1577_11610215

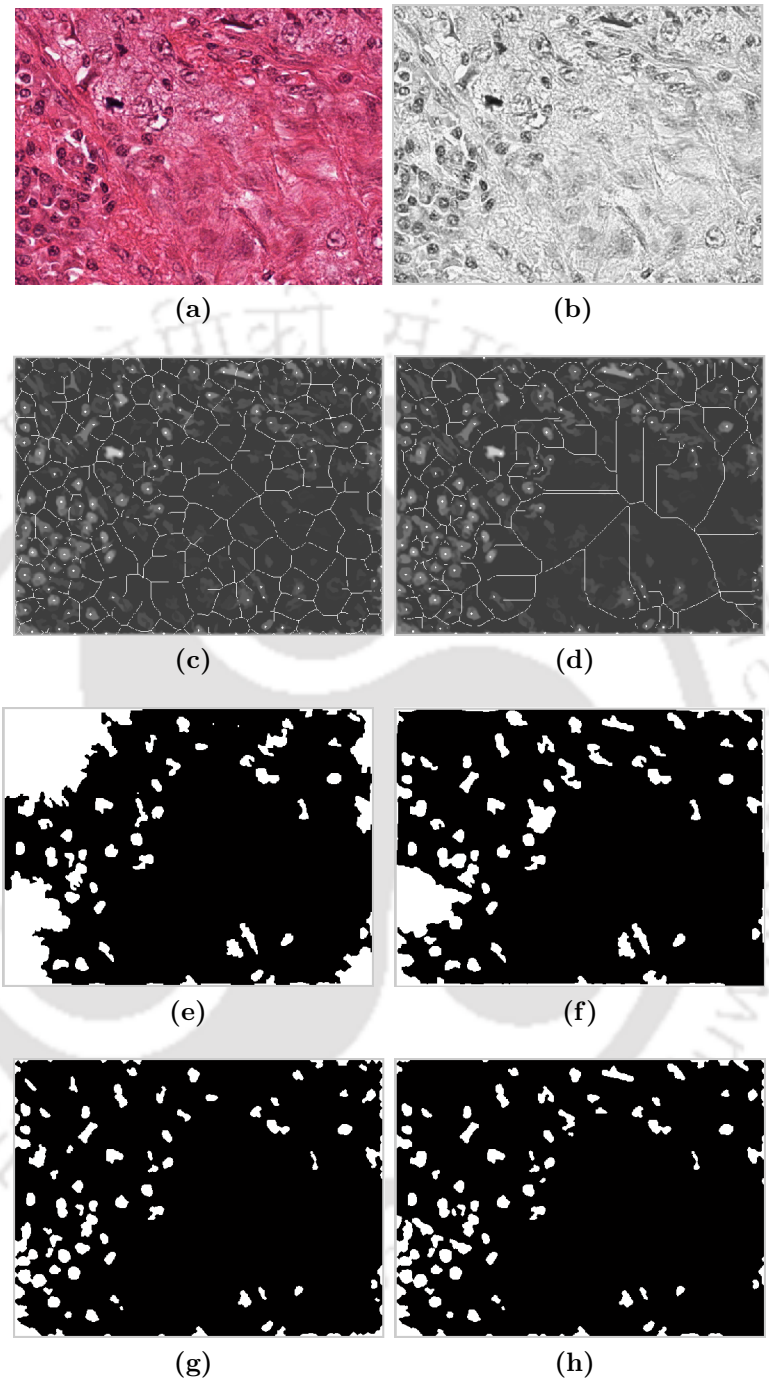
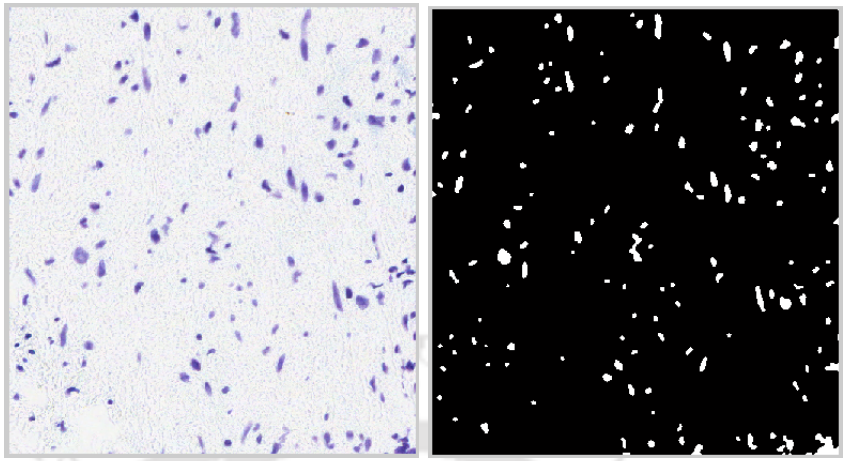


Figure 4.4: Results of the proposed algorithm (a) Input color image, (b) R channel, (c, e, g)- are the results for the background markers obtained from foreground markers scheme, and (d, f, h)- are the results for the proposed image driven background markers scheme. (c, d) - enhanced image + foreground markers (dots) + background markers (lines on the background), Some background markers in (c) crosses objects of interest unlike by proposed image driven background marking shown in (d), which leads to oversegmentation. (e, f)- marker controlled watershed segmentation output, (f) shows less false detections, and (g, h) - post-processed output, in between (g) and (h), (g) shows some nuclei missing unlike by proposed algorithm in (h).

TH-1577_11610215



(a) (b)

Figure 4.5: Segmentation results on the Feulgen stained image: (a) Input Feulgen stained image, (b) Segmented output by proposed method.

Illinois at Chicago (UIC). The visual results on H&E and Feulgen are shown in Figures 4.4 and 4.5. Figure 4.3 shows importance of Gaussian smoothing.

The effect of Gaussian smoothing on the proposed framework was analyzed. It helps to segment the blurred nuclei (marked in the Figure 4.3d) unlike segmentation without Gaussian smoothing Figure 4.3b. An improvement in the segmentation through the proposed image-driven background marking scheme is shown in 4.4. The foreground marker (FM) based background marking scheme was found inaccurate as it increases over-segmentation (Figure 4.4e) unlike the proposed image driven BMs (Figure 4.4f). The prostate tissue Feulgen stained images have comparatively uniform background (non-nuclei structures). Therefore, the best setting of morphological filtering parameters for opening and closing were lowered and are respectively 20 and 5. The results in Figure 4.5 shows that our algorithm effectively works on the Feulgen stained images too.

The quantitative results were obtained on 50 intestinal microscopic images since the ground truth (expert-marked nuclei set) was available. Table 4.2 lists the recall, precision, over-and-under segmentation performance. The recall and precision were high. However, under-segmentation was high since our framework was not able to detect some of the nuclei because of they are not detected either by FRST or iterative non-max suppression.

TH-1577_11610215

Table 4.2: Performance measures on intestinal microscopic images

Recall	Precision	Oversegmentation	Undersegmentation
93.33%	90.79%	8.63%	15.49%

4.2.3 Summary and future directions

Seeded watershed scheme can be a good choice for the multi-object segmentation, such as nuclei in histological images, due to its low computational complexity in comparison with image energy based models for segmentation. However, the watershed model is prone to over-segmentation. We improved its performance for histological images by introducing specific enhancements, and improved foreground and background marking. A tentative foreground map was used to avoid false detections. The problem of segmenting partially occluded nuclei remains unsolved here. Further improvements are required to make the proposed nuclear segmentation algorithm fully automated and more robust. For example, one can try using local or adaptive enhancement techniques and exploring image dependent marking schemes like the proposed background marking scheme. In the next section, we introduced deep CNN based nuclear segmentation which works for overlapping resolution.

4.3 Deep CNN based three class nuclear segmentation

Object recognition makes essential use of machine learning methods. Recently deep learning methods were shown to improve the state-of-the-art performance in speech recognition, object recognition and detection. Pioneering breakthroughs in images, video, speech and audio have been possible using the deep CNN [59]. We utilized CNN for the task of nuclei segmentation.

In the literature, CNN was used in one of the approaches to initialize the shape of the nuclei i.e. as a nuclei marker [60]. It was based on deep learning for shape initialization and dictionary learning for shape inference. A deep CNN prediction was morphologically post-processed to separate the overlapping such that it will yield two markers for the overlapped nuclei. Once the shape (marker) is initialized, final shape is inferred through the learned shape dictionary and statistics of the local neighborhood around marker. A different model was trained for

TH-1577_11610215

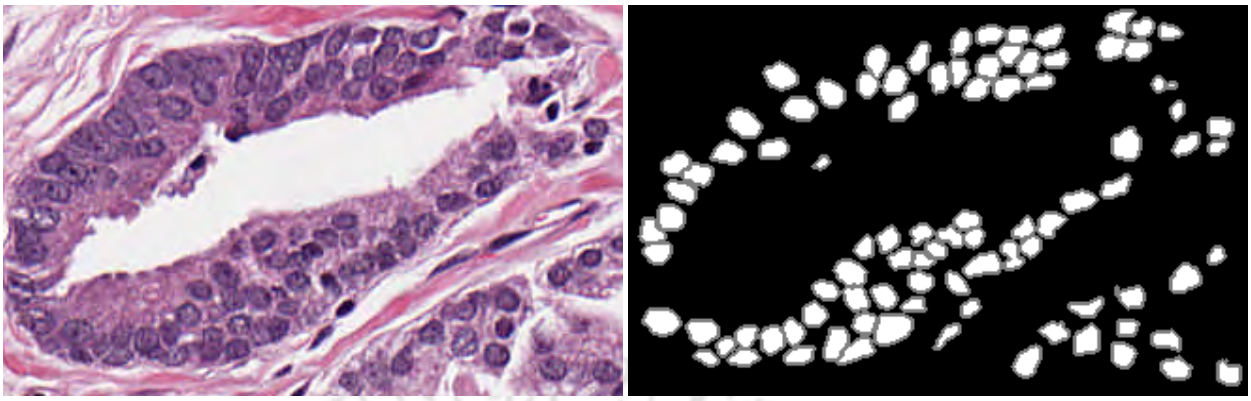


Figure 4.6: Prostate H&E stained image patch (left) and corresponding three class annotations (right).

different tissue and staining type, hence it is not generalized across the tissue type. The deep convolutional neural network (CNN) was found very accurate to initialize the shape i.e. to detect the nucleus. A superior performance was achieved when compared with the state-of-the-art.

We propose utilizing CNN for nuclei detection and overlapping resolution. For this, our novel idea is to use a CNN to label three classes of pixels: inside any nucleus, outside every nucleus, and on the boundary of any nucleus. It is the introduction of the third class that allowed us to separate touching nuclei. Also, our approach gives probabilistic detection of the boundaries such that a simple post-processing method such as region growing on the detected nuclei produces the desired nuclei segmentation.

We labeled the nuclei boundaries in our datasets and obtained a ternary mask for each train and test sets. The three classes of pixels were named inside, outside and boundary class. A typical annotation for the H&E stained prostate image and corresponding three class map is shown in Figure 4.6. These classes were formed by keeping the ± 1 pixel boundary class such that it invades equally in nucleus (1 pixel) and its background (1 pixel) to form a spatial complete ring. The two remaining parts i.e. nuclei inside and outside, were formed such that they do not invade the boundary class.

Given the three class output, the input to the CNN was a patch centered at the pixel to be classified. We trained the CNN architecture to learn three class discrimination i.e. it was

TH-1577_11610215

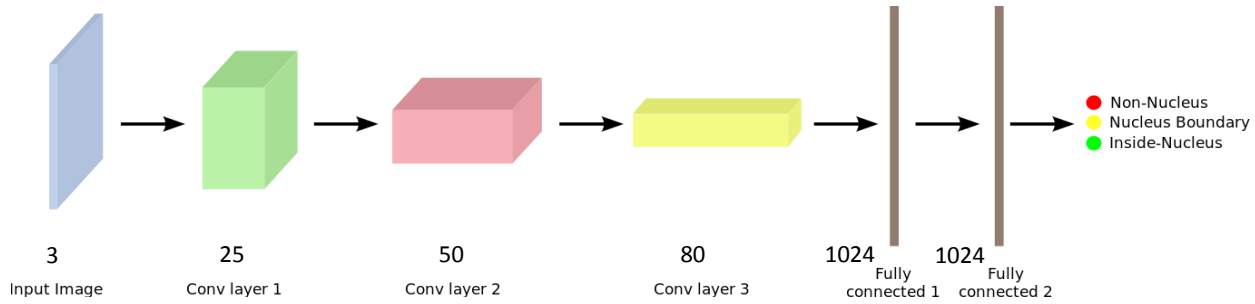


Figure 4.7: CNN architecture

trained to predict the probability of each pixel to belong to inside, outside and boundary class such that the sum of all three probabilities added to 1. Hence, we kept the last layer to be softmax activation with three nodes. We determined that three convolution layers and two fully connected layers were best suited for our problem. The patch around the central pixel helped the CNN learn meaningful information about the central pixel class. We extracted patches from 20 different images (one per patient). We trained CNN on $171,000$ number of training patches of size $51 \times 51 \times 3$ and validated on $45,000$ samples for 115 epochs. Inspired from Nair and Hinton [61] and Krizhevsky et al. [62], we used Rectified Linear Unit (ReLU) as a non-linear activation function in all convolutional and fully connected layers. This is because ReLU is non-saturating and speeds up training. We also used dropout [62] but with different probability (p) of dropping units in each layer. We increased p while going deeper in network because higher p in the initial layers results in information loss from the image while it acts as a good regularizer in deeper layers which avoids over-fitting. Table 4.3 shows the typical values of the CNN architecture and feature dimensions. Figure 4.7 shows the CNN architecture. Once the CNN was trained, it produced the probability map for all three classes.

