Shu, Jie (2014) Immunohistochemistry image analysis : protein, nuclei and gland. PhD thesis, University of Nottingham.

IMMUNOHISTOCHEMISTRY IMAGE ANALYSIS: PROTEIN, NUCLEI AND GLAND

JIE SHU, BEng, MSc

**Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy**

17th October 2014

Abstract

This thesis focus on the analysis of digitized microscopic image, especially on IHC stained colour images. The corresponding contributions focused on the automatic detection of stain colour and glands, the segmentation and quantification of cell nuclei, the analysis of liver cirrhosis and the development of a semi-automatic toolbox.

Colour is the most important feature in the analysis of immunostained images. We developed a statistical colour detection model for stain colour detection based on the histograms of collected colour pixels. This is acting on the approach "what you see is what you get" which outperforms the other methods on the detection of several kinds of stain colour.

Verifying the presence of nuclei and quantifying positive nuclei is the foundation of cancer grading. We developed a novel seeded nuclei segmentation method which greatly improves the segmentation accuracy and reduces both over-segmentation and under-segmentation. This method has been demonstrated to be robust and accurate in both segmentation and quantification against manual labelling and counting in the evaluation process.

The analysis of gland architecture, which reflects the cancer stage, has evolved into an important aspect of cancer detection. A novel morphology-based approach has been developed to segment gland structures in H-DAB stained images. This method locates the gland by focusing on its morphology and intensity characteristics, which covers variations in stain colours in different IHC images. The evaluation results have demonstrated the improvements of accuracy and efficiency.

For the successive development of three methods, we put them in a semi-automatic toolbox for the aid of IHC image analysis. It can detect different kinds of stain colour and the basic components in an IHC image. The user created models and parameters can be saved and transferred to different users for the reproduction of detection results in different laboratories.

To demonstrate the flexibility of our developed stained colour detection technique, the tool has been extended to the analysis of liver cirrhosis. It is a novel method based on our statistical colour detection model which greatly improves the analysis accuracy and reduces the time cost.

.

List of Publications

1. **Jie Shu**, Guoping Qiu, Mohammad Ilyas and Philip Kaye. Biomarker Detection in Whole Slide Imaging Based on Statistical Color Models. *The MIDAS Journal-Computational Imaging Biomarkers for Tumors (CIBT), http://www.midasjournal.org/browse/publication/758,* 2010 [Chapter 3]

2. **Jie Shu** Guoping Qiu, Mohammad Ilyas, G. Dolman. A Semi-Automatic Image Analysis Tool for Biomarker Detection in Immunohistochemistry Analysis. *International Conference on Image and Graphics (ICIG).* 2013 [Chapter 3]

3. **Jie Shu**, Guoping Qiu, Philip Kaye and Mohammad Ilyas. Segmenting Overlapping Cell Nuclei in Digital Histopaholgy Images. *35th IEEE Engineering in Medicine and Biology Society (EMBC).* 2013 [Chapter 4]

4. **Jie Shu**, Hao Fu, Guoping Qiu and Mohammad Ilyas. An Efficient Gland Detection Method Based on Texture and Morphological Transformation. *Medical Image Understanding and Analysis (MIUA).* 2013 [Chapter 5]

5. Hao Fu, Guoping Qiu, Mohammad Ilyas and **Jie Shu**. GlandVision: A Novel Polar Space Random Field Model for Glandular Biological Structure Detection. *In British Machine Vision Conference (BMVC),* 2012 [Chapter 5]

6. Hao Fu, Guoping Qiu, **Jie Shu** and Mohammad Ilyas. A Novel Polar Space Random Field Model for the Detection of Glandular Structures. *IEEE TRANSACTIONS ON MEDICAL IMAGING (TMI), vol 33(3): 764 − 776,* 2014 [Chapter 5]

7. Dolman Grace, **Jie Shu**, Hawkes C, Zaitoun AM, Fallowfield JA, Irving WL, and Guha IN. Quantification of Elastin As a Predictor of Clinical Outcomes in Cirrhosis Caused by Chronic Hepatitis C (CHC) Infection. *British Association for The Study of The Liver (BASL).* 2013 [Chapter 6]

8. Hoad Caroline, Palaniyappan Naaventhan, Kaye Philip, Chernova Yulia, Bawden Stephen, Stephenson Mary, Dolman Grace, **Jie Shu**, Guoping Qiu, James Martin,

Costigan Carolyn, Austin Andrew, Marciani Luca, Gowland Penny, Guha Indra Neil, Francis Sue, Aithal Guruprasad. Multi-Modal Quantitative Magnetic Resonance for the Estimation of Fibrosis, Inflammation, Fat and Iron Accumulation in the Whole Liver. *ID: HEP-14-0273 Submitted to the Journal of Hepatology on 03-Feb-2014* 2014

Acknowledgements

Someone asked me why I endeavor to work on my PhD. That was a dream when I was a child, I answered. I have been living my childhood dream in the past 31 years. Approaching the dream, I realized, it was not my own but also the dream of those people who supported me to complete this doctoral thesis.

I would like to express the first and deepest appreciation to my supervisor: Prof. Guoping Qiu, for his valuable guidance, patients and consistent encouragement I received throughout the research work. Without his supervision and constant help this thesis would not have been possible. He who teaches me for one day is my father for life. I would also like to express my grateful to my second supervisor: Prof. Mohammad Ilyas who provided me the microscopic images and all the biological and medical knowledge.

Special acknowledge to doctor Philip Kaye and Grace Dolman who provide me the microscopic images and the biological and medical knowledge as well. To my two examiners: Prof. Jon Garibaldi and Prof. Nasir M. Rajpoot, whose careful proofreading and insightful suggestions have led to an improvement of the quality of this thesis. To the Prof. Peter Blanchfield and Prof. Rong Qu who reference me for the PhD application to the University of Nottingham.

Special thanks to my colleagues: Dr. Hao Fu, Bozhi Liu, Qian Zhang, Yujie Mei, Wenwen Bao, Orod Razeghi, Min Zhang, Mercedes Torres, Salvador Garcia Bernal, Dr. Poay Hoon Lim, Dr. Ligang Zheng, and Nadine Holmes.

Thanks to the organiger and all the members in IMA and VIPLAB who gave me the opportunities to present my research and share experiences with each other. To all the friends who help me a lot in the daily life. They are Yi Ding, Yongqing Wen, Xing Jiang, Guanshi Wang, Jinglang Huang, Cheng Zhong, Qi Shen, Siqian Wu, Meiyi Zhu, Rui Zhou, Yusu Wang, Zhilan Xu, Shixin Wang, Zijian Cai, Xiaokai You, Jianqiao Liu, Wentao Cui, Rong Cang, Jing Ma, Yiwei Lv, Chang Liu, Yuquan Zhang, Lulu Zhang, Guanghui Chen, Qian Zhou, Yifei Cheng, Yunxiao Li, Weibin Lin, Yizhen Shi, Yuanchao Peng, Ziqi Sun, Zhe Liu, Hui Guo, Zhibao Mian, He Wang, Yunhao Lin.

This thesis is dedicated to my parents, who support me the study of PhD, without them this study won't be possible. To my grandfather, grandmother, uncle and aunt who have passed away. Wish you rest in peace in the heaven.

Contents

List of Figures

List of Tables

To my family...

Chapter 1

Introduction

Pathologists routinely explore tissue slides through a microscope and produce diagnostic and prognostic results based on the observations. The increasing number of tissue slides, and the importance of this type of examination in both clinical medicine and biological research, make this visual work tedious and ineffective.

With the introduction of Digital Pathology [5], tissue slides can be converted into digital files, and pathologists can use a variety of image analysis techniques to explore the digital slides in order to aid the diagnosis of disease.

Computerised pathology slides may help doctors make faster and more accurate diagnoses, and have the potential to revolutionise current pathology practices of diagnosis and prognosis. However, compared to other medical imaging modalities, such as X-ray, magnetic resonance imaging (MRI) and compute topology (CT), digital pathology [6] is a relatively new area of medical imaging. Although there have been efforts to develop dedicated image analysis techniques for the analysis of digital slides, many challenges still exist.

This thesis investigates image analysis techniques for the processing of digital slides, and has developed new methods and solutions that may aid pathologists in the examination of digitised tissue slides.

1.1 Background & Motivations

In this section, there is a brief description on the background of digitized IHC images. This includes a description of the IHC stains that are concerned with in this thesis and a description of the components detected in IHC images.

1.1.1 Immunohistochemistry

Immunohistochemistry (IHC) is a process of detecting targeted antigens (proteins) in tissue sections by the use of labelled antibodies through antigen-antibody interactions [7]. It was discovered by Coons and Jones [8], who used this technique for the detection of bacteria, and has since become a standard tool in diagnostic pathology [9]. The target antigens or proteins, called biomarkers, can be coloured by different chromogens and visualised through microscopes by pathologists. The effect of staining is to highlight the region of interest and provide contrast against the background.

Tissue slides are normally prepared with one or more stains. For example, Diaminobenzidine (DAB) is one of the most commonly used stains in IHC staining and gives a dark-brown colouration to biomarkers (see Fig. 1.1). A counterstain, the heamatoxylin (H), which stains the background tissue blue, normally follows. In this thesis, the stained protein contained in the cell nuclei subjected to DAB staining is Protein 53 (P53). P53 is a tumour suppressor protein that is expressed predominantly in the nucleus of the cell. Inspecting the distribution of P53, which has been demonstrated to be over-expressed in malignant tumours,
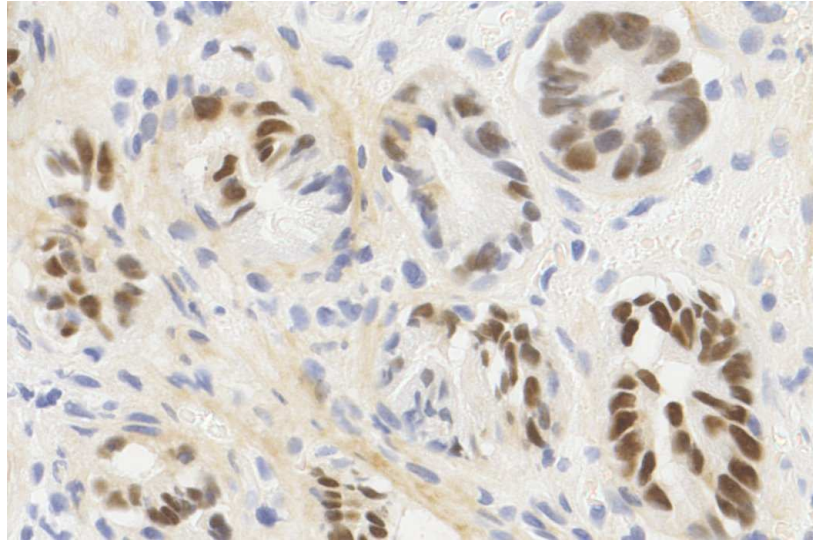
FIGURE 1.1: An example of H-DAB stained digital slide images.

aids the diagnosis of colorectal cancer [10], for example, the grading of colon cancer. A typical H-DAB stained digital slide image is shown in Fig. 1.1.

Sirius Red (SR) is another commonly used stain in the diagnosis of liver cirrhosis (see Fig. 1.2). It colours the biomarker (collagen) pink and is followed by a counter-stain that stains the background pale yellow. This stain is used to detect the distribution of collagen fibrils in liver cirrhosis biopsies. The quantification of collagen fibril expression in the tissue section has been demonstrated to correlate with the stages of liver cirrhosis [11].

Another type of stained tissue sample used in this thesis is brown-coloured elastin in liver biopsies (see Fig. 1.3). Elastin is the main component of elastic fibres. These elastic fibres are produced by the fibroblasts and smooth muscle cells in arteries and have been linked to the maturation of liver fibrosis. These fibres can also be used as a signal of liver cirrhosis for they increase in number in liver fibrosis and cirrhosis [12].

Ever since the cell was recognised as the fundamental building unit of the human body, pathologists have sought to exploit the underlying principles of diseases, especially for cancer. The use of computers to analyse cells can be dated back
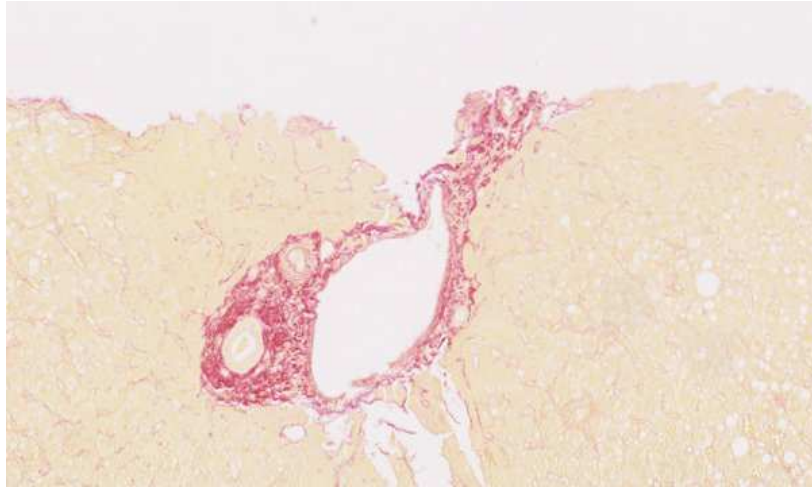
FIGURE 1.2: An example of SR stained digital slide image of liver cirrhosis.
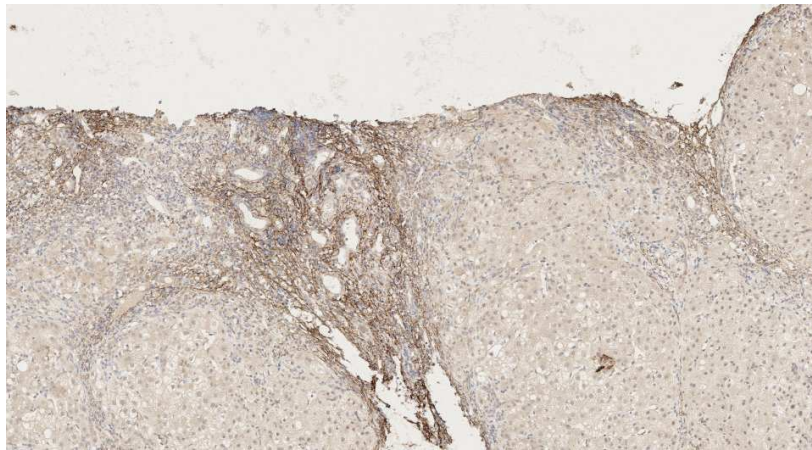
FIGURE 1.3: An example of brown stained elastin fibres image of liver cirrhosis.

to the mid-1950s. However, automated segmentation and analysis of cell nuclei is generally a difficult problem due to the great variability in cell nuclei and the complexity of the data.

1.1.2 Nuclei segmentation and gland detection

The diagnosis of disease in IHC images requires identifying the nuclei that contain positively stained biomarkers. In H-DAB stained images (see Fig. 1.1), the P53 biomarker will react to the DAB stain and appear brown, and will be termed

positively stained. In contrast, other biomarkers will react to the haematoxylin (H) stain and appear blue, and will be termed negatively stained. In DAB stained IHC images, the cancerous nuclei, which will appear in a brown colour, can be observed and quantified to determine the stage of cancer [13].
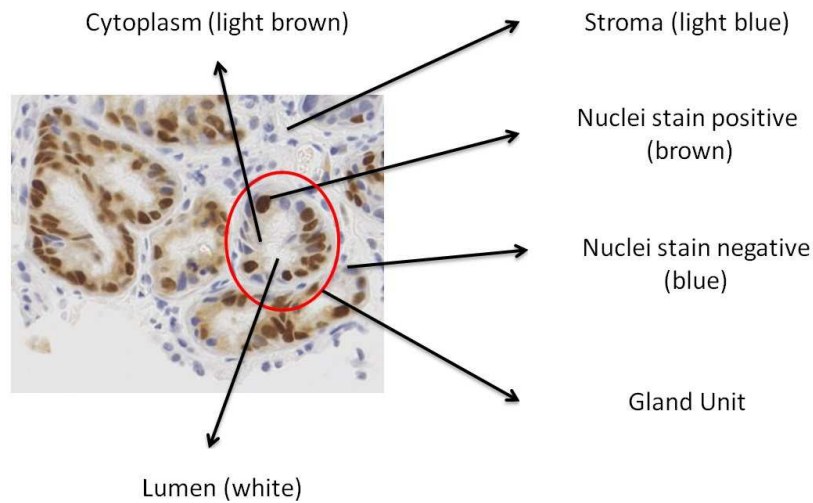
FIGURE 1.4: The nuclei and gland in H-DAB stained sample.

As illustrated in Fig. 1.4, a gland is a more complicated structure than a collection of nuclei in an IHC image. It is formed by a group of nuclei that are in a closed or partially closed chain. There are four components in a glandular structure, including the chain of boundary nuclei, the cytoplasm, the lumen and the stroma. The lumen is the centre of the gland and the chain of nuclei is the boundary. The nuclei that form the boundary are in a round shape and spaced more closely than the other nuclei. The cytoplasm is contained between the boundary nuclei and the lumen, while the stroma is outside the boundary nuclei.

The presence of glands, their extent, size, shape and other morphological traits, are all regarded as important indicators for identifying the severity of disease [14]. From benign to malignant, the glands in the biopsy sample may change from well-differentiated with similar shapes to fused and mixed with irregular shapes. Thus, the diagnosis of the specimen is important mainly to explore changes

in gland architecture and the distribution of cancerous nuclei within the gland structure.

Therefore, to make a diagnosis with digital slide images, it is necessary to be able to identify positively stained and negatively stained biomarkers, and to recognise nuclei and glands. To automatically accomplish these tasks using computer algorithms is difficult and presents many challenges.

1.2 Challenges

In this section, the challenges of achieving the goal of the research pipeline, such as stain colour detection, nuclei segmentation and gland detection, are stated and discussed.

1.2.1 The challenges for stain colour detection
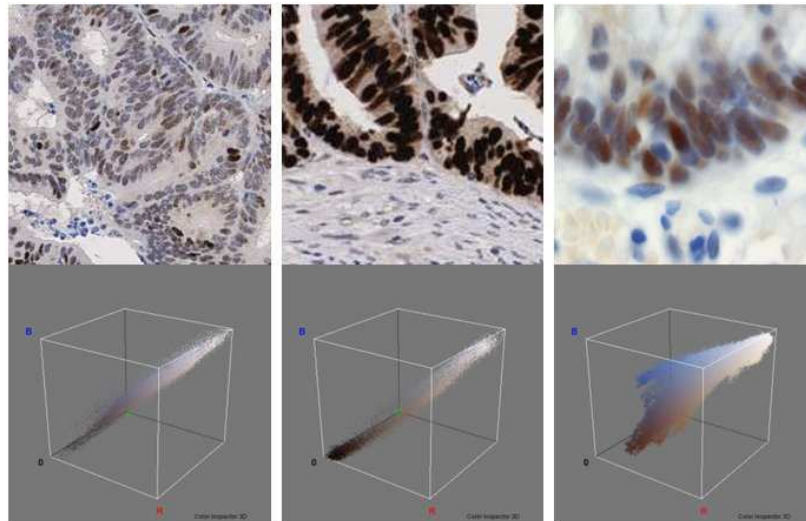
FIGURE 1.5: The colour distribution in different tissue samples stained with the same immunostaining (P53) varies significantly. Top row: image samples. Bottom row: plots of the colour distribution in the RGB colour space.

Fig. 1.5 shows three biopsy samples stained with P53 immunostaining. Visually, it can be seen that there are a range of different shades of brown pixels even within a slide, and that they vary hugely across different samples. Thus, the challenges of detecting colour pixels in stained images can be summarised as follows.

Firstly, there may be variations in staining across different samples or in a single sample caused by problems in tissue slide preparation. These may include variations in stain concentration, the duration of staining, and the thickness of the cut tissue sections.

Secondly, the stain colours may have overlapping spectra, and these coloured pixels are hard to classify as a specific colour. For example, black and dark brown are hard to distinguish even by eye.

Thirdly, the colours in a digitised slide image may not be linear due to variations in the stain's absorption characteristics, i.e., the brown reactive colour after staining with DAB is scatter distributed and does not follow a linear absorption algorithm [15].

Another problem associated with digitised IHC image diagnosis is the huge image size. A single slide can be as large as $50,000 \times 50,000$ pixels. Sieving though such an enormous number of pixels manually is a huge task. What is desired is an automatic solution that will guide pathologists to inspect region of potential interest rather than having to go through every pixel of the image.

1.2.2 The challenges for nuclei segmentation and quantification

Nuclei segmentation is the first step in nuclei quantification. Nuclei may be either isolated or clumped, as presented in microscopic images. Segmenting or

separating them is difficult, especially when dealing with overlapping or densely clustered nuclei against a complex background. For example in Fig. 1.5, the nuclei have different shapes, sizes and illumination. Such variations among nuclei are not only present across samples, but also within a single sample. These variations may easily lead to over- or under-segmentation problems. Over-segmentation arises from the false segmentation of a singular nucleus into multiple fragments. Under-segmentation means the reverse, segmenting multiple nuclei into a single nucleus. To resolve these problems is an extremely challenging task, for the following reasons.

Firstly, the tissue sections and colour stains are usually unevenly prepared. Different coloured nuclei in an IHC image, both the positively and negatively stained, have large variations in illumination. Pixels from the background may also be stained with similar colours as nuclear pixels, resulting in low contrast and weak borders. The non-uniform inter-illumination also makes it difficult to extract the nuclei from the grey-scaled image.

Secondly, a large intra-intensity variation in nuclei is always present. Gaps or high gradient values may occur inside nuclei and be difficult to classify, as compared to real edges. The strong variations may form apparent holes in nuclei when converting to a binary image, or make the nuclear pixels sparser and unevenly distributed. The noise inside a nucleus may easily lead to an over segmentation problem.

Thirdly, images of histological sections represent a 2-dimensional view of a 3-dimensional entity. Thus, individual nuclei, although in different planes of the section, may appear to be aligned and overlapping when viewed from one perspective. The overlapping areas of aligned nuclei and the number of aligned nuclei also vary across samples, which increases the difficulties of separation and may cause an under-segmentation problem.

1.2.3 The challenges for gland detection

A gland is a complicated structure in a tissue slide. The segmentation of glands is a challenging problem, for the structure is a combination of several components rather than a single object in a tissue slide. The increasing complexity of the background may also tremendously increase the segmentation difficulty. Although a gland seems like a closed chain of nuclei, the chain may in fact be either closed or partially closed. The thickness of the boundary, as well as the shape and size of different glands may have large variations. In Hematoxylin-Eosin (H&E) stained samples, the components of a gland are stained with different colours that transform gland detection from a binary to a colour image, thus reducing the difficulty of segmentation. Unfortunately, this colourisation adds another challenge, the variation in stain colour, to the segmentation process. In H-DAB stained IHC images, more problems arise in gland segmentation, and are summarised in the following three aspects.

(a)   (b)   (c)
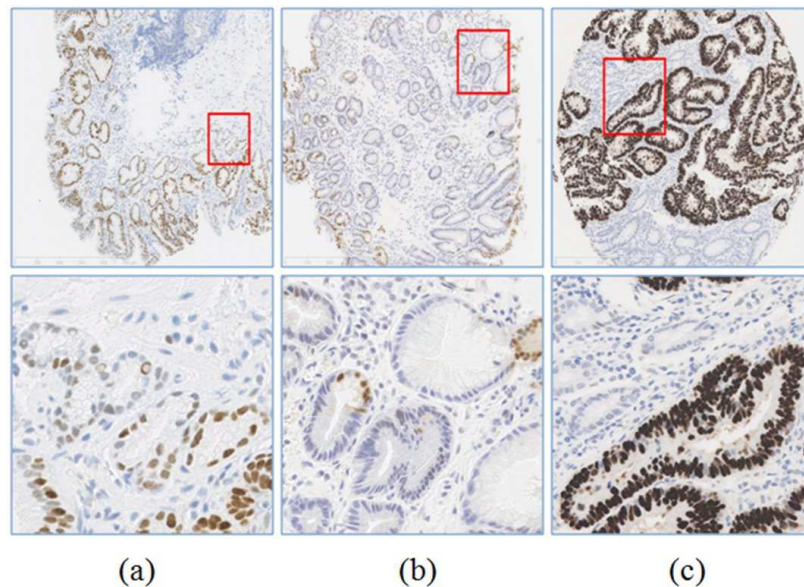
FIGURE 1.6: Examples from our data set. (a) a sample of glands that have vacant artefact regions on the boundary; (b) a sample of glands that have coloured luminal regions; (c) a sample of glands with different coloured nuclei.

Firstly, the vacant artefact regions may appear in the boundaries of glands and the nuclei may be loosely distributed within the gland boundary. The gland nuclei boundary may not be closed due to these vacant artefact regions and large gaps between loose nuclei. Stromal nuclei may be densely distributed around the gland and be hard to classify.

Secondly, the colour information in H-DAB stained IHC images may be useless for gland detection. It is hard to detect the components of a gland depending solely on colour (see Fig. 1.4 and Fig. 1.6).

Thirdly, the image may include large variations in intensity between blue coloured nuclei and brown coloured nuclei. The intensity of blue stained nuclei is much greater than that of brown stained nuclei, which may result in overlooking the lighter ones. These challenges are illustrated in Fig. 1.6.

1.3   Aims & Objectives

This project aims to develop digital slide processing and analysis techniques to tackle the challenges of digital slide image analysis outlined in Section 1.2. The aim of this thesis is to develop new and better techniques for detecting stain colour, and segmenting nuclei and glands in H-DAB stained samples, and then expand to other IHC image analysis. Specifically, it aims to achieve the following objectives.

- **Stain colour detection**

The aim of developing the stain colour detection method is to extract the positive stain from the background and provide quantitative data on stained positive colour for pathologists and researchers. Previous methods have regarded the stain colour as a single colour, or assumed that light was absorbed linearly by

the stains. However, the stains are prepared in a multi-step process on tissue slides, which makes the positive colour a mixed colour, and the light crossing the stain may be scatter absorbed. The aim is to develop flexible techniques that can cope with these variations.

- **Nuclei segmentation and quantification**

The aim of developing automated nuclei segmentation is to isolate the nuclei in a tissue slide and provide quantitative data for pathologists and researchers. The segmented nuclei, either isolated or separated from clusters, can be quantitatively analysed with regard to counting positive or negative nuclei, measuring the distribution of nuclei in an area and measuring characteristics of each nucleus such as size, shape and intensity.

Previous work in this research area failed to separate clustered or overlapping nuclei in areas with severe clustering, which can lead to inaccurate quantification results. Over-segmentation and under-segmentation always occur in the segmentation process, due to the huge variations in background situations and nuclei characteristics across samples. The aim is to develop methods that can segment the nuclei accurately and thus avoiding over- and under-segmentation.

- **Gland detection**

The aim of developing automated gland segmentation is to detect the location and boundaries of glands in a tissue slide and provide quantitative data for pathologists and researchers. As compared to pixels and nuclei, the gland is a higher-level structure constructed from both of them, and contains more information. Despite unique characteristics of its own, such as its size, shape and intensity distributions, a gland can be quantitatively analysed further in several respects. These include the density of the nuclei boundary, the thickness of the

boundary, the distribution of tumour and normal nuclei and the separation of tumour and stromal nuclei.

In previous methods, the segmenting of glands has usually depended on the detection of differences in colour staining in each component of a gland's structure (see Fig. 1.4). However, these methods are hard to transfer to DAB-stained samples, which do not have contrasting colours in these components. The aim is to develop gland segmentation techniques based on the intensity changes and morphology characteristics of gland structure, rather than the colour information.

1.4 Contributions

There have made several contributions to the field of digital slide image analysis, which are summarised as follows.

- **A novel stain colour detection technique**

Colour is the most important feature in the analysis of immunostained images. A statistical colour detection model has been developed for stain colour detection based on the histograms of collected colour pixels. This, in turn, is based on the approach "what you see is what you get", which outperforms the other methods on the detection of several kinds of stain colour. To demonstrate the flexibility of the developed stain colour detection technique, the method has been extended to the analysis of liver cirrhosis. Results show that this statistical colour detection model greatly improves the analysis accuracy and reduces the time cost.

- **A novel nuclei segmentation technique**

Verifying the presence of nuclei and quantifying positive nuclei is the foundation of cancer grading. A novel seeded nuclei segmentation method has been proposed which greatly improves the segmentation accuracy and reduces both over-segmentation and under-segmentation. This method has been demonstrated to be robust and accurate in both segmentation and quantification against manual labelling and counting in the evaluation process.

- **A novel morphology-based technique for gland segmentation**

The analysis of gland architecture, which reflects the cancer stage, has evolved into an important aspect of cancer detection. A novel morphology-based approach to segmenting gland structures in H-DAB stained images has been proposed. This method locates the gland by focusing on its morphology and intensity characteristics, which covers variations in stain colours in different IHC images. Results show that this new method improves accuracy and efficiency.

- **A digital slide analysis toolbox.**

These newly developed colour detection, nuclei segmentation and gland segmentation techniques have been put in a semi-automatic software toolbox that analyses IHC images. The toolbox, implemented in an open source platform, ImageJ, has been made publicly available to researchers.

(http://rsb.info.nih.gov/ij/plugins/ihc-toolbox/index.html)

1.5 Organization of the Thesis

- In Chapter 2, previously proposed detection and segmentation techniques for the analysis of IHC images in stain colour detection, nuclei segmentation and gland detection are reviewed.

- In Chapter 3, the newly developed statistical stain colour detection method for the detection of stain colours in immunostained images is introduced. This statistical colour detection model has been compared with two popular methods used in recent studies on different kinds of IHC stained images, against the manually labelled ground truth.

- In Chapter 4, a new nuclei segmentation and quantification approach is introduced. This approach has been compared with several previous methods against manually labelled ground truth.

- In Chapter 5, a novel morphology-based gland segmentation approach in H-DAB stained images is introduced. It is then been compared with several previous methods against manually labelled ground truth.

- In Chapter 6, an extension of the proposed research pipeline is introduced for the calculation of the Collagen Proportionate Area (CPA) in liver cirrhosis and for the classification of normal stained samples and over-stained liver samples.

- In Chapter 7, the contributions in this thesis and the suggestions for future works are discussed.

- In Appendix A, a digital slide analysis tool-box incorporating the newly developed analysis techniques and implemented in the open source platform ImageJ is presented.

- In Appendix B, the user manual for the developed IHC analysis tool-box is introduced.

Chapter 2

Literature and Methods

The analysis of IHC stained colour images involves the detection and extraction of the basic components contained in digitised slides. These components include pixel level objects, such as the stain colour and its cellular or nuclear location, and the higher-level object that is constructed from these pixel level objects, for example, the structure of a gland. Variations in these components and the complexity of the background in the IHC image make it a challenge to detect or segment them.

This thesis reviewed previously proposed detection and segmentation techniques for the analysis of IHC images in three specific domains: stain colour detection, nuclei segmentation and gland detection. The studied methods with their contributions and weaknesses in these different domains were presented according to the research hypothesis.

This chapter is organised as follows. Stain colour detection methods are reviewed in single colour detection, multi-colour separation and colour correction, and presented in Section 2.1. Nuclei segmentation methods are reviewed in Section 2.2 under the nuclei extraction and splitting scheme. A review of

gland segmentation methods either with or without colour information is presented in Section 2.3. Finally, the techniques that have been used in this thesis are presented in Section 2.4, and the discussion in Section 2.5.

2.1 Techniques For Stain Colour Detection

In this section, literature on techniques that have been used in immunohistochemistry image analysis for stain colour detection and separation are reviewed. These can be divided into four main categories: colour model transformation, colour clustering, different colour separation and colour correction.

2.1.1 Colour model transformation based colour detection

Methods based on colour model transformation and colour spectrum separation have been the most widely used colour detection methods in the analysis of colour stained histological images. This is because the stained tissue slides were usually digitised through conventional light microscopy and RGB digital cameras [6]. Normally, pixels were classified as a specific colour by thresholding the single colour channel of Hue Saturation Intensity (HSI) space [16], the blue channel of RGB colour space [17], CMYK colour space [18] or the normalised RGB colour space [19]. Therefore, the colour stained samples, for example, the DAB stained brown colour samples, can be converted into grey scale images by using the single channel of the transformed colour model. By applying an appropriate threshold to this grey scale image, background pixels can be eliminated. The remaining pixels, which have higher values, may either be analysed quantitatively or recognised as cell pixels for further analysis [20].

The aspect that may affect the detection accuracy of colour transformation methods is the selected threshold. The threshold values for some of these methods were determined by examining the stain colour on a set of stain-controlled tissue images [16, 21]. The rule for manually selecting the threshold is to maximise the positively stained colour pixels and minimise background colour pixels at the same time. Another way to select the threshold is to use the average intensity value on a set of stain-controlled tissue images [17, 22]. However, obtaining a threshold based on manual selection or the average intensity from stain-controlled tissue sample may result in low accuracy due to the stain variations among different samples. In addition, the transformation based method regarding the stain as a specific colour may result in the failure of stain colour detection, because the tissue slides are always multi-stained.

2.1.2 Clustering based colour detection

Rather than using a manually set threshold, it is better to classify the desired colour following training from labelling experts [23]. Colour pixels in a stained image can be classified into different clusters in different colour spaces. They can be classified by classical computer vision classifiers such as SVM, C4.5 and Naive Bayes [24], trained by the neural network classifier [25] or assigned scores based on the distance of the shades in HSI colour space [26].

As there was no standard colour space suitable for all kinds of stain colour detection methodology on different stains, the developed methods were usually evaluated on different colour models to determine the best performance [24]. Since the classification based methods were reliant on the training samples, the detection accuracy may vary with the manual labelling discrepancy and the colour model used. As a result, it is hard to select a fixed combination of an algorithm and a colour space that will perform well on all kinds of samples.

2.1.3 Multi-colour separation based colour detection

Despite the prevalence of colour model-based methods, the colour deconvolution method [27] was developed based on the light absorption spectra of different colour stains. It is a separation method based on the absorption characteristics of the individual stain. For each stain, the light is differently absorbed by its red, green and blue channels through a transmit function, shown as follows:

$$I_{out} = I_{in} \times e^{s \times f} \tag{2.1}$$

The intensity of incident light I_{in} is absorbed by the amount of stain s with its absorption factor f. The output intensity of light I_{out} is detected after the light pass the stain. The relation between the input light and the output light is described by the Lambert-Beer law. To make it linear, the formulation is then transferred and defined as the optical density (OD) for each RGB channel in a following way:

$$OD = -\ln(\frac{I_{out}}{I_{in}}) = s \times f \tag{2.2}$$

Each pure stain with its optical density can be set into a 3 by 1 OD vector matrix of three RGB channels. The OD matrix has been normalised and orthogonally transformed to get the independent information of each stain. The colour deconvolution matrix is then obtained by the inverse of normalizsd OD matrix. As a result, the OD intensity levels for each stain can be obtained by the sum of OD values in its Red, Green and Blue channels, calculated from corresponding stain vectors [27].

This method performs colour stain separation much better than colour space-based methods on H&E stain [28], H-DAB stain [29] and other commonly used colour stains [30]. However, this method is based on a linear light absorption

algorithm, which may reduce the detection accuracy if the light is not linearly absorbed by the stain, as is the case with DAB stain [15].

2.1.4 Colour correction based colour detection

Despite the fact that these general methods mentioned above can achieve good results in the quantitative analysis of IHC images, the existing variations of stain colour may lead to the reduction of colour detection accuracy. Possible solutions of reducing colour variations are colour normalisation or colour correction [31].

In general, colour can be normalised and corrected through colour transfer and colour deconvolution methods. Colour transfer converts the input image into a colour space which has low or no correlations between each channel, for example the $\zeta\alpha\beta$ [32]. Considering about the mean and standard deviations between the input image and the target image on each channel, the input image can be corrected,

$$\zeta_{out} = \frac{\zeta_{input} - \bar{\zeta}_{input}}{\tilde{\zeta}_{input}} \times \tilde{\zeta}_{target} + \bar{\zeta}_{target} \tag{2.3}$$

$$\alpha_{out} = \frac{\alpha_{input} - \bar{\alpha}_{input}}{\tilde{\alpha}_{input}} \times \tilde{\alpha}_{target} + \bar{\alpha}_{target} \tag{2.4}$$

$$\beta_{out} = \frac{\beta_{input} - \bar{\beta}_{input}}{\tilde{\beta}_{input}} \times \tilde{\beta}_{target} + \bar{\beta}_{target} \tag{2.5}$$

where $\bar{\zeta}$, $\bar{\alpha}$, and $\bar{\beta}$ are the channel means and $\tilde{\zeta}$, $\tilde{\alpha}$, and $\tilde{\beta}$ are the channel standard deviations. This method has been used in the colour corrections of H&E stained samples [33]. Although this method has been demonstrated reasonable constancy in the stain area, unsatisfactory results have been shown regarding the background, due to the largely covered Eosin stain area [4]. This stain dominates the statistics that change the white background to pink in result image.