CNN was trained using Torch [63] on a single GPU - NVIDIA Tesla K20c. Initial learning rate for the networks was 0.01 and was decreased by a factor of 10 when the validation error was approximately constant. For testing, we input the pixels in parallel to the GPU to reduce the prediction time.

The probability maps were post-processed to produce the nuclei segmentation. Nuclei can be detected easily by thresholding the inside class probability map. These detected nuclei acts

TH-1577_11610215

Table 4.3: CNN architecture dimensions

Layer	Filter size	Dimension	Dropout
Input image	-	51x51x3	0
Conv layer 1	4x4	48x48x25	0.1
Max-pool layer	2x2	24x24x25	-
Conv layer 2	5x5	20x20x50	0.2
Max-pool layer	2x2	10x10x50	-
Conv layer 3	6x6	5x5x80	0.25
Max-pool layer	2x2	3x3x80	-
Fully connected 1	-	1024x1	0.5
Fully connected 2	-	1024x1	0.5

as a marker in our proposed method similar to shape initialization in [60]. However, we have also predicted the nuclei boundary that helps for ease segmentation which was not the case in [60] (They required additional processing stages such as shape dictionary learning). The detected markers were iteratively grown one pixel till a spatial location was found such that the average boundary class probabilistic score of pixels on the boundary of dilated region was highest. As we grow the boundary by 1 pixel in each iteration, the average probability of the set of pixels in this grown region (p_{grown}) is compared with the average of non-zero values of the (CNN) predicted boundary probability ($p_{boundary}$) and we stopped if $p_{grown} \leqslant p_{boundary}$. Thus, we stop growing when there is high probability of 1 pixel grown area to belong to the boundary class. Also, this growing insured that there will be no overlapping due to 1 pixel growth in each iteration. This is because if overlapping happens, we just removed the growth in that iteration by 1 pixel erosion. In this way, the nuclei overlapping was avoided. A typical diagram for the post-processing is shown in the Figure 4.8.

4.3.1 Experiments and results

We used H&E stained prostate (CPCTR) and breast (TDLU) datasets to access the performance of our algorithm. We compared proposed technique with cell profiler pipeline[1] for nuclei segmentation [64]. In an interesting way, we trained our algorithm on the CPCTR breast histological images and found favorable segmentation results on independent TDLU dataset.

[1]The cell profiler pipeline is available at http://cellprofiler.org/examples.html/HumanCells

TH-1577_11610215

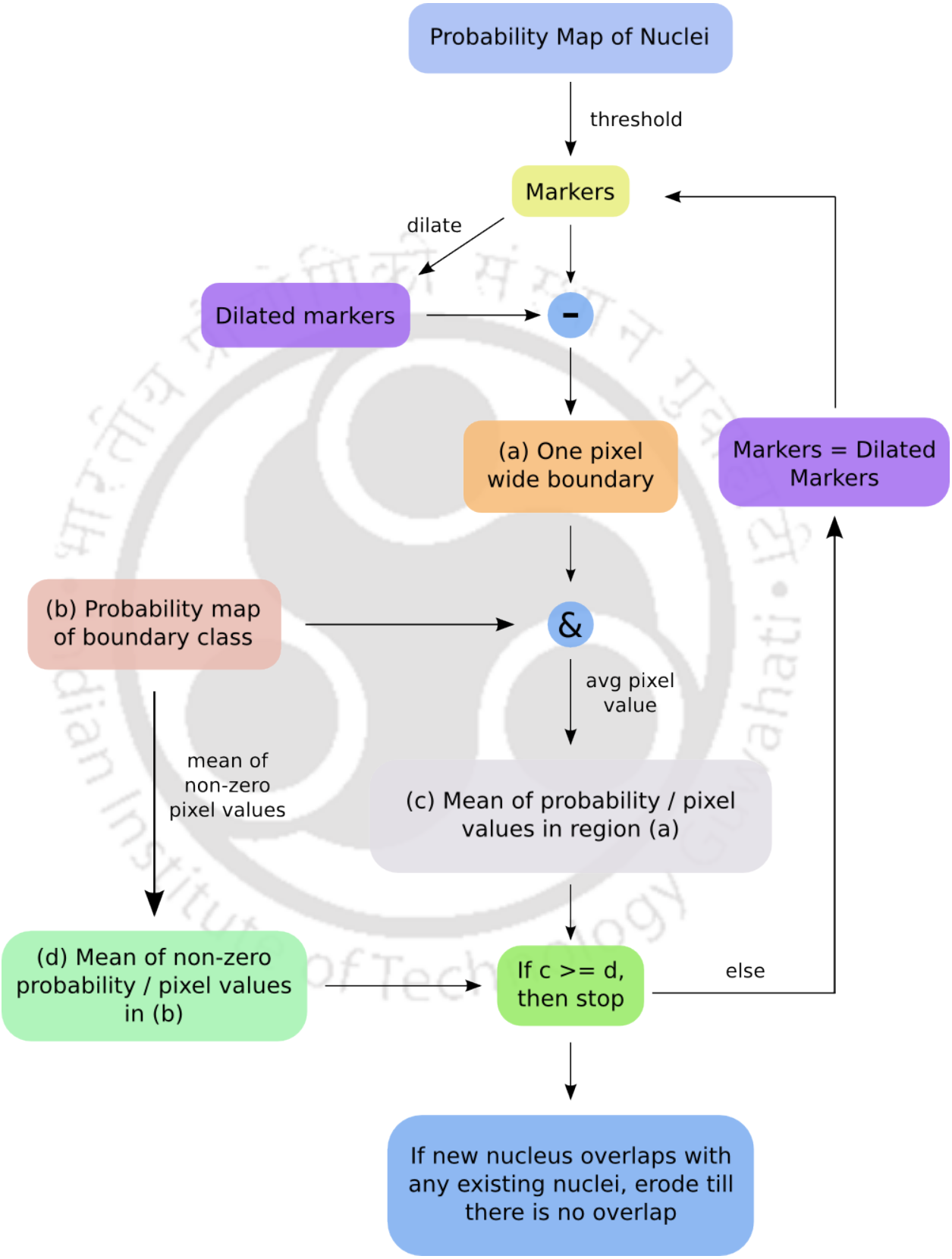


Figure 4.8: Flow diagram of post-processing

TH-1577_11610215

4.3.1.1 CPCTR

The Figure 4.9-4.10 shows visual result on different nuclei appearance prostate histological images. It shows that proposed algorithm works well on the variety of nuclei appearances proving its generalization ability. It was desired since the training images were selected such that they had diverse nuclei appearance. The proposed algorithm also resolves the nuclei overlapping. In Figure 4.11, a (red) circle indicates that overlapped nuclei were separated out by proposed algorithm.

We also performed quantitative validation on the CPCTR dataset. We compared proposed algorithm with the cell profiler nuclei segmentation pipeline. F1-score and object Dice coefficient metrics were used to access the nuclei segmentation performance. We annotated 1000 nuclei across 10 test images in the CPCTR dataset. The F1-score for proposed method and cell profiler were 0.80 and 0.75 respectively. The object Dice coefficient for proposed method and cell profiler were 0.70 and 0.62 respectively. It indicates that proposed algorithm outperformed the cell profiler nuclei segmentation pipeline.

4.3.1.2 TDLU

We evaluate the performance of the proposed algorithm on the independent (not used for training) other organ (breast) to assess the generalization ability of proposed algorithm. Figure 4.12 shows impressive segmentation results on the breast histological image even when breast tissue type was not part of training. This excellent performance was possible because of the color normalization since it suppresses unwanted color variation which could have been a bottleneck for training a generalized model previously [16][32]. However, the major contributor to our algorithms accuracy is likely to be the three class formulation of CNN output. We also credit to the training database which covered most type of nuclear variations in prostate tissue made available in the huge CPCTR database. Another example is shown in Figure 4.14 with marked nuclei overlapping resolution.

TH-1577_11610215

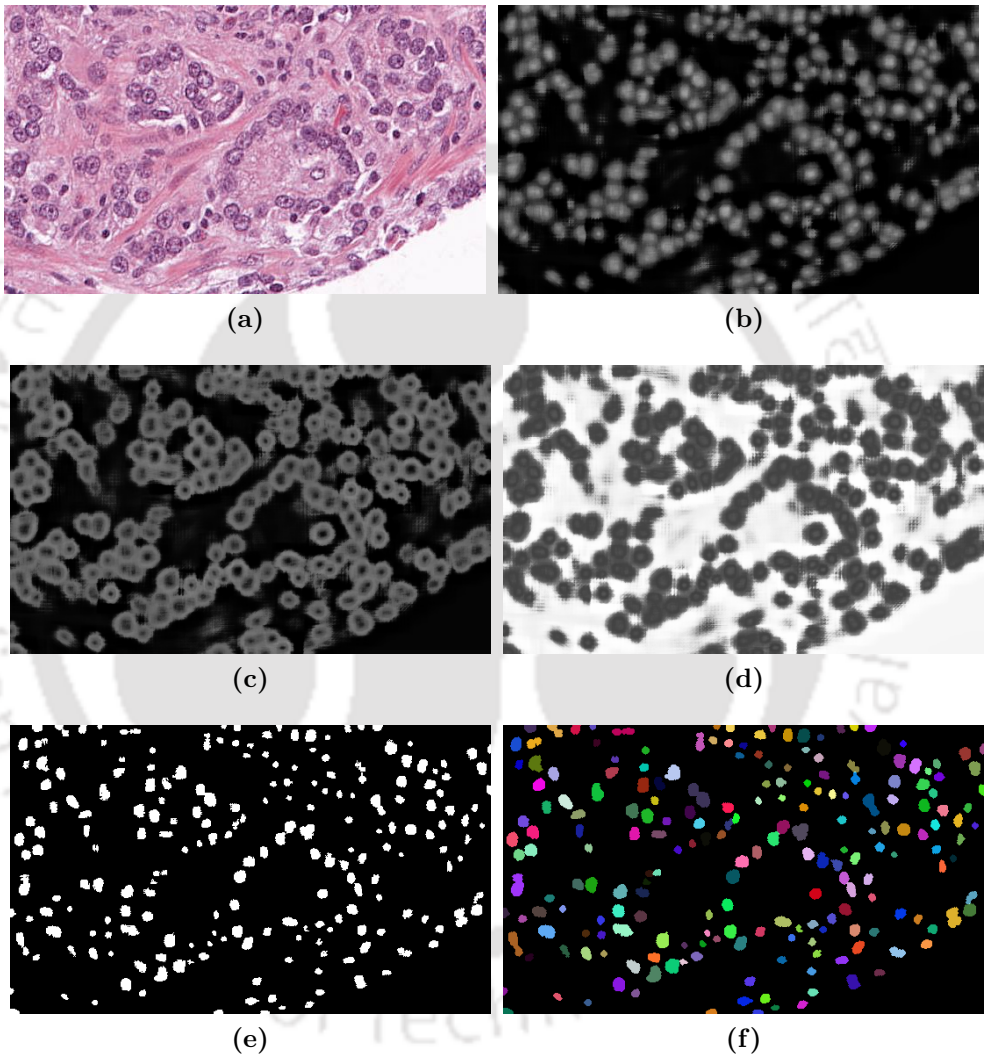


Figure 4.9: The CNN is given an input (a) which produce output (b, c, d) corresponding to inside, boundary and outside class respectively. (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f).

TH-1577_11610215

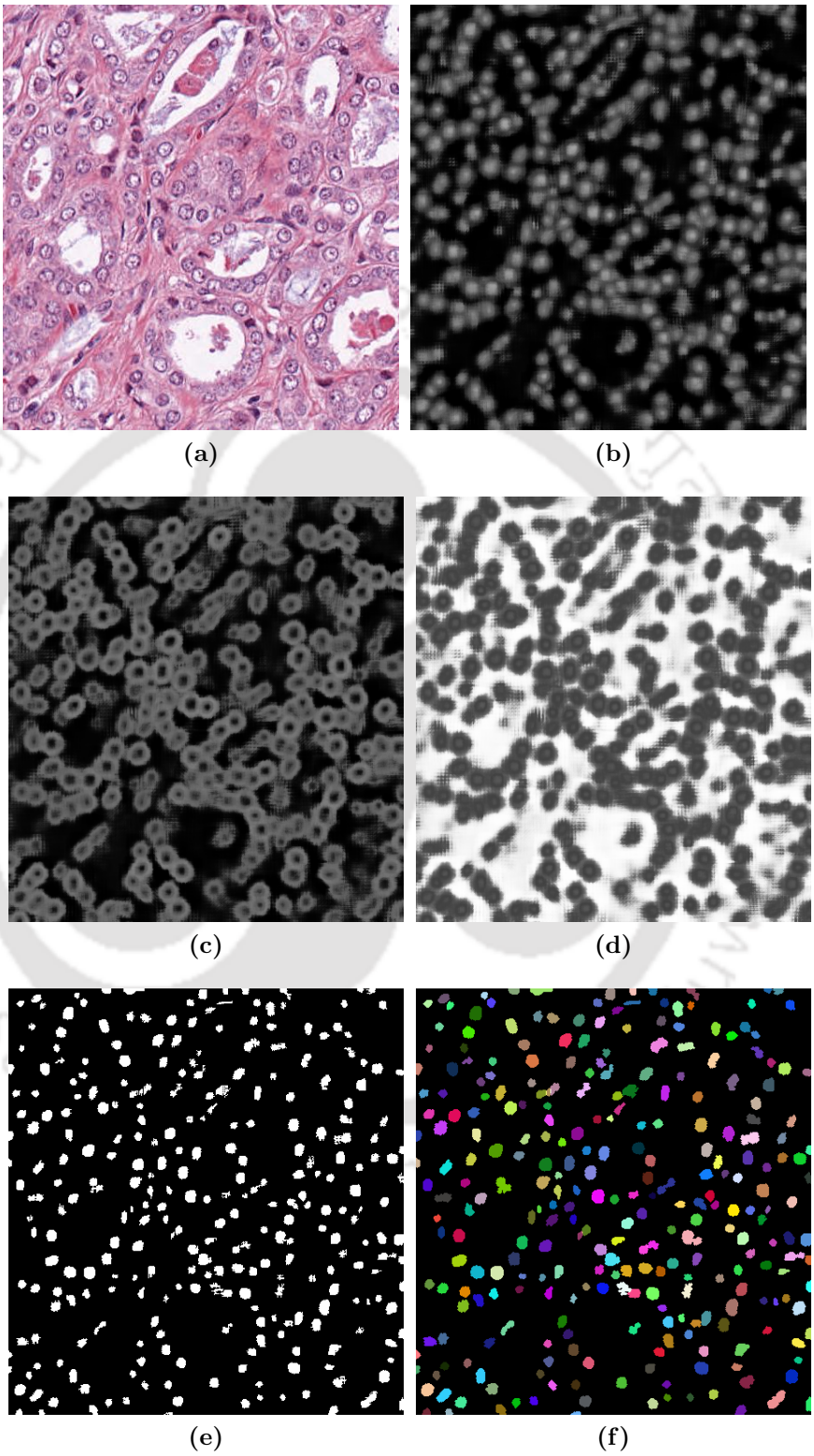


Figure 4.10: The CNN is given an input (a) which produce output (b, c, d) corresponding to (inside, boundary and outside class). (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f).

TH-1577_11610215

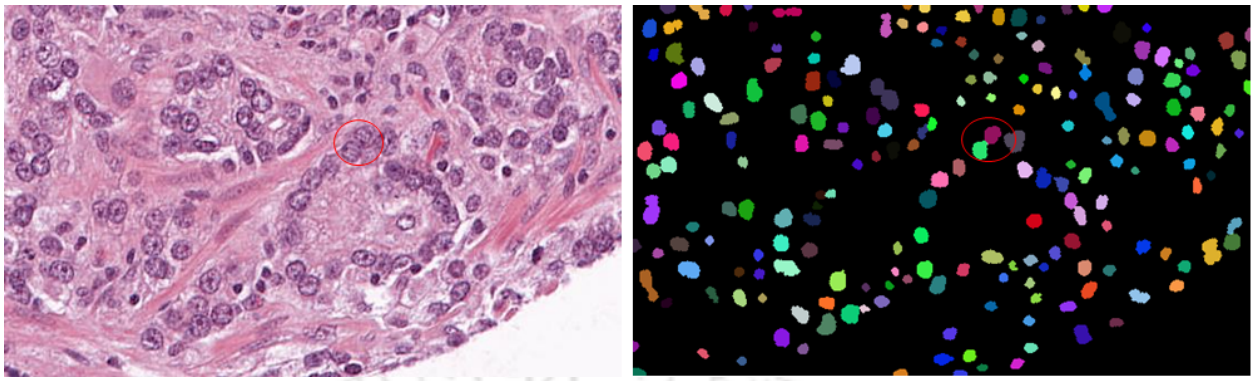


Figure 4.11: The overlapped nuclei were distinctly segmented

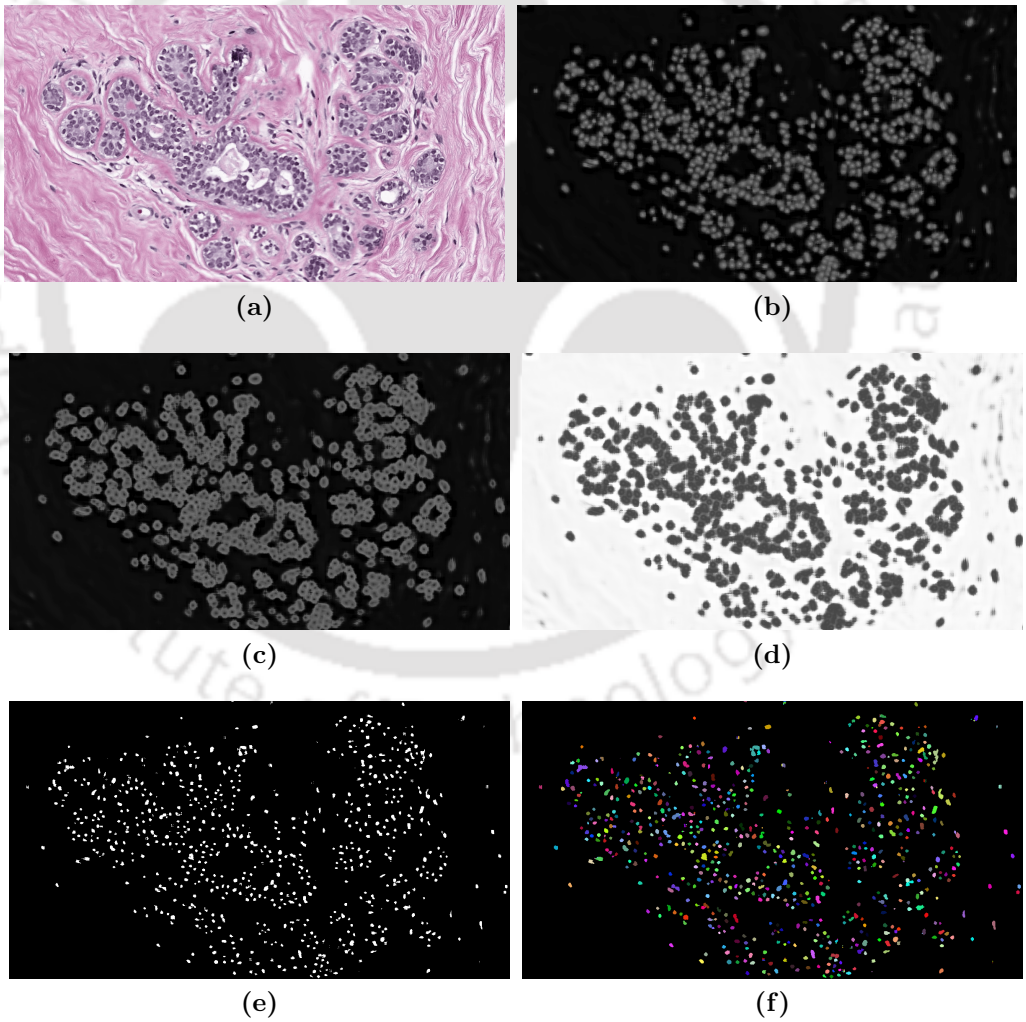


Figure 4.12: The CNN is given an input (a) which produce output (b, c, d) corresponding to (inside, boundary and outside class). (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f).

TH-1577_11610215

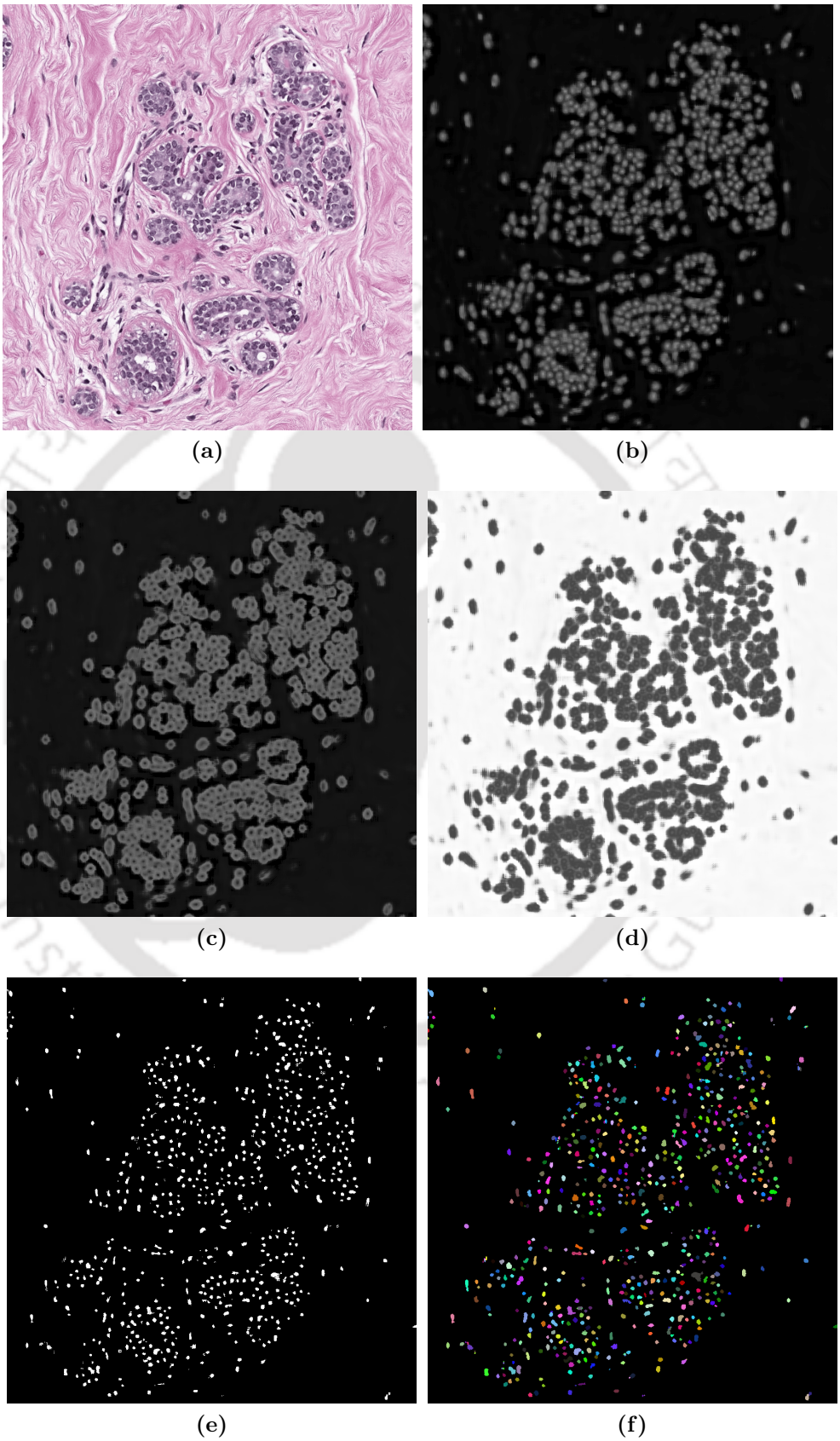


Figure 4.13: The CNN is given an input (a) which produce output (b, c, d) corresponding to (inside, boundary and outside class). (b) was thresholded to produce marker (e). Marker (e) was iteratively grown till the desired boundary to produce nuclei segmentation shown as color map (f).

TH-1577_11610215

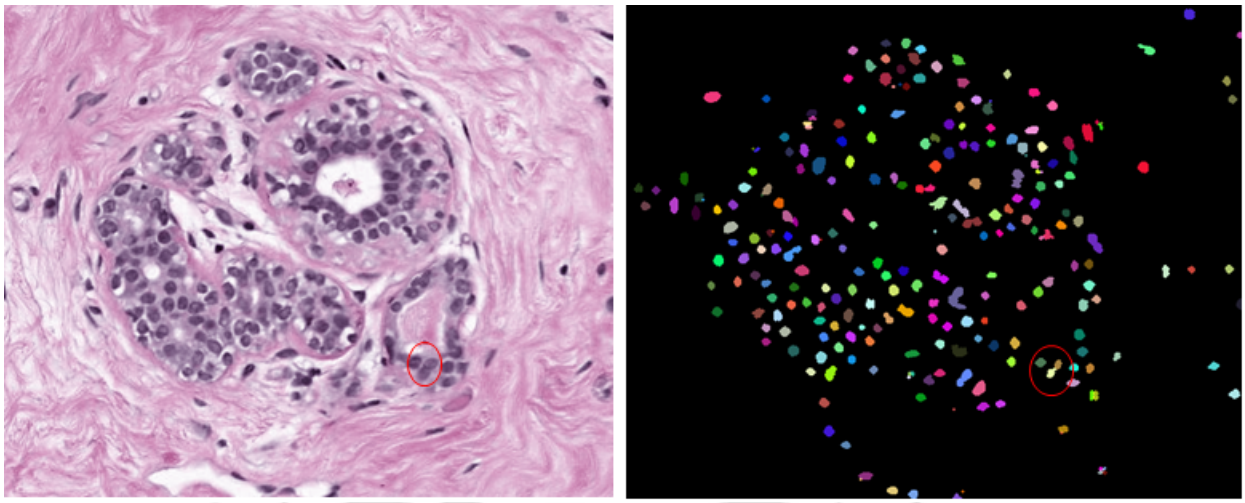


Figure 4.14: The overlapped nuclei were segmented separately.