The alternative correction method has been shown in Magee et al. [34] was based on the colour deconvolution. Within colour deconvolution based method, the input image has been decomposed into three greyscale images with different stains. Then the correction can then be made by using singular value decomposition (SVD) [35] and non-linear mapping method [36]. Other methods that have been used in colour correction are: gradient in LUV colour space [37], polynomial surfaces based method [38], automatic stain vector estimation based method [39] and the combination of non-negative matrix factorization (NMF) and a colour transfer based method [40]. However, the current normalisation or correction methods have been generally limited on the correcting of H&E stain samples [31]. The reasons for the correction failure in DAB stained samples may be due to the saturated nuclei, non-uniform colouring DAB areas, and significant staining of the background [34].

2.2 Techniques For Nuclei Segmentation

Looking into the literature published in this domain, it is very difficult to sort it into one or two specific techniques. Nevertheless, the majority of works are based on a few basic approaches such as intensity thresholding, feature detection, morphological filtering, region accumulation and deformable model fitting [41]. Thus, the techniques of nuclei segmentation are classed based on a combination of these approaches to separate nuclei in this section.

2.2.1 Intensity thresholding and morphological filtering

Computer-aided nuclei segmentation has been studied for over 50 years. However, automated segmentation and analysis of cells is generally a difficult problem, due to the great variability in cells and the complexity of the data [41].

Intensity thresholding, the conceptually simplest method and computationally cheapest method, was applied first in nuclei segmentation and is still in use even in recent studies. Auto-threshold methods such as the Otsu [42], Isodata [43], etc. were developed to calculate threshold values on image histograms to help threshold nuclei. Further segmentation processes can be conducted in the thresholded nuclei area [13, 44–47].

Although the nuclei areas have higher or lower intensity values than the background, the intensity has huge variations across the sample. Because the single threshold-based method might not always produce satisfactory results, clustering-based methods were considered that can also achieve the same purpose for nuclei region extraction: for example, the k-means [48], the graph cut [49, 50] and the Bayes classifier [13]. For those images that have a low contrast between nuclei and background, differential features, for example, the edge, have been realised as another way to segment nuclei efficiently and may provide useful information, such as the segmentation based on the gradient change in Luv colour space [37], the gradient change in original image [51], the combination of mean gradient value and gradient change [52], and the linking of pixels that have high gradient changes in a log polar transformed image [53]. However, the gradient values still suffer from a problem with intensity variations, being that the obtained nuclei area may contain lots of noise due to the complexity of the background and non-uniform stains. The morphological transform and mathematical processes were considered to refine the thresholded results or reduce the noise before thresholding, by means such as filling holes [54], open-by-reconstruction and close-by-reconstruction [55], and histogram enhancement [56]. These methods are always regarded as pre-processing in the nuclei segmentation.

2.2.2 Region accumulation with distance transforms

Many of the nuclei are clustered or overlapping in digitised images; therefore, thresholding-based segmentation may not satisfy the requirements of quantified analysis. Separating these clustered or overlapping nuclear areas has become a challenging problem in recent studies. A region growing based watershed transform has been used to solve this problem [57]. However, as this transform frequently lead to over-segmentation, a marker-controlled watershed was employed to reduce the likelihood of this happening [58]. The marker for each nucleus refers to the regional minimum or regional maximum only at the specific location of the nucleus. The classical watershed (topographical relief) has then been transferred to the growing of such specific points (centre).

Many methods discuss the methodology of extracting singular markers. The simplest way to obtain the markers is by extracting the regional minimum from a greyscale image [59] or the Ultimate Erosion Points (UEP) from a Euclidian distance transformed binary image [56]. The regional minimum, based on the original image, may suffer from the intensity variations. To improve the detection accuracy and reduce the intensity variations, Rogojanu et al. [47] obtained the markers by extracting the regional minima in the extended-minima transformed image, and formed the splitting line of touching nuclei based on a template matching method. Orti et al. [49] obtained seed markers from a distance transformed image and then considered a graph-cut based method with which to from the splitting line, by calculating the distance values between voxels and forming the minima cut. Combining the intensity with the distance map, Baek et al. [60] obtained the seed markers from cell region.

Locating the UEP in a Euclidean distance map can only split those overlapped or clustered nuclei that have narrow waists, the small overlapping area, in the overlapping or clustered area. Instead of using the traditional Euclidean distance transformation, many modified transformations have been studied to obtain seed

markers where the connected nuclei have general waists. Malpica et al. [61] obtained the seed markers from the h-dome processed Euclidean distance transformed image by subtracting an appropriate h value of the transformed image, in order to reduce the noise peaks and improve the detection accuracy. The accuracy is controlled by the selection of h value. Based on the h-dome extracted seed markers, Cheng et al. [62] used the outer distance transform to draw the splitting line between overlapping nuclei by finding the least distance between the pixels in the nucleus and the centre. Yang et al. [63] used the conditional erosion to extract the seed markers by considering two cell-like masks and finding the last eroding points. The results may be sensitive to the size of erosion structures.

The distance, modified distance transform and intensity regional minima based marker-controlled watershed have all largely improved the nuclei segmentation accuracy, but they may still fail to split the overlapped or clustered nuclei that have a largely overlapping or clustering area. The combinations of distance transform and other methods have been proposed in recent studies to solve this problem. Zhang et al. [64] have used the expanded concave curve map and the distance transform to split the overlapping nuclei. The curve at the edge of the nucleus was expanded along the line of the direction of curvature, and converged at a centre point of a limited circular area. Markers were then obtained from the distance transformed separated particles. Based on the gradient direction and distance transform, Surut et al. [65] assume that the directions should be constrained to lead the pixels in a single nucleus to the centre, or direct the pixels in clustered nuclei to their destination. These methods can deal with the separation of overlapped nuclei which have a large overlapping area, but may be sensitive to the smoothness of nuclei edges, and the results may vary with the complexity of the background.

2.2.3   Region accumulation with shape features

The segmenting of clamped nuclei in a severely clustering area from a complex background is quite a challenge. Lin et al. [55] have used a radial symmetry method by searching the centre of the round like object to segment nuclei from the morphological filtered complex background. They combined this method with a morphological pre-processing and watershed in their four-step framework. An image with a complex environment and unevenly prepared stain was filtered with open-by-reconstruction and close-by-reconstruction to smooth the nuclear regions, which have a darker intensity. Next, a radial-dependent symmetry transform processed the image to generate nuclear markers with maximum values in a disc plate. In the disc plate, the pixels inside the nucleus have values being increased accordingly with the increasing of the disc range, while the pixels outside the nucleus have values being decreased at the same time. The watershed then segmented the nuclei based on the detected markers. A number of texture and shape features were considered in the post-processing step to eliminate noise particles.

This four-step method was improved in a later publication [28] by transforming the single scale with a single marker to a multi-scale with different markers. To determine a disc size for morphological pre-processing across entire images seems difficult; thus, Veta et al. [28] determined multiple values for the morphological transform. At each scale, the nuclei in the images have a marker that could be used for nuclei segmentation. During the post-processing, the segmented nuclei were assessed under different scales in order to classify the segmented nuclei as correctly segmented or over-segmented. The over-segmented nuclei were then selectively merged based on some criteria. Although this morphological filtering based method yields good results, the transformed image has removed many characteristics for the splitting of overlapped or clustered nuclei

and, by smoothing, has also overlooked the nuclei that have lower background contrast.

Another method used to solve this problem is Laplacian-of-Gaussian (LoG) operation, which used a blob detection filter to extract the seed markers [2]. This filter is given by:

$$LoG(x, y; \sigma) = \frac{\partial^2 G(x, y; \sigma)}{\partial x^2} + \frac{\partial^2 G(x, y; \sigma)}{\partial y^2} \qquad (2.6)$$

where σ is the scale value, and $G(x, y; \sigma)$ is a Gaussian with 0 mean and scale σ. It generates a peak response at the centre of each nucleus and is robust to the chromatin texture, which has smaller value than the nucleus. The LoG method is often used to detect candidate nuclear seeds [66], and has shown to be an accurate method of seeding the image [67]. For a better detection of nuclear seeds using the LoG method, an Eigenvalues-of Hesssian blob detection and distance transform have been combined to reduce the non-uniformity in the nuclear channel [68]. The seeds were then detected by LoG from the pre-processed binary image and constrained by a distance-map-based scale selection. However, this method may lead to high over-segmentation problem due to the highly textured intra-chromatin of the nucleus and variations in nuclear structures. Further classification [66] or regional accumulation methods, for example, watershed [67], can then be applied based on the detected seed points, or in consideration of the angular information between nucleus normals and the pixel in the proximity of detected seed points [69].

Instead of marker extraction to improve the segmentation accuracy, watershed has been regarded as the initial separation of severely clustered nuclear areas followed by different kinds of merging strategies, such as the h-minimal [70], the feature and compactness score [71], morphology and intensity [72], the measurement of homogeneity colour pixels [73], optimisation strategy [74], and

shape and size constraints [75]. The criteria used in these methods are aimed at the segmentation of nuclei from the specific sample, which may be varied across different samples.

This region accumulation based method can obtain much better segmentation results than the thresholding based methods for quantitative analysis. However, the watershed method only deals with the splitting of overlapping or clustered nuclei instead of considering about the contour of segmented nuclei. Thus, the segmented nuclei may have inaccurately labelled contour, and it is often combined with other methods to refine the segmented results, for example, the combination of active contour [76] and [77].

2.2.4 Deformable model fitting

Despite the region accumulation nuclei segmentation method and the deformable model fitting method, for example, the active contour has been used to more accurately delineate nuclear contours [44]. It is a powerful segmentation algorithm but suffers from high computation costs and the initialisation problem. The improvement in contour accuracy may come at the cost of increased computation time [78]. To cover these two problems, many methods have been proposed. Hu et al. [45] improved the traditional active contour by adding growth energy to cover the initialisation problem of conventional "snake" method, by restricting the contour movements along radial directions to reduce the computation cost. Dzyubachyk et al. [76] proposed a new approach by locating the initial position of the level-set function depending on the final position in the previous frame for the initialisation, in order to reduce the time cost. Based on a seeding process from a distance map, Pise et al. [79] obtained the active contour without user interaction, by using the obtained seed markers from distance transform as the initial contour; they have demonstrated that this achieves better segmentation results than the watershed method. Nevertheless, the initialization and time

cost remain quite challenging problems, and it is hard to balance the cost of computation against the contour accuracy.

2.2.5 Feature detection with circle fitting and concave points linking

Given the prior knowledge that nuclei show circle-like shapes, many studies have used circle-like masks, for example the ellipse, to match the objects in the image [80]. The goal of this matching is to identify the best-fitting ellipse, determined by the segmentation angle, area moment and circumference. Nuclei may also be detected by using a blob detector, such as the Gaussian based method presented in [81], which converted the overlapped separation into a cluster analysis problem. These shape descriptors can be better used when the shapes of nuclei are invariant. Otherwise, it is better to use edge detection followed by concave points linking.

The methods for obtaining of concave points may differ from the linking protocol. Saeger et al. [52] have computed the convex hull and regard the concave points as the pixels at the nuclei contour and between two convex points; the pair of points was linked when they have least distance and are perpendicular to the convex hull. Cong et al. [46] recognised the concave points by using a polygonal approximation on the original contour, and fitted the crease boundary with a closed curve function. Also using the polygonal approximation, Raman et al. [82] combined it with a measuring of angular change to compute the curvature values, the points being linked when they have maxima positive curvature values and least cost. The concave points were found by computing the maximum curvature along the contour, and they were used as the vertices of Delaunay triangulation to form the splitting line under three geometric constraints [83]. Based on the detected concave points, Plissiti et al. [51] have calculated the least squares fitting to find all the possible ellipses that fit to the

candidate separated nuclei, in order to reduce the false concave points caused by noise. Considering the concave points for clustered nuclei separation may have better results for the nuclei area, which have explicit difference between the normal curve and the concave points at the connected area.

Rather than classifying the curve and the concave at the edge of nuclei, an iterative spatial voting method has been proposed based on the moving of a filtering function along all edge pixels to detect the cell nuclei seeds in severely clustered or overlapped nuclei area [1]. It uses an oriented Gaussian kernel on the gradient magnitude image to infer saliency. The orientation of the kernel is relocated and reoriented to the gradient direction of each maximal voting pixel iteratively. For each iteration, the centre of the kernel moves toward the centre of the nucleus along the radial line of nuclear shape as it relates to symmetry. Using this geometric method, Jung et al. [81] proposed to segment nuclei by growing the voted centres to reach the detected edges that are thresholded from the gradient image. Similar to Bai et al. [84], Wienert et al. [52] used the voting method to extract nuclear seed points, and Voronoi tessellation to decompose the nuclear region. Further actual segmentation is performed with an active contour model based on additional intensity and geometric constraints.

However, use of the voting method in [84] created false nuclear centres in overlapping areas, due to the edges existed in these areas. To solve this problem, two modifications have been made in Filipczuk et al. [85]. They intend to move the convolution Gaussian kernel centre to the centre of the voting area rather than the edge pixels, and to use mean-shift instead of voting in the final iteration to capture the centre points. The detected nuclear centres are regarded as the initial position of the level-set to contour the nuclei. The method presented in [49] improved the voting speed of [85] by using single-pass voting for accurate and faster seed detection. This voting method is based on the detected edge of the nucleus, which has a very high accuracy in the locating of centre point, but may

fail in the detection of nuclei that have weak borders in a complex background with uneven stains.

2.3 Techniques For Gland Segmentation

The segmentation of gland structures is quite a challenge because glands have variable appearances and are embedded in complex backgrounds. From the shape, it seems easy to detect candidate glands by detecting closed regions of nuclei from controlled iterative morphological closing and region growing [86]. However, this is only suitable for detecting glands in tissue samples with fewer stroma nuclei, because when dealing with a complex sample it will miss small glands and incur false positives.

The colours of objects, such as the nuclei, stroma, cytoplasm and lumen, afford possibilities for segmenting glands in a reasonable combination scheme. In colour H&E stained prostate cancer images, gland structures were segmented based on the colour information in [13]. Each image consists of four components with different colour stains, such as the blue stained nuclei, pink stroma, purple stained cytoplasm and white lumen. Using a trained Bayes classifier in RGB colour space, these colours can be recognized separately. Gland structures were then segmented depending on the detected components and the level-set evolution function. With H&E stain, Naik et al. [87] applied the same gland segmentation approach as in [13] on breast cancer images as well. Also, using the level-set, Maqlin et al. [88] combined it with grid analysis of the surrounding nuclei to differentiate true and false lumen areas. Based on the detected lumen-like area through colour pixels clustering, Xu et al. [89] proposed a geodesic active contour to segment gland structures. The energy function of this active contour was dependant on the colour gradient and curve evolutions. Nguyen et al. [90] proposed a gland segmentation method in prostate cancer images based

on the colour information similar to that used in [87] and [13]. In this method, the boundaries of gland structures were extracted by connecting detected cytoplasm and nuclei that were sufficiently close. Next, the lumen areas were iteratively grown until the constraints were met. The same gland detection work has been done in [91] and tested in classifying grading levels with different feature extraction methods. By using the k-means clustering method, the pixels in the image have been classified into four different clusters and employed to construct the circle-like object graphs for gland detection in [4]. The features between objects can then be used to construct graph-based gland segmentation.

These colour characteristics used in H&E stained samples may not be transferred to H-DAB stained samples for the gland detection, due to the different prepared stains on tissue slides. The stain in the H-DAB sample consists of only two different colours, blue and brown. It is hard to classify those four components mentioned above based solely on colour or cluster characteristics. Therefore, the features used in this kind of sample have been replaced by the shape features and intensity changes. Farjam et al. [3] have proposed a gland segmentation method based on shape and texture information. They used the Gaussian function as a feature extractor on a greyscale image, and variance filtered the image separately. Pixels in these two processed images were then clustered into lumen, stroma and nuclei through k-means. The exclusion of nuclear pixels from images that containing lumen and stroma was used to obtain the glandular regions. Fu et al. [92] transformed the original image into polar space, and then a pair-wise algorithm was applied to link adjacent nuclear pixels to find the boundary of the gland. However, these shape and intensity features based methods suffer the problem of high computation costs, and may be sensitive to the intensity change due to the processing on low-level information.

2.4 Methods Used in the Thesis

In this section, the methods have been used in this thesis are introduced. It includes the statistical colour detection model, region growing, watershed nuclei segmentation, morphological gray scale reconstruction, PHOG feature and SVM.

2.4.1 Statistical colour detection model

A statistical model is a collection of probability distribution or density functions that describes the relationships between the variables. A histogram-based Bayes statistical classifier is a popular non-parametric modelling approach that is widely used in skin colour detection [93]. It treats the pixels as either skin or non-skin and represents them in histograms. The experiment in [94] demonstrates a Bayes statistical classifier with a histogram technique outperforms other tested classifiers for human face detection. In recent studies, Region-based statistical skin detection has been applied in [95], where the colour models used were similar to those in [93].

As described in [93], statistical colour models can be constructed using the histograms of the colour images. The probability of a colour, represented in a colour space RGB, can be obtained by first quantising the colour into colour bins [rgb] and then counting the number of pixels that fall into each bin, i.e.

$$Prob(rgb) = \frac{\#[rgb]}{\#TotalPixels} \tag{2.7}$$

To construct a statistical model for positively and negatively staining pixels, the labelled positive and negative pixels will be used, respectively, i.e.

$$Prob(rgb|S) = \frac{\#S[rgb]}{N_s}, Prob(rgb|\overline{S}) = \frac{\#\overline{S}[rgb]}{N_{\overline{s}}} \tag{2.8}$$

where S represents the number of positive pixels in a colour bin (rgb) and $\overline{S}$ represents negative staining, respectively, and N_s is the total number of positive and $N_{\overline{s}}$ the total number of negative pixels.

Based on the statistical models of positively and negatively stained pixels, it is possible to use the likelihood ratio approach to build the classifier. A pixel is classified as positively stained if

$$\frac{Prob(rgb|S)}{Prob(rgb|\overline{S})} \geq \Theta \tag{2.9}$$

where $0 \leq \theta \leq 1$ is the threshold. The value of θ trades off correction detection and false positives, which is the most important property of the classifier. One possible way to determine the threshold value can be determined as follows

$$\Theta = \frac{c_p Prob(S)}{c_n Prob(\overline{S})} \tag{2.10}$$

where c_p and c_n are the application-dependent costs associated with false positives and false negatives. One possible way to compute $Prob(S)$ is

$$Prob(S) = \frac{N_s}{N_s + N_{\overline{s}}} \tag{2.11}$$

2.4.2 Region growing

Region growing, which is based on a region or pixel, is one of the basic segmentation methods in image processing. The process starts with the selection of a set of initial seed points. It grows the regions by adding neighbouring pixels to the initial seeds. The neighbourhood pixels are examined to determine whether they satisfy the predefined criteria. Pixels are iteratively added into the seed regions until no neighbouring pixels are valid.

Seed points are selected by a criteria defined by users, such as the intensity of pixels or gradient values. More than one seed may be selected in an image for region partition or region accumulation. Each region should have at least one seed point in order to accumulate similar pixels in the neighbourhood.

For the growing phase, pixels in the 4- or 8-connected neighbourhood are examined in each iteration. The neighbouring pixels are marked if they are similar to the seed pixels based on the criteria. These marked pixels are then added into the region grown from the seed pixels and regarded as the new seeds for the next iteration. Growth stops if there are no pixels added into the region or the edge or boundary has been reached.

2.4.3   Watershed nuclei segmentation

Watershed is a region growing based segmentation algorithm derived from topographic and hydrological concepts [58]. The goal of this algorithm is to detect the watershed lines in an image. The greyscale image is considered as a topographic surface, with water flooding the surface from the catchment basin that corresponds to the regional minimum. The flooding water from different sources meets at crest lines that form watershed lines.

Normally, the topographic surface is formed from a grayscale image, a gradient image, or a morphological transformed distance map. To reduce noise and over-segmentation, the original image is often pre-processed with morphological filtering and Gaussian filtering. Extracting the correct markers corresponding to the regional minimum is considered to be the most efficient way to improve segmentation accuracy.

The watershed markers in a greyscale image or distance map can be determined by a threshold-based method; for example, extended minimum [47], h-dome [64], etc. In gradient image, the regional minimum theoretically corresponds to the homogeneous grey level regions.

Over- or under-segmentation often occurs after the original watershed segmentation or marker-controlled watershed segmentation. A post-processing step is required to resolve these problems. In general, merging or splitting methods in post-processing have been developed based on characteristics of intensity, shape information and a priori knowledge.

2.4.4 Morphological gray scale reconstruction

Morphological reconstruction is a mathematically-based operator that extracts marked components connected in a binary image. It can also be applied to a greyscale image for the tasks of filtering, segmentation and feature extraction [96]. Open-by-reconstruction and close-by-reconstruction first erodes and then dilates, or first dilates and then erodes, respectively.

For open-by-reconstruction, two greyscale images should be prepared. Image I is the original grey scale image, and image J is the marker image obtained from finding the minimum value in a pre-defined disc D.

$$\varphi(J) = (J \bigoplus D) \bigwedge I \qquad (2.12)$$

where "$\bigoplus$" is the dilation and "$\bigwedge$" is the minimum, which means the pixels in the dilated region were assigned the minimum value in the same region in I. Thus, the pixel values in J in the same domain were less than the pixel values in I. This assumes that the pixels that have intensity values larger than

the threshold are white and the rest of the pixels are dark. This simulates the opening of the white region in disc D.

The reconstruction of I from J occurs in these changed regions by finding the maximum value in a pre-defined disc D.

$$\varphi(J) = (J \bigodot D) \bigvee I \tag{2.13}$$

where "$\bigodot$ is the erosion and "$\bigvee$" is the maximum, which means that the pixels in the eroded region were assigned the maximum value in the same region within J. This simulates the closing of the white region in disc D.

From the minimum filtered marker J, the image I is reconstructed as follows:

$$\rho_I(J)(p) = max\{k \in [0, N-1] | p \in \rho_{T_{k(I)}}(T_k(J))\}, \forall p \in D_I(r) \tag{2.14}$$

where p is the pixel in domain D of image I, ρ is the reconstruction process, and k is the threshold value from 0 to $N-1$.

2.4.5 PHOG feature and SVM

Pyramid histogram of Oriented Gradient (PHOG) is a feature descriptor that is combined with a support vector machine (SVM) in computer vision for object recognition. Histogram of Oriented Gradients (HOG) has been demonstrated to outperform other features in human detection [97]. It has been extended for use in a spatial pyramid to describe objects [98]. This has then been used in medical image analysis in [92] for gland segmentation.

The image is divided into small connected patches called cells, using the same sized square or radius. Gradient directions or edge orientations are computed for each cell. The histograms are compiled from the gridded image cells and the cells are constructed into larger regions known as blocks. These histograms are normalised to reduce variations in illuminations and shadows. The normalised histograms are then fed to the SVM classifier based on supervised learning for objects classification.

The blocks used in [97] are rectangular blocks and circular blocks. Rectangular blocks are square grids and circular blocks are log-polar grids with two kinds of central cells, a single central cell and central cells that are divided into angular parts. With the appropriate parameters set for both, these two blocks can yield good results.

2.5 Discussions

- **Stain colour detection**

Colour is the important feature in the IHC image analysis for the classification of positive and negative samples. The previous methods studied the colour detection or separation by using the colour model transform [18], the pixels clustering [24], the colour correction [33] and the lights absorption characteristics [27]. The colour model transform is based on the assumption that each stain presents as a specific colour that can be obviously detected in a single colour channel, for example, the DAB stain in Y channel of CMYK model [18]. Although the DAB stain normally reacts as brown colour in digital image, the dark stained areas may appear as different. The colour correction is based on the matching of the input image according to the selected target sample to reduce the colour variations across different samples. However, this method is currently only available

for H&E stained samples [40], due to the unevenly prepared stains for H-DAB samples [34]. Meanwhile, the linear light absorption function based colour deconvolution, which was the most popular method in the colour detection domain, may still lead to the inaccurate detection of DAB colour due to the fact that the light is not linearly absorbed by the DAB stain [15]. Since the project presented in this thesis begins with the detection of brown colour in H-DAB stained samples, the clustering based method may be the optimal choice. Therefore, the statistical model has been considered. It is easy to implement and the accuracy can be controlled by the pixels accumulation.

- **Nuclei segmentation and quantification**

From the pipeline of research, the colour represents the pixel level information, and the next step is to detect the pixel-constructed nuclei and the nuclei-constructed gland structures that contain more information for the analysis. For the nuclei segmentation, the simplest way of detecting the nuclei region is the threshold, but using solely the threshold may not satisfy the quantitatively analysis. It is necessary to consider a separation method for splitting the overlapped or clustered nuclei. Therefore, the classical method, watershed algorithm has been proposed and considered as a possible solution, with its low computation cost and efficient segmentation scheme. Instead of considering the active contour [44], concave points linking [46] and circle fitting [80], watershed method is ideal because the stained images have complex background and a large number of nuclei, and most of them vary from each other in shape, size and gradient changes, etc. The active contour is better at delineating the contour of singular nucleus or a small number of overlapped or clustered nuclei, due to the high computation cost and the initialisation problems [78]. The searching for concave points is quite challenge in such a complex background, due to the severely clustered nuclei and the unsmooth nuclei edge caused by the uneven stain. Circle fitting is better to segment those nuclei that have invariant shapes in

a non-complex background. In addition, watershed is a flexible region growing based method, which can be combined with various marker extraction methods or fragments merging methods to achieve the segmentation and quantification goals in different kinds of stained samples.

- **Gland detection**

The gland structure detection in H-DAB stained samples contains less colour information than H&E stained samples. Therefore, the colour characteristics based methods were not considered. The morphology based methods have been regarded as a possible solution for gland detection [3, 92]. Lumen extraction has been demonstrated as an efficient detection of gland structures in the first step in the previous studies. It is easy to be observed that, in H-DAB stained samples, the lumen and the stroma are hard to differentiate based on colour, intensity, size or shape, due to the large variations of glands and complex background. The morphological transform has been used to enhance the differences between lumen and stroma areas by computing against the inner or outer side of gland boundary respectively. The stroma areas with explicit differences from lumen can be easily ignored from the transformed image. However, the candidate lumen obtained after transform still contain the stroma areas due to the large number of stroma nuclei may form like closed structures. Thus, the SVM classifier can be used to further classify glands as true or false.

Chapter 3

Stain Colour Detection

Colour is the most important feature in the analysis of immunostained images. The starting of the whole project is to detect the positive stain colours in IHC images. This chapter introduces the newly developed statistical stain colour detection method for the detection of stain colours in immunostained images. In the evaluation, this statistical colour detection model has been compared with two popular methods used in recent studies on different kinds of IHC stained images. The detection results are assessed against the manually labelled ground truth.

3.1   Introduction

Different observers may quantify the same slide differently due to the subjective nature of the process, thus causing inter-observer discrepancies. In order to minimize these inter-observer discrepancies, a semi-quantitative method based on visual scores has been proposed. Although the semi-quantitative method can reduce the discrepancies, these subjective scores have been demonstrated to have low reproducibility and low consistency [99]. Therefore, a computer based

objective quantification method is required to improve the detection accuracy and reduce the variations introduced by different observers. In interpreting an IHC image, it can regard the detection of stained pixels as colour detection in computer vision. Thus, the positively stained colour pixels can be extracted from the background by applying a colour detection method.

The method proposed here for stain colour detection in digital IHC images is a statistical colour detection model. This model was created from a huge collection of colour pixels that contain both positively and negatively stained colour pixels in the image. Based on statistical models of the positive and negative pixels, a maximum likelihood classifier is used to automatically classify pixels in digital slides into positively and negatively staining pixels. This statistical colour detection model was tested in the following schemes, shown in Fig. 3.1. First, to select the appropriate colour space, the created models from different colour spaces were assessed, such as RGB colour space, YCbCr colour space, and colour spaces without luminance like opponent colour space (rg-by) and CbCr colour space on the same dataset. Second, the models with appropriate colour space created by different users were tested to assess their robustness. Third, the detection accuracy of the proposed method was evaluated on different kinds of IHC stained images. For example, H-DAB stained oesophageal and colorectal biopsies, SR stained collagen and DAB stained elastin in liver cirrhosis samples. Finally, this stain colour detection model was compared with several previously proposed colour detection methods on the stain colour detection. The results obtained from these experiments have demonstrated the robustness and effectiveness of this statistical colour detection method.
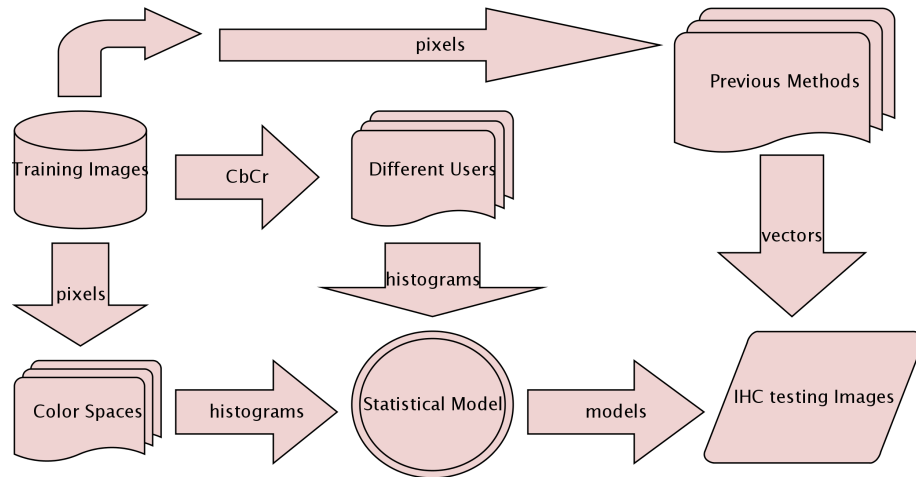
FIGURE 3.1: The evaluation scheme of statistical model.

3.2 The Building of Statistical Colour Models

Colour has been extensively used in computer vision for object detection. In order to detect the stain colour, a very large number of labelled pixels are collected through a semi-automatic tool (see Appendix A). Statistical models for positively and negatively stained pixels are constructed separately, and a maximum likelihood classifier is used to classify these colour stained pixels. Since colour plays the most significant role in biomarker identification, a model was built which only used the chromaticity signals. It has been shown that using chromaticity signals not only reduces the model's complexity but also is a more effective way of characterizing and modelling the immunostaining colours.

As when building any automatic machine classification algorithm, the first thing needed is training samples. The positively stained pixels and background pixels were collected separately and used to create the colour statistical models.

3.2.1 Statistical colour model construction

The statistical model presented in [93] has been used for the detection of positive immunostain colour. Generally, it is known that Red, Green and Blue (RGB)

colour space is not suited for image analysis. One reason is that chromaticity information and brightness (luminance) information are mixed together in this colour space, and it is often desirable to be able to process the chromatic and luminance signals separately. Therefore, separating the chromaticity signal from the luminance signal is commonly used in practice. In immunostaining detection, it is the chromaticity signal or the colour spectrum that is of interest rather than the absolute brightness. The chromaticity signals encode the spectral information of the immunochemical and therefore can be used to detect positive staining. From a computational perspective, using a 2D chromaticity space will make it easier to model the probability density function.

Four colour models have been tested in this chapter, including RGB, opponent colour model, YCbCr and a model without luminance, CbCr. The red-green (rg) and blue-yellow (by) chromaticity signals were derived from the original RGB input as follows

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B}, rg = r-g, by = \frac{r+g}{2}-b$$
(3.1)

The Cb and Cr chromaticity signals were derived from the original RGB space as follows

$$Cb = -0.1687 \times R - 0.3313 \times G + 0.5 \times B, Cr = 0.5 \times R - 0.4187 \times G - 0.0813 \times B$$
(3.2)

The colour statistical models in the chromaticity space can now be constructed as

$$Prob((rg, by)|S) = \frac{S[rg, by]}{N_s}, Prob((rg, by)|\overline{S}) = \frac{\overline{S}[rg, by]}{N_{\overline{s}}}$$
(3.3)

$$Prob((Cb, Cr)|S) = \frac{S[Cb, Cr]}{N_s}, Prob((Cb, Cr)|\overline{S}) = \frac{\overline{S}[Cb, Cr]}{N_{\overline{s}}}$$
(3.4)

where S represents the number of positive pixels in a colour bin and $\overline{S}$ represents

negative staining, respectively, and N_s is the total number of positive and $N_{\bar{s}}$ the total number of negative pixels. The model in other colour spaces can be similarly constructed.

The calculation of probability for each colour bin has been illustrated in Fig. 3.2.
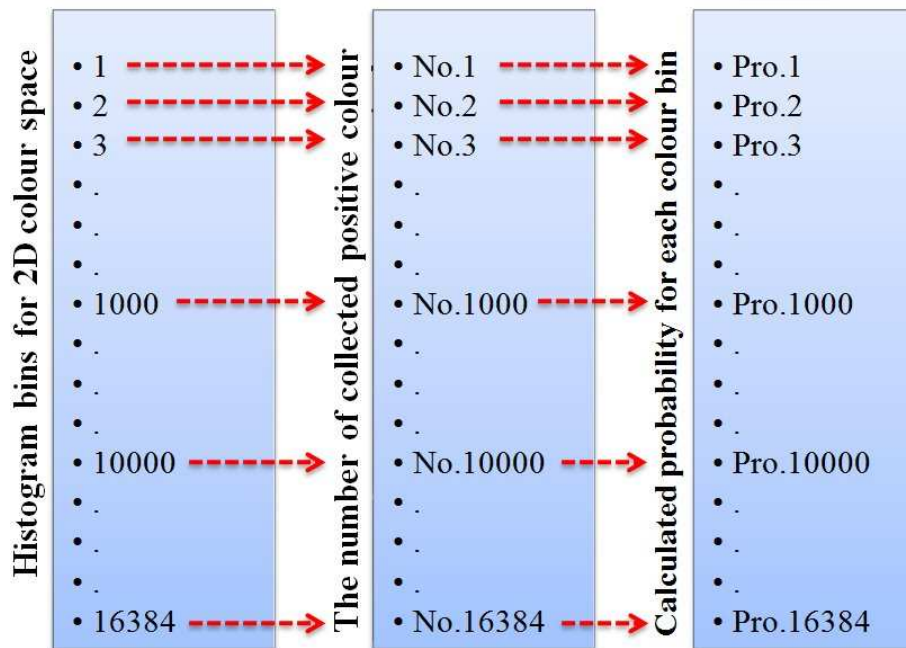
FIGURE 3.2: The illustration of probability calculation. The left column is the index of colour bin for 2D colour space with 128×128 bins. The middle column is the number of collected colour pixels quantised in each colour bin. The right column is the calculated probability value for each colour bin (see Equation 3.4).

3.2.2 The evaluation of statistical model with different colour models

The technique was tested on two sets of digital slides from oesophageal and colorectal biopsies. The slides were scanned using a Hamamatsu scanner. A semi-automatic tool was first used to manually label positively stained pixels,

as described in Appendix A. The data for building the statistical colour models included 20 images with resolution 6720×4200. The models were then tested on another set of 75 images with resolution 6720×4200. Both the training and testing images were captured under 20× magnitude. I manually prepared the ground truth of the test dataset. The purely brown images were obtained after eliminating the background pixels.

The obtained results were plotted with Receiver Operating Characteristic (ROC) curves with a true positive ratio and false positive ratio. These ratios were calculated as follows:

$$TP = \frac{Correct}{Total}, FP = \frac{Detected - Correct}{Background} \tag{3.5}$$

where TP is the True Positive ratio, and FP is the False Positive ratio. Only positively stained colour pixels were shown in the detected result images and converted to greyscale. These two ratios were calculated at each intensity value on these result images (see Fig. 3.3).

	64 bins in each axis			128 bins in each axis			254 bins in each axis		
	S	$\overline{S}$	OL	S	$\overline{S}$	OL	S	$\overline{S}$	OL
rg-by	449	588	202	1296	1368	403	3149	3099	496
Cb-Cr	56	155	18	181	516	49	514	1308	129
RGB	6553	13735	2082	35915	79150	9001	61372	139005	14354
YCbCr	1991	4122	698	11038	23182	2972	44033	97609	10549

TABLE 3.1: THE EXPERIMENTAL RESULTS WITH DIFFERENT COLOUR MODELS. The number of bins occupied by positively staining pixels (S), background pixels ($\overline{S}$) and overlapping (OL) bins in colour histograms of 4096, 16384, or 65536 bins (rg/by and Cb/Cr) and 262144, 2097152 or 16777216 bins (RGB and YCbCr)

As expected, the colour distribution occupied a relatively small part of the colour space. Table 3.1 shows the number of non-empty bins for different sized colour histograms in 4 different colour spaces. Two spaces with only chromaticity channels are included, the opponent colour space and CbCr space. Also shown
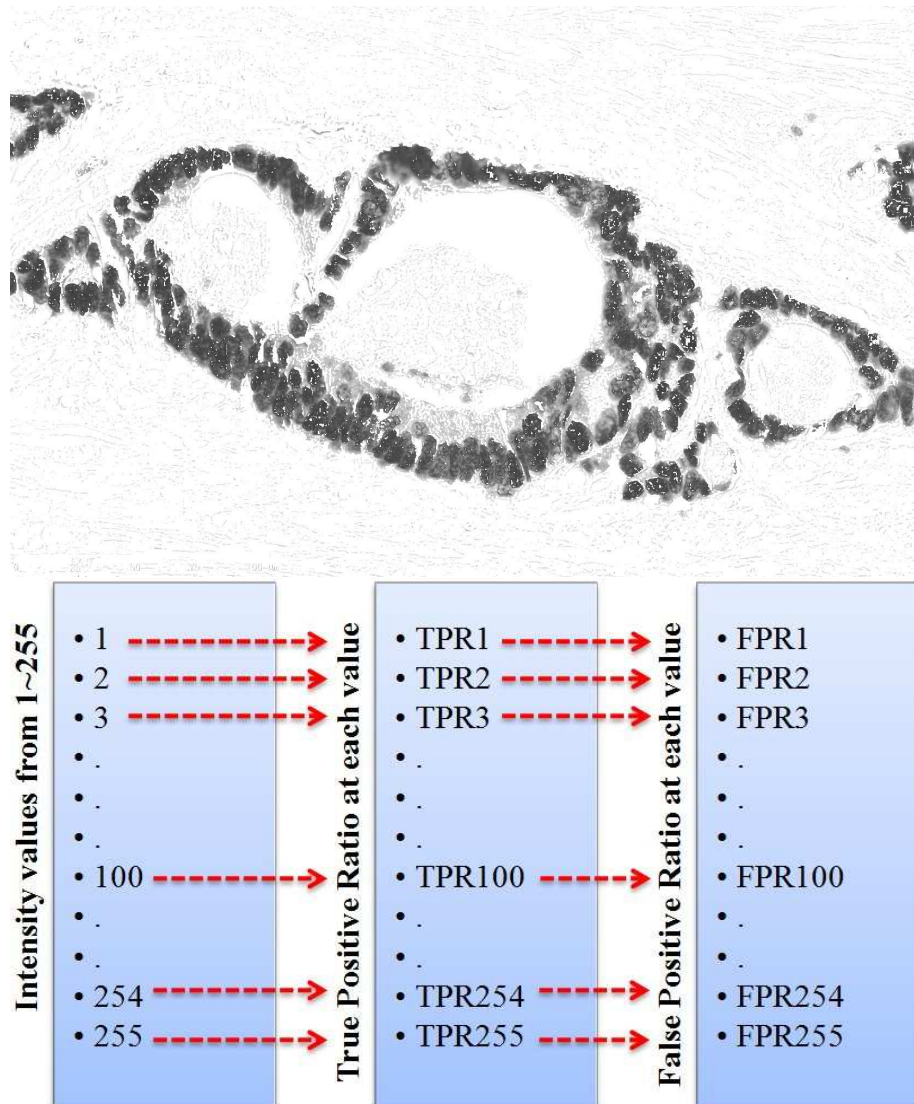
FIGURE 3.3: The illustration of evaluation process. The top image is the greyscale result image with detected stain colour pixels; background pixels are set to 255. The bottom image is the illustration of the true positive ratio and false positive ratio calculation. The ratios are calculated against ground truth at each intensity level from 1 to 255 respectively; these can then be summed to calculate the average ratios.

are the number of overlapping bins, i.e., bins that are occupied by both positive and negative pixels. It is interesting to observe that the CbCr chromaticity space has the smallest number of overlapping bins and the experimental results confirm that this space gave the best performance.

Fig. 3.4 shows the ROC curves of models of different colour spaces with different number of histograms bins. The plotted true positive ratios and false positive ratios were obtained at the intensity level from 1 to 255 (see Fig. 3.3). ROC curve compares the two criterion changes, true positive ratio (TP) and false positive ratio (FP), at each intensity level as the operating characteristics. The result with highest TP and lowest FP would be the optimal model. Each colour model was evaluated with three different numbers of colour bins.

It can be seen that although all models achieved very good results, the CbCr chromaticity model achieved better results than models that included both chromaticity and luminance. This indicates that chromaticity is sufficient and luminance is a distraction in terms of building the model. As mentioned before, 2D chromaticity signals simplify the model, make it faster to compute, and demand less memory. The optimal choice for the number of histogram bins is 128, obtaining better results than 256 bins with a smaller computation cost.

3.3 The evaluation of statistical model with different users

Since the construction of a statistical model is based on collecting colour pixels, variations in the detection results may occur between models that are constructed by different peoples. Thus, it is necessary to evaluate the robustness of the created statistical colour detection models. The robustness of the statistical colour models created by different people was evaluated by measuring its detection accuracy and variations.

Four people were asked to join in this experiment to detect the brown colour from DAB stained IHC images. Only one - myself - possessed a pathology background; the other three had non-pathology backgrounds. The evaluation

was treated as the detection of all brown colour pixels. Therefore, the participants were not required to have any pathology background. They were asked to use the same training dataset to create models. These models were then tested on the same test datasets, which were distinct from the training sets. Since different people may regard different colours as brown, the standard deviations in their detected true positive ratio and false positive ratio were calculated separately. The comparison is against the manual ground truth.

The prepared training dataset was captured from two kinds of DAB stained images, 10 Whole Slide (WS) images with resolution 6720×4200 and 10 Tissue Microarrays (TMA) images with resolution 5120×4096. These two kinds of images were re-sorted into three types of training samples. Each type of training sample consisted of 10 images, such as 10 WS images, 10 TMA images or 10 mixed images (5 WS images and 5 TMA images). Each people were required to build three statistical models for the detection of brown colour. The colour pixels used were collected separately from these three types of DAB stained training samples. As a result, four peoples were required to create 12 models with the collected colour pixels.

For the test datasets, these included two datasets of DAB stained images from WS images and TMA images. Both DAB stained datasets contain 50 images that were captured under 40× magnitude with resolution 1680×1050. The models based on TMA training samples were tested on TMA test samples and the models constructed from whole slide training samples were tested on whole slide test samples. The models created from mixed images, which covered a wider range of colours than the previous two groups, were then tested on both test samples.

The results obtained from the test dataset by the 12 models are shown as ROC curves in Fig. 3.5. The true positive and false positive ratios were calculated in the histogram bins from 0-255. All 12 models obtained good results with high accuracy and low false positive ratios, and the 4 user-created models performed

better in the whole slide image dataset than in TMA images. The curves with a true positive ratio of less than 0.8 are nearly the same, for the models obtained detection accuracy is very high at each intensity level.

Although most of the models achieved similar results, there were still variations between them requiring examination in further detail. The average values of the detected results corresponding to the three training datasets were calculated separately. The deviations in each model were then calculated against the corresponding averages in different histogram bins. The formulations are shown as follows

$$R_{avg}(i) = \frac{\sum_{u=1}^{4} R(i)}{4}, i = [0, 200], STD = \sqrt[2]{\frac{\sum_{u=1}^{4} [R(i) - R_{avg}(i)]^2}{4}} \quad (3.6)$$

where $R(i)$ is the average true positive or false positive ratio obtained from each user at threshold i, n is the number of involved users, and the intensity threshold has a range from 0 to 200. $R_{avg}(i)$ is the average true positive or false positive value of 4 users at threshold i. This calculation is based on the average true positive and false positive rates at each threshold. From the result shown in Fig. 3.6 it is obvious that each model, constructed by different users, had small variations in its true positive and false positive ratios on both whole slide and TMA test images when the corresponding training images were used. For example, at a threshold of 150, the four user-constructed models all had the same true positive ratios (92.1%) and the false positive ratios varied between 1.1% and 1.2%.

As a result, the models created by different people from the same dataset may all provide the satisfactory results with similar level of detection accuracy. Also, the models based on the accumulated histograms may be able to reduce the subjective discrepancies between people.

3.4 The comparison of statistical colour detection model with other methods

In this section, two popular colour detection methods proposed in previous studies were compared with the statistical stain colour detection model. They were trained and tested on the same datasets. The methods were assessed and compared in terms of detection accuracy, and the separation of stain colours.

3.4.1 The comparison of detection accuracy

The dataset used in the comparison experiment was the same as that used in the robustness evaluation in Section 3.3. The previously obtained average results from 4 users were compared with colour deconvolution and CMYK.

	WholeSlide		TMA	
	model(WS)	model(Mix)	model(TMA)	model(Mix)
User1	98.6%	94.8%	97.0%	95.4%
User2	98.6%	98.6%	97.0%	97.0%
User3	94.5%	96.1%	98.0%	98.3%
User4	98.4%	94.5%	96.3%	96.3%
AVG	**98.5%**	**97.9%**	**97.6%**	**97.6%**
CD	88.7%		90.2%	
CMYK	90.8%		84.0%	

TABLE 3.2: THE AUC VALUES OF ROC CURVES. WS = whole slide test images; TMA = TMA test images; ST(WS) = statistical model constructed on 10 whole slide training images; ST(TMA) = statistical model constructed on 10 TMA training images; Mix = statistical model constructed on 10 mixed training images (5 WS images and 5 TMA images); AVG = average values of four user-detected results; CD = Colour Deconvolution.

The experiment was applied on ImageJ. Colour deconvolution was programmed by Landini as a plugin of imageJ [100] based on the National Institutes of Health Image macro. The built-in vector H DAB (red = 0.26814753, green = 0.57031375, blue = 0.77642715) was used to discriminate the brown colour

from the background, for there is no single dyed samples prepared to obtain the vectors on these two datasets. The CMYK model was also developed in ImageJ based on the functions mentioned in [101].

As shown in Fig. 3.7, all runs achieved a very high accuracy, with close to 100 percent correct detection compared to the ground truth. Using the 50 whole slide test images, the average value of the statistical models obtained the best result with a false positive ratio of 5.3% as compared to CD, which had a false positive ratio of 7.7%. The CMYK model attained the lowest correct detection ratio of 95.4%. With the test on 50 TMA images, the best result was also attained by the average value of the statistical models, i.e. a 5.9% false positive ratio as compared to CD, which had an 11.6% false positive ratio. The CMYK model still had the lowest correct detection ratio of 91.4%.

To clarify the results, the calculation of AUC (area under ROC curve) through integral calculus was applied in the evaluation process. Table 3.2 represents the percentage of calculated area values under each curve against the whole area. The statistical colour models obtained the best results, occupying the largest sized space under the curve.

Colour deconvolution obtained better results than CMYK on TMA test images but was not as good on whole slide test images. From this table, it can also be seen that the user-created models had small variations in terms of their detection accuracy. For example, for brown colour detection, the lowest AUC was 94.5% and the maximum was 98.6%. As a consequence, models constructed by different users can achieve the desired accuracy and are sufficiently robust.

3.4.2   The comparison of stain colour separation

In the experiment, the normal brown colour could be easily detected and separated from the background by all three methods. However, the detection of

brown colours in a dark-stained slide is quite challenging. For example, dark brown colour pixels were under-counted in the results obtained from CMYK, as shown in Fig. 3.8(b) & (e). The dark-brown pixels in dark-stained regions are missing in the CMYK detected result image, the yellow image. Meanwhile, CD also detected dark brown as the colour blue, as shown in Fig. 3.8(a) & (d). The evaluation of CMYK has proven that a colour space-based method performs less accurately in stain colour detection [102]. Classifying colours in the colour space may suffer an overlap problem. The CD method also suffers from this problem due to the overlapping light spectrum and non-linear transform. The statistically-based colour detection models detected dark brown and blue accurately, i.e. "what you see is what you get". The wider the range covered by the collected pixels, the more colour can be detected by the statistical model.

3.5 Discussion and conclusion

The IHC stains have been regarded as the important indicators, which have obviously characteristics, for the quantitative analysis of many biopsies, for example, colon cancer and liver cirrhosis. Stain colour detection has been studied by many previous papers of using image analysis techniques to give objective and accurate results. In this chapter, the statistical colour detection model has been applied for stain colour detection.

Instead of detecting a specific colour, the detection of stain, as has been demonstrated, is based on the approach "what you see is what you get". To create the model, colour pixels were visually determined and collected. This was the visual determination of the positive colour and negative colour in one slide. However, the accuracy of the model might have been affected by the selected colour space. Four colour spaces were compared in the DAB colour detection experiment, which demonstrated the detection accuracy for each colour space. RGB

colour space and the absolute luminance channel was discarded to reduce both computation costs and the requirement for computer memory. Therefore, this model was created based on the 2D chromaticity signals, CbCr colour space, which have a higher accuracy than the rg-by opponent colour space and other 3D spaces (see Section 3.2.2).

Although the visually determined colour pixel collection process can improve the detection accuracy, human variations may still result in different detection accuracies. To evaluate the variations between different people, 12 models were tested on the same test dataset. Models were created by four peoples from different training images. The obtained detection accuracy only showed tiny variations between peoples, and inter-observer discrepancies were not obvious. This means that the accumulated histograms from multiple training samples can provide stable results with small variations (see Section 3.3). It offers the ability to reduce the subjective human discrepancies.

Since the samples are prepared with multi-stains which may have overlapping bins in the colour space, using the statistical model to differentiate the brown colour and blue colour in H-DAB stained samples has clearly shown the colour separation (see Section 3.4.2).

It is clear that stain colour detection is similar to normal colour detection in computer vision. A statistical model can also generate good results in the medical image analysis domain. With the DAB stained tissue samples in particular, the statistical model combined with an interactive human training process can obtain better results than either the CD or CMYK methods (see Section 3.4.1). It is recommended to collect sufficient colour pixels to improve the detection accuracy. The pixels that are collected should cover the whole range of colour shades as much as possible.
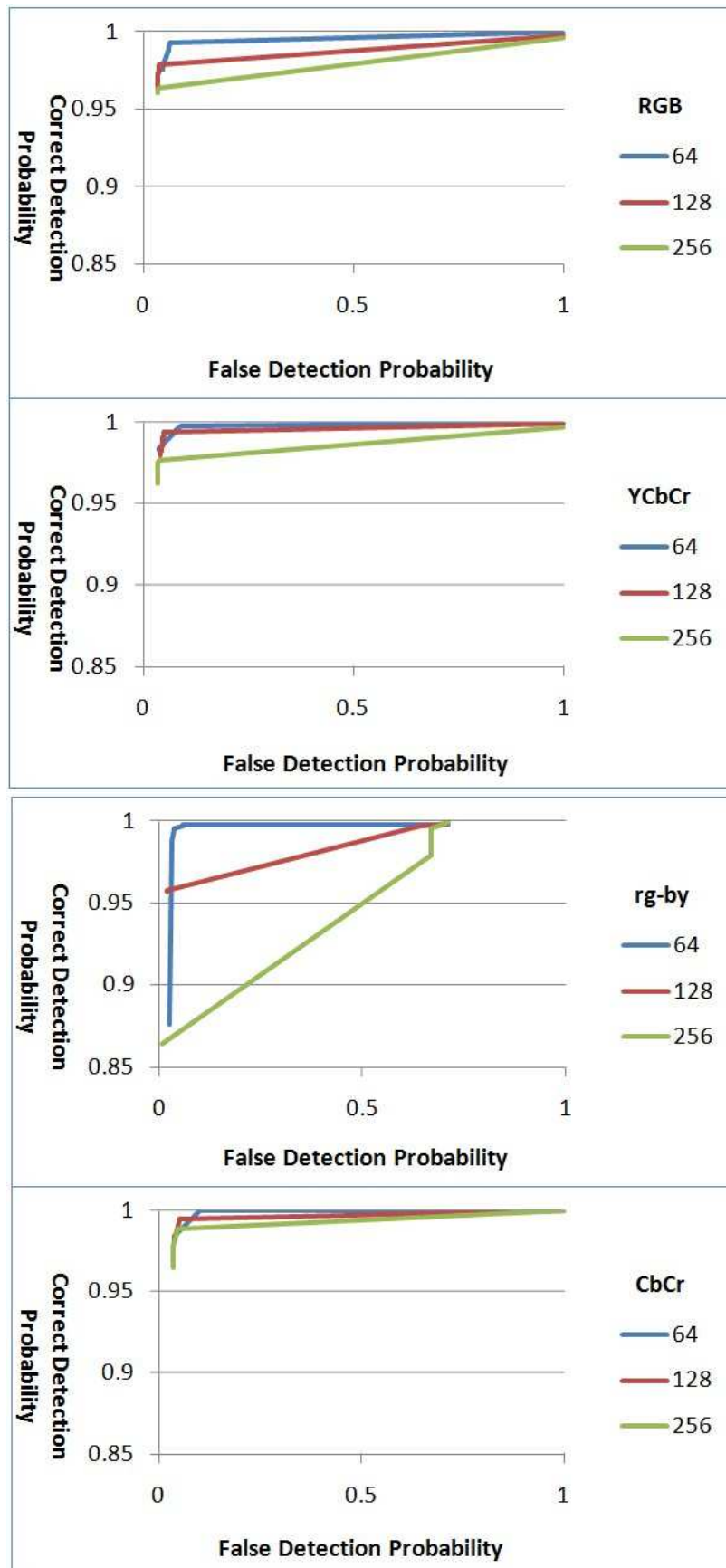
FIGURE 3.4: The ROC curves of statistical colour detection models with different colour spaces. The blue curve shows the ROC of models created with 64 histogram bins; the brown curve shows the ROC of models created with 128 histogram bins; and the green curves shows the ROC of models created with 256 histogram bins.
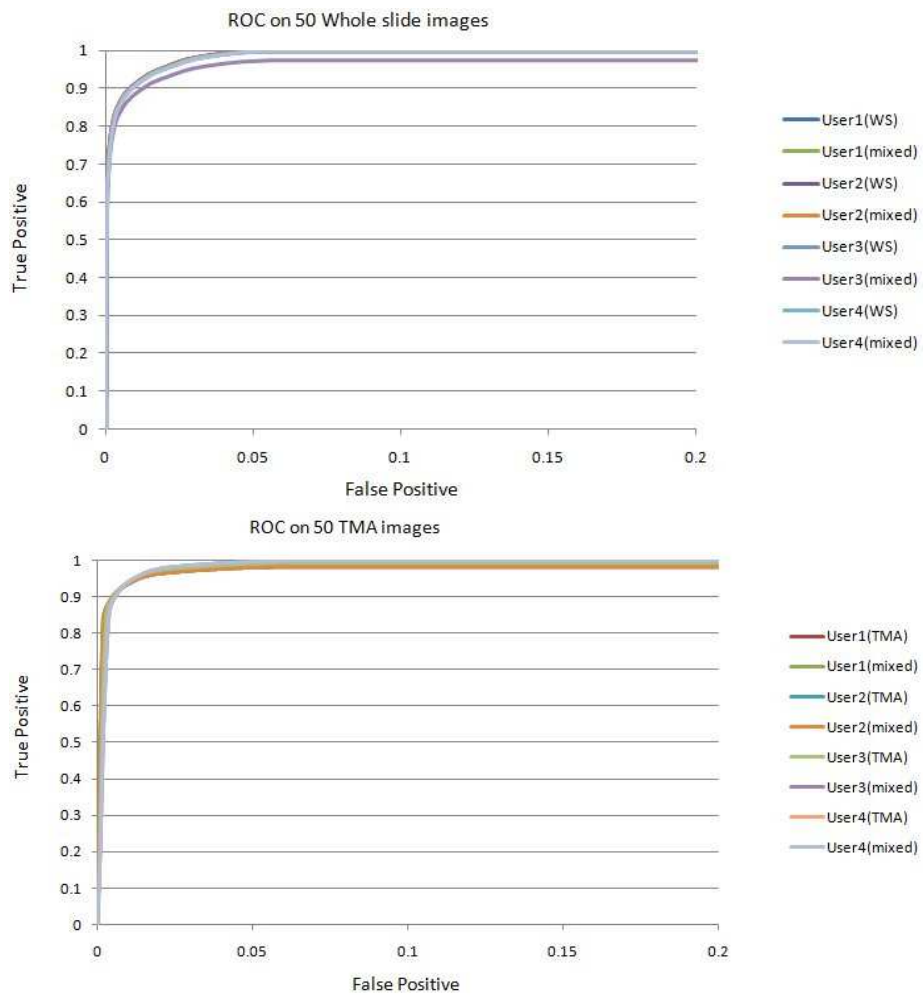
FIGURE 3.5: ROC curve for different user constructed statistical models. Top row: ROC curve for User constructed statistical models (based on 10 whole slide images and 10 mixed images) on 50 whole slide test images; Bottom row: ROC curve for User constructed statistical models (based on 10 TMA images and 10 mixed images) on 50 TMA test images.
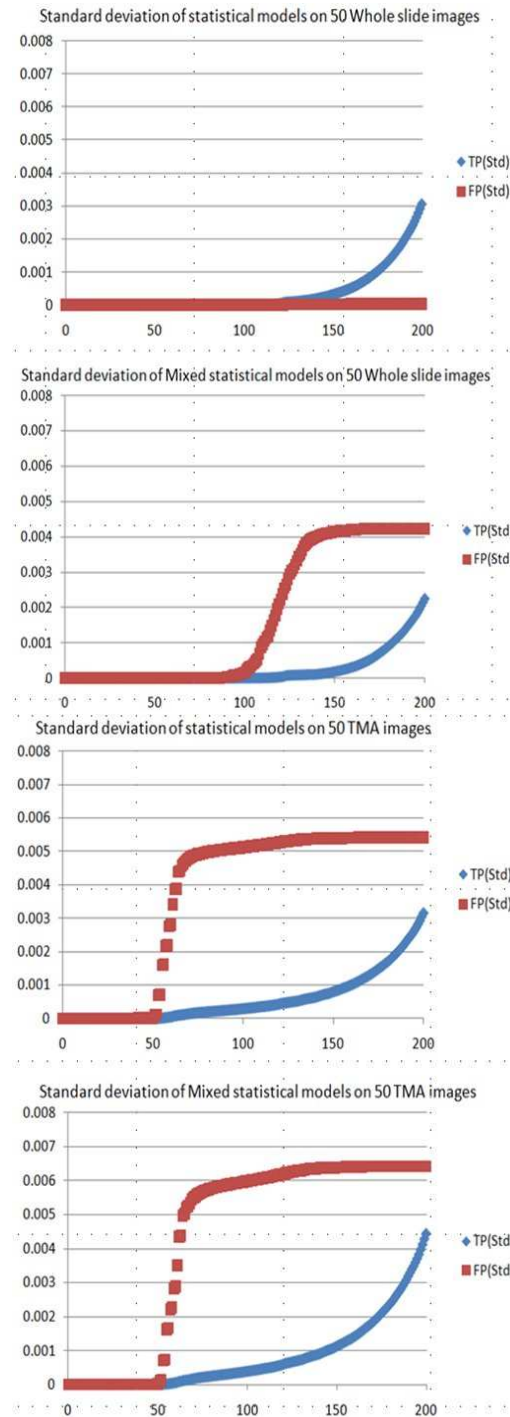
FIGURE 3.6: Standard deviation in statistical models for the evaluation of model variation based on the intensity threshold. Top two: standard deviation in statistical models trained from whole slide images and mixed images which were tested on 50 whole slide images, Bottom two: standard deviation in statistical models trained from TMA images and mixed images which were tested on 50 TMA images. The horizontal axe is the intensity value, and the vertical axe is the standard deviation.
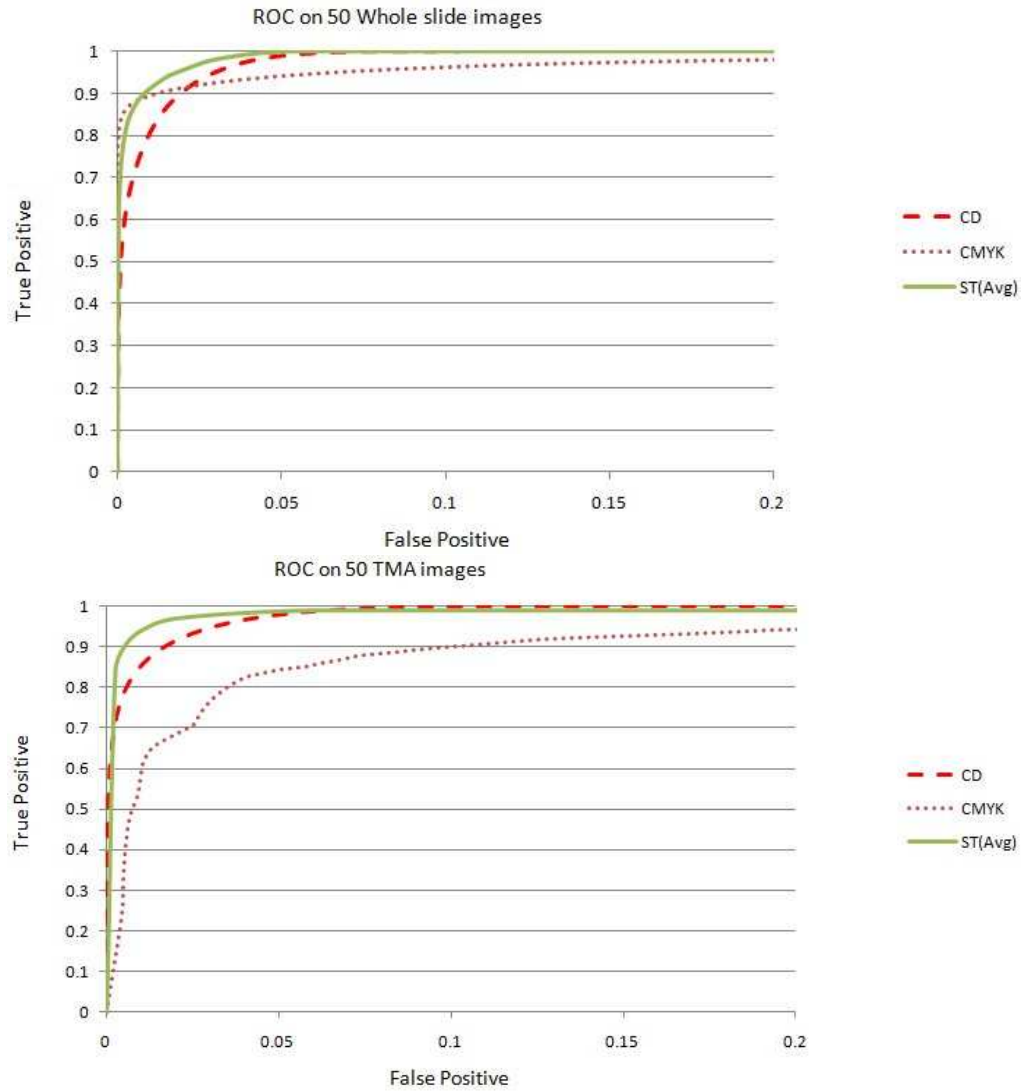
FIGURE 3.7: ROC curves for colour deconvolution, CMYK and average value of statistical models. The plot on the top shows the ROC curves for colour deconvolution (red dashed line), CMYK (brown dotted line) and the average value of the statistical models (smooth green line) on 50 whole slide test images; the plot on the bottom shows the ROC curves for colour deconvolution, CMYK and the average value of the statistical models on 50 TMA test images.
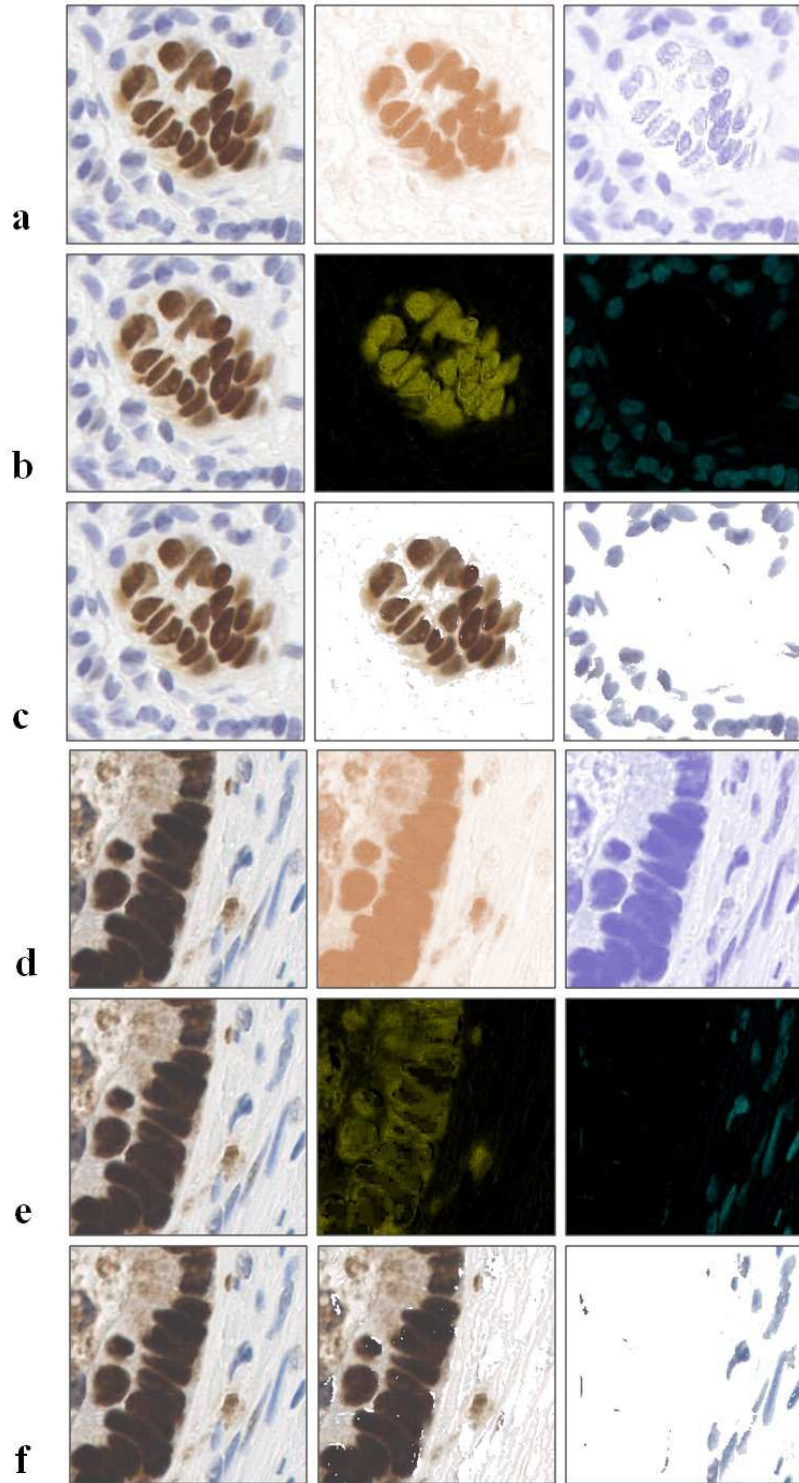
FIGURE 3.8: Stain colour detection on dark-stained slides. Left column is the original image, middle column is detected brown colour, and the right column is detected blue colours. (a) & (d) are obtained from CD, (b) & (e) are obtained from CMYK, (c) & (f) are obtained from the statistical model.

Chapter 4

Nuclei Segmentation

Verifying the presence of nuclei and quantifying positive nuclei is the foundation of cancer grading. Studies on nuclei analysis can be dated back to the 1950s [41]. It has been studied in a vast array of papers covering several research domains such as cell counting, the identification of cell types or phases, cell migration and interaction quantification, and intracellular structures [103]. In this chapter, a new nuclei segmentation and quantification approach is introduced.

4.1 Introduction

A multistage watershed-based approach for segmenting severely clustered and overlapping cell nuclei has been proposed, which can reduce the problems mentioned in Section 1.2.2. The severely clustered nuclei are those nuclei that are clumped and difficult to be split by most of the previous methods. The proposed approach can also simultaneously reduce both over-segmentation and under-segmentation and improve detection accuracy. The first process proposed was to detect initial seeds by utilising morphological features. Based on these detected

seeds, a region growing-based method was developed to obtain the correct seeds. It was then followed by a watershed algorithm to yield the final nuclei segmentation. This approach is easy to implement, relatively lightweight, and very efficient for nuclei segmentation and quantification. Experiments on a newly proposed dataset have shown that this approach outperformed several popular methods in previous studies of nuclei segmentation. After the segmentation accuracy had been evaluated, the quantification results were measured on another dataset with large-scale images and vast numbers of positively stained nuclei.

4.2   Materials and Methods

All experimental studies were performed on images of human colorectal cancers. These images were IHC stained images prepared using Diaminobenzidine (DAB) as the chromogen. They were digitized from both Whole Slide Images (WS) and Tissue Micro Array (TMA) images with complex background and areas of severely clustered nuclei.

The first step in this work was developing the nuclei segmentation approach and evaluating its accuracy. The segmentation process was first evaluated on a set of small-scale test images. An example of these test images is shown in Fig. 4.1. The nuclei are severely clustered and there is great variation in shape among different nuclei. The ground truth of this test dataset consisted of manually contoured nuclei samples that I prepared by myself (see Fig. 4.5). Next, the segmentation approach was applied on another set of DAB stained images. These images, which were of a larger scale and contained severely clustered nuclei, were used to evaluate the accuracy of the quantification. An example of test images for nuclei quantification is shown in Fig. 4.13; it includes vast numbers of nuclei and a complex background. A ground truth for the quantification

evaluation was also prepared manually by a pathologist and me, based on the "cell counter" plugin in ImageJ (see Fig. 4.13).

The proposed approach consists of the following three steps. First, the foreground regions that may contain nuclei are extracted. Second, a new seeded watershed algorithm combined with a region growing step is applied to the foreground regions for nuclei separation and segmentation. Third, redundant pixels are eliminated and an ellipse is fit to each segmented nucleus. This whole process is outlined in three parts, Algorithm 1, Algorithm 2 and Algorithm 3.

4.2.1 Foreground & background Classification

In general, the intensity of the foreground, i.e., the pixels belonging to the nuclei, can be easily distinguished from the background by a global threshold method such as Isodata [43]. However, in DAB stained TMA images, the nuclei boundaries could easily be misclassified using a global threshold method, as the mean intensity levels vary across the background of the entire image. The global threshold method obtains the threshold value based on the intensity or histogram of the whole image. Therefore, a global threshold method cannot provide satisfactory results, as shown in Fig. 4.1(b).

Instead, a combination of global threshold and local threshold can be used to tackle this problem. Local thresholding allows an adaptive threshold value to be applied in a local region and can efficiently counter the issue of intensity variations across the background [86]. The local threshold is applied to pixels that are classified as foreground by the global threshold method. This same automatic thresholding method [43] has been used for both global and local thresholding. The size of the local thresholding regions (window) and the step-size of the moving window will affect the accuracy of the local thresholding results. In order to classify the foreground from the background in a local window, the
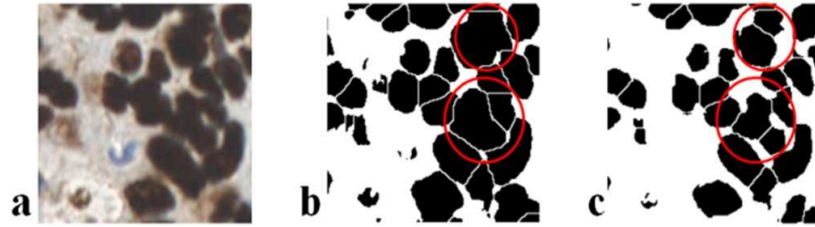
FIGURE 4.1: Watershed on extracted foreground regions. (a) original image. (b) watershed performed on the global thresholded images. (c) watershed performed on the combined global and local thresholded images. The falsely segmented nuclei are highlighted with red circle.

size of this window should be larger than the size of the nuclei in the image. Smaller windows would falsely eliminate pixels belonging to the nuclei and lead to over-segmentation, while larger windows would cause under-segmentation. Therefore, it is important to set an appropriate size for the local window. Results obtained with different window sizes are shown in Fig. 4.2(b)-(d). The determination of window size is also discussed in Section 4.5. The horizontal
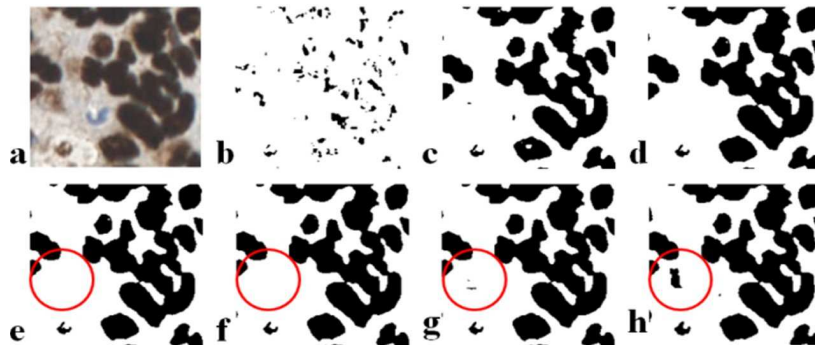
FIGURE 4.2: Local window movement.((a) original image. (b-d) different window sizes, 10×10, 30×30, 50×50. (e-h) different moving distances per iteration, the shown moving distances include 1 pixel, 5 pixels, 15 pixels and 25 pixels. The missing detected nucleus is highlighted with red circle.

distance that the window moves is determined by the size of the window, which means $1 \le d \le s$, where d is the horizontal moving distance and s is the size of the moving window. As shown in Fig. 4.2, nuclear pixels are largely eliminated when the moving distance is small, especially when d = 1. Although smaller moving distances (from 1 pixel to 20 pixels) outline the boundaries of clustered nuclei very well, a weakly stained object (circled area) is missed. The moving

distance is set to be 25 pixels, which can reduce the computation cost and obtain the best result.