4.4 Discussion

Given the fact that CNN can be used to learn meaningful features, we trained it to learn the discrimination between nuclei inside, boundary and outside class. Proposed method was successful in segmenting distinct as well as overlapped nuclei due to its effective three class formulation, use of color normalization for standardization and sufficient variety of nuclei variations in the training dataset. In future, we would like to train on few tissue types and learn a generalized nuclei segmentation model which works on most of the other tissue types not used for training.

TH-1577_11610215

5

Breast Cancer HER2 Score Prediction

Contents

5.1 HER2 and its clinical scoring of breast cancer 71
5.2 HER2 predictor 73
5.3 Complete membrane staining estimator 77
5.4 Experimental Results 83
5.5 Discussion and Conclusion 85

Cancer is a leading cause of death worldwide, accounting for 15% of all cancer death [65]. There has been a significant fall in the breast cancer mortality rate during the last two decades (from 33.1% to 21.5%) because of awareness about self-examination, breast cancer screening programs and improvements in early detection systems. The advancement in development of computational pathology algorithms would improve the detection and diagnosis rate further. Clinical screening can be done by breast self-examination (BSE), mammography (MMG), magnetic resonance imaging (MRI), and pathology. MMG is the X-Ray of breast to detect from very small to large breast tumors and it is recommended annually for women above 40 years of age. A meta-analysis showed that MMG decreases the breast cancer mortality rate by 25-30% in women below 40 years [66]. BSE helps in early detection of cancer and is recommended for

TH-1577_11610215

Table 5.1: HER2 staining score, pattern and assessment [2]

Score	Pattern	Assessment
0	No membrane staining observed or faint staining in less than 10 % of tumor cells	Negative
1	Faint, weak membrane staining in some proportion of tumor cells	Negative
2	Feeble, non-uniform membrane staining in greater than 10% of tumor cells (recommended FISH testing)	Equivocal (Weakly Positive)
3	Vigorous, intense membrane staining is observed in more than 10% of the tumor cells	Positive

women above 20 years once in 6 months. MRI is preferred only in high risk individuals and it can detect breast cancer most accurately. American cancer society recommend breast cancer screening tests for all women above 40 years to improve the detection and survival rate [67]. Digital pathology is not yet clinically accepted for accurately grading breast cancer since the tissue biopsies of affected region are screened by experts at nuclear and glandular level.

5.1 HER2 and its clinical scoring of breast cancer

Human epidermal growth factor receptor 2 (HER2) proteins are produced by HER2 gene. HER2 proteins are receptors on the breast cells. These receptors aid to control the healthy breast cell life functions like growth, division, and repairing. However, in about 20-25% of breast cancers, the HER2 gene produces many copies of itself (known as HER2 gene amplification). All these extra HER2 genes direct the breast cells to make many HER2 receptors and this condition is known as HER2 overexpression [68]. This makes breast cells grow and divide in an uncontrolled way. HER2 overexpression almost exists in 25% of invasive breast cancer cases. Scoring of HER2 overexpression is an important in prognostic and predictive cancer model design. In clinical practice, pathologists observe and analyze morphological features to give a HER2 score (0 ,1, 2+, 3+) to each breast cancer slide stained using HER2neu stain. The criteria for HER2 scores used by pathologists is shown in Table 5.1.

Immunohistochemistry (IHC) refers to the process of detecting antigens (e.g. proteins like HER2) in cells of a tissue section by exploiting the principle of antibodies binding specifically to

TH-1577_11610215

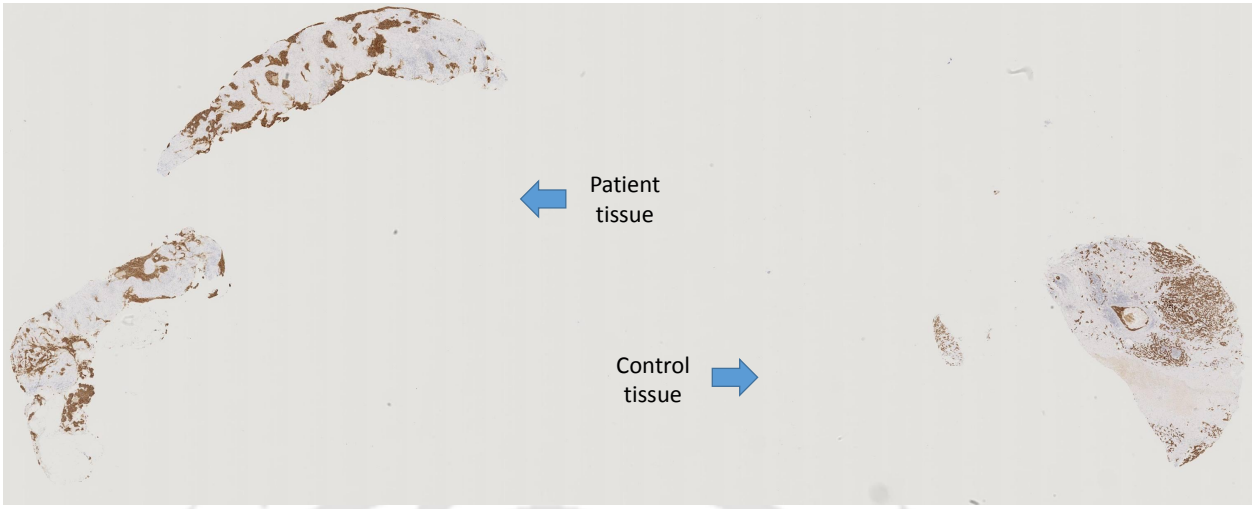


Figure 5.1: A typical whole slide image in our dataset includes patient and control tissue [2]. Control tissue was HER2 3+ for all WSI. Note that from now onward we showed down sampled version of the whole slide image for better visibility. The downsampling factor was 32.

antigens in biological tissues. Determination of HER2 receptor status is done by a pathologist who scores HER2neu stained slide of a patient to inform an oncologist whether to use anti-HER2 therapies such as trastuzumab (Herceptin®). Such assessment suffers from inter and intra (human) observer variations in diagnosis of tissue biopsy samples [9][69][70]. Color variation is inevitable in histology [32] as discussed in Chapter 3. There are other factors which cause more color variations in immunohistochemistry. These include ischemic time, tissue fixative and fixation time, tissue processing, efficiency of epitope retrieval, selection of antibody or its clone and detection system [71]. Hence, IHC staining control is required which can suppress such variation [71] since it acts as a standard or benchmark staining reference for grading the patient IHC WSI. In our HER2 contest dataset [2], the control tissue was HER2 3+. An example of a typical image with control and patient tissue is shown in Figure 5.1.

As shown in Table 5.1, based on the attributes of IHC complete membrane staining in whole slide image (WSI), HER2 score (0, 1, 2+ or 3+) is evaluated by pathologists. Scores 0 and 1 are accessed as clinically negative. However, scores 2 and 3 are positive and are treated with anti-HER2 therapies.

As far as we know, there is no literature on the automated estimation of HER2 score

TH-1577_11610215

and percentage of complete membrane staining (PCMS) as of now. In fact this problem was organized as a HER2 scoring contest challenge [2] and the results of the contest are not yet published. We propose a machine learning framework that predicts the HER2 score as well as simple image processing framework which estimate PCMS.

5.1.1 Contributions

- We trained a deep CNN to learn the patch HER2 score prediction and accumulate the patch scores to predict the HER2 score for WSI through a novel empirical criteria. In addition to the CNN learned features, we input the hand crafted average control tissue information to first fully connected layer so that CNN can normalize the variations in IHC.
- PCMS is estimated through a simple image processing approach which estimates the normalized tumor and IHC stained area and compute the percentage by simple division of these two factors.

Validation of the proposed algorithm on the independent testing images demonstrate the effective prediction of HER2 score and PCMS. Even in difficult cases of WSI where pathologist gave wrong HER2 scores, our algorithm predicted accurate HER2 score.

5.2 HER2 predictor

Deep learning methods, especially deep CNNs, have shown state-of-the-art performance in speech recognition, object recognition and object detection [59]. Additionally, CNNs can learn meaningful discriminating filters for the classification task. Because of these motivating factors, we first trained a CNN to predict HER2 score of a patch. Then, the HER2 scores of the patches are aggregated to predict the HER2 score of the IHC whole slide image through an empirically derived criteria. Flow diagram for HER2 prediction is shown in Figure 5.2.

Training any deep CNN architecture directly on the huge WSI is not possible because of memory restrictions of the current low configuration computers as well as increase in the input

TH-1577_11610215

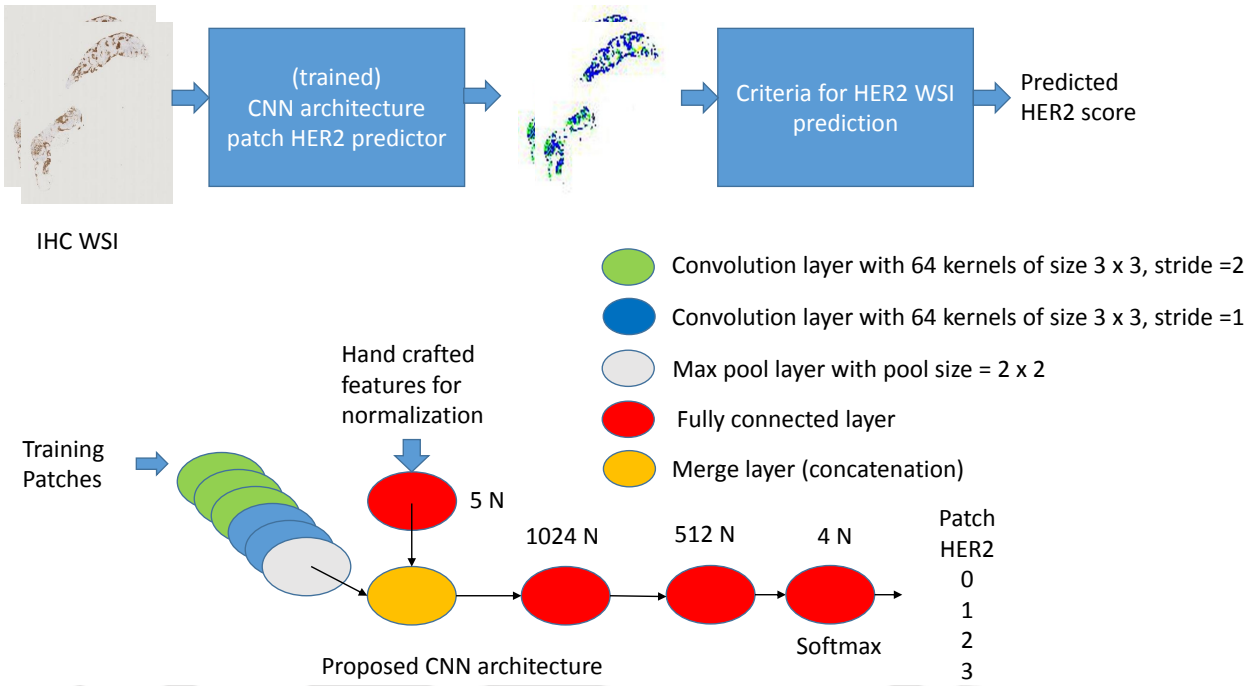


Figure 5.2: Flow diagram of the HER2 predictor

dimensions as we try to process large images in one go. Hence, we trained a patch based CNN to learn the HER2 score of the patch. A typical image in HER2 dataset [2] is shown in Figure 5.1. It shows that a control tissue is usually put on the same slide as the patient tissue in each image. Hence we need to separate the control tissue from patient tissue which is done by simply taking a mid-vertical line to separate the control from patient tissue. Also we calculate the control tissues average stain intensity. It is used as a handcrafted feature to the first fully connected layer of the CNN architecture in addition to the CNN learned features. These hand crafted features were included so that different IHC variations can be normalized such as color [32] and other IHC variations [71].

5.2.1 Preprocessing and patch sampling of WSI

As the WSI have stained region (relevant) and non-stained background (non-relevant), we removed the non-informative background in the pre-processing in all stages our proposed algorithms. In each WSI (like shown in Figure 5.1), the control tissue is segmented out by vertically dividing the WSI into two equal parts (the other part being patient tissue). Average control

TH-1577_11610215

stain information is computed from the control tissue part and was used as hand crafted feature for normalization. First we remove the non-stained background using Otsu thresholding on the gray scale image. Then we took separate average of the r, g, b channels and this length 3 vector was average control tissue information. This was used as one of the input to the fully connected layer in the CNN architecture as shown in the flow diagram Figure 5.2. Each WSI is segmented into non-overlapping patches of size $512 \times 512 \times 3$. Then, relevant patches were obtained by the following criteria: If the gray scale mean of every patch is less than 220, it is categorized as useful since unstained background has high intensity. This operation removes 60-70 % of non-relevant patches from training. Obtained patches were downscaled to the size of 224x224x3 to reduce the computations by reducing number of pixel with negligible decrease in quality. We sampled non-overlapping 40,671 training patches from 52 training WSI and divided them into 80 % for training and 20 % for validation. The sampled patches from single WSI were annotated with the same HER2 score as that of the WSI. These patches were used to train the CNN architecture shown in Figure 5.2. We used 28 independent WSI for testing.

5.2.2 CNN patch HER2 estimator

The CNN is trained on the training patches to predict their HER2 score. The typical values of the CNN architecture are listed layer wise in Table 5.2. There are five convolution layers followed by 2×2 pooling layer and two fully connected layers. In addition to the CNN feature maps obtained at the output of pool layer, we added average control tissue information as a hand crafted feature to normalize the variations in IHC histological images. The hand crafted features were input to a separate five neuron layer with ReLU activations which produce five outputs. These five outputs are concatenated with the pooling layer outputs. The last layer is four output softmax layer to predict the probability maps for the HER2 scores i.e. 0, 1, 2+, and 3+. ReLu activations were used in all layers except the output layer. Patches of size 224x224x3 were given as input to the first layer of the architecture and average control tissue information was given as input to dense layer (Layer 8) as listed in the Table 5.2.

We trained the CNN on 40,671 training samples. It took 15 hours to train on 2 GB NVIDIA

TH-1577_11610215

Table 5.2: CNN architecture typical values

Layer	Type	Input dimension dimension	Filter	Stride	Pad	Output dimension
1	Conv	224x224x3	3x3x3x64	2	1	112x112x64
2	Conv	112x112x64	3x3x64x64	2	1	56x56x64
3	Conv	56x56x64	3x3x64x64	2	1	28x28x64
4	Conv	28x28x64	3x3x64x64	1	0	26x26x64
5	Conv (dropout=0.25)	26x26x64	3x3x64x64	1	0	24x24x64
6	max Pool	24x24x64	2x2	1	0	12x12x64
7	Dense	3x1	None	None	None	5x1
8	Merge 6 & 7	None	None	None	None	9221
9	Dense	9221	None	None	None	1024
10	Dense (dropout=0.5)	1024	None	None	None	512
11	Dense (dropout=0.5)	512	None	None	None	4

GPU in Keras for 65 epochs. Keras is a deep learning library for theano and tensorflow [72]. Average testing time on WSI was 6 minutes.

5.2.3 Criteria for WSI HER2 estimation

The CNN predicts the HER2 score of each patch in WSI. We observed that CNN labels the patches of single WSI from 0 to 3. This is because of the heterogeneity of structures in WSI. We proposed novel criteria to predict WSI HER2 score from the HER2 scores of its patches. Let CNN_{score0}, CNN_{score1}, CNN_{score2}, CNN_{score3} are the number of patches from a single WSI which are predicted as HER2 score 0, 1, 2+ and 3+ respectively. Let N be the total number of patches of WSI after removal of irrelevant patches. Then we proposed following criteria to select the grade.

1) If $CNN_{score3} > 8\%$ of N, WSI was given a score 3.

2) If $CNN_{score2} > 40\%$ of N and $CNN_{score3} < 8\%$ of N, WSI was given a score 2.

3) If $CNN_{score1} > 14\%$ of and $CNN_{score2} < 40\%$ of N and $CNN_{score3} < 8\%$ of N, WSI was given a score 1.

4) If all the above fails, WSI was classified as score 0.

TH-1577_11610215

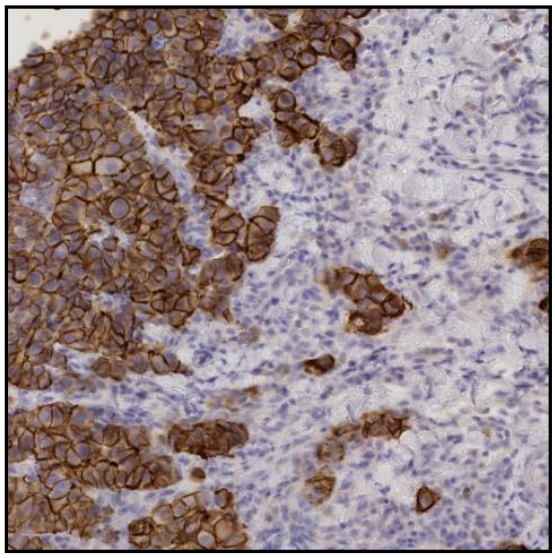


Figure 5.3: IHC image patch that have tumor cells with complete membrane staining (in brown)

This criterion gave the HER2 score to WSI from its patches HER2 score. Typical values of the above thresholds were empirically evaluated and motivated from how pathologist grade the WSI [73].

5.3 Complete membrane staining estimator

The percentage of complete membrane staining (PCMS) is percentage of tumor cells which have complete membrane staining. An example patch of the IHC WSI having complete membrane staining is shown in Figure 5.3.

According to the updated guidelines for HER2 assessment [73], there are two crucial factors in determining the score of HER2 over-expression. They are the percentage of tumor cells exhibiting complete membrane staining and intensity of membrane staining. Although HER2 score and percentage of membrane staining are correlated, it would be worthwhile to report the percentage tumor cells exhibiting complete membrane staining (irrespective of intensity) along with the computer-predicted HER2 score. This will aid the pathologist to evaluate the reliability and effectiveness of the automatic HER2 prediction as well as validate their PCMS score.

TH-1577_11610215

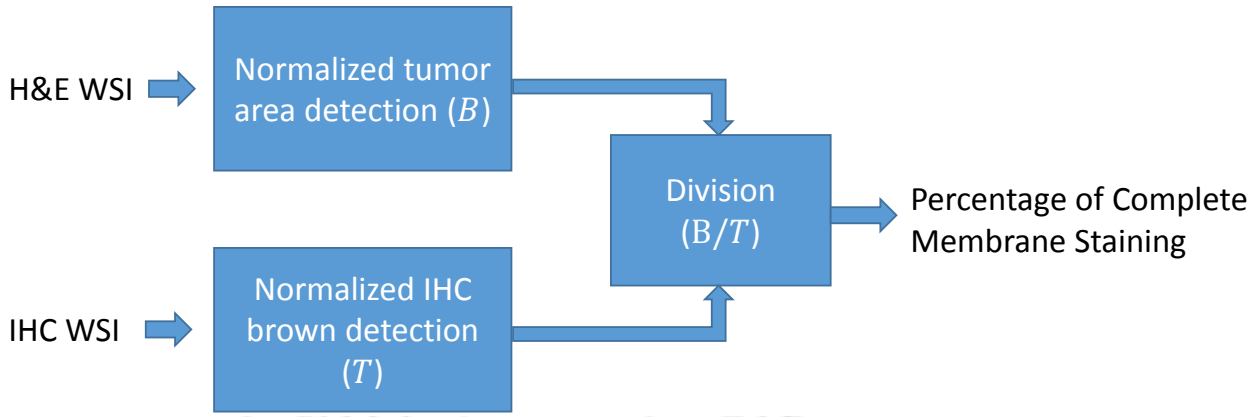


Figure 5.4: Flow diagram of percentage of complete membrane staining (PCMS)

We proposed a simple image processing approach to estimate the percentage complete membrane staining (PCMS). The flow diagram of PCMS is show in Figure 5.4. It includes the computation of normalized tumor area in H&E image and normalized brown stained area of the IHC image. Then, ratio of normalized tumor to the normalized brown resulted in the PCMS. As it is impractical to attempt image processing on the massive whole slide images using normal configuration of resource, we employed the patch processing approach. As the WSI image is of huge size, we can save the computational power of a resource by down sampling WSI with negligible effect on visual quality of histological structures. Thus, we downsampled the WSI image by factor 2 and extracted patches of size 1000×1000. In similar line of HER2 prediction, we neglected the patch whose $mean > 230$ AND less edge pixel (<10% of total pixels in a patch). The edge pixels were detected by canny edge detector [74]. We propose three-step approach to estimate the PCMS. It include H&E tumor detection, IHC brown staining detection and computation of PCMS.

5.3.1 Tumor detection (T)

The first logical step is to determine the extent of tumor regions in a given patch so that we can locate potential tumor cells in that patch. Although, tumor region was not annotated, careful observation H&E WSI revealed the subtle visual differences between nuclear structures

TH-1577_11610215

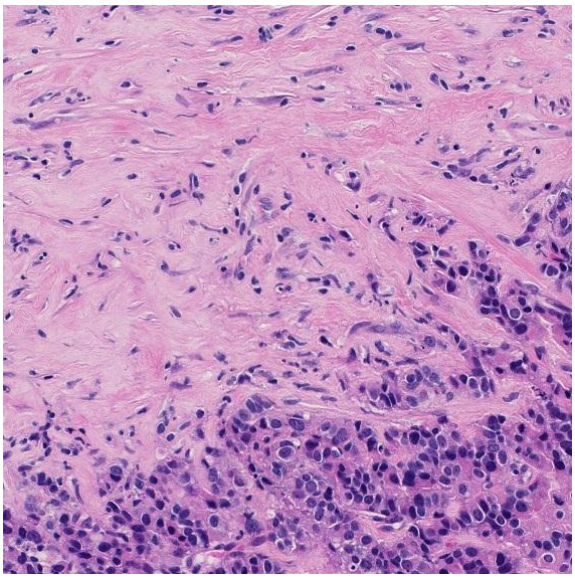


Figure 5.5: H&E image of tumor-affected tissue. Different nuclear structures indicate a high likelihood of tumor presence at the lower bottom of the image patch.

of tumor and healthy tissue regions[1]. Tumor cells can be separated by their distinct appearance. For example, nuclear structures of tumor cells were round, enlarged, blue-black blobs, in contrast to the small, less round structures of normal cells. Also, in the tumor active region, the rate of cell division is uncontrolled and thus areas of high nuclear spatial density were detected as potential tumor regions. An example image with high tumor likelihood is depicted at the bottom of the image in Figure 5.5.

A flow diagram of the potential tumor region is shown in Figure 5.6. We isolate the blue nuclei structures into a separate image for better processing. This is accomplished using SNMF stain separation [32]. It decomposes H&E stained image into two structure preserved image components corresponding to H and E. As H binds to the nuclei, we select the H channel using the color appearance vector W as done in [32]. Then, a simple Otsu thresholding on this H component gave us the binary image of the nuclear structures. We used contour-detection algorithm to create contours around every individual nuclear structure [75]. Identified structures are filtered based on area (> 250 pixels) and eccentricity (< 0.7) to retain only the round, engorged blobs which act as estimated potential tumor markers that most likely are tumor

[1]Each IHC stained WSI was accompanied with adjacent tissue section of H&E stained WSI [2]

TH-1577_11610215

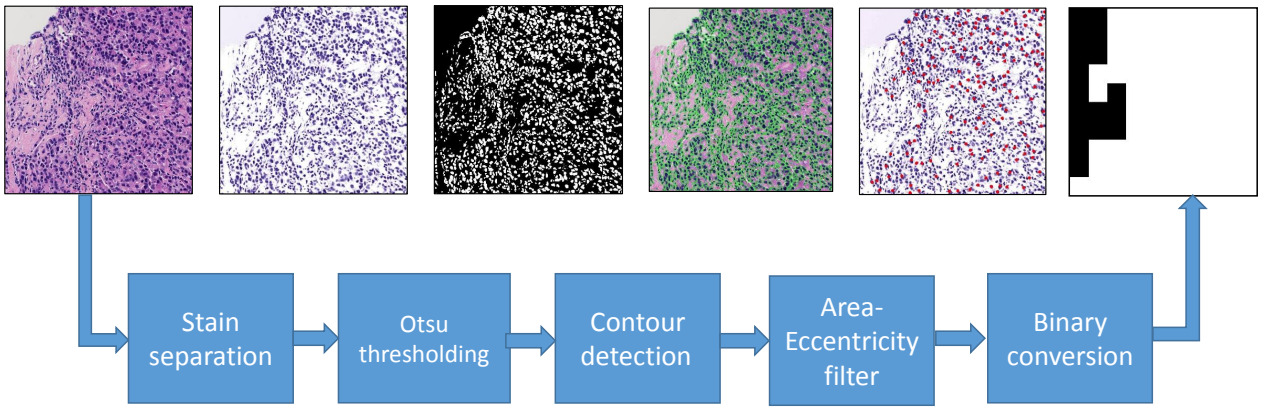


Figure 5.6: H&E tumor detection

regions. We divide the image patch further into small patches (boxes here onwards) of size 100×100. If there is potential tumor marker (red dot) in each box, then we assign that box as tumor i.e. the cells contained in the box are tumor cells (shown as logical 1 (white) in the Figure 5.6-right most image). This estimated binary probability map indicates the percentage of potential tumor region (ratio of potential tumor pixels to the total pixels). For example in Figure 5.6, the algorithms returns the number $t_i = 0.83$, implying that 83% of the patch is covered by potentially tumor-invaded regions. The same operation was performed on every relevant patch we sampled. The effective tumor region T per WSI is the average of the t_i across all the selected patches.