4.2.2 Seeded watershed

Although the combination of local threshold and global threshold can clearly detect the boundaries of clustered nuclei, the largely eliminated intra-pixels may lead to over-segmentation. Filling in the holes offers the potential to minimize the over-segmentation effect. Meanwhile, it may also fill the inter-nuclei gaps that may lead to under-segmentation. Thus, the normal watershed may not produce the desired results. Therefore, it is better to use a seed-controlled watershed to segment clustered nuclei.

Algorithm 1 The process of the proposed seeded watershed. α is a threshold value, which is the larger value between the average size of particles (Avg) and the minimum size constraint (C_s). S_n is the size of nuclei, and C_n is the centre of nuclei.

Require: RGB image I
Require: Detected brown colour image I_{brown}
  Convert I to grey scale image I_{grey}
  Obtain $mask1$ by Auto-threshold I_{grey}
  Obtain $mask2$ by Local-threshold $mask1$
  Apply EDM watershed on $mask2$
  $\alpha = Max(Avg, C)$
  for i = 1 to n **do**
    if $S_n < \alpha$ **then**
      Remove S_n
    else
      Calculate C_n
    end if
  end for
  Conditional region growing outlined in Algorithm 2
  Seeded watershed on $mask2$
  Post-processing outlined in Algorithm 3

Distinct from previous seed-controlled watersheds, the seeds used in this approach were obtained through the following two steps. First, two binary masks,

$mask1$ and $mask2$, were prepared, see Algorithm 1. Binary $mask1$ is obtained by the global threshold [43] and prepared as the foreground mask. Binary $mask2$ is obtained using the combined global and local threshold method described in Section 4.2.1 that is used for nuclear seed extraction. $mask2$ is then transformed to a Euclidian distance map (EDM) [104]. The initial seeds are obtained by finding the Ultimate Eroded Points (UEP) in this transformed mask, as shown in Fig. 4.3(a). The watershed splits the areas with overlapping nuclei into smaller particles. Instead of focusing on segmentation accuracy, the overlapping or clustered nuclei are over-segmented, which can split the clustered area as much as possible.

Larger sized particles are more likely to belong to nuclear objects, while smaller particles may belong to noise. The removal of noisy particles depends on the mean size range of the nuclei. However, the sizes of nuclei vary across different images, even within a single image. And severely clustered areas may result in inseparable particles that have a large size. The average size of nuclear particles may be larger than the size of correctly segmented small nuclei. Therefore, a minimum size constraint should be added to prevent over-elimination. This can be formulated as follows:

$$N = P | f(P_i) \geq Min(Avg, C), N \in P \tag{4.1}$$

Where N are the nuclear particles after removal, P are the nuclear particles before removal, $f(P_i)$ is the size of particle i in P, Avg is the average size of all particles in P, and C is the minimum constraint value. During the removal process, particles whose sizes are smaller than the mean size or the minimum size constraint are regarded as noise, as shown in Fig. 4.3(c). For the size variation in nuclei, an image that has a large number of small particles may result in a smaller average value than the constraint. And an image that has a large number of large particles may result in a larger average value than the constraint. This

constraint value is a balance critierion, for these variations may occur in test images.

Algorithm 2 Conditional region growing. In $mask1$ and $mask2$, the intensity value of pixels that belong to objects is 0, and the value of background pixels is 255. N_i is the ith growing nuclei particle. P_i are the pixels belong to nuclei N_i. $Intensity_i$ and $Intensity_G$ are the average intensity value of pixels belong to N_i, and the intensity value of growing pixels G. G_{mask1} and G_{mask2} are the intensity value of growing pixels G in $mask1$ and $mask2$. $Intensity_{mask2}$ is the intensity value of pixels in $mask2$.

Require: $mask1$ and $mask2$
  $P_i \in N_i$
  for $i = 1$ to n **do**
    if $G_{mask2} = 255$ AND $G_{mask1} = 0$ **then**
      Add G to P_i
      Calculate $Intensity_i$
    else
      if $G_{mask2} = 0$ AND $Intensity_G < Intensity_i$ **then**
        Set $Intensity_G = 100$
      end if
    end if
  end for
Require: $mask2$
  if $Intensity_{mask2} = 100$ **then**
    Set $Intensity_{mask2} = 255$
  end if

Second, the remaining particles in filtered $mask2$ are regarded as seeds for region growing as shown in Fig. 4.3(d). These particles are candidate seeds for nuclei segmentation that need to be further classified by the following process, see Algorithm 2. This region growing process that is based on $mask1$ aims to retrieve missing nuclei that are falsely eliminated by local thresholding from $mask2$ due to their weakly stained colour intensity. This is similar to a watershed algorithm, which floods water from catchment basins (the seeds) to the dams (the separation lines) where the waters coming from different basins meet each other. Unlike a classical watershed, cutting "necks" between nuclei are added instead of creating separation dams due to previous over-segmentation. The "necks" are generated in favour of measuring the intensity of dam pixels,
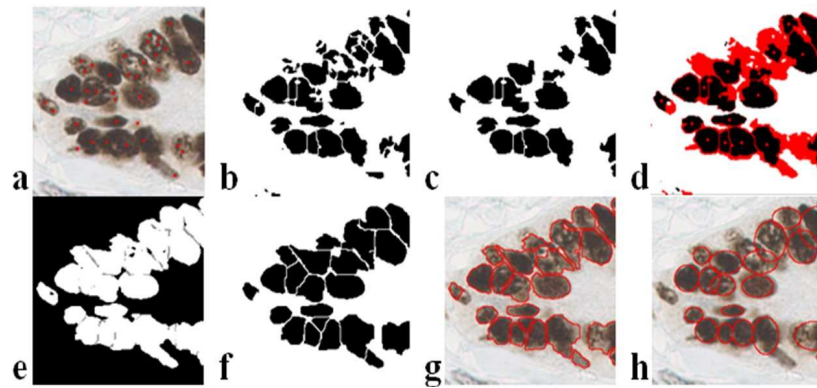
FIGURE 4.3: Region growing and nuclei segmentation. (a) original image, red points in (a) are initial seeds obtained from UEP. (b) the combined local and global thresholded results. (c) small particles in (b) are removed according to their mean nuclei size and minimum size constraint. (d) seeds used for region growing are highlighted in the center of each particle, and the surrounded red pixels are the pixels that will be grown. (e) the neck pixels are set to gray during the region growing process. (f) final segmentation results of the watershed. (g) the contour of each nucleus based on (f). (h) ellipses approximate final segmentation results.

whose intensity is larger than the mean intensity of the touched nuclei. Each growing nuclear particle is recorded and the "necks" are formed by the gray pixels, see Fig. 4.3(e). Adding the "necks" offers the watershed the ability to separate clusters and the potential to merge over-segmented nuclei as well. Falsely separated nuclei can be merged if the dam pixels have an intensity less than or equal to the average intensity of the growing nuclei that flooded from different catchment basins, see Fig. 4.4(a)-(d). The pixels belonging to "necks" have a high potential to be background and should be eliminated after the growing process. Based on this neck-eliminated binary $mask2$, the final seeds can be extracted by finding regional minimal points in the EDM transformed image. The final segmentation results are then generated by the seed-controlled watershed, as illustrated in Fig. 4.3(f) & (g).

4.2.3 Post processing
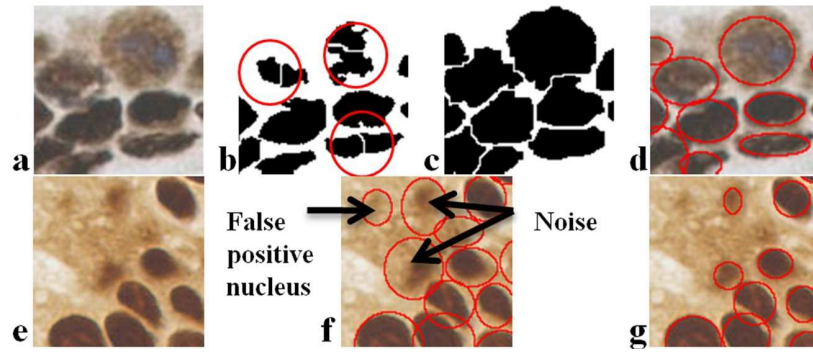
FIGURE 4.4: Post processing. (a) & (e) original image. (b) over-segmented nuclei in the initial segmentation. (c) merged nuclei after region growing process. (d) final segmentation. (f) falsely identified nuclei before post-processing. (g) false positive nuclei are eliminated after post-processing. The falsely splited nuclei before region growing are highlighted with red circles.

Algorithm 3 Post-processing. S_n is the size of nth nuclei after auto-threshold.

Require: $mask2$
  $P_i \in N_i$
  for $i = 1$ to n **do**
    if $G_{mask2} = 255$ AND $G_{mask1} = 0$ **then**
      Auto-threshold P_i
      Calculate S_i
      if $S_i < \beta$ **then**
        Fitting ellipse on N_i
      end if
    end if
  end for

As shown in Fig. 4.4(f), the segmentation results contain noise pixels inside some particles. These noise pixels are usually stained cytoplasm pixels surrounding the nuclei. Therefore, a post-processing step to eliminate the noise is needed, see Algorithm 3. To achieve this, each particle is examined separately and an intensity histogram is generated for those pixels inside the particle. An auto threshold method [43] is performed on the local intensity histogram. Pixels in each particle are considered to be nuclear pixels when their intensity is less than the local threshold. Particles are recorded with location and size information. The size of the particle is renewed after noise pixels are wiped off. Smaller particles are thresholded and removed as noise. Finally, an ellipse is fitted to the

remaining nuclei particles.

	AS	CD	OS	US	Miss	FP	AR	OR
Global [43]	904	664	13	207	381	20	51.5%	52.5%
E-min [105]	925	787	30	78	370	30	61.6%	62.2%
Voting [1]	810	439	74	214	538	83	38.4%	34.9%
LoG [2]	1641	837	338	44	46	422	64.4%	66.2%
Ours	**1213**	**1015**	**53**	**75**	**122**	**71**	**81.3%**	**80.3%**

TABLE 4.1: MEASUREMENT OF SEGMENTATION RESULTS AGAINST MANUALLY LABELED GROUND TRUTH. AS=Auto Segmentation, the number of segmented nuclei result. CD=Correct Detection. OS=Over Segmentation. US=Under Segmentation. Miss=Miss segmented nuclei. FP=False Positive. AR=Average accuracy Rate, the average correct detection rate. OR=Overall accuracy Rate, OR=CD/Ground truth.

4.3 Experimental Results for Nuclei Segmentation

The DAB stained sample dataset included 52 images of 200×200 pixels. These images contained areas of heavily clustered nuclei cut from 14 TMA images and 4 WS of colorectal cancer. I manually labelled 1265 nuclei in this dataset. The parameters used in this thesis are: the window size is 50×50 pixels, the moving distance is 25 pixels, and the minimum size constraint is 200 pixels. These parameters are empirically decided; although a systemic method to obtain these would be highly desirable, it is still a very challenging task (all data and the ImageJ plugin of our method are available at at the plugin website of ImageJ, http://rsb.info.nih.gov/ij/plugins/ihc-toolbox/index.html). This approach has been compared with a classical watershed performed on global thresholded [43] images and extended minima transformed images (E-min) [106]. In addition, the iterative voting method presented in [1] was also tested on the same dataset using their publicly available software. The method developed by Al-Kofahi et. al. [2], which has been applied as a toolkit in FARSIGHT [107], has also been compared. To measure performance, similar evaluation criteria to Al-Kofahi et al. [2], were adopted in the experiment including correct detection,
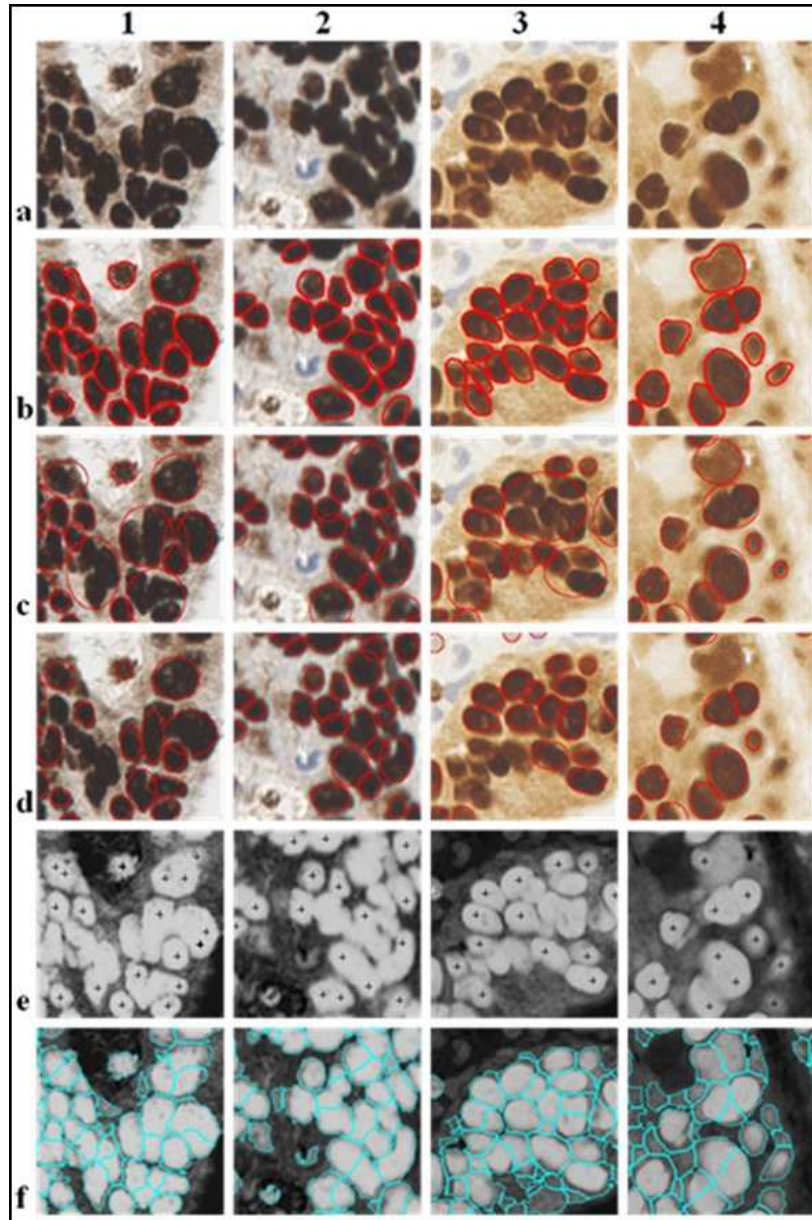
FIGURE 4.5: Segmentation results. In row (a) original images from 1-4. (b) manually labeled ground truth. (c) watershed on global threshold. (d) watershed on extended minima transform. (e)iterative voting. (f) LoG.

over-segmentation, under-segmentation, missing nuclei and false positives. Correct detection means the number of correctly segmented nuclei. The number of missed nuclei is calculated as the number of nuclei included in both the under-segmented areas and the foreground areas. The number of false positive nuclei is calculated as the number of false positive nuclei in over-segmented areas. The
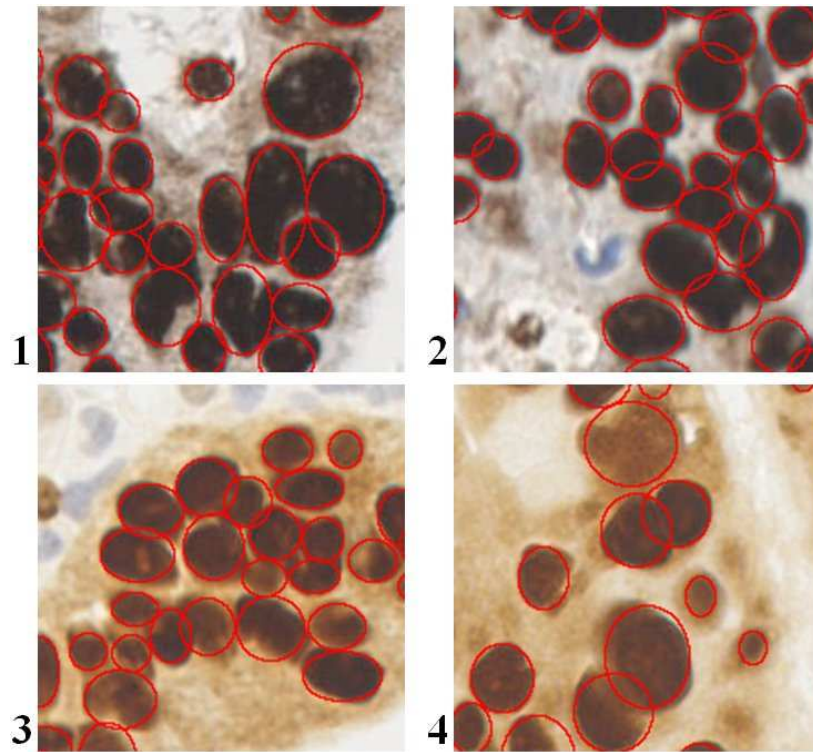
FIGURE 4.6: Segmentation results. (our proposed method.)

results generated by the proposed approach together with those obtained by the previous methods are compared in Table 4.1.

From Table 4.1 it can be seen that using a global threshold generates more under-segmentation and less correct detection. The extended minima transform reduced the under segmentation; however, many unevenly stained nuclear particles were falsely eliminated and the correct detection number was still low. The iterative voting method locates the centres of nuclei based on the border and gradient values of each nucleus. The weak borders between nuclei in the clustered areas made this method susceptible to under-segmentation. The method used in [2] led to a higher over-segmentation rate than the others due to the highly textured intra-chromatin of the nucleus and variations in nuclear structures. The method introduced here, which balanced over- and under-segmentation, significantly increased correct detection and reduced missed detection. Some segmentation results are shown in Fig. 4.5and Fig. 4.6.

4.4 DICE Coefficient Evaluation

Another evaluation applied was the Dice coefficient. This measures the overlapped region between segmented and ground truth against their union. It is commonly used for evaluation of segmentation techniques. It is defined as:

$$D(X,Y) = 2 \times \frac{|X \cap Y|}{|X| + |Y|} \tag{4.2}$$

Where X and Y are the two segmented regions; one is generated by method and the other one is ground truth (see the illustration in Fig. 4.7). The same dataset as presented in Section 4.3 has been used. In the evaluation, each test image includes a single labelled nucleus and is accompanied by the corresponding ground truth images (see Fig. 4.7). Therefore, I have evaluated 1,335 testing images against the same number of ground truth images. Two segmented and labelled nuclei should be compared depending on the size of overlapped regions. The segmentation results are obtained before the post-processing (see Fig. 4.3(g)). The post-processing is used to remove noise pixels, which may also remove the pixels that belong to the nucleus. A broken nucleus is not suitable for segmentation evaluation.

The Dice coefficient evaluation results are shown in Fig. 7. The number of segmented nuclei with a Dice coefficient higher than 0.5 is 718, and their mean value is 0.73. The number of nuclei with a Dice coefficient higher than 0.8 is 198, and their mean value is 0.84. Although the Dice coefficient is not good enough, the accurate contouring of the nuclei in such a complex background and clustered area is quite a challenge.

FIGURE 4.7: The illustration of the Dice coefficient evaluation. The left image is ground truth, the middle image is the segmentation result and the right image is the evaluation result. The grey labelled nucleus in the left and middle image is the evaluated one. The grey labelled pixels in right image denote the overlapped pixels between the ground truth and the segmentation result. The white pixels together with grey pixels in the right image denote the union pixels. Therefore, the Dice coefficient is the grey pixels in the right image, divided by the union pixels.
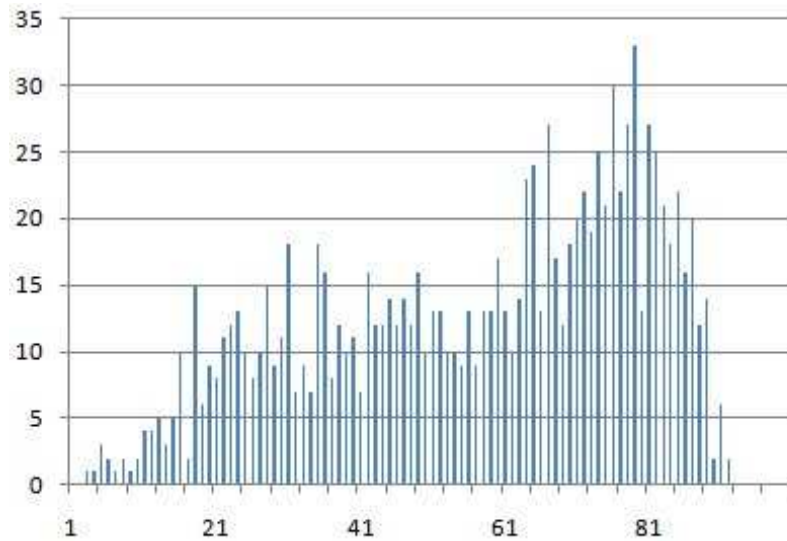
FIGURE 4.8: The Dice coefficient evaluation results. The horizontal axis is the Dice coefficient, and the vertical axis is the number of nuclei.

4.5 Parameter Estimation

Three parameters were engaged in the developed watershed-based nuclei segmentation approach, including the size of the local thresholding window, the moving distance of the window, and the minimum constraint on seed size. The moving distance of the local thresholding window, which only affects the time cost to some extent, was not evaluated in this estimation process and set to be a

constant value per movement. In this section, the other two parameters are discussed and tested with different values on the same dataset used in Section 4.3.
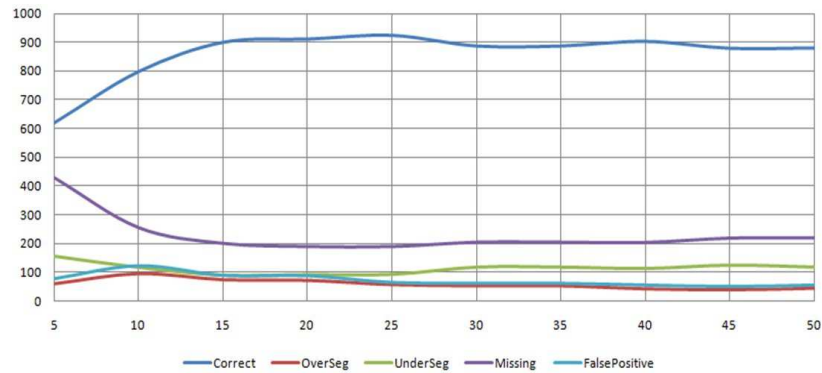
FIGURE 4.9: Plot of segmentation results with different window sizes. The horizontal axis is the changing size, and the vertical axis is the number of nuclei.
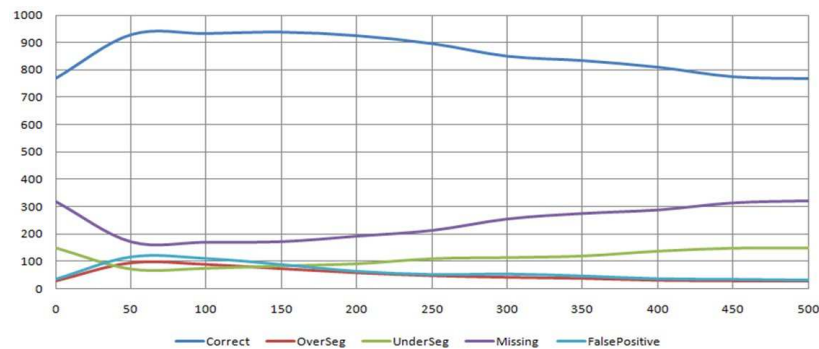
FIGURE 4.10: Plot of segmentation results with different seed sizes. The horizontal axis is the changing size, and the vertical axis is the number of nuclei.

In estimating a parameter, the other two parameters were set to constant values. For example, the local thresholding window size was the first estimated parameter and the seed size constraint was set to 200 pixels. The tested half window sizes included 10 values from 5 to 50 pixels. Then the best performing window size was obtained and set to be constant for the estimation of seed size constraint. The tested seed size constraints ranged from 0 to 500 pixels. Each tested

value yielded a segmentation result that includes all the criteria used in the evaluation process in Section 4.3. The results are separately plotted in Fig. 4.9and Fig. 4.10.

Exploring these two plots, it can be seen that the generated results have small variations in correct segmentation. The increasing of size, either window size or seed size, may lead to an increase in the number of under-segmented nuclei and a decrease in the number of over-segmented nuclei. For the window size parameter, the best result was obtained at 25 pixels (half size of the window), which is the mean size of nuclei (see example in Fig. 4.11). The results showed little variation when the values were larger than or close to the size of the nuclei, such as greater than 25 pixels but less than 40 pixels. However, the correct segmentation number decreased when the window size was larger than the size of two nuclei, for example 40 pixels, due to the intensity variations between two nuclei.
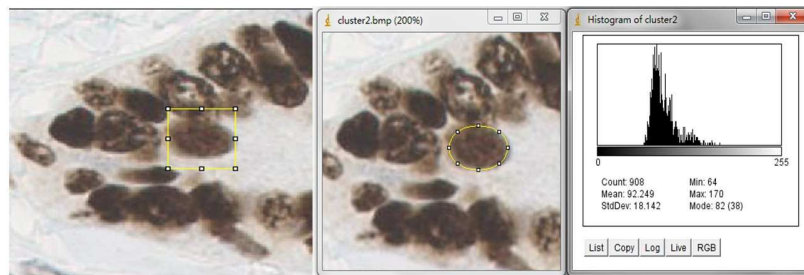
FIGURE 4.11: The square size and total size of nuclei on test image. The image on the left is fitted by a 44×44 square, and the image on the right is fitted by a ellipse which have 908 pixels inside.

The parameter of seed size constraint gave the best result at 150 pixels as compared to the size of the nuclei, which varied from 300 pixels to 1000 pixels (see example in Fig. 4.11). Sizes larger or smaller than this value around 100 pixels had similar correct segmentation results. This constraint value is used to prevent over-elimination of small particles that are generated by the initial separation. Thus, the values should be less than a fifth the size of the nuclei. Both these

tests indicated it is possible to select a value around close to the optimal value in order to obtain satisfactory results.

4.6 Quantification Evaluation

The developed nuclei segmentation approach, which can clearly detect individual nuclei in areas of clustered or overlapping nuclei, is the a priori step in nuclei quantification. In nuclei segmentation, the over- and under-segmentation problems are significantly controlled and the background noises are greatly reduced. Quantification can then be applied once accuracy in automatically segmenting nucleus is achieved.

Nuclei are quantified and proportionately measured in IHC stained images, which offers the potential of applying a growing number of biomedical applications [108]. Positively stained nuclei that have high potential disease association are quantified in this evaluation process. To evaluate the accuracy of positive nuclei quantification, the quantification process was divided into the following two steps. First, the positive stain colour, for example the brown colour should be detected. A statistical colour detection model (see Chapter 3) was applied to detect all positively stained pixels in the image. Each pixel in the image was classified and recorded as either positive or negative. Second, the nuclei were segmented and the positively stained nuclei were classified and quantified. The segmentation process was performed by using the above-mentioned seeded-based approach. Positive nuclei were then classified by measuring the brown pixels in each nucleus that had a proportion larger than 20%, as a proportionate value larger than 20% could largely minimise the counting of blue nuclei. Each positive nucleus was fitted by a red ellipse (see Fig. 4.13).

One potential use of the quantification of positive nuclei is the grading of cores in TMA images. Instead of manually counting, the automatic quantification

can reduce the time cost and generate objective results. Rather than considering about the F1-score evaluation, it is better to evaluate the correspondence between the automatically generated results and manual counting results. This is because the annotated ground truth for this dataset contains only the counted number of positive nuclei in each image (see Fig. 4.13), and it is impossible to annotated such a huge number of nuclei in such a complex background.

The datasets was used to evaluate nuclei quantification was also the DAB stained colon tissue samples. The proposed new dataset contained 11,826 manually counted positive nuclei in 33 images with resolution 1280×1024. These images had a complex background and areas of severely clustered nuclei similar to the dataset used in the segmentation evaluation (see Fig. 4.13). The evaluation was performed by measuring correlations between the tool-generated results and manual counting results. The ground truth of positive nuclei quantification was manually counted by two persons (myself and a pathologist) who had experience in histopathology. The correspondence between the human evaluators and the computer is presented in Fig. 4.12. The parameters used for this quantification evaluation were those that obtained the best results in Section 4.5. Compared to the previous methods [1, 2], the proposed method here performed much better, with higher correspondence.

4.7   Discussion and Conclusion

Quantification of positive nuclei is routinely used in diagnostic pathology to determine individual therapeutic strategies. In order to reduce the subjective bias in quantification, an objective method is of the utmost importance. Nuclei segmentation is regarded to be the prerequisite for nuclei quantification. The problems that arise in segmentation are under- and over-segmentation problem. In this chapter, an automated method for segmenting overlapping or clustered nuclei

in colorectal cancer has been proposed, which reduces both of these two problems. Experiments on a large dataset show that the new approach works very well and outperforms several popular techniques in the splitting of overlapping and clustered nuclei.

The first step of nuclei segmentation is the nuclei area extraction, the separating of nuclei pixels from the background. Several approaches have been presented in the literature in Section 2.2. Since the threshold-based method is the simplest and cheapest one, the auto-threshold algorithm has been used to obtain the threshold value automatically and reduce the time cost. To cover the intensity variations, the nuclei pixels were extracted from the complex background based on the histograms both from the whole image and from a moving sub window. The variations of segmentation accuracy according to the sizes of this sub window have been evaluated. The selection of an appropriate size is depending on the size of the nuclei, which has been shown in Section 4.5.

Accurate seed detection is the key to accurate nuclei segmentation, especially for the splitting of overlapping or clustered nuclei. The proposed novel seed detection method involves the nuclear characteristics of shape, size and intensity. It is intended to reduce the watershed algorithm over-segmentation problem in nuclei segmentation. The seeds were obtained after the removal of false initial seeds with size filtering and the refinement with intensity based region growing. The size filtering is controlled by two parameters: one is the automatically calculated mean nuclei size, and the other one is determined manually. This human-determined parameter is regarded as a complement to cover the nuclei size variations and reduce the over-segmentation. Region growing with intensity measurement is also intended to reduce this problem, by merging the over-segmented nuclei. In addition, the seeds were obtained from EDM, which may be sensitive to the neck size between nuclei; thus, the region growing method can also reduce the under-segmentation problem by shrinking the neck size.

However, while this method was proposed for splitting the overlapping or clustered nuclei and produce quantification results based on the novel seed detection method, it may not provide satisfactory contour of split nuclei due to the complex background and the large shape variations of nuclei. Thus, the Dice coefficient evaluation is not good enough. In addition, the contouring results used were not post-processed, which also gives raise to the low Dice coefficient.

As with similar techniques, there are several parameters that need to be determined empirically. However, these parameters can be determined according to the size of the nuclei in an image with specific magnification and resolution. It was shown that using the best performing parameters on this dataset yielded much better results on another dataset with the image captured under the same magnification. The dataset used for the evaluation of quantification was quite complex and difficult to label visually. The preparation of the ground truth with visually labelling was time-intensive and lasted for two weeks. As a result, the image analysis approach, which only takes several minutes, can attain highly correlated results as compared to this ground truth.
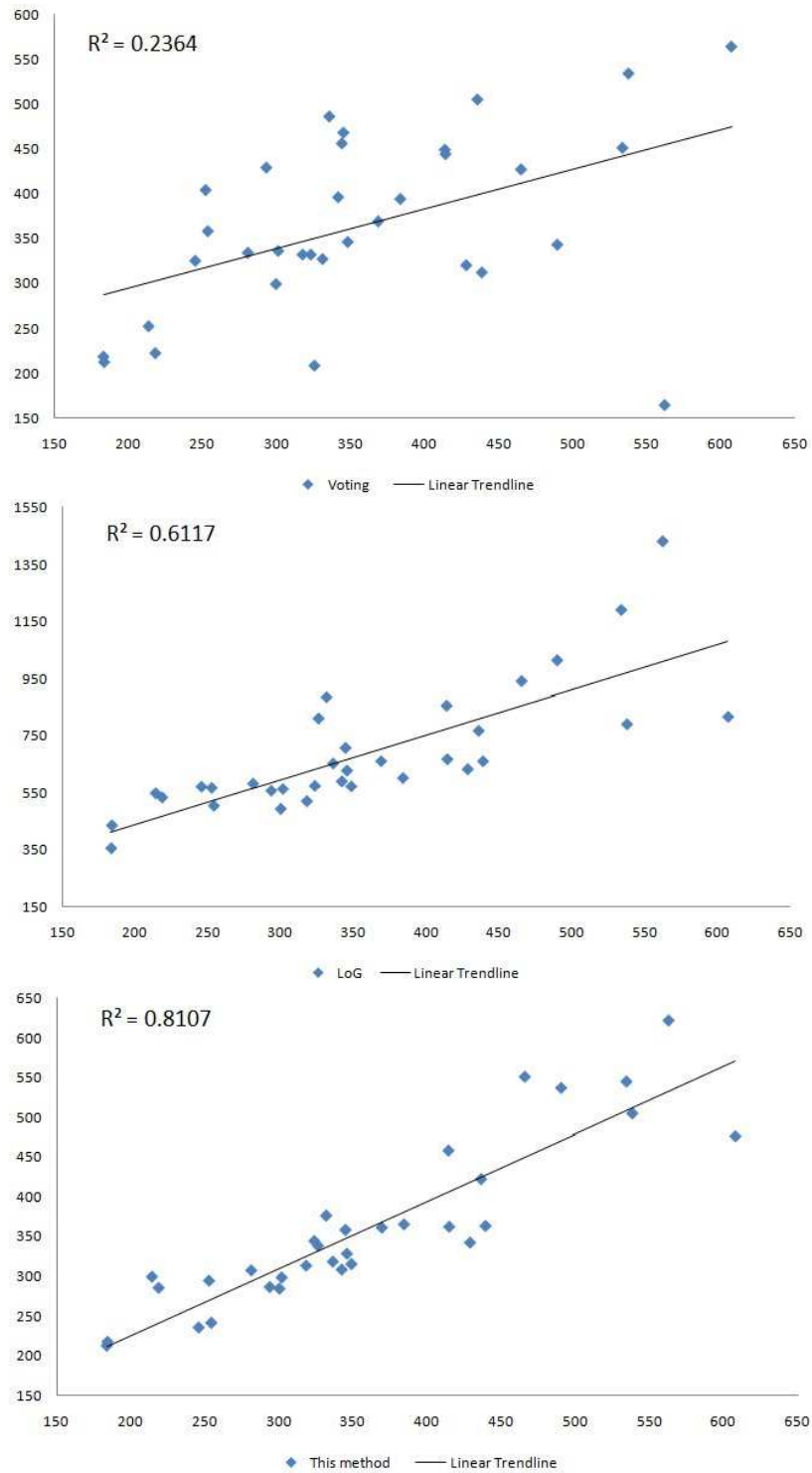
FIGURE 4.12: The quantification evaluation results. Scatter plot of quantification results between manually counting and automatic segmentation methods, Voting [1], LoG [2] and method presented in this chapter. The horizontal axis shows the manually counted number of nuclei; the vertical axis is the automatically counted number of nuclei.
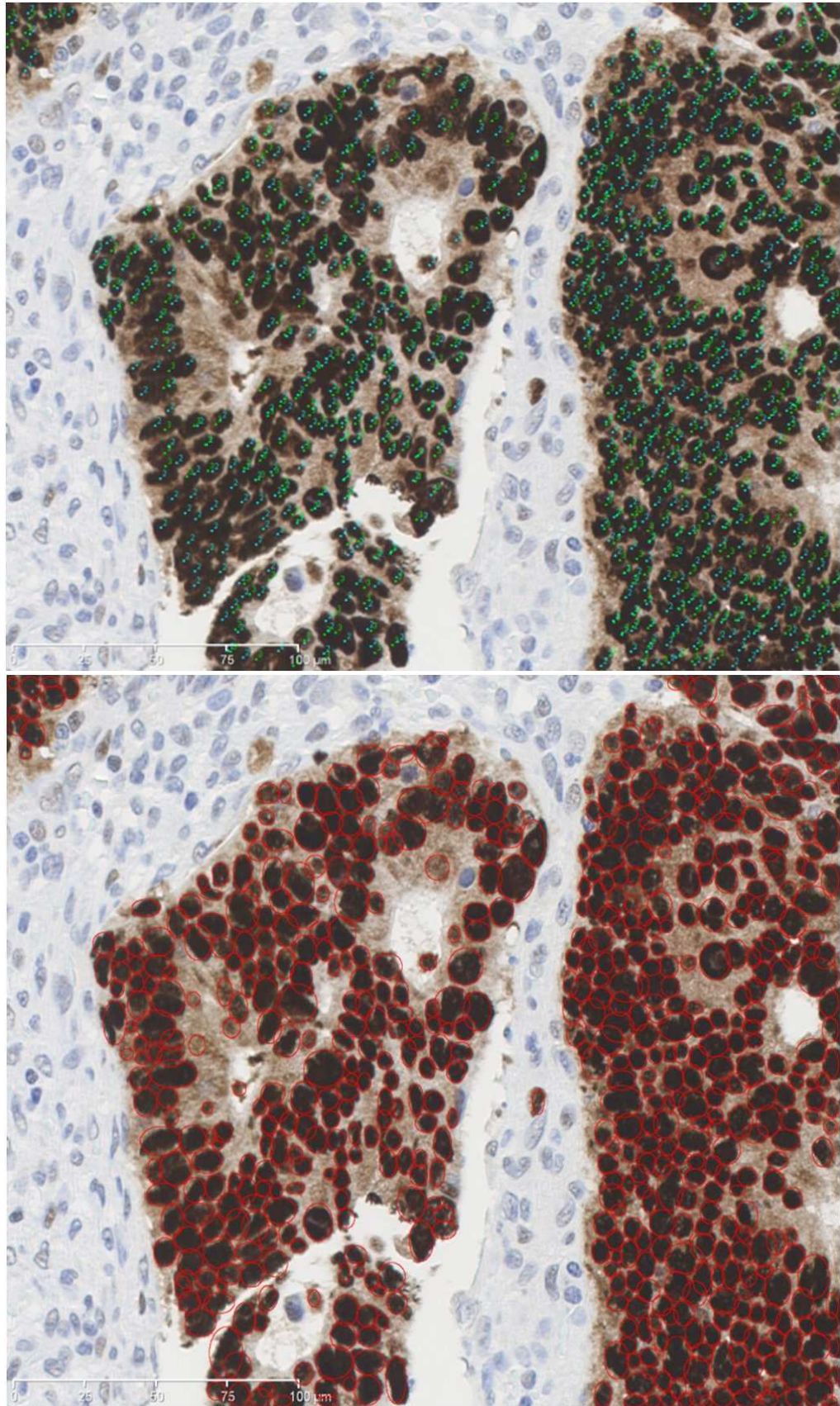
FIGURE 4.13: Quantitative evaluation of samples and manual counting of the ground truth. The upper image is the ground truth image which was manually marked by two persons using the "cell counter" plugin in ImageJ. The lower image is the nuclei quantification by our method.

Chapter 5

Gland Detection

Biopsies are prepared for further examination when cancer is suspected. The stained samples are explored and interpreted by the pathologist to verify the malignancy. The analysis of gland architecture, which reflects the cancer stage, has evolved into an important aspect of cancer detection [3]. Computerized gland detection is a recently developed area of study that has been researched for no more than 10 years. These recent studies have encouraged pathologists to use computer-aided methods in the cancer examination process. In addition, these objective methods may reduce visual interpretation problems. In this chapter, a new gland segmentation method is introduced.

5.1 Introduction

Most previous papers have segmented glands in Hematoxylin and Eosin (H&E) colour stained images. In H&E stained images, the components of a gland are coloured with distinguishable colours; for example, nuclei are blue, stroma is pink, cytoplasm is purple and the lumen is white [90]. Therefore, colour is considered as an important feature in gland segmentation and is normally detected

or classified into different clusters at the start of the segmentation. Unfortunately, this is not suitable with Hematoxylin-Diaminobenzidine H-DAB stained tissue images, as it only colours nuclei into two classes, cancerous and normal. No differentiable stain applied on cytoplasm and stroma. Therefore, the previous colour dependent gland detection methods might not work on this kind of stained image, see Fig. 5.1.

Accordingly, texture or feature-based methods are more flexible and can be applied on different kinds of colour stained tissue images. A novel morphology-based approach has been developed to segment gland structures in H-DAB stained images that can tackle the challenges mentioned above. This approach can be performed on grey scale images; thus, it is also applicable to H&E stained images. Novel methods are proposed to extract potential luminal regions and their enclosing chain of nuclei. Complete glands are then formed by growing each luminal region until it reaches its relative nuclear boundary. Finally, a support vector regressor is utilized to judge whether the proposed gland is true or not. Applying this methodology on a publicly available dataset has shown that this approach outperforms the state-of-the-art.

5.2 Method

The lumen is the central region of a gland, and has been demonstrated to be useful in locating glands in many previous papers [3, 4, 109]. Thus, the gland detection in H-DAB stained images is performed via the following three steps. First, the lumen-like regions are extracted. Second, each nucleus is assigned to its correlated luminal region and bounding boxes are fitted. Third, each of these predicted candidate glands is classified as a gland or non-gland by a Support Vector Regressor [92].
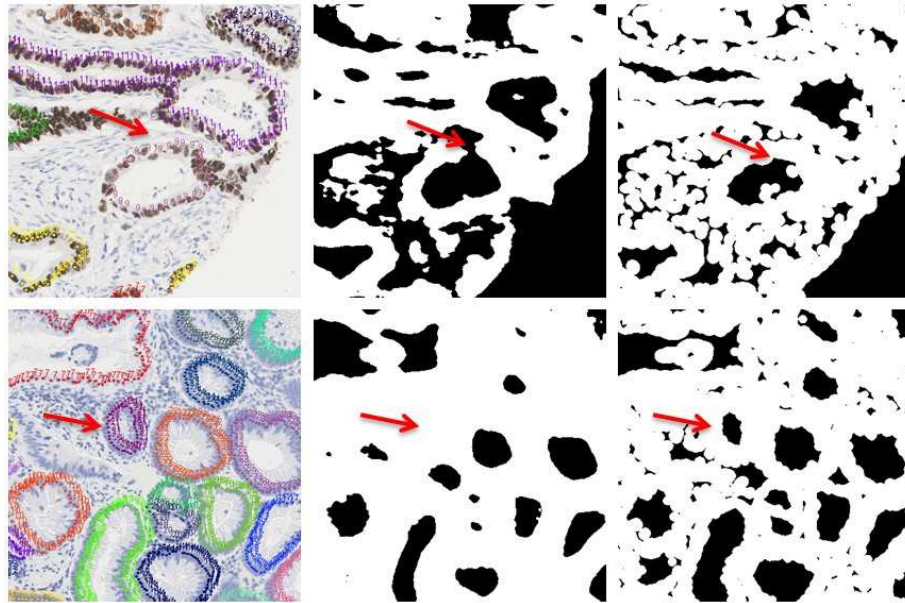
FIGURE 5.1: The candidate lumen regions detection. The left column is the manually labelled ground truth, the middle column is the lumen regions detected by Farjam et al. [3], and the right column is the lumen regions detected by the proposed method. The dark pixels in the detected images are lumen regions and the white pixels are the background. The red arrows in the top row point to the vacant artefact at the gland boundary, and the red arrows in the bottom row point to the falsely removed lumen which has a small size with closed-chain of boundary nuclei.

5.2.1   Luminal region extraction

Just as mentioned before, one of the challenges in gland detection is the inconsistent luminal colour. Thus, it is better to consider morphological information rather than colour information to detect the luminal regions. Morphologically, the nuclei are distributed like a closed chain and the lumen is the region inside this chain. The detection of the lumen is then equivalent to detecting the closed chain of nuclei. However, many glands may have unclosed nuclei chains due to artifacts, such as vacant areas at the boundary of a gland or aggregates that challenge gland detection. Thus, the luminal regions should be extracted from a new feature space instead of closing unclosed chains of nuclei.

As shown in Fig. 5.2(b) & (d), visually, the intensity of pixels inside the gland
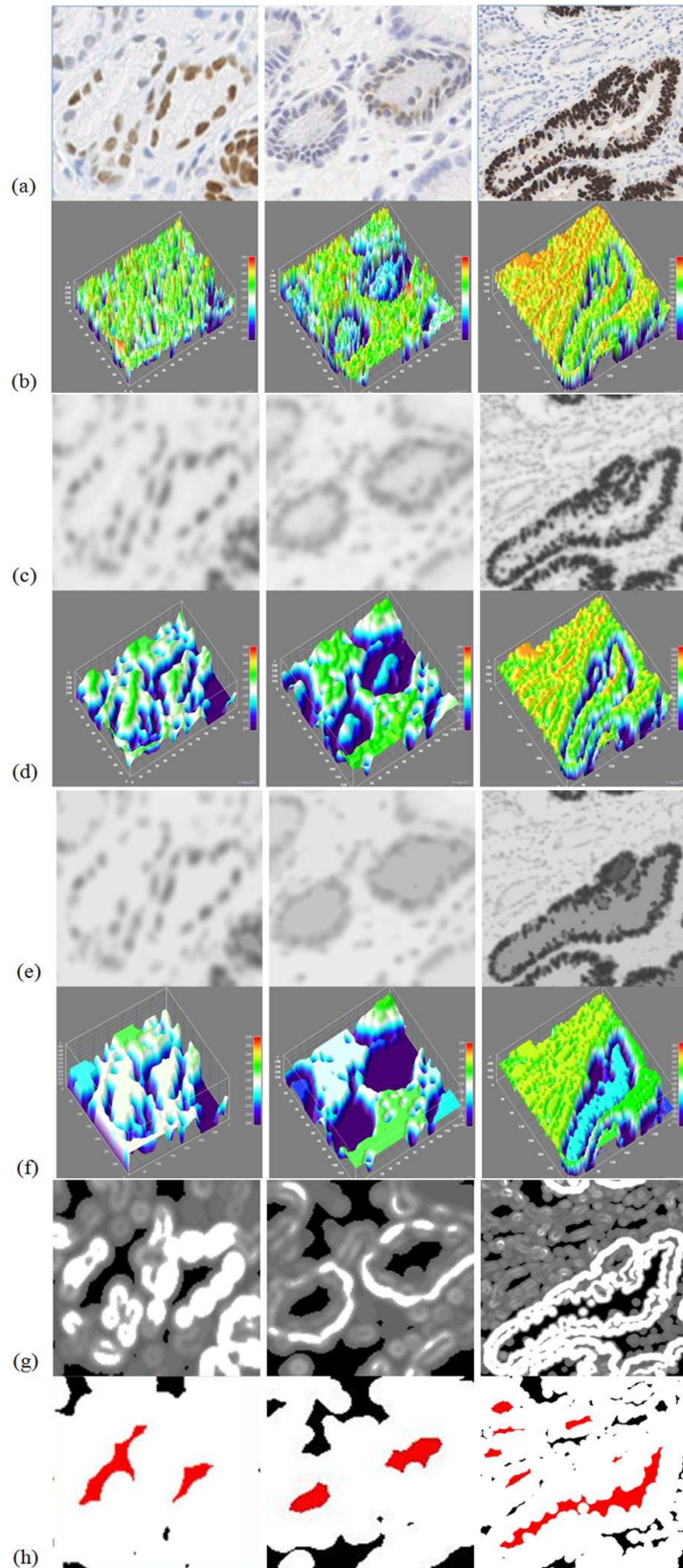
FIGURE 5.2: Steps in candidate luminal region extraction on H-DAB images. (a) original sample images; (b) the 3D surface plot of (a); (c) the Gaussian blurred image; (d) the 3D surface plot of (c); (e) the open-by-reconstruction applied on (c); (f) the 3D surface plot of (e); (g) variance feature space of (e); to make them visible, pixels with non-zero variance values are enhanced; (h) binary image after a simple threshold applied on (g); red pixels are true luminal regions and black pixels are false luminal regions.

is much higher than in the surrounding nuclei and appears like a hill or peak, while the boundary pixels surrounding the gland appear like a valley and the pixels in the stroma appear like ridges. The valley occurs between the hill and the ridges and exhibits a sharp edge. This change in acute intensity is a distinct feature of glands. Utilizing the higher intensity in the lumen and this sharp edge feature, Farjam et al. [3] detected the lumen and nuclei boundary separately. They applied a feature extractor on these two areas to enhance the sharp edge and close the nuclei chain. The enhanced the sharp edge in the extracted nuclei boundary and then used the closed nuclei chain to detect the lumen. However, it is hard to close the boundary in the presence of large vacant artifacts and small glands may be missed, see Fig. 5.1.

From another point of view, the sharp edge pixels at the boundaries of glands have larger variations than the other structures. Based on this phenomenon, it can be assumed that the luminal region is surrounded by pixels having a larger variation. This means the gland boundary, whether a closed or partially closed chain of nuclei, would form a closed sharp edge chain. Accordingly, variance filter was adopted to highlight the edges and transform the original grey scale image into a variance based texture space.

$$Var(I) = \delta^2 = \frac{1}{n}(\sum_{i=1}^{n}(I_i^2)) - (\frac{1}{n}\sum_{i=1}^{n}I_i)^2, \{I|I \in C(r) = 1 \cdots N\} \quad (5.1)$$

where I is the image intensity at location $(x_i; y_i)$, C is the size of filtering window with radius r, and N is the number of pixels in this filtering window. A pixel that has a rapid intensity change in its neighbourhood will have a higher variance value. However, such rapid intensity changes occur across the whole image, even in colour-polluted lumenal regions. This makes it hard to distinguish genuine luminal regions. Instead of extracting the sharp edges from such a complex environment, a morphological process that combines a Gaussian blur

followed by an open-by-reconstruction method [96] to highlight the luminal regions was proposed. This process reduces the variations in luminal regions and makes the pixels in these regions significantly different from the others.

Morphological reconstruction can be conceptually regarded as repeated dilations of the seed image to fit the mask image. Open-by-reconstruction is a procedure of erosion followed by a morphological reconstruction that aims to preserve the shape of the remaining image components after erosion. In a binary image, it can erode the components with a pre-defined disk to filter out undesired objects. This can then be transferred to a grey scale image to smooth the regions with intensity changes based on a round disk. Using this morphological process, it can reduce the intensity of regional maximum pixels and smooth the intensity changes (see Fig. 5.2(d)). After this process, the variance of lighter pixels, which usually correspond to the luminal region, are smoothed to zero and the between the valley and the peak are reserved. The luminal regions can then be detected through a simple threshold operator. The detected luminal like regions are shown in Fig. 5.2(g) and the details for the whole gland detection process are shown in Algorithm 4, Algorithm 5, Algorithm 6.

5.2.2 Nuclei region detection

The result obtained from the previous section consisted of candidate luminal regions that might correspond to true glands or not. To differentiate true glands from the false ones, it is better to utilise the information in the surrounding nuclei chain. In H-DAB colour images, nuclei are coloured blue or brown; the brown nuclei have a much lower intensity than the blue ones. This makes it difficult to detect all the nuclei regions in such an image, as can be seen in Fig. 5.3(a). The extraction of nuclei regions, the pixels that belong to nuclei, by applying a colour-based method [27], K-means clustering method or Laplacian of Gaussian (LoG) method may all fail to detect blue nuclei due to their weak intensity (see

---

Algorithm 4 Candidate lumen extraction for gland detection. I_{pixel} is the intensity value of pixels in image I. Particles with size $S_{particle}$ smaller than the threshold α are removed. The reserved particles are recorded as candidate lumen L_i.

Require: DAB stained RGB image I
  Pre-processing of image I with Gaussian blur
  Apply Open-by-Reconstruction on pre-processed image I
  Apply Variance filter on reconstructed image I
  $I_{pixel} \in I$
  if $I_{pixel} > 0$ **then**
    Set $I_{pixel} = 255$
  end if
  $S_i \in I$
  for $i = 1$ to n **do**
    if $S_i < \alpha$ **then**
      Remove S_i
    else
      Record as L_i
    end if
  end for

Fig. 5.3(b) & (c) & (d)). It can be found that using a combination of histogram equalisation and Laplacian followed by the autothreshold method [43] can obtain surprisingly good results, as shown in Fig. 5.3(e). The detail of this process is shown in Algorithm 5. Although the LoG method obtained good result, the characteristics between nuclei are difficult to use in differentiating between the gland nuclei chains and the stroma nuclei chains. Thus, the pixels that belong to nuclei should be detected completely, which can benefit the differentiation. This method may also offer the ability to accurately locate the bounding box for each gland.

5.2.3 Candidate gland preparation

Overlapping or clustered nuclei are largely assembled at the gland boundary. However, stroma nuclei are clustered as well. Relying on clustered or assembled nuclei particles may result in falsely locating the gland and misidentifying

---

Algorithm 5 Nuclei particles detection for gland detection. Nuclei particles with size $S_{particle}$ smaller than threshold β are removed. The reserved nuclei particles are recorded as N_j.

Require: DAB stained RGB image I
  Convert RGB to gray scale image I_g
  Enhance I_g with histogram equalization
  Filter I_g with laplacian
  Auto-threshold processed I_g
  if $I_{pixel} > \beta$ **then**
    $I_{pixel} = 255$
  end if
  Watershed segmentation on I
  for $j = 1$ to n **do**
    if $S_j < \gamma$ **then**
      Remove S_j
    else
      Record as N_j
    end if
  end for

---

the gland boundary. After the nuclei regions were extracted, the watershed algorithm was applied to separate the connected regions into a set of small particles. This splitting operation can beneficial to locating the bounding box for each gland. Each nucleus was assigned a label and implemented iterative region growing process to connect most of them. The grown pixels from each nucleus were assigned the same label as the original nucleus. The number of growing iterations can be determined by the largest gap sizes between nuclei at the gland boundary, except in the case of large artifacts. This step connects each nucleus with its neighbours, which can then be used as a barrier to stop the lumen growing.

Next, the image with connected nuclei was treated as the mask image, and performed another seeded region growing for each previously detected luminal region. Two processes were included in this region growing, including the corresponding nuclei boundary detection and the merging of falsely split luminal regions. Following the region growing, the grown seeds, the luminal regions and
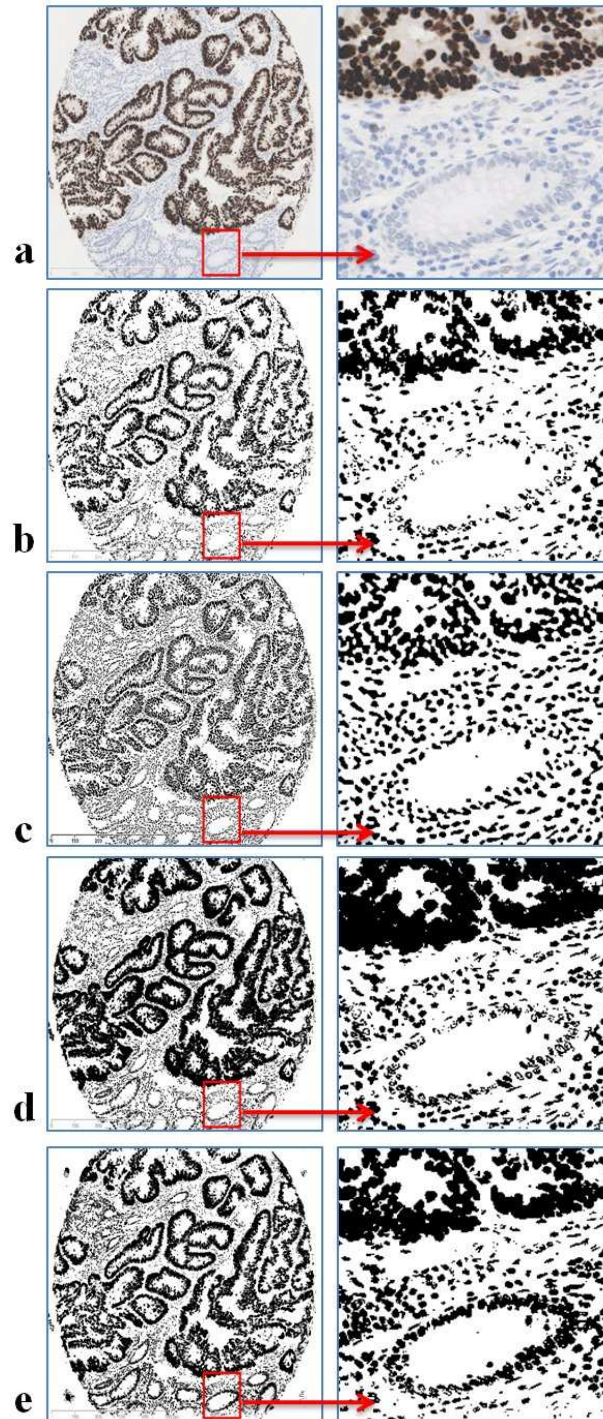
FIGURE 5.3: Nuclei region detection. (a) is the original image with high intensity variations in nuclei; (b) is the binary image obtained from colour deconvolution; (c) is the binary image processed by Laplacian of Gaussian; (d) is the binary image obtained from K-means with K=3; (e) is the binary image obtained by our proposed approach.

the correlated nuclei were jointly considered to be candidate glands. A bounding box was then fitted for each candidate gland, see the Algorithm 6.

Algorithm 6 Candidate gland preparation for gland detection. The grown pixels G_{pixel} of L_i are those pixels that not belong to any nuclei particles N_j. L_p are lumen that touched by the current grown lumen L_i. The nuclei particles N_j, covered by the enlarged L_i and L_p with size r, are recorded as belonging to a gland with lumen L_i.

Require: Seeds L_i
Require: Mask N_j
Require: Radius r
  for $j = 1$ to n **do**
    Grow N_j
    Calculate the centre C_j of N_j
  end for
  for $i = 1$ to m **do**
    while $G_{pixel} \notin N_j, j \in [1, n]$ **do**
      Grow L_i with G_{pixel}
      Record touched N_j by L_i
      Connect L_i with $L_p, p \in i$
    end while
    Calculate D_i between L_i and touched C_j
    Enlarge L_i and L_p with size r
    if Any C_j in r **then**
      Record N_j as touched by L_i
    end if
    Fitting bounding box for $L_i \cup N_j$
  end for

In region growing, the nuclei touched by each luminal region were recorded separately. These nuclei are regarded as the gland boundary nuclei for the touched luminal region. The size of the gland can also initially be obtained from the distance between the luminal region and the recorded nuclei. After this process, similar to [4], a ring-like outer domain was defined, aiming to simulate the chain of epithelial nuclei. The nuclei on the inner side of the gland may close and halt the growth of the seed lumen to touch the nuclei on the outer side (see Fig. 5.4). Therefore, the luminal region can only touch some of the nuclei at the boundary. To obtain the missing nuclei a ring-like region is proposed. The ring is fitted to the boundary of touched nuclei and extended to a pre-defined range according
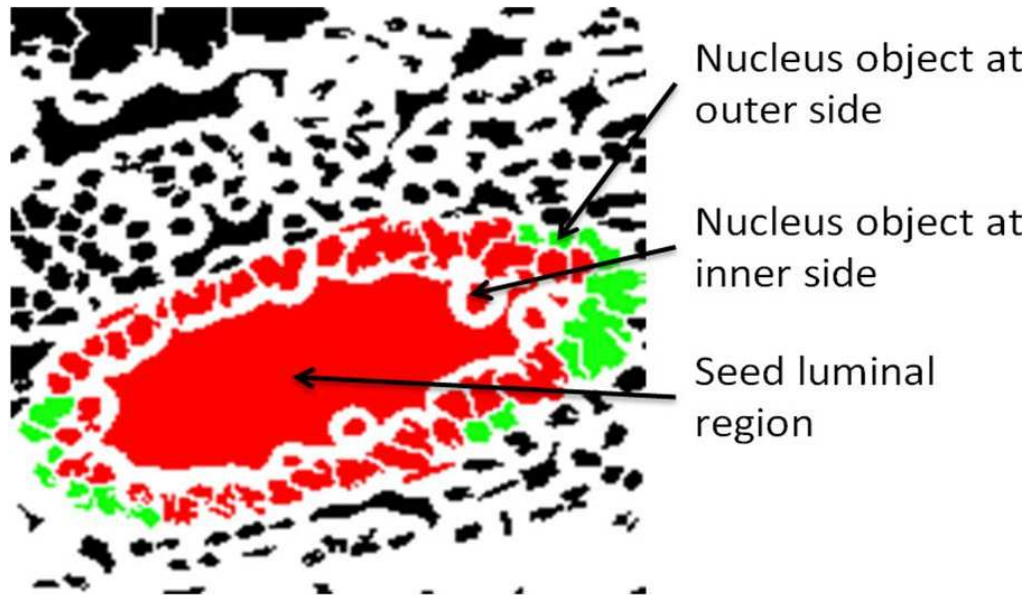
FIGURE 5.4: Detecting the gland nuclei boundary. The large red region is the seed luminal region, which grows to touch the red nuclei. The green nuclei are detected after applying the extended ring-like region.

to the thickness of the gland boundary. In addition to the nuclei reached by this grown seed, nuclei whose centres were located inside this domain were also allocated to this seed.

The luminal regions may be falsely split into several small regions during the lumen extraction step due to variations in gland shapes and the inner distances of boundary nuclei. Separate luminal regions may also be merged when they touch each other during region growing. The nuclei touched by each seed lumen region are then combined to form the boundary of the merged lumen.

5.2.4 Gland classification

The possible glands were contoured by minimum bounding boxes. The bounding boxes we fitted to the candidate glands were depending on the distance between the centre of boundary nuclei and the edge of detected lumen. To verify whether these proposed candidate glands are true glands or not, following [92],

we designed another visual feature-based verification module to decide if the boxes detected correspond to the true glands or not. This classification process is illustrated in Fig. 5.5.
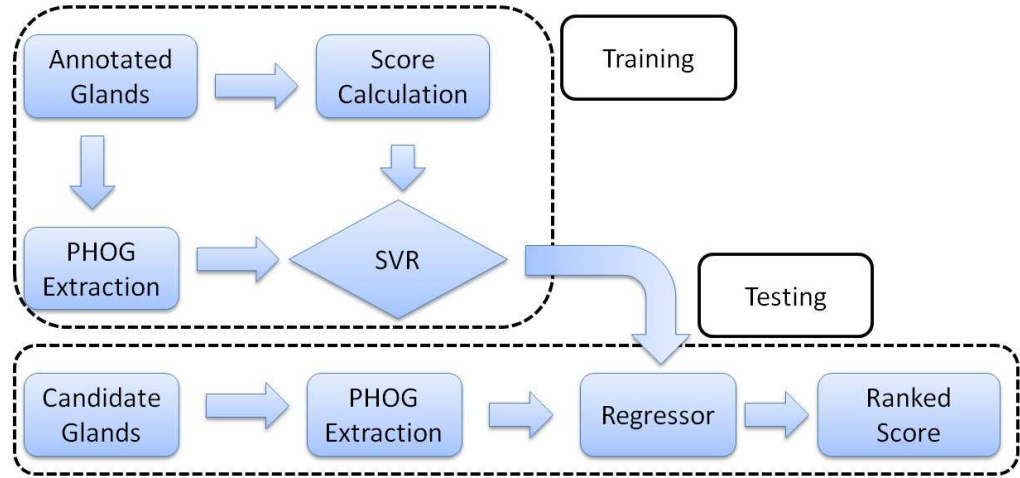
FIGURE 5.5: The gland classification scheme.

For each of these bounding boxes, the Pyramid of Histograms of Orientation Gradients (PHOG) visual feature [98] was extracted from the smallest bounding box enclosing it. The PHOG feature was chosen owing to its efficiency in capturing the shape structure, which is considered to be the most discriminative feature of the gland [92]. A score was calculated to measure the quality of the contour. This score is Dice coefficient, which is defined as the maximum overlap between the detected contour S and any ground-truth glands S_i:

$$score = max_i \frac{|S \cap S_i|}{|S \cup S_i|} \tag{5.2}$$

Then a regressor was trained to regress this score using the PHOG features and utilize the popular LIBSVM toolbox [110] to learn a nonlinear Support Vector Regression (SVR). The output of this SVR is considered to be a probability representing how likely it is a bounding box contains true gland. It is then used to rank all the bounding boxes.

5.3   Experimental Results

The experiments were conducted on the proposed dataset containing 20 high resolution 1280×1024 pixel H-DAB stained microscopic images of human colon tissues. The dataset consists of tissues with preinvasive (i.e. adenomatous) tumours and well or moderately differentiated invasive tumours. In these situations, the glandular architecture is preserved to varying degrees and thus heterogeneous morphologies can be interrogated. The poorly differentiated tumours, and these - by their very definition - have lost their glandular architecture were not involved in the test. The images we used come from different tumours (i.e. different patients). They were captured on a scanner (Hamamatsu nanozoomer) with $10\times$ standard mode and on a digital camera attached to a microscope (Olympus). The staining was done at a variety of different times.

All the cases selected for this project have been reviewed by a fully trained diagnostic pathologist whose task is to confirm the presence of the glands in every tissue section. Then all the glands are annotated and further confirmed by the pathologist. In total the dataset contains 1,072 glands, all of which have been manually annotated by me with pixel accuracy. Fig. 5.7 shows some images in this dataset, where the glands are manually labeled by the "cell counter" plugin in ImageJ. The whole dataset and the ground-truth data have now been made publicly available at http://www.viplab.cs.nott.ac.uk/download/Nott-Gland.html.

In the experiments, these 20 images contained in the dataset were split into two disjoint sets. Each image was assigned a number ranging from 1 to 20. While the odd-numbered images were used for training, the even-numbered images constituted the testing set. The training set was used to train the SVR model. Next, the testing set was then tested by the trained SVR model.

This task was considered as a gland detection problem. A predicted bounding box was treated as a true positive if its overlap with any ground truth gland
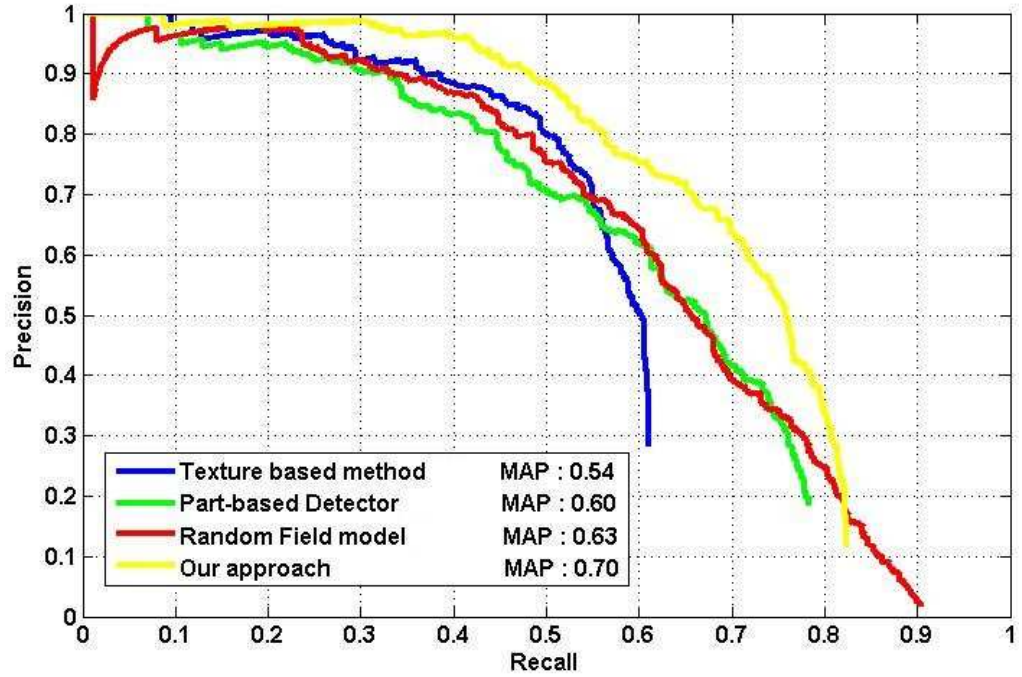
FIGURE 5.6: The comparison of experimental results obtained by different methods. The results are plotted in a Precision-Recall graph.

exceeded 0.5. For the assessing of experimental results, the ROC curve was not considered due to the ground truth contains only the annotated true glands. Thus, it is impossible to calculate the false positive ratio without the annotated false glands in ground truth. Mean Average Precision (MAP) [92], which is a widely adopted criterion in object detection literature, was utilized as the performance measure (see Equation 5.3).

$$Precision = \frac{TP}{TD}, Recall = \frac{TP}{GP}, MAP = \frac{\sum_{q=1}^{Q} AveP(q)}{Q} \qquad (5.3)$$

where TP is the number of detected true positive glands, TD is the total number of detected glands, GP is the number of labeled glands in ground truth, Q is the number of tested images, and $P(q)$ is the precision of Qth image. The proposed method generated results together with those previously reported, such as Texture based method [3], Random Field model [92], and Part-based Detector [92] are shown in Fig. 5.6.
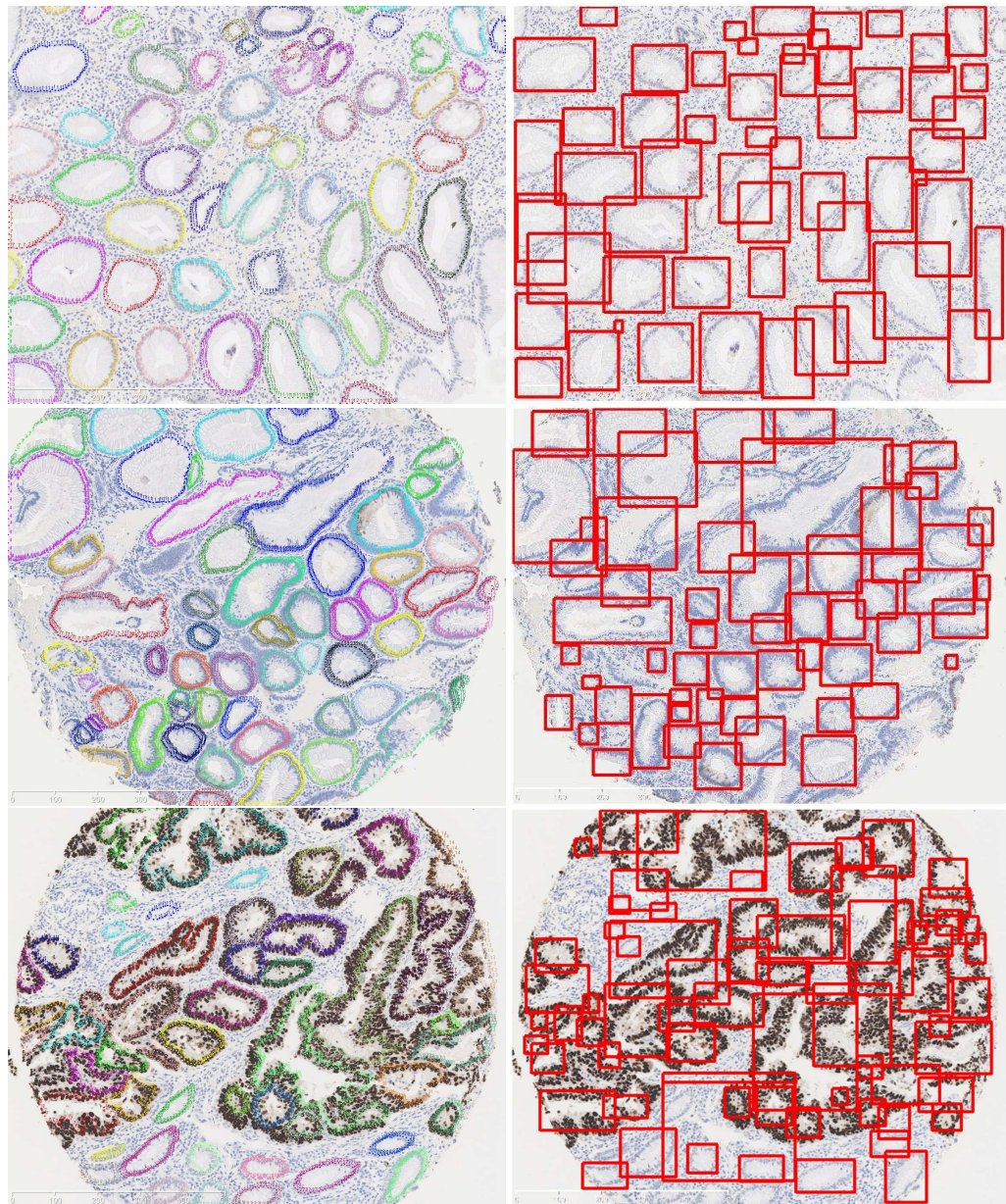
FIGURE 5.7: The gland detection results. For each row, the image on the left
is the annotated ground truth, and the image on the right is the detected result,
each gland is fixed with a red bounding box.
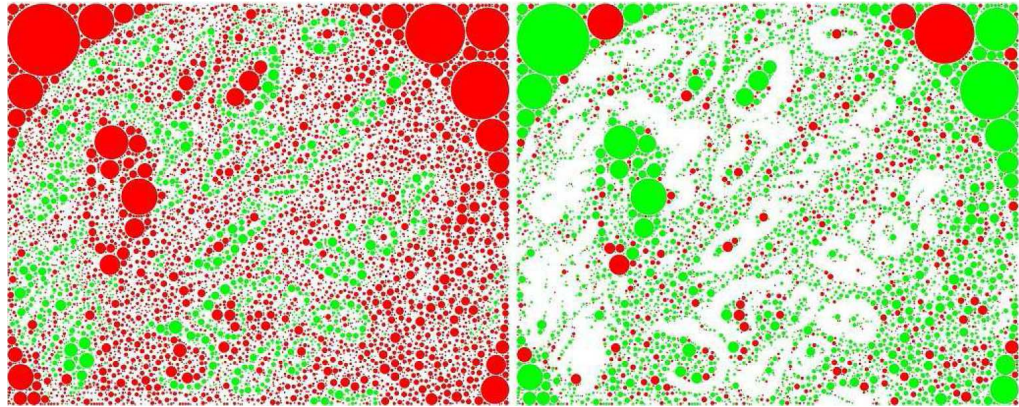
FIGURE 5.8: The results of the first two steps in [4]. Clusters can hardly be observed in the right image.