$$T = \sum_{i=1}^{Nsamples} t_i \tag{5.1}$$

5.3.2 IHC-brown stain identification (B)

The IHC stains the cell membranes of tumor cells by shades of brown. As we have potential tumor region T, we can compute the brown colored regions so that we can identify the proportion of tissue cells exhibiting complete membrane staining. An example of complete and incomplete membrane staining is shown in Figure 5.7. The SNMF stain decomposition would be perfect for isolating the brown stained regions [32]. However, we chose a computationally cheaper alternative to identify all the brown pixels in the image by empirically designing a threshold. The IHC brown stain was found to have the characteristic RGB composition of the

TH-1577_11610215

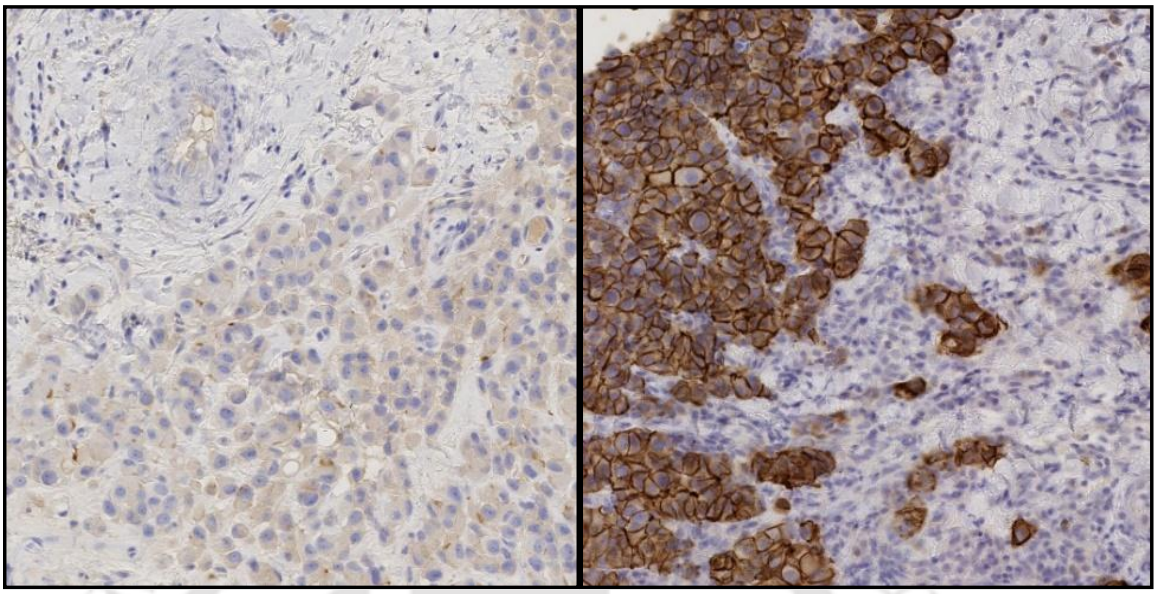


Figure 5.7: Incomplete (left) and complete (right) membrane staining

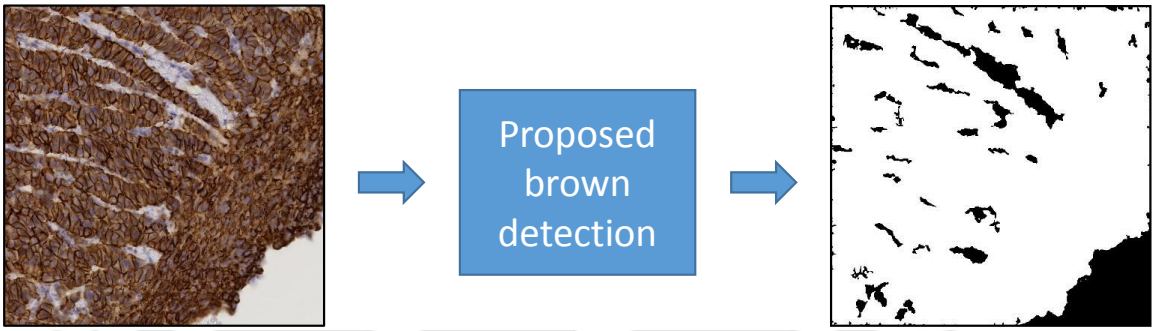


Figure 5.8: IHC image patch (left), detected brown region (middle) and detected contours (right)

form $[x, x+30, x+45]$, where the value of x determines the shade of the brown stain. Hence, for each pixel (r,g,b), we compute v1= [r-r, g-r, b-r]. Then we estimate the Euclidean distance d between v1 and v2 = [0, 30, 45]. Then we threshold d by an empirical threshold (¡30) to retain the pixel which were IHC stained. A typical example of IHC brown detection on a patch is shown in Figure 5.8. This operation is repeated on each patch to obtain the binary WSI indicating the IHC stained pixels and regions.

Sometimes there could be weakly brown stained area in background area due to spillover of stains. Tissue folds and other artifacts can also be stained brown. We should ignore such false positives while estimating the staining percentage. Also, the pixels with very light brown

TH-1577_11610215

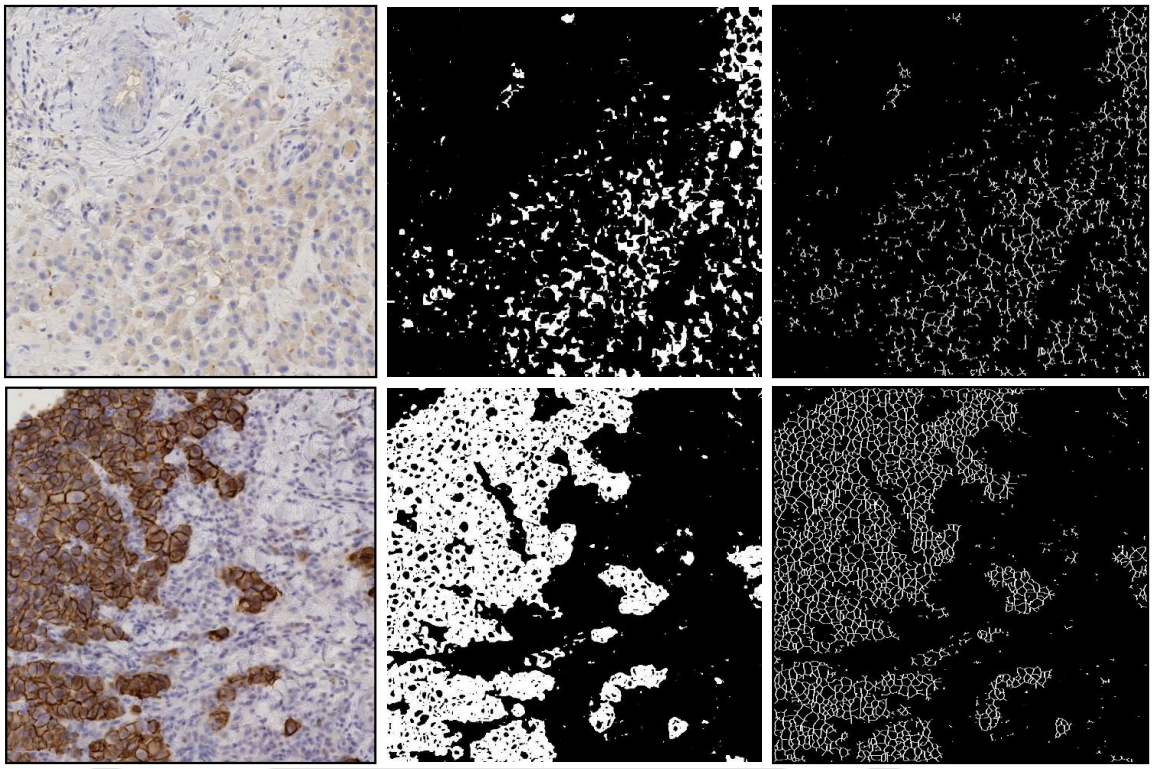


Figure 5.9: IHC image patch (left), detected brown region (middle) and detected contours (right)

shade are incomplete membrane staining as depicted in Figure 5.7. To remove the unintentional staining artifacts and incomplete membrane staining patches, we used the morphological skeleton operation in determining the completeness of membrane staining on the binary mask corresponding to brown stained region [76]. Since complete membrane staining forms the chicken-wire structure [73], the skeleton of IHC binary image is expected to closely resemble the chicken-wire. The spillover to the background is isolated from the usual texture of IHC stained histological structure, hence the skeleton would appear extremely fragmented and broken. Figure 5.9 shows the skeletons on the complete and incomplete membrane stained patches. The skeleton indicates that images with chicken wire are complete and broken fragmented structure are incomplete membrane staining. We detect that whether the structures in the binary image are broken or chicken-wire based on the reconstruction error or no overlap region. We compute a logical XOR of the binary image (output of binary conversion stage in Figure 5.8) and hole-filled binary skeleton to compute region of no overlap. If the region of no overlap is more than

TH-1577_11610215

20% of the whole patch, we called the patch as incomplete membrane stained. In this way, we ignored the patches having incomplete membrane staining. The binary mask corresponding to each patch p_i which is labeled complete was used to compute the PCMS. Let b_i be the patch PCMS. b_i is computed as ratio of detected brown pixel to the total pixel. We compute the average of all b_i across patches to compute the percentage of IHC staining per WSI. We call this quantity B.

$$B = \sum_{i=1}^{Nsamples} b_i \tag{5.2}$$

As the algorithm is based on patch processing and we normalized individual t_i and b_i, we do not need to register the consecutive tissue of H&E and IHC WSI for the correspondence of regions. This is a significant benefit of proposed algorithm.

5.3.3 Percentage of complete membrane staining (PCMS)

The problem of estimating the percentage of tumor cells exhibiting complete membrane staining now devolves into a simple division of two numbers. The quantity T indicates the normalized tumor area and the quantity B represents the normalized staining area. Hence, the fraction $\frac{B}{T}$ characterize the percentage of tumor cells having complete membrane staining.

$$PCMS = \frac{B}{T} \times 100 \tag{5.3}$$

5.4 Experimental Results

We validated the algorithm quantitatively and visually on the independent test data provided in HER2 contest [2]. We start with the quantitative validation.

Quantitative results: As shown in Table 5.1, HER2 score 0 and 1 are negative and hence we treat the both classes as one common class and evaluate the accuracy for HER2 prediction. The HER2 prediction algorithm accurately predicted HER2 score of 26 out of 28 WSI i.e. with 92.85 % accuracy. In some of cases, our algorithm outperformed the experts. A test WSI is shown in the Figure 5.10. There were three new experts that accepted to grade the testing WSI for its HER2 score and gave HER2 score 1, 1 and 0 to the WSI shown in Figure 5.10. The ground

TH-1577_11610215



Figure 5.10: Test WSI (HER2 3+). Three new experts gave HER2 score as 1, 1 and 0 while our algorithm accurately predicted 3+.

Table 5.3: Guidelines for the PCMS scoring [2]

Ground Truth	Criteria	Points
0	None	None
1	If the ground truth is $\geq$ 3% and algorithm reports % $\pm$ 2% of the ground truth	3
	If the ground truth is $<$ 3% and algorithm reports $<$ 3%	1
2	Within $\pm$5%	5
	Within $\pm$10%	2.5
3	Within $\pm$5%	5
	Within $\pm$10%	2.5

truth was 3+. Our algorithm predicted it accurately. This is a significant breakthrough such that experts can now get a second opinion, and the non-experts can rely on such an automated system.

PCMS was in the range from 0 to 100 indicating the percentage of tumor cells having complete membrane staining. We followed Table 5.3 as suggested in HER2 contest [2] to evaluate the PCMS detection. In that, a maximum score is given when the predictions were closer to the ground truth or exactly equal. Accordingly, we estimate the PCMS for the 28 testing WSI and provide a score to each predictions according to criteria mentioned in the Table 5.3. We obtained the score of 22.5. It indicates that we have a fair prediction of the PCMS since the score was not too low.

Visual results: In Figure 5.11, we show an example of CNN patch prediction on typical WSI.