According to the plot, the texture-based method can only detect 62% of annotated glands with very low precision. Although the random field model achieved better results, it obtained a lower precision than this new feature-based method when they detected the same number of glands. That means that the random field model have higher false positive number of glands. The mean results of the random field model attained more accurately contoured glands than others; for example 16% glands were above an overlap value of 0.83. However, the newly proposed approach detected many more true glands and a higher MAP, greatly outperforming all previous approaches. Some qualitative results are shown in Fig. 5.7.

The object-graph algorithm introduced in [4] was also implemented on this dataset. Taking one of the testing images for example, the first step of [4] is to decompose the image into a set of circular objects: nuclei, stroma and lumen. The result is shown in Fig. 5.8 (left), where the nuclei are colored in green, stroma and lumen are colored in red. These stroma and lumen circles are then clustered into two classes based on some geometric features. These features are specifically designed in the hope that the two clusters may correspond to the inner gland part and outer gland part. Although this clustering step works fine in [4], it does not work on our dataset. As can be seen in Fig. 5.8 (right), it is

hard to observe any structures of these two clusters (colored in red and blue). Therefore, the applying of this approach could not be continued..

5.4 Discussion and Conclusions

In this chapter, an efficient gland detection method for H-DAB images was presented, which can also be applied on H&E stained images (see Appendix B). This colour-free method was proposed based on the characteristics of glands. It covers variations in level of staining, gland shape and size. Rather than missing the small glands or the glands which have vacant artefacts at the boundary, as has been the case in previous methods, the proposed method can effectively extract the luminal regions. It can detect the glands with closed-chain or partial closed-chain, and the glands with large size or small size lumen. This colour-free method can also boost the detection speed. For example, the test on the dataset used in this experiment required only several minutes, including the SVM classification process.

In addition, nuclei regions can also be detected, assisting the accuracy of locating each gland. The combination of intensity enhancement and Laplacian algorithm works quite well for nuclei extraction. The previous methods, which used colour-based or k-means based nuclei pixels extraction, all failed in the H-DAB samples due to the large intensity variation. Although the LoG method can produce good results without missing weak stained nuclei pixels, it still reduced the differences between boundary nuclei and stroma nuclei. The proposed method covers the large intensity variations and preserves the differences.

Using the different ring-like region from [4], this ring is the enlarged lumen edge, which produces more accurate contour of a gland than when simply using a polygon shape. This proposed ring-like region can be automatically adjusted

by the detected distance between the boundary nuclei and the lumen edge, in order to detect the glands with thin or thick boundary.

Thus, the accurate positioning of bounding boxes on the detected candidate glands improves the accuracy of later classification. The subsequent auto-threshold method makes it possible to detect nuclei automatically. By incorporating another PHOG based SVM, the proposed method outperformed all previous approaches on a publicly available gland dataset.

Chapter 6

The Analysis of Liver Cirrhosis

This chapter introduces the method developed for the analysis of liver cirrhosis by calculating the Collagen Proportionate Area (CPA). The proposed statistical stain colour detection model in Chapter 3 is used in this analysis.

6.1   Introduction

In the past, the liver biopsy scoring system depends on the description of changes in the architecture without being concerning about the assessment of collagen, fibrosis, or the distribution of the changes in tissue slides. However, liver fibrosis contains important information for the diagnosis and prognosis of liver cirrhosis [111]. The extent of fibrosis marked with Sirius red stain has been assessed using a quantitative analytical approach [112]. The quantification was then converted by a simple calculation to the proportion of collagen in a single tissue section [104]. This collagen proportionate area (CPA) calculation has been demonstrated to be a representative approach to cirrhosis staging [113].

The CPA can also be represented as the percentage of target stain colour pixels occupying a prepared tissue slide. In other words, this is a quantification of

the pixels with the desired colour as well as all the pixels that constructed the tissue slide. The noise made during slide preparation, which derives from two aspects, also needs to be removed from the quantification process. The first source of noise is tissue fragments and shadowed areas close to or connected to the digitized tissue slide (see Fig. 6.1(a) & (b)). The second source of noise is the edge of the tissue slide; alternatively, the background may contaminated by the stain colour (see Fig. 6.1(c)). These artefact noises may be present in many tissue samples and affect the CPA calculation. In addition, over staining problems always occur at one end of an elongated tissue slide that contains a large number of stained colour pixels.

As a consequence, developing an automatic calculation of CPA for liver cirrhosis samples should resolve these problems. The developed method requires three steps;. The stain colour is detected by the statistical colour model first, and then the slide is contoured automatically together with its colour and morphology characteristics; finally, the CPA is calculated and the normal stained samples and over-stained samples are classified using a simple and efficient method.

6.2   Method

The statistical colour detection is described in Chapter 3. The detection of coloured collagen pixels or coloured elastin pixels is similar to the detection of brown colour in DAB stained colon samples. The CbCr colour spaces are used for the model construction, see Equation 3.4. In this section, the main tasks introduced are the contouring of the tissue slide, which removes the noise and is described in Section 6.2.1, and the calculation of Quad CPA, which uses the minimum bounding box method as presented in Section 6.2.2.
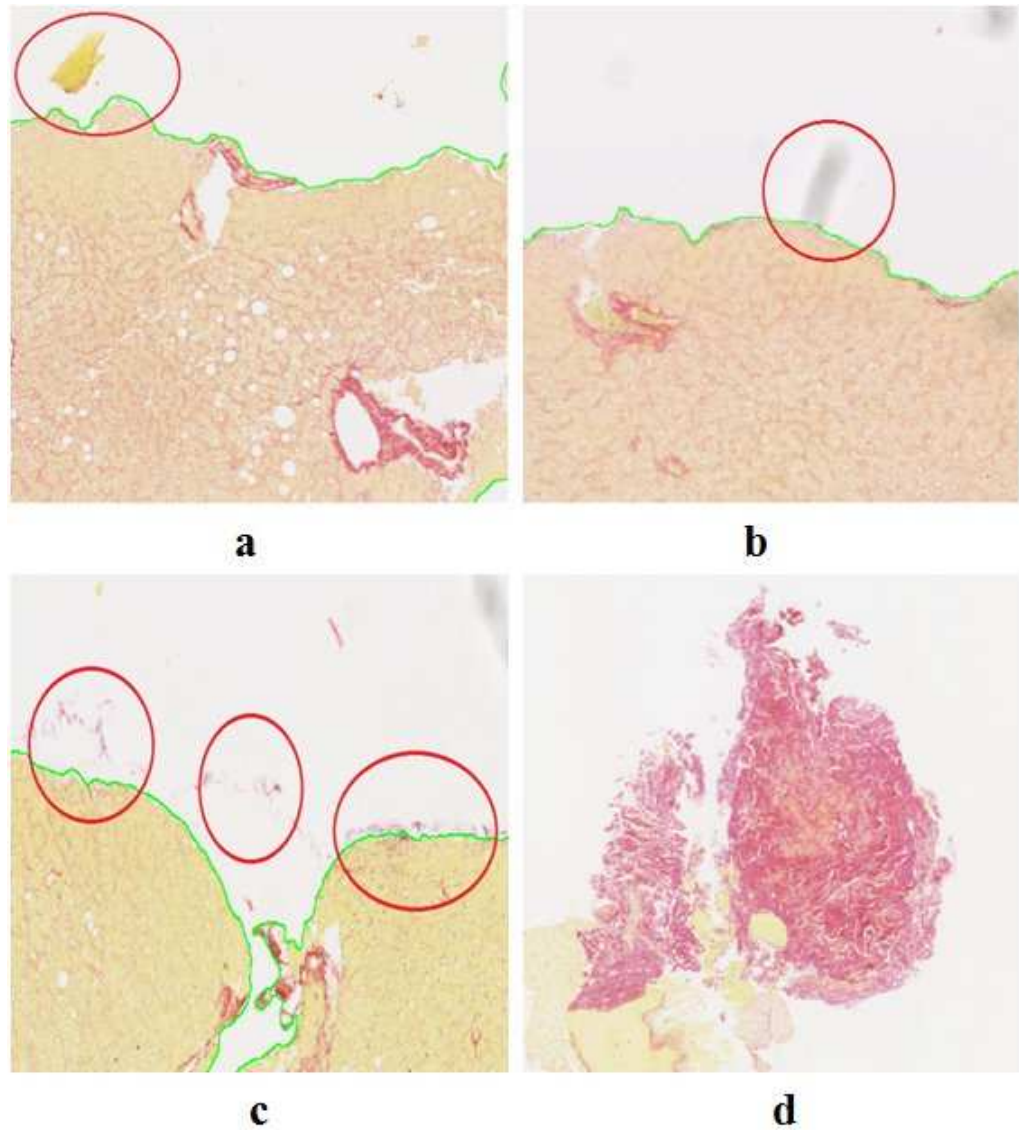
FIGURE 6.1: Sirius Red stained liver cirrhosis biopsies. Slides (a), (b) and (c) are tissue slides with noises due to tissue fragments, shadowed areas, and stain contamination, respectively. The noisy areas are red circled, and the green lines are the manually labelled edges of the tissue slide. (d) is an over-stained tissue slide with large portal tracts and no fibrosis.

6.2.1 Contouring of the tissue slide

The tissue slide contour serves to identify those pixels that belong to valid tissues in the CPA calculation. Visually, the whole tissue slide is stained a pale yellow, which is different from the grey-appearing background in the samples. It is easy

to discriminate objects and background based on colour or intensity information, such as global threshold. The problem here is objects that have similar intensity or colour may represent artefact noise and need to be removed before tissue slide contouring. According to the colour and morphological characteristics of the tissue slide and the artefact noise, different methods have been proposed as follows.

Since shadow pixels have a darker intensity than the background, using a global threshold may not achieve the goal. The original image of Fig. 6.1 (b) is plotted into 3D RGB colour space. The pixels from shadow and the pixels from tissue can then be easily separated by eye (see Fig. 6.2 (a)). The shadow pixels are clumped in a small and different area from the tissue pixels. Thus, the colour information is considered to be a possible solution to distinguish shadows from the coloured tissue slide. Using a colour detection method such as a statistical model or colour model transform can obtain good results. However, the images that are used for CPA calculation contain a huge number of pixels, so these methods may result in high computation costs and consume a great deal of memory .

A histogram of pixels from the background, shadowed area, and tissue area are plotted separately in the Red, Green and Blue channels in Fig. 6.2(b). From the histogram plots, it can be seen that the background pixels and shadow pixels have tiny variations in each RGB channel, very different from tissue pixels. The large variations in the Red and Blue channels make the tissue pixels easy to distinguish from the others. This is computed as follows:

$$(V_c(r) - V_c(b)) \times (V_c(r) - V_c(b)) > T \tag{6.1}$$

Where $V_c(r)$ and $V_c(b)$ are the values for a pixel on the red and blue channels, respectively, and T is a constant threshold. This process can effectively discriminate between a shadow and the tissue slide. Then the artefact fragments are
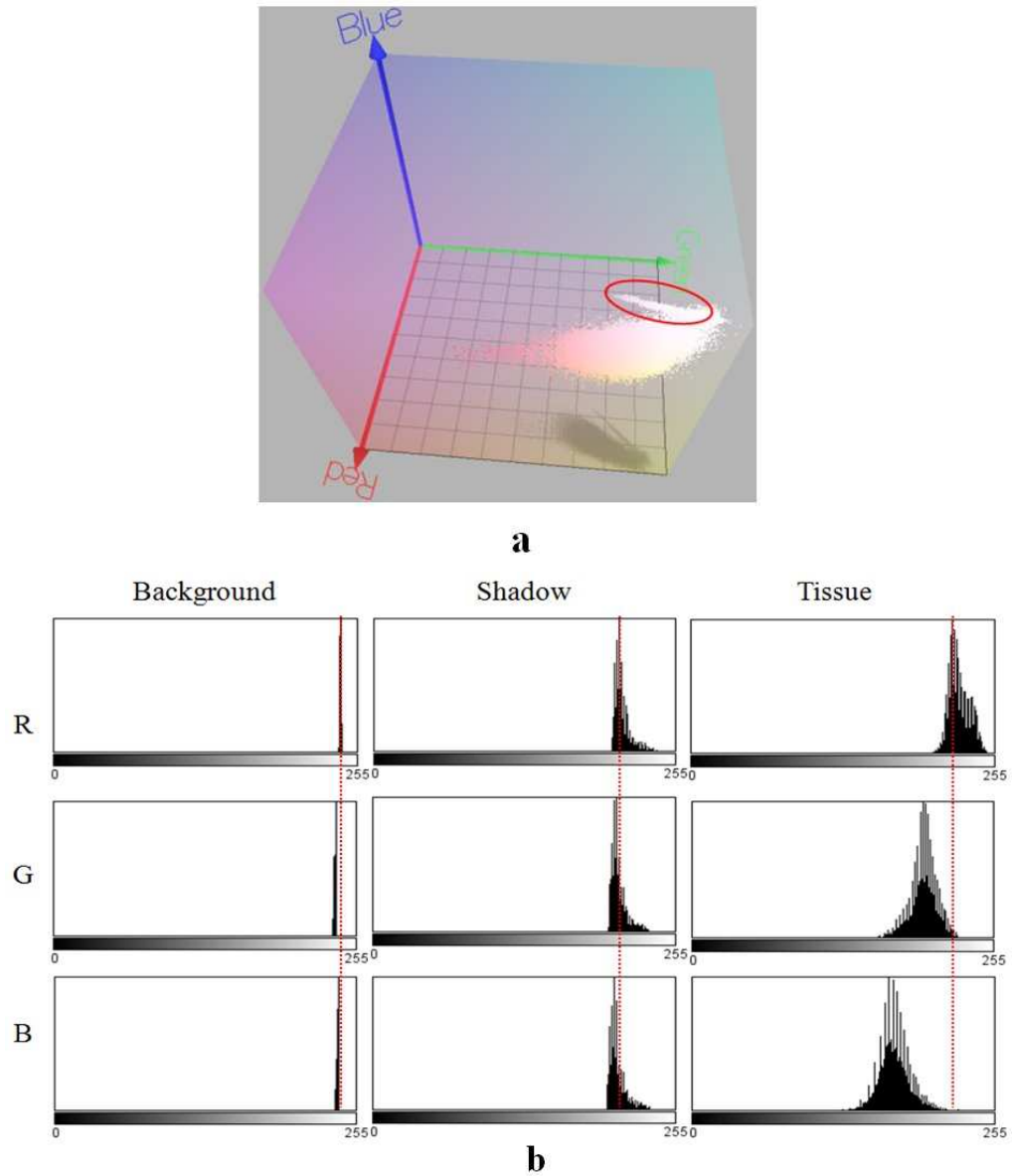
FIGURE 6.2: The plot of colour pixels in image Fig.6.1(b). This image is plotted in 3D RGB colour space in (a), and in the histogram of each RGB channel in (b). The red circle in (a) denotes the shadow pixels in 3D RGB colour space. The red dashed line is the position of histogram peak in R channel; it shows the variations between each channel of RGB colour space.

removed by particle size analysis. The processing result is shown in Fig. 6.3 (b) and the final result is shown in Fig. 6.3 (c) after filling holes and filtering out small particles.
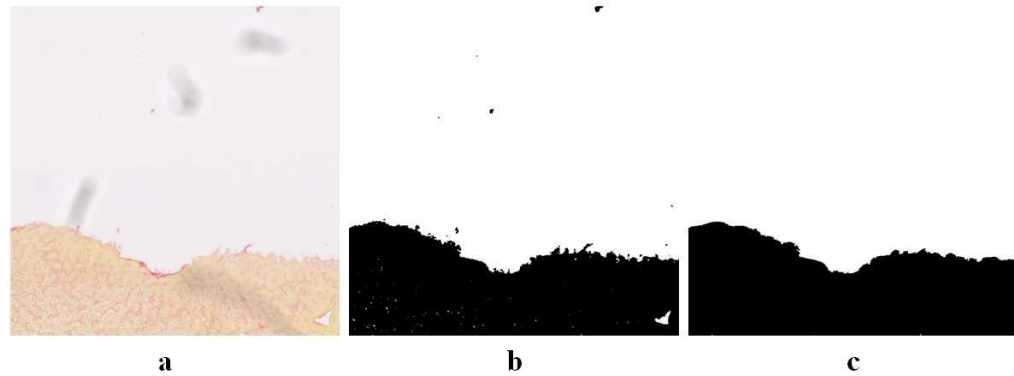
FIGURE 6.3: The removal of artifact noises. (a) the original image with shadowed areas, (b) the detected tissue slide with colour channel variations, (c) the final result.

6.2.2 The Quad CPA

Large portal tracts may cause the pixels around them be over-stained. These areas have no fibrosis, which may lead to inaccurate calculation of CPA. Detecting these over-stained samples is difficult, for they may have an appearance similar to correctly stained samples with large fibrosis bridges. A simple method was considered, the QCPA, to filter out these over-stained samples by computing the standard deviation in CPA values in four areas of the tissue slide. This method can result in two findings, samples where large portal tracts caused over-staining, and samples where the whole slide is over-stained. The whole process is illustrated in Algorithm 7.

Where T is a threshold for CPA, and T_s is a threshold for standard deviation Std. This means a large CPA from a whole image may reflect a greater potential for over-staining on the entire slide. And a small CPA may be regarded as a normal stained slide or indicate large stained areas with no fibrosis.

There are two problems with dividing slides into four areas. One is an image may contain multiple slides; the other is the tissue slide present in the image may not be laid horizontally. We calculated the minimum bounding rectangle for each slide and then divided it equally into four rectangles. The minimum

---

Algorithm 7 The QCPA calculation algorithm.

Require: Sirius red stained RGB image I
Require: Detected pink colour pixels N_{pink}
Require: Detected tissue pixels N_{tissue}
  Calculate CPA = N_{pink} / N_{tissue}
  Draw minimum bounding box for each slide in the image
  Divide each slid into 4 pieces along horizontal
  if CPA $\geq T$ **then**
    Mark the image as Over-stained
  else
    for i = 1 to 4 **do**
      Calculate CPA value V_i for each piece
    end for
    Calculate standard deviation Std
    for i = 1 to 4 **do**
      if $V_i \geq T_s$ **then**
        Mark the slide as large portal tracts
      else
        Mark the slide as normal
      end if
    end for
  end if

---

bounding rectangle is drawn using the method proposed in [114]. This method uses the edge and boundary points of a segmented image to draw the bounding rectangle, which is simpler and faster than most other methods. The fitting of a bounding rectangle is presented as follows, see Fig. 6.4.

The process starts by finding the centroid point of an object, which is the intersection of the diagonals of the bounding rectangle. Suppose the segmented object has boundary points (x_i, y_i), where i=1, 2, n. Then the centre of the object (C_x, C_y) is calculated by:

$$C_x = \frac{1}{n} \sum_{i=1}^{n} x_i, C_y = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{6.2}$$

Crossing the centroid point, the major axis and the minor axis each have an angle with the horizontal line. Let θ be the angle between the major axis and the
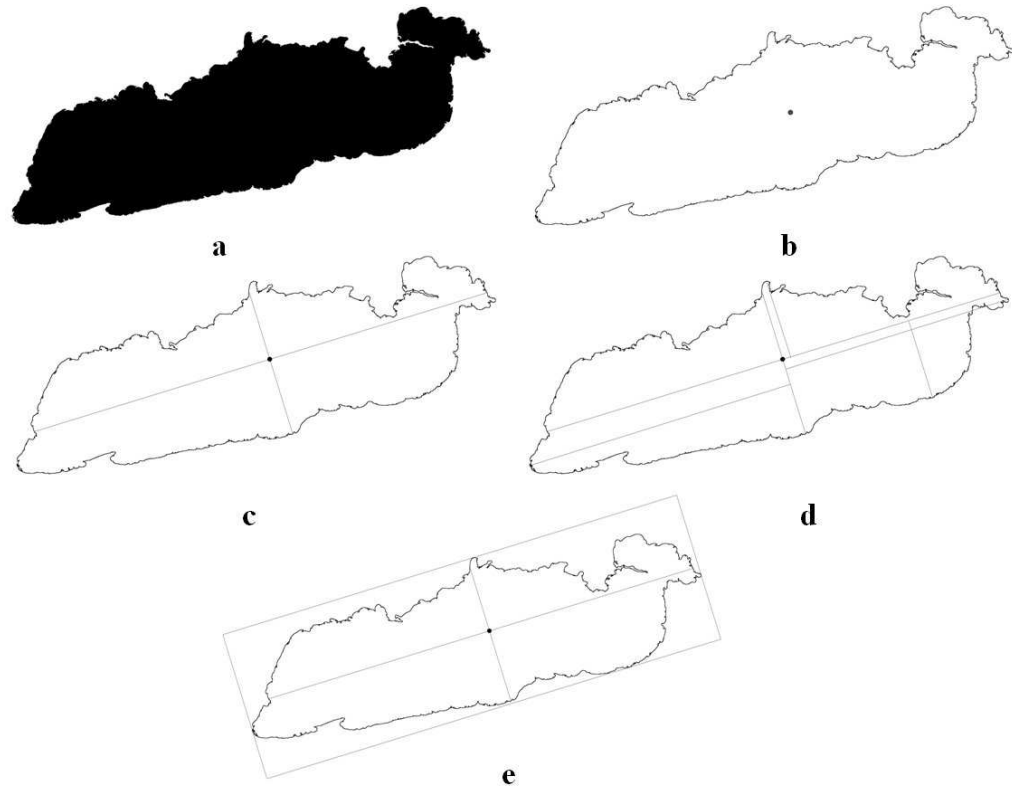
FIGURE 6.4: The illustration of drawing the minimum bounding box. (a) is the input image; (b) is the centre point and the edge of the object; (c) is the major and minor axis; (d) is the upper and lower furthest points; (e) is the minimum bounding box.

horizontal line; then the major axis can be described by the function:

$$y - C_y = \tan\theta(x - C_x) \tag{6.3}$$

It is known that the sum of the square of the perpendicular distance from the boundary points to the major axis is the minimum. The distance from boundary points i=1, 2, n to the line in Equation 6.3 is:

$$d_i = (x_i - C_x)\sin\theta - (y_i - C_y)\cos\theta \tag{6.4}$$

The sum of the square of the distance in Equation 6.4 is

$$D = \sum_{i=1}^{n} [(x_i - C_x)\sin\theta - (y_i - C_y)\cos\theta]^2 \tag{6.5}$$

In order to obtain the minimum value of D, Equation 6.5 becomes the angle of the major axis. Therefore, D can be minimized with respect to angle θ, $\frac{\vartheta D}{\vartheta \theta}$, which gives

$$\tan 2\theta = \frac{2\sum_{i=1}^{n}(x_i - C_x)(y_i - C_y)}{\sum_{i=1}^{n}[(x_j - C_x)^2 - (y_i - C_y)^2]} \tag{6.6}$$

Calculating Equation 6.6 yields the value of θ. Then the major axis described in Equation 6.3 is the best fitting line to this object, which is invariant under the scaling. The minor axis crossing the centroid and perpendicular to the major axis can also be obtained. The major axis and the minor axis determine the orientation of the object and the bounding rectangle as well.

To form the bounding rectangle, we should find the four vertices that are at the furthest distance from the major or minor axis. The intersections of the boundary of the object and the major or minor axis may not be the furthest points. The furthest points include the upper left, the upper right, the lower left, and the lower right with respect to the major axis. Therefore, the edge points should be first determined to be above or below the major axis and the minor axis, separately. These can be calculated through major and minor line equations as follows:

$$V_{major} = (y_i - C_y) - \tan\theta(x_i - C_x), V_{minor} = (y_i - C_y) + \cot\theta(x_i - C_x) \tag{6.7}$$

where (x_i, y_i) is the edge point where i=1,2,n, and θ is the angle between the major axis and the horizontal line. If the value of V_{major} or V_{minor} is larger than 0, this point is above or to the left of the major or minor axis. If the value of V is smaller than 0, this point is below or to the right of the major or minor axis. If the value of V equals 0, it is on the axes. Then the furthest points above, below,

on the left side and on the right side can be obtained.

Assume the furthest four points are obtained as (x_{up}, y_{up}), (x_{low}, y_{low}), (x_{left}, y_{left}) and (x_{right}, y_{right}) with respect to the major and minor axes. Then the line crossing the upper point (x_{up}, y_{up}) and lower point (x_{low}, y_{low}) which are parallel to the major axis can be described as :

$$(y - y_{up}) - \tan\theta(x - x_{up}) = 0, (y - y_{low}) - \tan\theta(x - x_{low}) = 0, \quad (6.8)$$

Similarly, the left point (x_{left}, y_{left}) and right point (x_{right}, y_{right}) parallel to the minor axis can be described as :

$$(y - y_{left}) - \tan\theta(x - x_{left}) = 0, (y - y_{right}) - \tan\theta(x - x_{right}) = 0, \quad (6.9)$$

The lines given in Equation 6.8 are parallel to the major axis, and the lines given in Equation 6.9 are parallel to the minor axis. Thus, the intersections of these four lines can yield the vertices of the rectangle, such as the top left point is the intersection of the upper line and the left line, which is:

$$tl_x = \frac{x_{up}\tan\theta + x_{left}\cot\theta + y_{left} - y_{up}}{\tan\theta + \cot\theta}, tl_y = \frac{y_{up}\cot\theta + y_{left}\tan\theta + x_{left} - x_{up}}{\tan\theta + \cot\theta}$$

$$(6.10)$$

And the top right point is the intersection of the upper line and the right line, which is:

$$tr_x = \frac{x_{up}\tan\theta + x_{right}\cot\theta + y_{right} - y_{up}}{\tan\theta + \cot\theta}, tr_y = \frac{y_{up}\cot\theta + y_{right}\tan\theta + x_{right} - x_{up}}{\tan\theta + \cot\theta}$$

$$(6.11)$$

And the bottom left point is the intersection of the bottom line and the left line, which is:

$$bl_x = \frac{x_{low}\tan\theta + x_{left}\cot\theta + y_{left} - y_{low}}{\tan\theta + \cot\theta}, bl_y = \frac{y_{low}\cot\theta + y_{left}\tan\theta + x_{left} - x_{low}}{\tan\theta + \cot\theta}$$

$$(6.12)$$

And the bottom right point is the intersection of the bottom line and the right line, which is:

$$br_x = \frac{x_{low} \tan\theta + x_{right} \cot\theta + y_{right} - y_{low}}{\tan\theta + \cot\theta}, br_y = \frac{y_{low} \cot\theta + y_{right} \tan\theta + x_{right} - x_{low}}{\tan\theta + \cot\theta}$$
(6.13)

Now, we obtain the bounding rectangle of this object. In order to calculate the QCPA, we still have to divide the rectangle into four equal parts with respect to the major axis. The drawing of the minimum bounding rectangle and the four divided parts are shown in Fig. 6.5. The pixels in each part are examined separately and the QCPA is calculated. As a result, five CPA values are obtained, including the whole CPA and four partial CPA values according to the QCPA method. The measurement of these five CPA values calculates the standard deviation among four QCPA values and the binary value, true or false (true for 1, and false for 0), obtained after thresholding the whole CPA. The total CPA is thresholded by a threshold T, which can determine whether the whole slide is over-stained. Then the true or false Boolean value can be obtained. Combining the Boolean value and the four QCPA values, the calculated standard deviation is then thresholded by a threshold T_s to determine whether the sample is over-stained due to large portal tracts or not. The threshold T and the threshold T_s are obtained from a set of training samples.

a   b   c
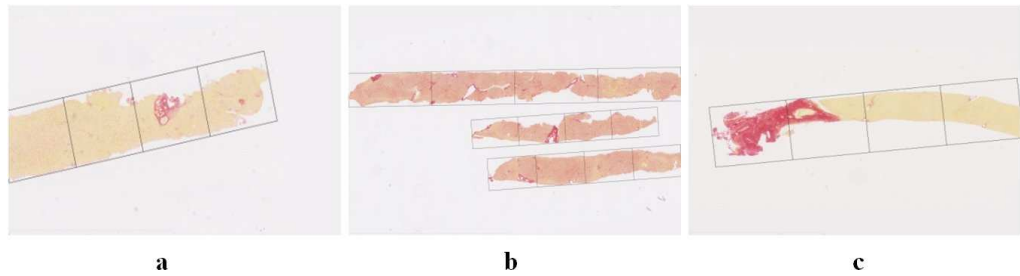
FIGURE 6.5: The minimum bounding box for each slide in the image. The left image is a normal stained liver sample, the middle image is a liver sample where the whole slide shows over-staining, and the right image is an over-stained image with large portal tracts. The dark lines present the minimum bounding box and the divsion of a slide into four areas with equal rectangular boxes.

6.3 Experimental Results

There consists of two experimental results. The detection results of SR stained collagen and the brown stained elastin of using statistical colour detection model are presented at first. Then, the QCPA method has been evaluated.

6.3.1 Evaluation of stain colour detection on liver cirrhosis biopsies

Collagen stained with Sirius red appears pink in colour. The detection of pink colour in liver biopsies was evaluated using the statistical colour detection model. The training dataset for the statistical colour model included 5 SR stained images with resolution 3360×2100, and the test dataset included 20 SR stained images that were captured under 5× magnitude with resolution 3360×2100.

This colour detection model was also applied to the detection of brown colour in stained Elastic fibres in liver cirrhosis biopsies. The training samples used to create the model comprised 10 images with resolution 5600×4200. The model was then tested on a large dataset that include 38 images with resolution 5600×4200. I have labelled the ground truth.

It is clearly seen that both of the detection results, on the SR stained dataset and the brown stained elastin dataset, have a high true positive ratio together with a low false positive ratio (see Fig. 6.7).

6.3.2 Evaluation of QCPA method for normal- and over- stain separation

In order to determine these two thresholds, T and T_s in Algorithm 7, the QCPA method has been trained on 20 Siruis Red stained liver cirrhosis samples which
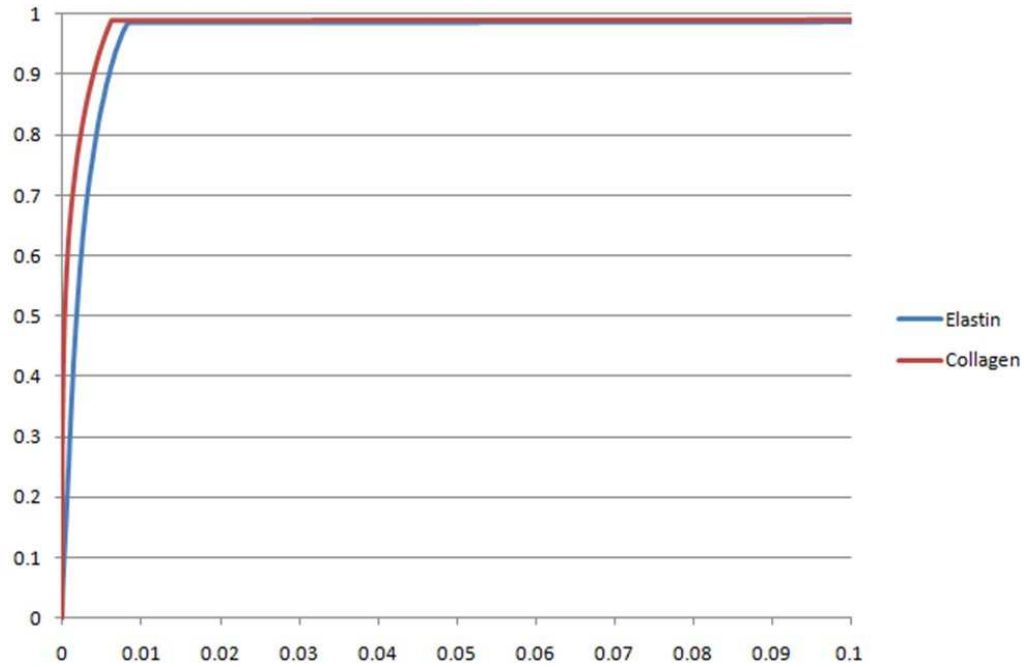
FIGURE 6.6: Experimental results shown as ROC curves on SR stained dataset and brown stained Elastin dataset. The horizontal axis is the false positive ratio and the vertical axis is the true positive ratios. The brown curve is the ROC of SR stain colour detection results. The blue curve is the ROC of brown stain colour detection results.

include 10 over stained samples and 10 normal stained samples. These training images which have different resolutions captured under $5\times$ magnification and only have single slide in the image are randomly selected. The generated results from training samples are plotted in Fig. 6.7. It is obvious that, calculating the standard deviation based only on the four partial CPA may not distinguish the normally stained samples from the over-stained samples. However, another plot clearly identified the normally stained and the over-stained. This plot wass obtained by first thresholding the whole measured CPA value into true or false with respect to over-staining and normal stain. Considering true=1 and false=0, the standard deviation was calculated based on this binary value and four partial CPA values. The combination of the whole CPA and four partial CPA clearly identified two kinds of images. For whole CPA, the threshold is $T \in (0.06\ 0.1)$, and the standard deviation, the $T_s \in (0.04\ 0.24)$. This means, if the whole CPA is larger than the value in the range from 0.06 to 0.1, the image is over-stained.
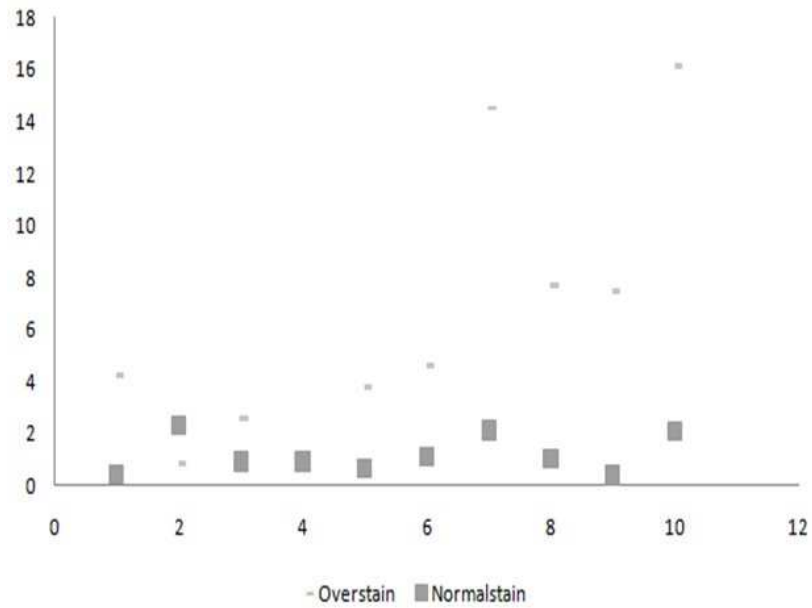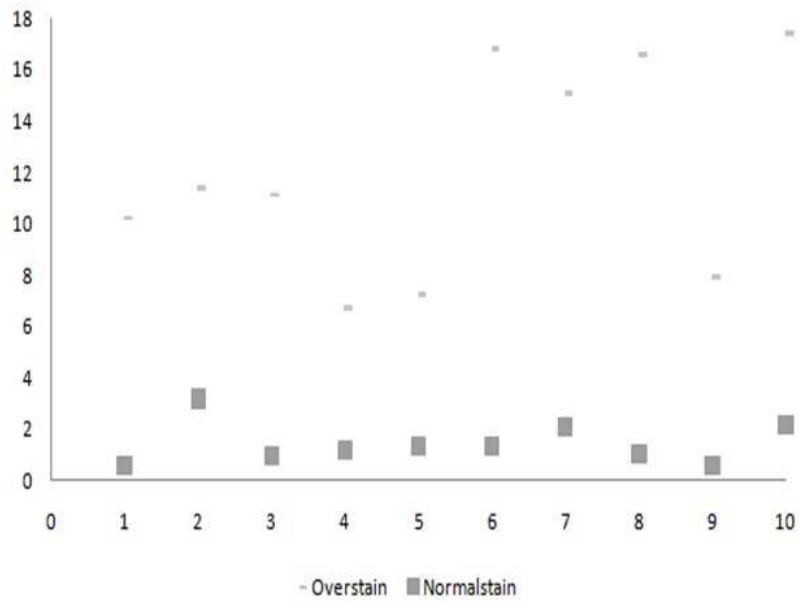
FIGURE 6.7: The scatter plot of CPA values from training samples. The horizontal axis is the images from 1 to 10, the vertical axis is the standard deviation value for each image from the QCPA method. (a) shows the standard deviation of four partial CPA values, (b) shows the standard deviation of the thresholded whole CPA , resulting in binary values (true=1, false =0) together with four partial CPA values.

And if the standard deviation is larger than the value in the range 0.04 to 0.24 then large portal tracts exist, resulting in over-staining. The obtained threshold range was determined by the CPA or QCPA maximum value in normally stained images and by the minimum value in over-stained images. This is due to the small number of training samples.

The testing dataset was Sirius Red stained liver cirrhosis samples. It included 50 images from $5\times$ magnification with resolution 5600×4200, which I manually classified into normally stained samples and over-stained samples. The over-stained samples include slides that are coloured pink everywhere and slides that are over-stained just on one side. The used model for stain colour detection is the statistical model presented in Section 3.2.1. The generated results contained the whole slide CPA calculation and the QCPA calculation together. For images that had multiple tissue slides and fragments, the slides larger than 10,000 pixels were measured one by one. The number of slides measured in the experiment was 123.

First threshold applied on whole slide					
	Manual	Detected	Correctly Detected	Correct Ratio	FPorFN
Over	46	32	26	81.2%	7.8%
Normal	77	91	71	76.9%	43.5%
Second threshold applied on each partial slide					
	Manual	Detected	Correctly Detected	Correct Ratio	FPorFN
Over	46	50	41	89.1%	11.7%
Normal	77	73	68	88.3%	10.9%
Total	123		109	88.6%	

TABLE 6.1: THE STAIN CLASSIFICATION RESULTS. FP = False Positive with respect to Over stain, FN = False Negative with respect to Normal stain, Over = Over-stained, Normal = Normal-stained.

The standard deviation values of all tissue slides are presented in Fig. 6.8. The tested results were measured with different threshold values in the range of those obtained in the training step. Fig. 6.8 shows the obtained QCPA values can be clearly divided into two types by setting the line from 0.04 to 0.1. This

a

b
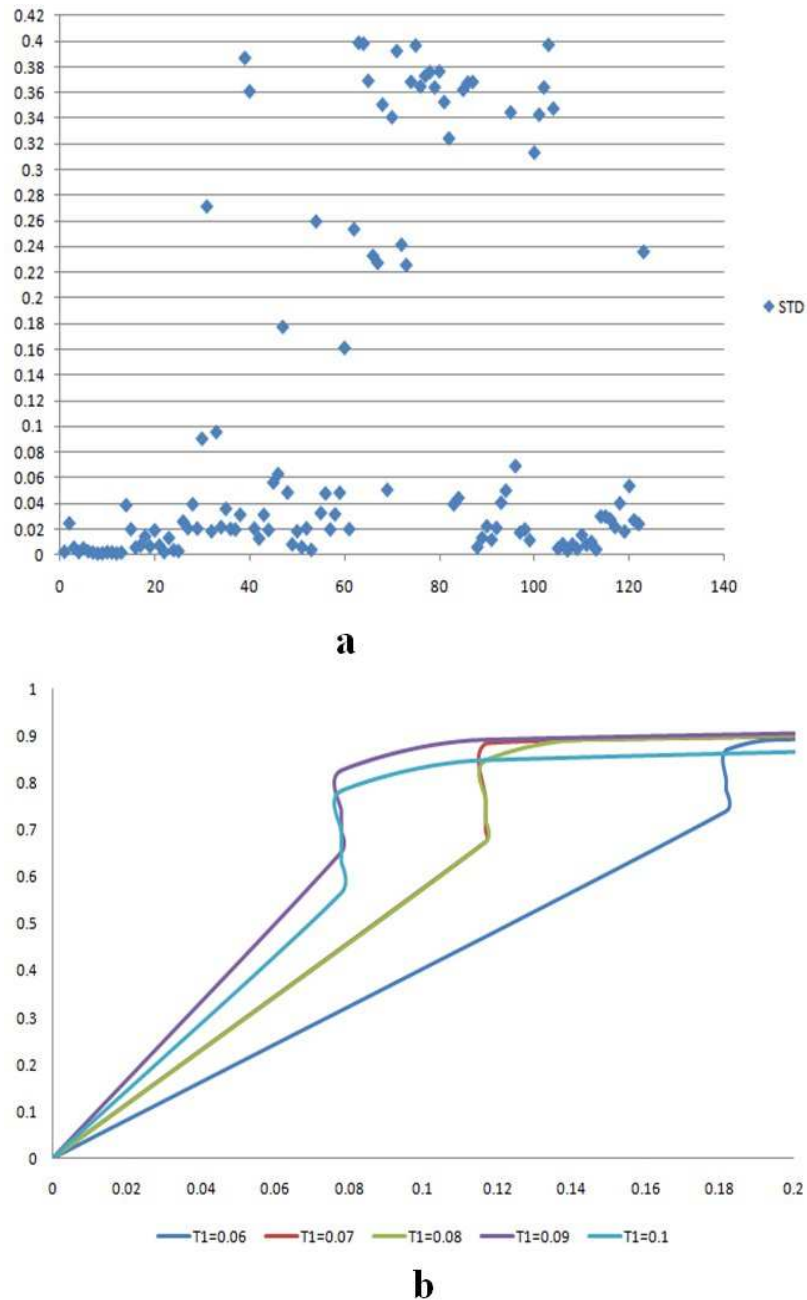
FIGURE 6.8: The experimental results. (a) is the distribution of the measured standard deviation for each single slide in the image, (b) is the tested results with different threshold T and threshold T_s T1=threshold T, which changes from 0.06-0.1, and threshold T_s changes from 0.04 to 0.05, 0.06, 0.08 and 0.1.

range is contained in the training obtained range. Therefore, in the evaluation process, T was set in the range from 0.06 to 0.1 and T_s was set in the range from 0.04 to 0.1. For example, $T = 0.06, T_s = (0.04, 0.05, 0.06, 0.08, 0.1)$ and

$T = 0.07, T_s = (0.04, 0.05, 0.06, 0.08, 0.1)$, etc. The evaluation results are also presented in Fig. 6.8 as a ROC curve. The true positive ratio decreases tremendously at T values greater than 0.09, and for T_s greater than 0.05. The best results were obtained when T=0.09 and T_s=0.04, with the highest true positive ratio and the lowest false positive ratio.

Using the best two threshold values obtained, the results are shown in Table 6.1. Supposing the Over-stained sample is positive and the Normal stained sample is negative, then the true positives, true negatives, false positives and false negatives can be calculated. The ratios after the first threshold and after the second threshold are calculated separately using these criteria. It is apparent that applying the second threshold increases the correct ratio and decreases the false positive ratio.

6.4 Conclusions

Using the applied semi-automatic statistical colour detection model, it is easy to calculate CPA in liver cirrhosis. The method for tissue contouring is simple, highly efficient and easy to implement. To deal with normal staining vs. overstaining, QCPA method was applied, instead of the normal CPA calculation for the determination. This classification was used to divide the images into two types: images with normal CPA and images requiring visual re-examination. From the images with normal CPA, the calculated values can be used directly for diagnosis. In contrast, the images classified as over-stained need to be re-examined by the pathologist. This classification process works to reduce the cost of human labour and improve the detection accuracy. More images will be tested in the future.

Chapter 7

Concluding Remarks

The main contributions are highlighted in this chapter. The effectiveness, robustness and improvements that these contributions represent are presented and discussed. Further analysis and development of future research directions related to these studies are presented as well.

7.1 Main Contributions

This thesis focus on the analysis of digitised microscopic images, especially as it applied to IHC stained colour images. The corresponding studies focused on the automatic detection of stain colour and glands, the segmentation and quantification of cell nuclei, the analysis of CPA in liver cirrhosis and the development of a semi-automatic toolbox. The main contributions are related to:

- Stain colour detection on different kinds of IHC stained images;

- Cell nuclei segmentation and quantification in complex backgrounds and severely clustered areas;

- Gland detection on H-DAB stained IHC images;

- CPA calculation of collagen fibrils and elastin fibrils on liver cirrhosis samples;

- Over-stained and normally stained classification of liver cirrhosis samples;

- Development of semi-automatic toolbox for the analysis of IHC stained images;

These contributions are summarized as follows.

7.1.1 Colour detection

A semi-automatic statistical colour detection model has been proposed for the stain colour detection. This model is visually trained for stain colour detection. Then it has been compared with two popular methods, such as colour deconvolution and the CMYK model. The detection results have shown this method outperforms them.

This semi-automatic colour detection model contains three main advantages as compared to other stain colour detection methods. In the detection of brown colour from H-DAB stained samples, the unevenly prepared stain, the saturated stained nuclei, and the mixed staining of background all create challenges. The DAB stain is not a single stain colour. It is always mixed with haemotoxylin, which stains the background. Thus, it is inappropriate to regard it as a single colour. Moreover, the colour correction methods used in H&E stained samples may also fail in the correction of DAB stain.

The first advantage of using the statistical colour detection model is the effective colour detection based on the histograms of colour pixels used for training. Instead of considering the other machine learning methods, this model is the simplest, with the lowest cost of processing the high-resolution digital medical images. The selected 2D CbCr colour space, instead of 3D space, also shows

the reduction of computation costs. The model is semi-automatic in that the desired colour pixels for training can be visually selected rather than determined by a channel of the selected colour space. Human intervention is considered to complement the selection of the desired colour pixels. The detection accuracy has been demonstrated by the experimental results in Section 3.2.2. The accumulated histograms from multiple training samples can provide stable results with small variations (see Section 3.3). The model offers the ability to reduce the subjective human discrepancies.

The second advantage lies in the detection approach of "what you see is what you get", which can be used to detect linearly absorbed as well as scattered absorbed (DAB stain) stain colour. The lights are not linearly absorbed in H-DAB staining, which nullifies the use of the linear absorption function to detect the brown colour. This is apparently shown in the detection evaluation in Section 3.4, and in the colour separation in dark stained samples in Section 3.4.2. The human involved training process is a possible solution to this problem. The variations between human-selected training pixels were assessed in this thesis. The experimental results have demonstrated their detection accuracy was higher in all cases than the lights-absorption based method.

The third advantage is that the detection process is automatic and the saved model can be sent to other people for reproduction of results. These advantages, and the results of the evaluation have shown the robustness and efficiency of this proposed method, which also resolves the problem of detecting non-linear absorbed stain colour.

7.1.2 Nuclei segmentation

A novel nuclei segmentation and quantification method was conducted on IHC images with complex backgrounds and areas of severely clustered nuclei. The

segmentation method is similar to other methods that include three steps: nuclei area extraction, clustered or overlapped nuclei separation, and post-processing. The quantification element consists of counting the segmented nuclei that have positively stained biomarkers.

In this segmentation method, the nuclei area is accurately extracted from a complex background through a local thresholding method. Clustered or overlapped nuclei are split using a novel seed detection method based on the commonly used watershed. The clustered and overlapping nuclei are separated and over-segmented nuclei are merged under a set of rules. This is a successful combination of watershed, region growing and a priori knowledge of intensity and morphological features.

The evaluations of segmentation and quantification are processed separately. For the segmentation evaluation, this method and several previous popular methods, including the classic watershed and the seeded watershed (regional minimum, LoG method and iterative voting method), were compared on the same testing dataset with severely clustered and overlapping nuclei. The results demonstrated that this segmentation method outperforms the others. The evaluation of the quantification was further tested on another dataset that had a huge number of clustered and overlapping nuclei. This method generated results that are comparable with manual counting. Previous seeded watershed methods (LoG method and iterative voting method) were tested on this dataset as well. The results showed that this method achieved much higher correspondence with human counting. Although the shown Dice coefficient evaluation may not provide satisfactory result, the high correspondence of quantification result has demonstrated the possibility of future improvement.

This new seed detection method is controlled by a priori knowledge regarding the size of the nuclei, which makes the segmentation and quantification stable and robust. The sizes of used sub-window in the local thresholding process

and the selection of initial seed size have been evaluated in Section 4.5. The selection of the size is depending on the size of nuclei. Thus, this method can be applied to the images with different resolutions.

The proposed novel seeds detection method for the nuclei in a complex background stimulates the abilities of splitting those nuclei in severely overlapping and clustered area. It reduces the over- and under-segmentation problems caused by the normal watershed method. The improved splitting results demonstrate a higher correspondence with the manually counting results (see Section 4.6). Therefore, this method can also be applied for quantification or complementary analysis to other medical images that have cell nuclei shapes similar to those examined in this thesis.

Its local threshold, which can cover intensity variations among nuclei, may fail to cover strong intensity variations in nuclei. However, this is a challenging problem similar to those mentioned in Section 1.2.1, which makes it difficult to develop a solution suitable for all situations.

7.1.3 Gland detection

Since the detection of the gland is based on the detection of the lumen, a novel lumen detection method was proposed through a morphological transformation that can detect potential lumens quickly and efficiently. It covers the problem of detecting the lumen regions that have different colours in H-DAB samples. Using this novel method, the partially closed glands and glands with small lumen regions can both be detected with accuracy (see Section 5.2.1).

The boundary for each gland is detected by a newly proposed boundary detection method. Instead of segmenting the nuclei, this method was proposed based on the roughly separated nuclei particles, which highly increased the detection speed. The combination of intensity enhancement and Laplacian followed by

auto-threshold filtering extracted the pixels belong to nuclei efficiently. This extraction preserves the differences between stroma nuclei and gland boundary nuclei. It also has the potential to accurately locate the bounding box for each gland. The proposed ring-like region is the enlarging of the detected lumen edge in order to find the gland boundary. It is used to cover the differences of thickness of gland boundary, which also have the potential to accurately locate the bounding boxes for different glands.

The detected potential glands with fitted bounding boxes were then classified as either true glands or false glands by a support vector machine. In the evaluation process, several previous gland detection methods, such as the object graph method, texture based method, and polar space based method, were compared. In order to show the significant improvement of this method, the part-based method in computer vision was applied as well. The results have shown that this method outperforms the others, and have also demonstrated its robustness and effectiveness. This method was proposed on greyscale images, meaning it can be used for other immunostaining images as well, such as H&E stained images (see Section B).

7.1.4 Semi-automatic tool

This tool was developed using a semi-automatic scheme that is suitable for different kinds of IHC image analysis. It can detect the basic components in an IHC image and is useful in conducting further analysis. The user created models and parameters can be saved and transferred to different users for the reproduction of detection results in different laboratories. This tool has been arranged for free download online.

7.1.5 CPA calculation and normal- or over- stained classification

This was a novel CPA calculation method based on the proposed statistical colour detection model. The detection of the contour of tissue slide is based on the morphological and colour features such as size, the density of pixels and the intensity values in each colour channel. This method is proposed to calculate the CPA efficiently and accurately. Experimental findings have demonstrated that the detection results have a high true positive ratio and a low false positive ratio against the manually prepared ground truth. Also, the comparison with diagnosis results, evaluated by a pathologist, has shown a high level of correspondence. This method has been used for CPA calculation in detecting elastin as well. Also, it has been demonstrated that the detection results have a high level of correspondence against the diagnosis results.

In order to classify liver stained samples as either normal- or over-stained, the QCPA method was developed. The standard deviation was calculated among the CPAs between a quartered tissue slide and the whole slide for an accurate and effective classification. The semi-automatic tool has been extended to include this CPA calculation, in order to assist pathologists with diagnosis.

7.2 Limitations and Future Research

The studies and the obtained results represented in this thesis indicate some interesting research domains that are capable of further improvement and analysis.

The developed statistical colour detection model can detect stain colours that have different chromogens, but may fail in separating stain colours that only have differences in luminance; for example, pink and purple. In addition, the separation of colours in multi-stained samples was not evaluated: for example,

samples with three or more stain colours. It is possible to improve this model for multi-colour separation and make it more useful. For the colour detection evaluation, the colour deconvolution has been compared with the vectors obtained from their laboratory, which may cause the detection variations on the examined dataset in this thesis. It is also needed to prepare a single DAB stained samples in our laboratory, to calculate the corresponding vectors and re-evaluate its detection accuracy.

For nuclei segmentation and gland detection, the parameters are set with a priori knowledge that may vary across different users. One possible solution, and a challenge for future work, may be the development of automatic parameter setting solutions from training samples. The developed nuclei segmentation method is dependent on the round shape of nuclei, and may fail to detect more elongated ones. A possible solution is to consider shape features like concavity values in the separation process. The contour of the segmented nuclei can also be improved by adding some post-processing steps, such as assessing the curve of edge, the removing inner holes, and locating the accurate edge according to gradient changes. The gland classification is based on a support vector machine, which is hard to apply as a plugin in imageJ. In order to make it simple and more effective in practical application, more classification methods can be considered.

This research pipeline is a basic framework for IHC image analysis. From colour pixels, to the nuclei and gland, this pipeline has covered the detection and segmentation of these components that can be observed. The varied combination of the methods discussed can achieve different goals of analysis. Furthermore, these methods are not limited on the datasets used in this thesis. They can be adapted to other similar datasets as well.

The semi-automatic toolbox is the realising of this research pipeline. It can be extended to satisfy many practical applications that could aid pathologists in

diagnosis and prognosis; for example, the CPA calculation presented in Chapter 6. Improvements to this toolbox will include integrating more applications and making it available to users who can then develop their own tools based on this toolbox.

7.3 Conclusions

In conclusion, a research pipeline has been successfully created based on the aims and objectives presented in Section 1.3. The colour pixels detection method has been proposed with high detection accuracy and low computation cost in Chapter 3. In Chapter 4, the nuclei segmentation method has been demonstrated low under- and over-segmentation results. The nuclei quantification results have also shown the robustness of this segmentation method, which has a high level of correspondence with manual counting results. The gland detection method presented in Chapter 5 has high efficiency and high accuracy in the detection of glands in DAB samples.

The detection and segmentation of the basic components of IHC images offers the ability to thoroughly understanding them. Although, the development of these detection and segmentation methods may be complex and difficult than using a machine learning based method, it may provide more explicit and obvious information to aid the pathologists in diagnosis, for example, the size, shape, and number of detected basic components in IHC images. Lastly, this research pipeline has been applied in liver cirrhosis analysis and obtained amazing results, which also have a high correspondence with diagnosis results shown in Chapter 6.

Appendix A

The Interactive Tool

Computer aided diagnosis systems are becoming increasingly importance in cancer detection. They provide instant results in an objective way and greatly reduce the time cost for quantified analysis. We developed a semi-automatic toolbox for the complete analysis of immunostained images by detecting the stained proteins (colour pixels), nuclei and gland structures. The scheme of detection and segmentation processes is considered a crucial step in various computational pathology frameworks [115].

A.1 Introduction

The interpretation of protein expression is a standard procedure for the diagnosis and prognosis of cancer and other diseases [116]. Increasingly, monitoring changes in targeted proteins has been used to measure cancer treatment response and disease progression. Routine diagnosis is predominantly performed by visual inspection and scoring. Many factors affect the accuracy and reliability of diagnosing from stained slides. These factors are the problems inevitably associated with visual diagnosis, which is subjective and relies heavily on the

experience of the observers. Large discrepancies between observers may occur due to these problems. The inter- and intra-observer variations generated in the results can be as high as 30% - 40% in cancer grading [117]. Thus, it is highly desirable to design an objective analysis method to aid diagnosis and prognosis.

Currently, the goal of exploring tissue samples quantitatively is to visually identify and count the markers that spread out in the images. Although this may generate more accurate results than qualitative analysis, the huge time cost and much lower efficiency have blocked its development. In order to cope with these problems, computer image analysis techniques have been considered.

IHC stained images contain three basic components, including coloured antigens, cell nuclei and gland structures, which are often used in the diagnosis and prognosis of different diseases. The detection of these basic components offers a way of automatically analysing an image according to the needs of the pathologist.

Although automated image analysis through computer-aided techniques is considered to be a possible solution toward standardization, fully automated techniques are still under development due to the immaturity of the technology at its current level. The results generated by automatic computerized methods may vary with morphological and stain colour characteristics in different tissue samples [7]. On the other hand, human examination is still regarded as the golden standard and is far from being replaced. Thus, human intervention in the analysis process is highly valued, rather than fully automated solutions [117].

Those developed methods, presented in this thesis, are all put in a semi-automatic toolbox for the aid of IHC image analysis and present them in this chapter. The processes in this tool have been developed to follow the examination scheme from protein to gland. The tool consists of three functions and is available for human intervention. The first is the colour detection function, which follows the "what you see is what you get" approach and uses an interactive tool to offer

the users the capability to visually and intuitively train the statistical colour detection model. Users are allowed to detect colours using one of several pre-built colour models or to build different ones for different stains. The second function is the nuclei segmentation function, which offers two ways of analysing nuclei, quantification and segmentation. Users can select different modes of nuclei analysis to meet different goals. The third function is the gland detection function, which can generate the coordinates of bounding boxes of candidate glands for further classification, such as by a support vector regressor. Both the nuclei segmentation and gland detection features can be adapted for the detection of different sized nuclei and glands and also be applied to detect similar structures on other types of stained images. Three applications of these methods have been conducted. The evaluation of these applications consisted of comparing the computer-generated results against the manual ground truth and clinical outcomes.

A.2   The Semi-Automatic Tool

Three previously developed methods have been integrated in a semi-automatic toolbox. This toolbox is coded as a plugin in ImageJ. The workflow is shown in Fig. A.1. It starts from the detection of coloured antigen and can be used to detect positively stained nuclei as well as the positive nuclei forming a gland structure. The goal of understanding the IHC image is to detect each of its components separately and then group two or more detection results for further analysis, for example, grading. The toolbox contains three functions, including stain colour detection, nuclei segmentation, and gland detection that are described in the following three sub-sections.
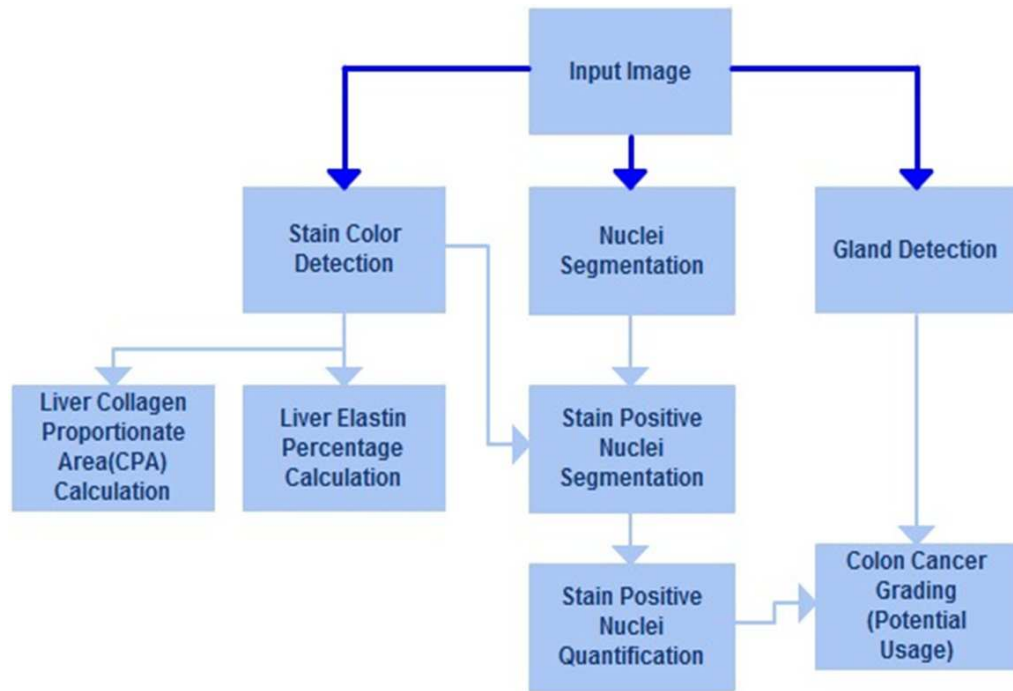
FIGURE A.1: The workflow of interactive tool and the integrated functions in the tool.

A.2.1   Toolbox for colour detection

The colour detection function is based on the statistical model presented in Section 3.2.1, which allows rapid colour detection from arbitrary IHC stained slides. It contains two phases, the training phase and a detection phase. Users can start training by selecting a Region of Interest (ROI) through a rectangular tool in imageJ. There are two components to this visual selection; one is the selection of the user's desired colour, and the other is the initial placing of a sliding bar in the scrolling panel, shown in Fig. A.2. Background pixels can be filtered out using the sliding bar and appear as 255 in the resulting image. A statistical model is constructed based on the histogram of the remaining colour pixels, which are quantified and collected. The training phase requires re-selecting the ROI on multiple training samples in order to obtain a wide range of shades of the target colour. When a new training sample is added, the model is then re-calculated automatically based on the accumulated histograms. At the end of collecting
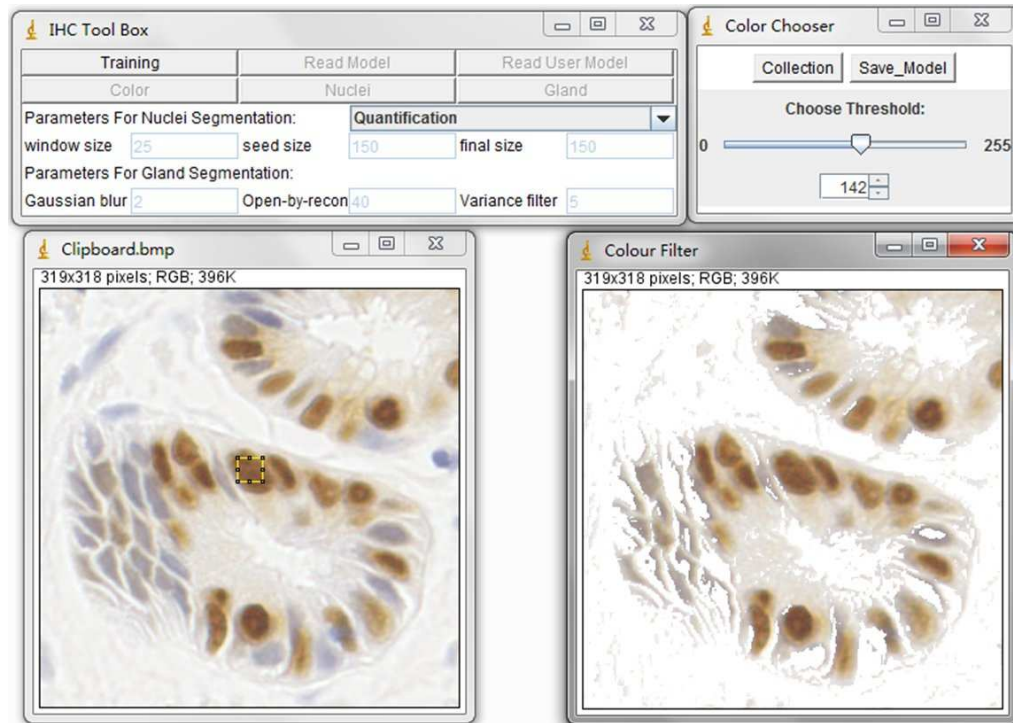
FIGURE A.2: Stain colour detection in the toolbox. This is an illustration of the training phase in colour detection. The left image is the original image. The right image is the resulting image. The colour pixels in the resulting image are colours similar to the colours in the selected ROI in the original image. The background pixels are removed by using the scrolling bar and are set to be 255 in the resulting image.

training samples, the created statistical model can be saved and reused for subsequent detection. For detection, the tool flexibly allows the user to either use the default DAB detection model that was obtained from our experiment, or the user-customized and saved model. Then similar staining colour in IHC images can be automatically detected (see Fig. A.3 & Fig. A.4).

A.2.2   Toolbox for nuclei segmentation and quantification

We implemented the nuclei and gland segmentation functions in a customized mode. The segmentation and quantification functions process the image depending on the parameters set up by the user. In nuclei segmentation, the function can generate two outcomes, a quantification result and a nuclei contour result,

FIGURE A.3: Examples of detected H-DAB stained colour image.
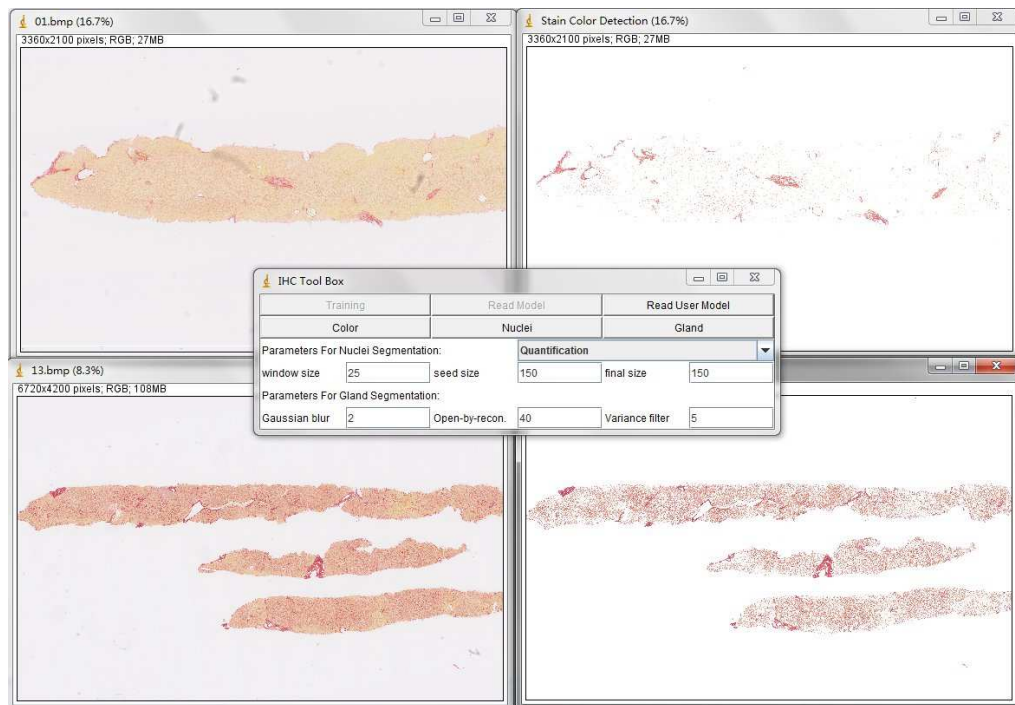
FIGURE A.4: Examples of detected SR stained colour image.

shown in Fig. A.5. The quantification consists of counting the positive nuclei and oval fitted nuclei segmentation. Instead, the contour is drawn at the edge of the detected positive nuclei. Unless customized parameters are set, this nuclei segmentation and quantification functions are automatically processed. Users
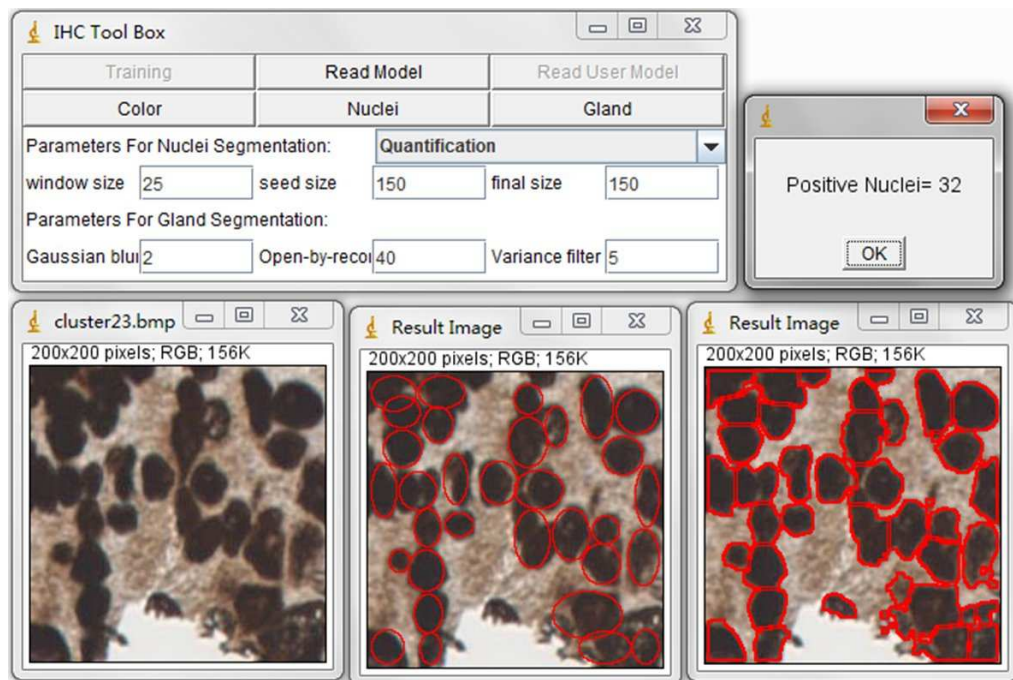
FIGURE A.5: Nuclei segmentation and quantification in tool box. This is an illustration of the nuclei segmentation and quantification results on H-DAB stained images. The quantification mode can generate quantification result and ellipse fitted result image. The segmentation mode can generate segmentation results.

can set parameters according to the particulars of the IHC stained image that is being interpreted. The parameter of window size refers to the size of the local threshold window, where a local threshold can be applied to remove the interval pixels between severely clustered nuclei. It also covers the intensity variations across the whole image for better nuclei region extraction than using a global threshold. The seed size parameter refers to the selection of initial seeds and filtering out noise particles for a region growing process. These noise particles may decrease the average value due to variations in the size of the nuclei. The final size parameter refers to the size filtering process, which can remove the

detected noise nuclei after the final watershed has been applied. For details of parameters determination please see Section 4.5.

A.2.3 Toolbox for gland detection

For gland detection, we can locate each potential gland in a bounding box and then export the coordinates of this bounding box, such as the coordinates at the top left and bottom right. Then these bounding boxes can be classified by other classification methods such as SVM.

A Gaussian blur parameter is used in pre-processing to smooth the image in order to reduce the intensity variations. A parameter for open-by-reconstruction can also be used as a morphological transform to reduce intensity changes in a region that has a higher intensity value than its surroundings. This is performed as an erosion of white regions larger than a set threshold in the binary image with a parameter-defined disk followed by dilation. Thus, this parameter is the diameter of a disk that should be larger than the thickness of the nuclei but smaller than the average diameter of the lumen. The variance filter parameter is determined by the thickness of the nuclei and is used to perform transform operations.

A.3 Experimental Results

This tool-box has been evaluated by analyzing three different kinds of stained IHC images, including 100 DAB stained images for the detection of P53 in colon cancer, 20 SR stained images for the detection of collagen relate to liver cirrhosis, and another set of DAB stained images for the detection of elastin in liver cirrhosis biopsies. The 100 DAB stained images for P53 detection included 50 Tissue Microarray Images (TMA) and 50 Whole Slide Images (WSI).

The evaluation process compared the automated detection results with the manual ground truth. The calculated criterions include the true positive ratio, false positive ratio and the correlations between tool generated results and manual labelling results for each colour stain detection. The results of this colour detection are presented in Table A.1. It shows the statistical model detected more true positive pixels than the rest of two methods. The ROC curve for colour detection on the dataset of 100 DAB stained slides is shown in Section 3.3. And the ROC curves for colour detection on pink stained collagen samples and brown stained elastin samples are shown in Fig. 6.7. The nuclei segmentation and quantification results are shown in Section 4.6 & Section 4.3. The gland detection results are shown in Section 5.3. All these evaluation results have demonstrated the accuracy of this developed tool-box.