TH-1577_11610215

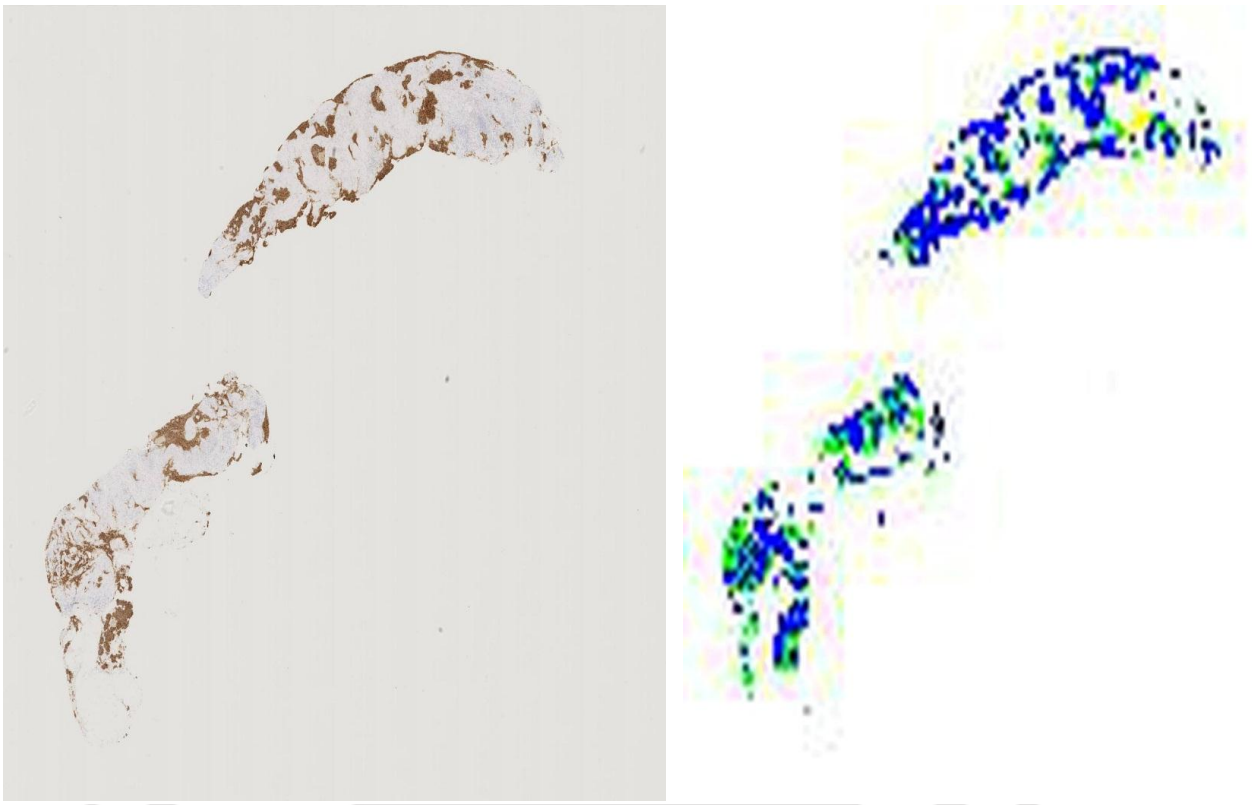


Figure 5.11: Proposed patch CNN prediction on WSI. The WSI was accurately scored as HER2 3+.

The prediction were color coded as yellow, red, green, and blue corresponding to HER2 score 0, 1+, 2+, and 3+ respectively. As shown, most of patches were labelled 2+ (green) and 3+ (red) by CNN and our criteria accurately assigns 3+ to this WSI. The visual result of meta-processing regarding PCMS is shown in Figure 5.12. We cropped approximately same patch from the H&E and IHC image for detecting tumor region and detecting the IHC stained area respectively. These two patches were processed to compute t and b. Figure 5.12 shows binary images we obtained before computing the t and b. The proposed algorithm for HER2 prediction and complete membrane staining was ranked first in the HER2 contest [2].

5.5 Discussion and Conclusion

HER2 prediction from breast cancer WSIs was a challenging problem because of the heterogeneity of tissue structure across WSI. Hence, HER2 over expression detection is essential in diagnostic and prognostic cancer model design because 25% of the breast cancer cases report

TH-1577_11610215

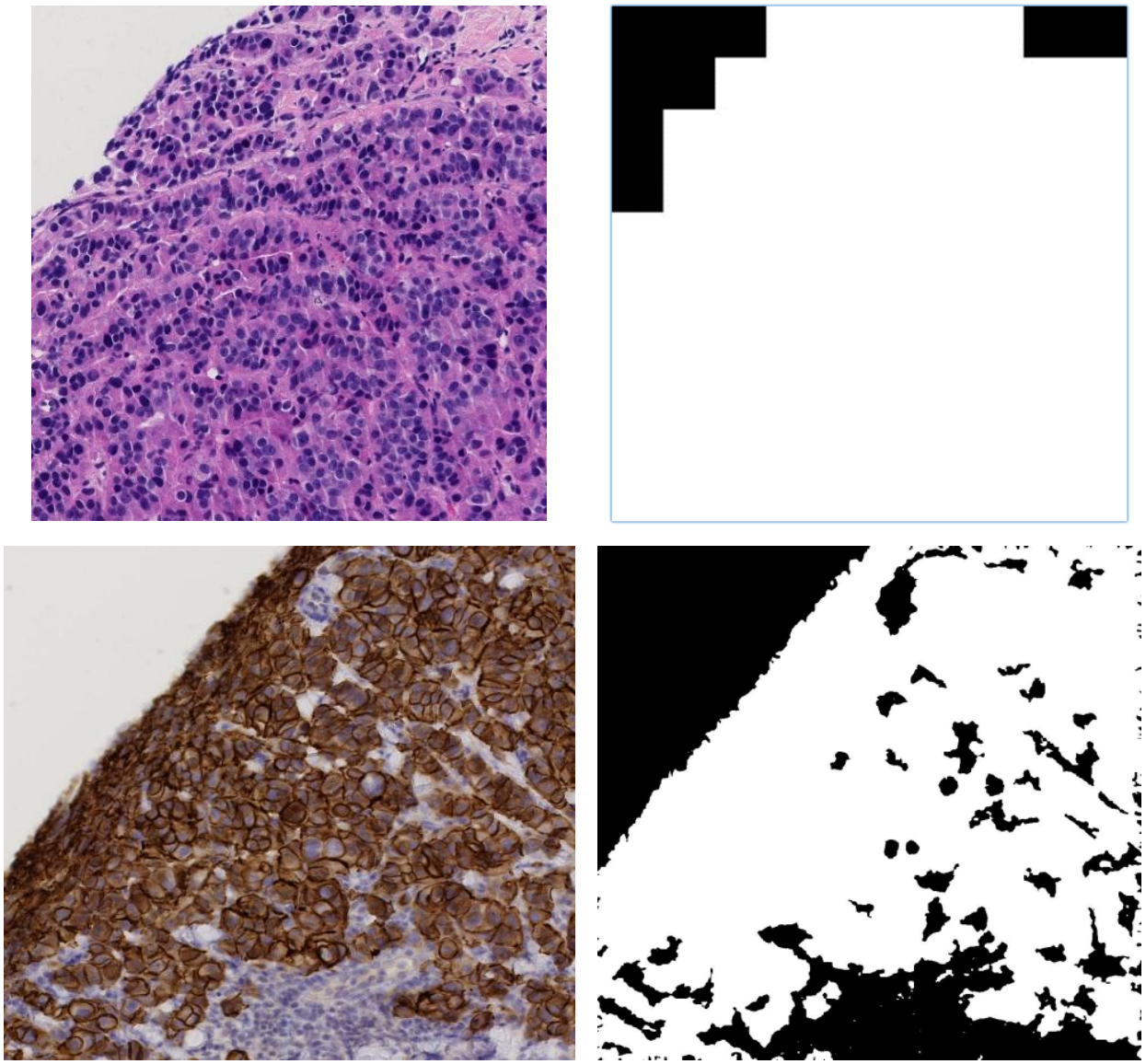


Figure 5.12: Tumor detection in H&E (first row) and brown detection in IHC image patch (second row).

TH-1577_11610215

HER2 over expression. Due to heterogeneity across WSI, experts find it difficult to accurately identify the HER2 score. We proposed a deep learning based framework to predict the HER2 score of a WSI. Deep CNN was used to predict the patch HER2 score and then our criteria identified the HER2 score of WSI based on the combination of HER2 labeled patches in that WSI. The criteria was motivated from guidelines described in [73]. The testing performance shows that our algorithm is better at predicting HER2 score than others that participated in the competition [2]. In some cases, we even outperformed the experts. This is due to the fact that experts have inter- and intra-variations in reporting HER2 score and heterogeneity of histological structures. We give credits to CNN and our criteria in accurate prediction of the HER2 score.

Percentage of complete membrane staining (PCMS) was estimated by a simple image processing approach. The percentage ratio of normalized brown pixels to the normalized tumor pixel gave the PCMS score. These factors were estimated from the H&E and IHC stained WSI. The normalization of factors was done to avoid registering H&E and IHC consecutive WSI. Tumor pixels were detected using the Otsu thresholding and morphological filtering on the stain separated H component. This processing locates large blobs with round shape which are potential tumor regions. Brown detector detects the pixels having brown shade through our proposed distance metric. The evaluation of PCMS can support the experts in their manual PCMS and validate the HER2 score.

In future, we would like to validate the generalization capability of proposed algorithm. We will get different breast tissue WSI from pathology labs from our collaborators and validate proposed HER2 predictor. Also, we want to try learning generalized model across different tissue types.

TH-1577_11610215

6

Conclusion

Digital pathology data is growing at the fastest rate in its history due to advances in microscopic acquisition technology and use of computer aided diagnosis (CAD) algorithms which require sufficient data. As the data is growing, it is becoming challenging for human to learn from such huge stack of the histological data. Hence, clinicians and experts are relying on machine learning and image analysis to learn meaningful patterns from this huge informative heterogeneous data. For example, Beck *et al.* [1] discovered novel stromal features which were out of consideration of experts. As discussed in Chapter 5, we have shown that it is possible to predict the HER2 score of the breast cancer whole slide image very accurately that surpasses lone human experts in some cases. Hence, such automated algorithms can be very useful for training medical students, and to provide second opinion to experts. Such methods do not suffer from inter- and intra-expert variations while making clinical decisions. In this thesis, we proposed solutions for the four different problems in computational pathology which include stain separation, color normalization, nuclei segmentation and HER2 score prediction.

In the stain separation, we modified the NMF cost function to include the biologically pertinent sparsity on the density maps that indicate stains ability to bind sparsely to the tissue. The proposed SNMF stain separation outperformed the state-of-the-art techniques

TH-1577_11610215

that were compared. Even experts preferred SNMF was accepted over NMF. Stain separation was covered in the Chapter 2. We used the SNMF stain representation in proposed color normalization discussed in Chapter 3. The technique relied on the accurate stain separation representation. In the stain separation representation, we estimate the basis of the color of the stains as a separate color appearance matrix and structure information i.e. density map. Since, the source image should be normalized to the target image chosen by an expert, we estimate the SNMF stain separation representation for both source and target. To color normalize, we replaced the color appearance of the source by that of target and linearly normalized the density maps of source with the target. We also validated the impact of the color normalization for the epithelial-stromal segmentation in prostate microarray cores.

Another problem we solved was nuclei segmentation. Nuclei segmentation was a challenging because of the diverse appearances of the nuclei within and across histological images. Overlapping of nuclei is another challenge for nuclei segmentation algorithm. We proposed an image processing approach which enhances the marker controlled watershed model through specific enhancements like Gaussian smoothing and use of markers to minimize the over-segmentation. However, nuclei overlapping was still unsolved. Therefore, we also proposed a deep convolutional neural network to resolve the overlapping using a three-class classification approach. This was covered in the Chapter 4. As color normalization helps to improve the performance, we used proposed color normalization for the color standardization of the training and testing images used.

HER2 overexpression is observed in the 25% of the breast cancer cases. Thus, detection of HER2 score of whole slide image (WSI) is important for diagnosis and prognosis in breast cancers. While participating in a competition [2], we proposed a deep CNN to predict the HER2 score of the patches and novel criteria to score the WSI based on the patches HER2 scores. We were successfully able to detect the HER2 score of the difficult WSI in which lone experts gave wrong HER2 score. The percentage of the membrane staining was estimated through a simple image processing pipeline. Proposed HER2 predictor was ranked first in HER2 scoring contest [2]. In this way, we proposed algorithms for histological images in computational pathology.

TH-1577_11610215

In the future, we are motivated to solve similar problems under computational pathology framework. We would like to use the proposed color normalization as a pre-processing step for various problems in histological images and compare the performance of it with non-normalized version to further popularize the impact of the color normalization among research community. Stain separation can be shown useful for the task of stain fusion where pathologist can visualize different biological markers on the same slide in one view which can simplify and could enhance experts diagnosis and prognosis accuracy. Also, we want to evaluate HER2 prediction algorithm on more IHC images available from our collaborators to validate the generalization HER2 predictor. We envision using the proposed algorithms in this thesis to solve the another important problem i.e. to predict the cancer recurrence in a patient based on its currently available histological images [77].

TH-1577_11610215

List of Publications that Resulted from this Thesis Work

International Journals

- **A. Vahadane**, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A. M. Schlitter, I. Esposito, and N. Navab, Structure-preserving color normalization and sparse stain separation for histological images", in *IEEE Transactions on Medical Imaging*, vol. 35, 2016.

- A. Sethi, L. Sha, **A. Vahadane**, R. J. Deaton, N. Kumar, V. Macias, and P. H. Gann, "Empirical comparison of color normalization methods for epithelial-stromal classification in h and e images", *Journal of pathology informatics*, vol. 7, 2016.

- N. Kumar, R. Verma, S. Sharma, S. Bhargava, **A. Vahadane**, A.Sethi,"A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology", *IEEE Transactions on Medical Imaging*, 2017.

International Conferences

- **A. Vahadane**, T. Peng, S. Albarqouni, M. Baust, K. Steiger, A. M. Schlitter, A. Sethi, I. Esposito, and N. Navab, Structure-preserved color normalization for histological images," in *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015, pp. 1012-1015.

- **A. Vahadane** and A. Sethi, Towards generalized nuclear segmentation in histological images," in *IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2013, pp. 1-4.

TH-1577_11610215

Bibliography

[1] A. H. Beck, A. R. Sangoi, S. Leung, R. J. Marinelli, T. O. Nielsen, M. J. van de Vijver, R. B. West, M. van de Rijn, and D. Koller, "Systematic analysis of breast cancer morphology uncovers stromal features associated with survival," *Sci Transl Med*, vol. 3, no. 108, pp. 108–113, 2011.

[2] W. University. (2016) Her2 scoring contest. [Online]. Available: http://www2.warwick.ac.uk/fac/sci/dcs/research/combi/research/bic/her2contest

[3] M. Veta, P. J. van Diest, S. M. Willems, H. Wang, A. Madabhushi, A. Cruz-Roa, F. Gonzalez, A. B. Larsen, J. S. Vestergaard, A. B. Dahl *et al.*, "Assessment of algorithms for mitosis detection in breast cancer histopathology images," *Medical Image Analysis*, vol. 20, no. 1, pp. 237–248, 2015.

[4] T. J. Fuchs and J. M. Buhmann, "Computational pathology: Challenges and promises for tissue analysis," *Computerized Medical Imaging and Graphics*, vol. 35, no. 7, pp. 515–530, 2011.

[5] 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 *MIAAB workshop*. Citeseer, 2007, pp. 1–8.

[6] A. Rabinovich *et al.*, "Unsupervised color decomposition of histologically stained tissue samples," in *NIPS*, 2003.

[7] M. Macenko *et al.*, "A method for normalizing histology slides for quantitative analysis." in *ISBI*, vol. 9, 2009, pp. 1107–1110.

[8] J. Xu, L. Xiang, G. Wang, S. Ganesan, M. Feldman, N. N. Shih, H. Gilmore, and A. Madabhushi, "Sparse non-negative matrix factorization (snmf) based color unmixing for breast histopathological image analysis," *Computerized Medical Imaging and Graphics*, 2015.

[9] H. Llewellyn, "Observer variation, dysplasia grading, and hpv typing a review," *American Journal of Clinical Pathology, Pathology Patterns Reviews.*, vol. 114, no. Suppl 1, pp. S21–S35, 2000.

[10] G. Thomas, M. Dixon, N. Smeeton, and N. Williams, "Observer variation in the histological grading of rectal carcinoma." *Journal of clinical pathology*, vol. 36, no. 4, pp. 385–391, 1983.

[11] J. Gu and C. R. Taylor, "Practicing pathology in the era of big data and personalized medicine," *Applied immunohistochemistry & molecular morphology: AIMM/official publication of the Society for Applied Immunohistochemistry*, vol. 22, no. 1, p. 1, 2014.

[12] L. Biosystem. (2016) Imagescope. [Online]. Available: http://www.leicabiosystems.com/digital-pathology/digital-pathology-management/imagescope

[13] J. H. Medicine. (2015) Breast cancer and breast pathology. [Online]. Available: http://pathology.jhu.edu/breast/grade.php

TH-1577_11610215

[14] C. C. Society. (2016) Prostate cancer. [Online]. Available: http://www.cancer.ca/en/cancer-information/cancer-type/prostate/grading/gleason-classification/?region=on

[15] A. Sethi, L. Sha, A. R. Vahadane, R. J. Deaton, N. Kumar, V. Macias, and P. H. Gann, "Empirical comparison of color normalization methods for epithelial-stromal classification in h and e images," *Journal of pathology informatics*, vol. 7, 2016.

[16] A. M. Khan, N. Rajpoot, D. Treanor, and D. Magee, "A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1729–1738, 2014.

[17] A. Vahadane and A. Sethi, "Towards generalized nuclear segmentation in histological images," in *Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on*. IEEE, 2013, pp. 1–4.

[18] L. Peter, D. Mateus, P. Chatelain, N. Schworm, S. Stangl, G. Multhoff, and N. Navab, "Leveraging random forests for interactive exploration of large histological images," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, 2014, pp. 1–8.

[19] M. Mete, L. Hennings, H. J. Spencer, and U. Topaloglu, "Automatic identification of angiogenesis in double stained images of liver tissue," *BMC bioinformatics*, vol. 10, no. 11, p. 1, 2009.

[20] R. B. Hazan, G. R. Phillips, R. F. Qiao, L. Norton, and S. A. Aaronson, "Exogenous expression of n-cadherin in breast cancer cells induces cell migration, invasion, and metastasis," *The Journal of cell biology*, vol. 148, no. 4, pp. 779–790, 2000.

[21] M. Veta, A. Huisman, M. A. Viergever, P. van Diest, and J. P. Pluim, "Marker-controlled watershed segmentation of nuclei in h&e stained breast cancer biopsy images," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium*. IEEE, 2011, pp. 618–621.

[22] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, "Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 97–114, 2014.

[23] M. Gavrilovic *et al.*, "Blind color decomposition of histological images." *IEEE TMI*, vol. 32, no. 6, pp. 983–994, 2013.

[24] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution." *AQCH*, vol. 23, no. 4, pp. 291–299, 2001.

[25] W. W. Parson, *Modern optical spectroscopy*. Springer, 2007.

[26] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[27] J. Newberg and R. F. Murphy, "A framework for the automated analysis of subcellular patterns in human protein atlas images," *Journal of proteome research*, vol. 7, no. 6, pp. 2300–2308, 2008.

[28] N. Trahearn, D. Snead, I. Cree, and N. Rajpoot, "Multi-class stain separation using independent component analysis," in *SPIE Medical Imaging*, 2015, pp. 94 200J–94 200J.

[29] A. R. Webb, *Statistical pattern recognition*. John Wiley & Sons, 2003.

[30] M. Niethammer *et al.*, "Appearance normalization of histology slides," in *Machine Learning in Medical Imaging*. Springer, 2010, pp. 58–66.

TH-1577_11610215

[31] A. Vahadane, T. Peng, S. Albarqouni, M. Baust, K. Steiger, A. M. Schlitter, A. Sethi, I. Esposito, and N. Navab, "Structure-preserved color normalization for histological images," in *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015, pp. 1012–1015.

[32] A. Vahadane, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A. M. Schlitter, I. Esposito, and N. Navab, "Structure-preserving color normalization and sparse stain separation for histological images," *early access IEEE Transactions on Medical Imaging*, 2016.

[33] T. Chen and C. Chefdhotel, "Deep learning based automatic immune cell detection for immunohistochemistry images," in *MLMI (MICCAI Workshop)*. Springer, 2014, pp. 17–24.

[34] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[35] T. T. Wu and K. Lange, "Coordinate descent algorithms for lasso penalized regression," *The Annals of Applied Statistics*, pp. 224–244, 2008.

[36] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani *et al.*, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[37] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *The Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010.

[38] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.

[39] M. T. McCann, J. Majumdar, C. Peng, C. Castro, and J. Kovacevic, "Algorithm and benchmark dataset for stain separation in histology images," in *Proc. IEEE Int. Conf. Image Processing*, 2014.

[40] E. D. Gelasca, J. Byun, B. Obara, and B. Manjunath, "Evaluation and benchmark for biological image segmentation," in *International Conference on Image Processing*. IEEE, 2008, pp. 1816–1819.

[41] C. Bayer, A. Kielow, D. Schilling, C.-A. Maftei, D. Zips, A. Yaromina, M. Baumann, M. Molls, and G. Multhoff, "Monitoring pai-1 and vegf levels in 6 human squamous cell carcinoma xenografts during fractionated irradiation," *International Journal of Radiation Oncology* Biology* Physics*, vol. 84, no. 3, pp. e409–e417, 2012.

[42] M. Gamble and I. Wilson, "The hematoxylins and eosin," *Theory and Pratice of Histological Techniques. New York: Churchill Livingstone*, vol. 2008, pp. 121–35, 2008.

[43] C. M. van der Loos, "Multiple immunoenzyme staining: methods and visualizations for the observation with spectral imaging," *Journal of Histochemistry & Cytochemistry*, vol. 56, no. 4, pp. 313–328, 2008.

[44] A. Veillard, N. Loménie, and D. Racoceanu, "An exploration scheme for large images: application to breast cancer grading," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 3472–3475.

[45] A. K. Jain, *Fundamentals of Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.

[46] D. Magee *et al.*, "Colour normalisation in digital histopathology images," in *OPTIMHisE (MICCAI Workshop)*, 2009, pp. 20–24.

TH-1577_11610215

[47] E. Reinhard *et al.*, "Color transfer between images," *IEEE Computer graphics and applications*, vol. 21, no. 5, pp. 34–41, 2001.

[48] P. A. Bautista, N. Hashimoto, and Y. Yagi, "Color standardization in whole slide imaging using a color calibration slide," *Journal of Pathology Informatics*, vol. 5, 2014.

[49] P. A. Bautista and Y. Yagi, "Staining correction in digital pathology by utilizing a dye amount table," *Journal of Digital Imaging*, pp. 1–12, 2015.

[50] A. Kolaman and O. Yadid-Pecht, "Quaternion structural similarity: a new quality index for color images," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1526–1536, 2012.

[51] L. Shamir, N. Orlov, D. M. Eckley, T. Macura, J. Johnston, and I. G. Goldberg, "Wndchrm–an open source utility for biological image analysis," *Source code for biology and medicine*, vol. 3, no. 1, p. 1, 2008.

[52] M. Kulikova, A. Veillard, L. Roux, and D. Racoceanu, "Nuclei extraction from histopathological images using a marked point process approach," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 831 428–831 428.

[53] E. Bengtsson, C. Wahlby, and J. Lindblad, "Robust cell image segmentation methods," *Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenit*, vol. 14, no. 2, pp. 157–167, 2004.

[54] E. Meijering, "Cell segmentation: 50 years down the road [life sciences]," *Signal Processing Magazine, IEEE*, vol. 29, no. 5, pp. 140–145, 2012.

[55] T. Mikolov, S. Kombrink, L. Burget, J. Černockỳ, and S. Khudanpur, "Extensions of recurrent neural network language model," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 5528–5531.

[56] P.-W. Huang and Y.-H. Lai, "Effective segmentation and classification for hcc biopsy images," *Pattern Recognition*, vol. 43, no. 4, pp. 1550–1563, 2010.

[57] M. E. Plissiti and C. Nikou, "Overlapping cell nuclei segmentation using a spatially adaptive active physical model," *Image Processing, IEEE Transactions on*, vol. 21, no. 11, pp. 4568–4580, 2012.

[58] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *Image Processing, IEEE Transactions on*, vol. 2, no. 2, pp. 176–201, 1993.

[59] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[60] F. Xing, Y. Xie, and L. Yang, "An automatic learning-based framework for robust nucleus segmentation," *IEEE transactions on medical imaging*, vol. 35, no. 2, pp. 550–566, 2016.

[61] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 807–814.

[62] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

TH-1577_11610215

[63] R. Collobert, S. Bengio, and J. Mariéthoz, "Torch: a modular machine learning software library," Idiap, Tech. Rep., 2002.

[64] M. R. Lamprecht, D. M. Sabatini, A. E. Carpenter *et al.*, "Cellprofiler: free, versatile software for automated biological image analysis," *Biotechniques*, vol. 42, no. 1, p. 71, 2007.

[65] C. R. UK. (2016) Cancer mortality for common cancers. [Online]. Available: http://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/common-cancers-compared#heading-Two

[66] V. L. Irvin and R. M. Kaplan, "Screening mammography & breast cancer mortality: meta-analysis of quasi-experimental studies," *PloS one*, vol. 9, no. 6, p. e98105, 2014.

[67] A. C. Society. (2016) Breast cancer prevention and early detection. [Online]. Available: http://www.cancer.org/cancer/breastcancer/moreinformation/breastcancerearlydetection/breast-cancer-early-detection-acs-recs

[68] breastcancer.org team. (2016) Her2 status. [Online]. Available: http://www.breastcancer.org/symptoms/diagnosis/her2

[69] J. Eaden, K. Abrams, H. McKay, H. Denley, and J. Mayberry, "Inter-observer variation between general and specialist gastrointestinal pathologists when grading dysplasia in ulcerative colitis," *The Journal of pathology*, vol. 194, no. 2, pp. 152–157, 2001.

[70] W. McCluggage, M. Walsh, C. Thornton, P. Hamilton, L. Caughley, H. Bharucha *et al.*, "Inter- and intra-observer variation in the histopathological reporting of cervical squamous in traepithelial lesion susing a modified bethesda grading system," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 105, no. 2, pp. 206–210, 1998.

[71] O. F. Rasmussen and R. Jørgensen, "Controls," *Immunohistochemical staining methods, 5th edn. Dako North America, Carpinteria*, vol. 127, 2009.

[72] F. Chollet. (2016) Keras: Deep learning library for theano and tensorflow. [Online]. Available: http://keras.io/

[73] E. A. Rakha, S. E. Pinder, J. M. Bartlett, M. Ibrahim, J. Starczynski, P. J. Carder, E. Provenzano, A. Hanby, S. Hales, A. H. Lee *et al.*, "Updated uk recommendations for her2 assessment in breast cancer," *Journal of clinical pathology*, pp. jclinpath–2014, 2014.

[74] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[75] S. Suzuki *et al.*, "Topological structural analysis of digitized binary images by border following," *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32–46, 1985.

[76] P. Maragos and R. Schafer, "Morphological skeleton representation and coding of binary images," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1228–1244, 1986.

[77] A. Sethi, L. Sha, R. J. Deaton, V. Macias, A. H. Beck, and P. H. Gann, "Abstract lb-285: Computational pathology for predicting prostate cancer recurrence," *Cancer Research*, vol. 75, no. 15 Supplement, pp. LB–285, 2015.

TH-1577_11610215