Dataset	Total	TP	FP	CORL
DAB(WSI)	88,200,000	99.9%	4.2%	0.960
DAB(TMA)	88,200,000	96.9%	3.9%	0.946
SR	225,792,000	95.9%	0.4%	0.998
Elastin	978,880,000	98.9%	0.1%	0.964

TABLE A.1: COLOUR DETECTION RESULTS.Total=Total image pixels, TP=True Positive ratios, FP=False Positive ratios, CORL=Correlation which is against the manually labeled ground truth

A.4 Conclusions

The main advantage of this tool is its semi-automated scheme, which integrates human knowledge to improve the accuracy of colour detection, nuclei segmentation and quantification, and gland detection. The tool is easy to use, and allows users to interactively and visually select the stain colours and the parameters. Rather than building a statistical model based on the selected ROI, the tool can be manually controlled to filter out pixels that are dissimilar to those in the selected ROI, at an interactive speed. The quantification and segmentation of

nuclei is selectable to achieve different interpretative purposes. The generated gland bounding box coordinates aid further classification and can be used on different kinds of stained images. This tool can also be extended to satisfy many other applications.

Appendix B

Instructions of Semi-Automatic Toolbox

- **Requires**

  ImageJ 1.31p or later, downloaded from http://imagej.nih.gov/ij/

- **Source**

  Contained in Immunostaining toolbox.rar, which can be opened using RAR and ZIP

- **Dataset**

  Nuclei Dataset: Include 52 images of 200×200 pixels. They are captured from Diaminobenzidine (DAB) stained Tissue Microarray (TMA) images and Whole Slide (WS) images. These images contain severely clustered nuclei.

  Gland Dataset: Include 20 images of 1280×1024 pixels. They are captured from DAB stained TMA images and WS images. These images contain normal glands and tumor glands.

- **Installation**
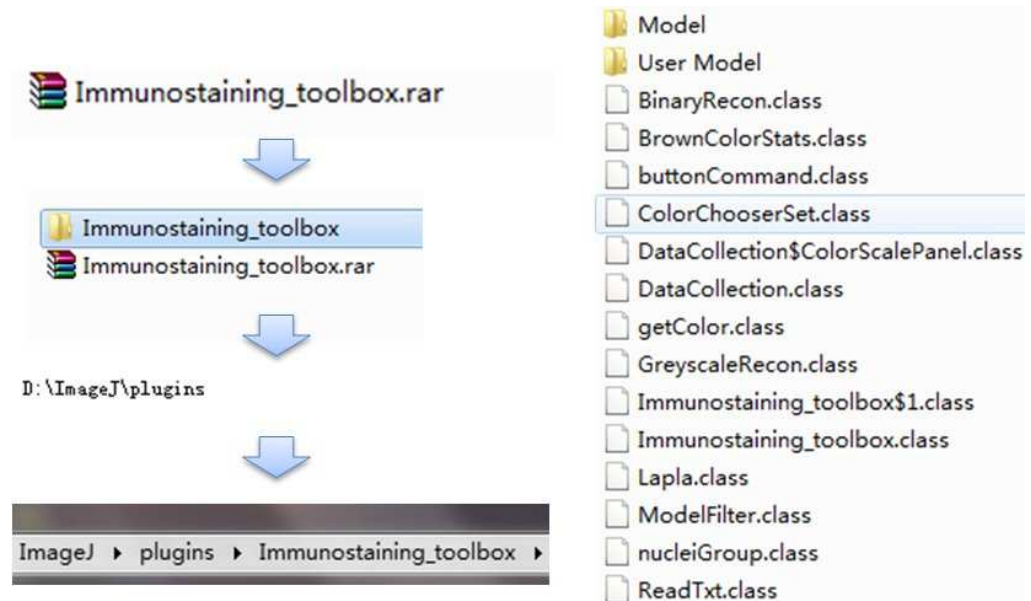
FIGURE B.1: Unzip the RAR and copy to plugin folder.
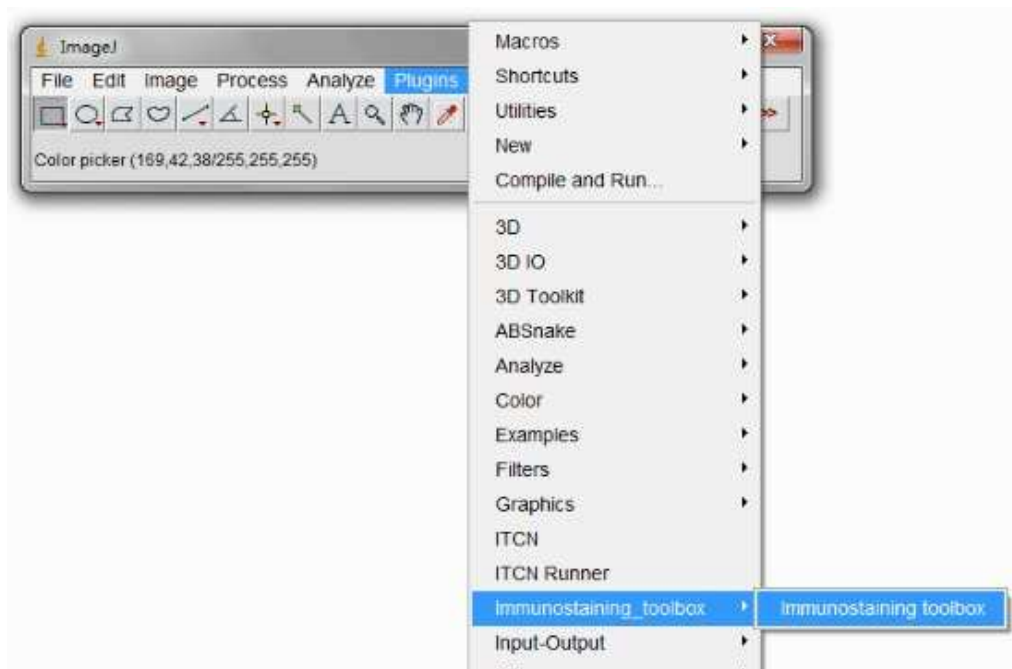
FIGURE B.2: The command on the plugin list.

Unzip Immuonstaining toolbox.rar to the ImageJ plugins folder, or sub-folder (see Fig. B.1). Then restart ImageJ, and there will be a new Plugins/Immunostaining toolbox/Immunostaining toolbox command (see Fig. B.2). Note: This plugin will only be shown on the Plugins list after you restart
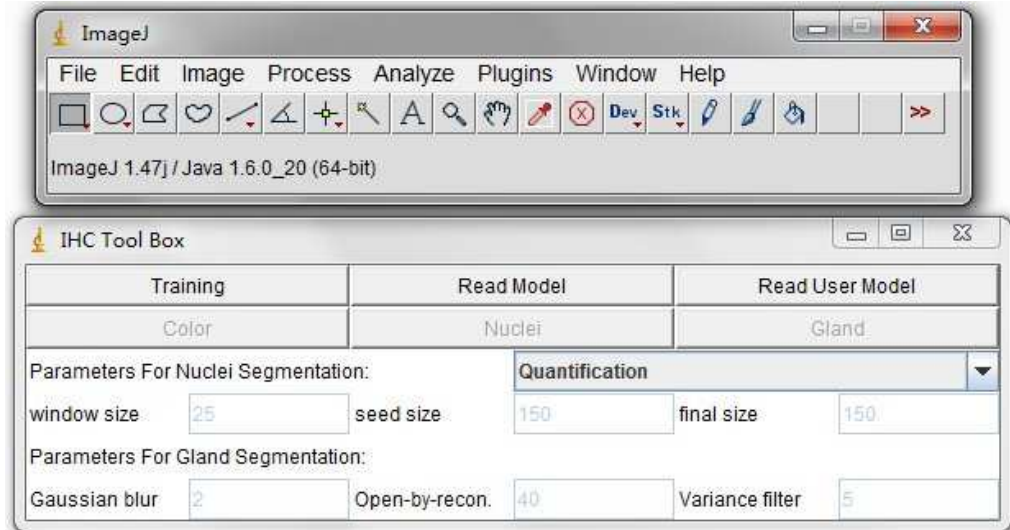
FIGURE B.3: The interface of toolbox.

ImageJ.

- **Descriptions**

  We developed this toolbox using a semi-automatic scheme that is suitable for different kinds of IHC image analysis. It can detect the basic components in an IHC image and is useful for researchers wishing to conduct further analysis (see Fig. B.3). The user created models and parameters can be saved and transferred to different users for the reproduction of detection results in different laboratories. The functions contained in this tool are semi-automatic colour selection, automatic statistical colour detection, automatic nuclei segmentation and automatic gland detection (locating bounding boxes to the candidate glands and requiring further classification, such as Support Vector Regression (SVR)).

B.1 User Manual

B.1.1 Colour detection

For the colour detection, the functions of semi-automatic colour selection and automatic statistical colour detection model are combined. Colour selection is used to select and reserve the positive colour pixels while the background colour pixels are eliminated (set to be 255). Then, the statistical colour model is created based on the histogram of these reserved positive colour pixels.

- **Training**

  This training process aims to train the statistical colour detection model based on visual selected colour pixels. You can start training by selecting a Region of Interest (ROI) through a rectangular tool in imageJ and pressing "Training" button on main panel (see Fig. B.4). The left image is original DAB stained colour image and the right image is the result image. Using the sliding bar presented in the sub-panel (Colour Chooser), you can visually determine the reserved colour pixels in the result image. Once all the desired colour pixels of an image are reserved, you can record them by pressing "Collection" button in the Colour Chooser panel. The model is then automatically calculated and created.

  In general, the training phase requires re-selecting the ROI in multiple training samples in order to obtain a wide range of shades of the target colour. You can close the collected image and re-open a new one. When a new training image is added, the model is then automatically re-calculated based on the accumulated histograms.

  At the end of collecting training samples, the created statistical model can be saved (Using the "Save Model" button in the Colour Chooser panel) and reused for subsequent detection. It is recommended to save the model
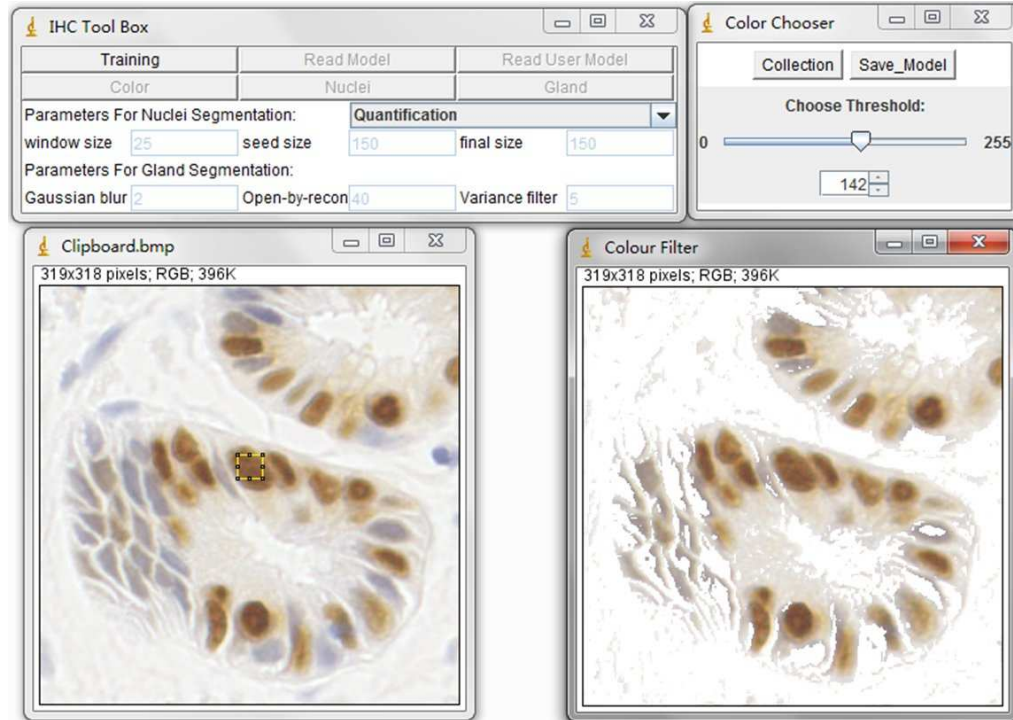
FIGURE B.4: Visually determine the reserved colour pixels.
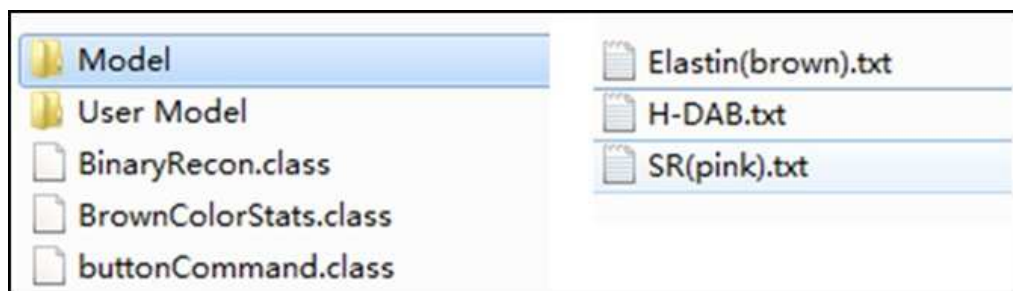
FIGURE B.5: Saved models.

as the ".txt" format. Please see the models we have created in the "Model" folder (see Fig. B.5).

These model files are pre-defined colour models for the detection of DAB stained brown colour, brown colour in elastin contained liver samples, and Sirius Read (SR) stained pink colour. Users can define their own models through the "Training" step and put them in the "User Model" folder.

By selecting these models or changing a new one, you can press the button "Read User Model" on the main panel to read them one by one. The

FIGURE B.6: Example for DAB colour detection.

"Read Model" button is set to read the default model "H-DAB.txt" for brown colour detection (see Fig. B.6).

- **Testing**

  Once you have read one colour detection model (for example, the H-DAB.txt), simply press the "Colour" button on the main panel. The tool will then generate the detected result in a new window.

B.1.2 Nuclei segmentation and quantification

The automatic nuclei segmentation is implemented in the detection of positive stained nuclei in DAB stained images, especially for the separating of severely clustered nuclei. The pre-defined parameters include a window size of 25×25 (half size of nuclei) pixels; the minimum size constraint is 150 pixels and the final size constraint is 150 pixels. These parameters are set for the segmentation
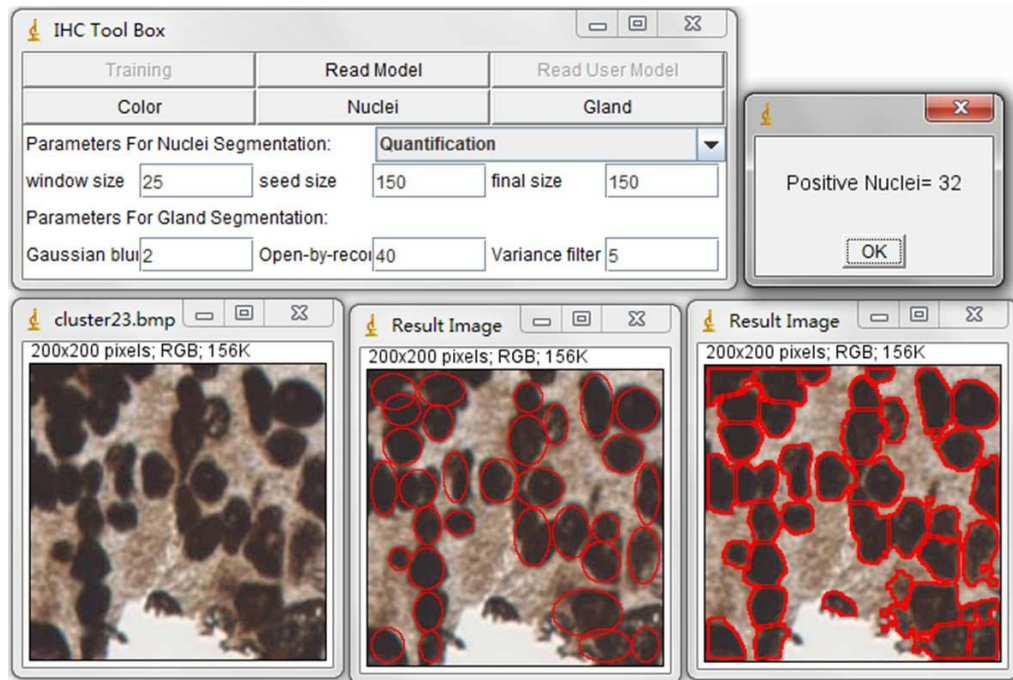
FIGURE B.7: Example for nuclei segmentation and quantification.

of clustered nuclei in our created dataset, the Nuclei Dataset. You can customise these parameters according to your datasets.

- **Nuclei Segmentation and Quantification**

  For nuclei segmentation and quantification, you can set those parameters before you press the "Nuclei" button (see Fig. B.7). The quantification consists of counting the positive nuclei and oval fitted nuclei segmentation, while the contour is drawn at the edge of the detected positive nuclei. Unless customised parameters are set, these nuclei segmentation and quantification functions are automatically processed. You can select the order from the drop-down menu to change between the processes of quantification and segmentation.

- **Parameter**

  In order to estimate the parameters for nuclei segmentation, you can use the square tool and ellipse tool in imageJ (see Fig. B.8). The image on
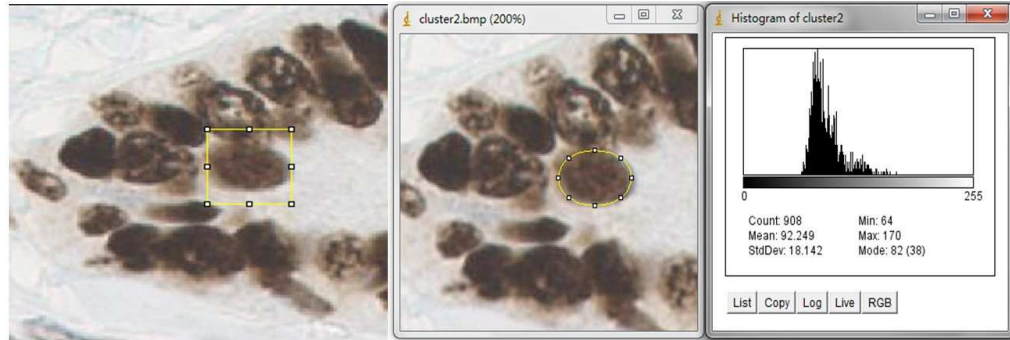
FIGURE B.8: Example for parameter determination.

the left is fitted by a 44×44 square, and the image on the right is fitted by an ellipse that has 908 pixels inside. For the window size parameter, it is better to set it as 25 pixels (half size of the window), which may cover the sizes of all nuclei. The seed size, corresponding to the minimum size constraint in the thesis, should be the half of smallest nucleus. In our dataset, the size of nuclei changes from 300 to 900 pixels, where we set 150 pixels as the seed size.

B.1.3 Gland detection

For gland detection, we can locate each potential gland in a bounding box and then export the coordinates of this bounding box, such as the coordinates at the top left and bottom right. These bounding boxes can then be classified by other classification methods, such as SVM. The pre-defined parameters include a Gaussian blur of 2 pixels; the open-by-reconstruction radius is 40 pixels and the variance radius is 5 pixels. These parameters are set for the detection of glands in our created dataset, the Gland Dataset. You can customise these parameters according to your datasets.

- **Gland detection**
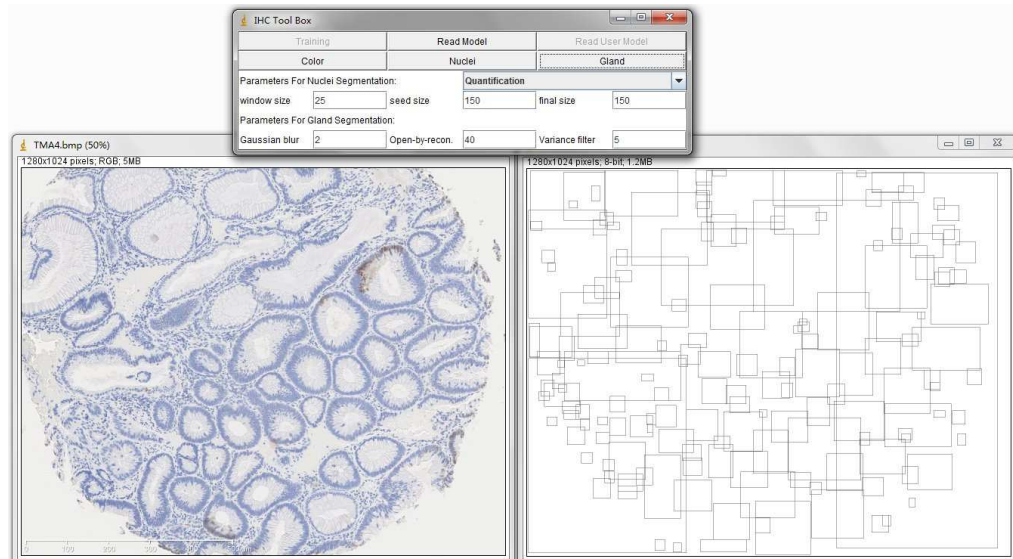
FIGURE B.9: Example for gland detection on H-DAB sample.

You can save the generated file in ".txt" format and name it as you wish, for example, "sample1.txt". This file contains the top left and bottom right coordinates of detected bounding boxes (see Fig. B.9). The final result image contains these bounding boxes and can be combined with the original image through ImageJ Process → Image Calculator → Operation: AND or XOR (see Fig. B.10). See the result image created by the operation AND as follows:

- **Parameter**

  In order to estimate the parameters for gland detection, you can use the square tool in imageJ. The square fitted in this gland is 108×94. The open-by-reconstruction radius depends on the size of the inner gland area. We set an average radius 40 pixels for our dataset. The radiuses of Gaussian blur and variance can be changed according to the specific environment.

  In addition, this gland detection method can be applied to H&E stained images as well, for example, see Fig. B.12 & B.13
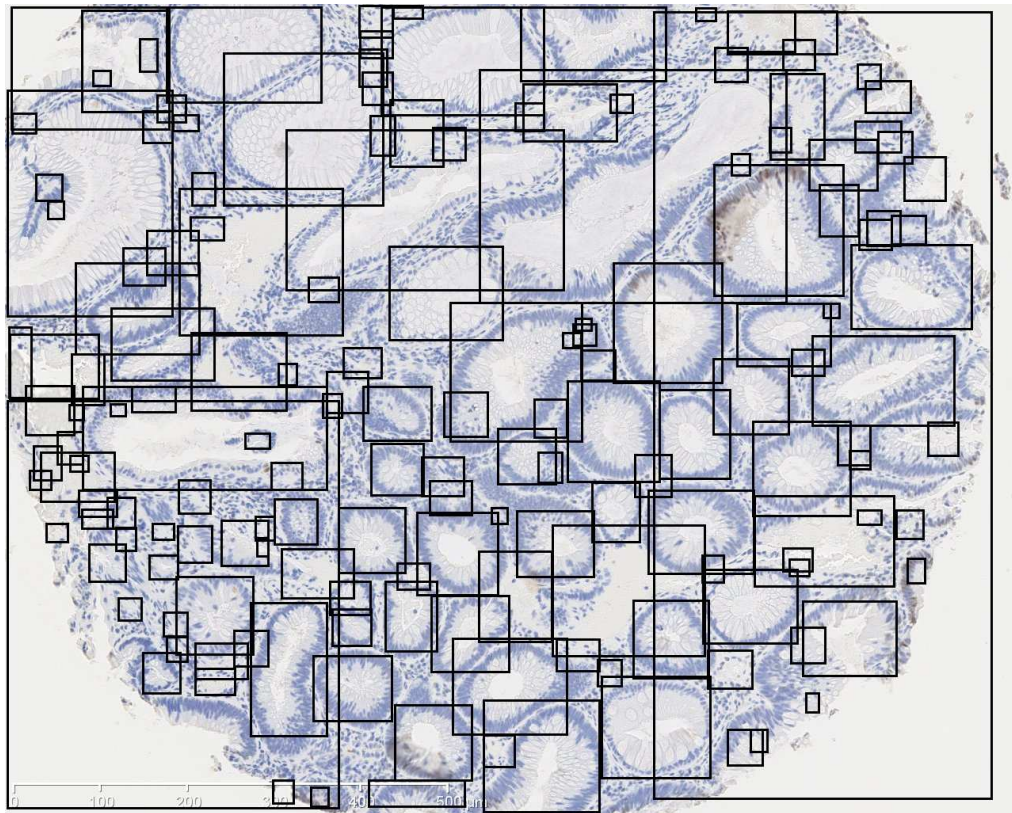
FIGURE B.10: Example for gland detection result performance.

FIGURE B.11: Example for parameter determination.
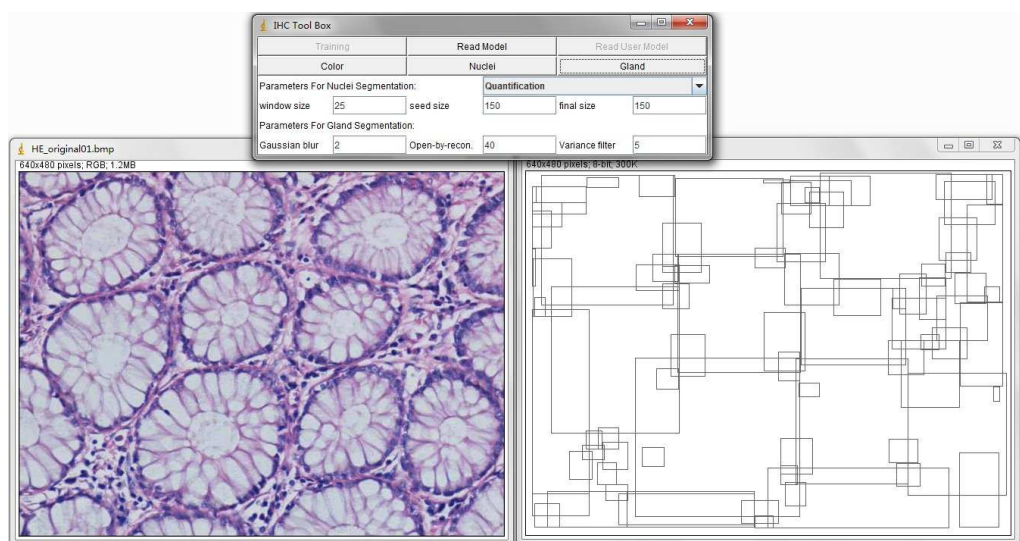
FIGURE B.12: Example for gland detection on H&E sample.
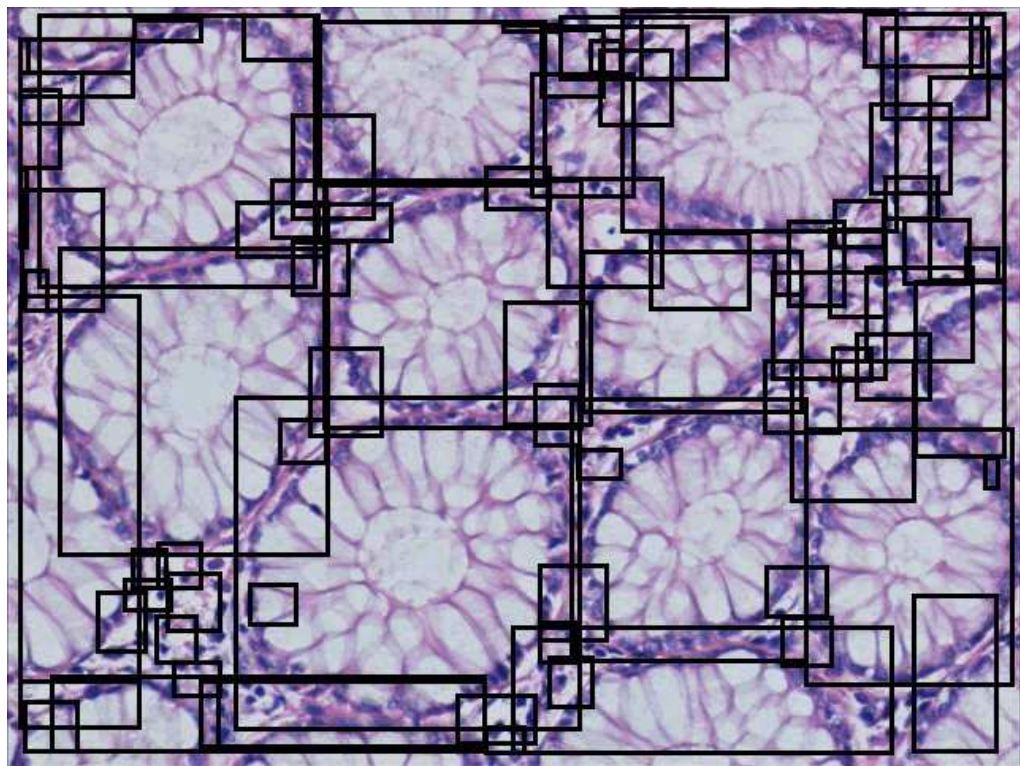
FIGURE B.13: Example for gland detection result performance.

Bibliography

[1] Parvin B, Yang Q, Han J, Change H, Rydberg B, and Barcellos-Hoff MH. Iterative Voting for Inference of Structural Saliency and Characterization of Subcellular Events. *IEEE Transactions on Image Processing*, 16:615–623, 2007.

[2] Yousef AK, Wiem L, William L, and Badrinath R. Improved Automatic Detection and Segmentation of Cell Nuclei in Histopathology Images. *IEEE Trasactions on Biomedical Engineering*, 57:841–852, 2010.

[3] Farjam R, Soltanian zadeh H, Jafari khouzani K, and Zoroofi RA. An Image Analysis Approach for Automatic Malignancy Determination of Prostate Pathological Images. *Cytometry*, 240(May 2006):227–240, 2007.

[4] Cigdem GD, Melih K, Akif B, and Cenk S. Automatic Segmentation of Colon Glands Using Object-Graphs. *Medical Image Analysis*, 14:1–12, 2010.

[5] Mike M. A better lens on disease. *Scientific American*, 302:74–77, 2010.

[6] Santa DC, Elisa F, and Enrico M. Computer-aided Techniques for Chromogenic Immunohistochemistry: Status and Directions. *Computers in Biology and Medicine*, 42:1012–1025, 2012.

[7] Walker RA. Quantification of immunohistochemistry-issues concerning methods utility and semiquantitative assessment I. *Histopathology*, 49: 406–410, 2006.

[8] Coons AH CH and Jones RN. Immunological Properties of an Antibody Containing a Fluorescent Group. *Proceedings of the Society for Experimental Biology and Medicine*, 47:200–202, 1941.

[9] Cregger M, Berger AJ, and Rimm DL. Immunohistochemistry and Quantitative Analysis of Protein Expression. *Archives of Pathology & Laboratory Medicine*, 7:1026–1030, 2006.

[10] Rodrigues NR, Rowan A, Smith ME, Kerr IB, Bodmer WF, Gannon JV, and Lane DP. p53 mutations in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 87:7555–7559, 1990.

[11] Sturm N, Marlu A, Arvers P, Zarski J, and Leroy V. Comparative Assessment of Liver Fibrosis by Computerized Morphometry in Naive Patients with Chronic Hepatitis B and C. *Liver International*, 33:428–438, 2013.

[12] Pellicoro A, Aucott RL, Ramachandran P, Robson AJ, Fallowfield JA, Snowdon VK, Hartland SN, Vernon M, Duffield JS, Benyon RC, and et al. Elastin Accumulation is Regulated at the Level of Degradation by Macrophage Metalloelastase (MMP-12) During Experimental Liver Fibrosis. *Hepatology*, 55:1965–1975, 2012.

[13] Naik S, Doyle S, Feldman M, and John T. Automated Gland and Nuclei Segmentation for Grading of Prostate and Breast Cancer Histopathology. *Surgical Pathology*, (c):284–287, 2008.

[14] Metin NG, Laura EB, Ali C, Anant M, Nasir MR, and Bulent Y. Histopathological Image Analysis: A Review. *IEEE Reviews in Biomedical Engineering*, 2:147–171, 2009.

[15] Chris M van der Loos. Multiple Immunoenzyme Staining: Methods and Visualizations for the Observation With Spectral Imaging. *Journal of Histochem Cytochem*, 56:313–328, 2008.

[16] Brey EM, Lalani Z, Johnston C, Wong M, McIntire LV, Duke PJ, and Patrick CW Jr. Automated Selection of DAB-labeled Tissue for Immuno-histochmical Quantification. *Journal of Histochemistry & Cytochemistry*, 51:575–584, 2003.

[17] Leal S, Diniz C, S C, Goncalves J, Soares AS, Rocha-Pereira C, and Fresco P. Semiautomated Computer-assisted Image Analysis to Quantify 3,3'-Diaminobenzidine Tetrahydrochloride-Immunostained Small Tissues. *Analytical Biochemistry*, 357:137–143, 2006.

[18] Nhu-An P, Andrew M, Joerg S, Sarit AR, Vladimir I, Ming-Sound T, James H, and David WH. Quantitative Image Analysis of Immunohisto-chemical Stains Using a CMYK Color Model. *Diagnostic Pathology*, 2, 2007. doi: 10.1186/1746-1596-2-8.

[19] Thomas AM, Kubilius MB, Holland SJ, Seidlits SK, Boehler RM, Anderson AJ, Cummings BJ, and Shea LD. Channel Density and Porosity of Degradable Bridging Scaffolds on Axon Growth After Spinal Injury. *Biomaterials*, 34:2213–2220, 2013.

[20] Bailey CM, Margaryan NV, Abbott DE, Schutte BC, Yang B, Khalkhali-Ellis Z, and Hendrix MJ. Temporal and Spatial Expression Patterns for The Tumor Suppressor Maspin and its Binding Partner Interferon Regulatory Factor 6 During Breast Development. *Development Growth & Differentiation*, 51:473–481, 2009.

[21] Li C, Takahashi C, Zhang L, Huseni M, Stankovich B, Mashhedi H, Lee J, French D, Anderson JE, Kim D, Howell K, Brauer MJ, Kowanetz M, Yan Y, Humke E, Ebens A, Hampton G, Lackner MR, Hegde P, and Jia S.

Development of a Robust Flow Cytometry-Based Pharmacodynamic Assay to Detect Phosphoprotein Signals for Phosphatidylinositol 3-kinase Inhibitors in Multiple Myeloma. *Journal of Translational Medicine*, 11, 2013. doi: 10.1186/1479-5876-11-76.

[22] Zhao S, Choksuchat C, Zhao Y, Ballagh SA, Kovalevsky GA, and Archer DF. Effects of Doxycycline on Serum and Endometrial Levels of MMP-2, MMP-9 and TIMP-1 in Women Using a Levonorgestrel-Releasing Subcutaneous Implant. *Contraception*, 79:469–478, 2009.

[23] Aymen M, Mounir S, Farhat F, Karima M, and Khaled BR. Automatic Image Segmentation of Nuclear Stained Breast Tissue Sections Using Color Active Contour Model and An Improved Watershed Method. *Biomedical Signal Processing and Control*, 8:421–436, 2013.

[24] Mete M and Topaloglu U. Statistical Comparison of Color Model-Classifier Pairs in Hematoxylin and Eosin Stained Histological Images. In *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pages 284–291, 2009.

[25] Andrew R, Sameer A, Casey L, Jeffrey HP, and Serge B. Unsupervised Color Decomposition Of Histologically Stained Tissue Samples. In *Neural Information Processing Systems (NIPS)*, 2003.

[26] Keerthana P, Avani T, Sandhya I, Gopalakrishna P, and Muktha P. Automation of Immunohistochemical Evaluation in Breast Cancer Using Image Analysis. *World Journal of Clinical Oncology*, 2:187–194, 2011.

[27] Ruifrok AC and Johnston DA. Quantification of Histochemical Staining by Color Deconvolution. *Analytical and Quantitative Cytology and Histology*, 23:291–299, 2001.

[28] Veta M, Huisman A, Viergever MA, van Diest PJ, and Pluim JPW. Automatic Nuclei Segmentation in H&E Stained Breast Cancer Histopathology Images. *PLoS ONE*, 8, 2013. doi: 10.1371/journal.pone.0070221.

[29] Rizzardi AE, Johnson AT, Vogel RI, Pambuccian SE, Henriksen J, Skubitz AP, Metzger GJ, and Schmechel SC. Quantitative Comparison of Immunohistochemical Staining Measured by Digital Image Analysis Versus Pathologist Visual Scoring. *Diagnostic Pathology*, 7, 2012. doi: 10.1186/1746-1596-7-42.

[30] Vilppu JT, Sanna R, Arttu V, Mervi J, and Jorma I. ImmunoRatio: A Publicly Available Web Application for Quantitative Image Analysis of Estrogen Receptor (ER), Progesterone Receptor (PR), and Ki-67. *Breast Cancer Research*, 12, 2010. doi: 10.1186/bcr2615.

[31] Mukesh S and Arya KV. Automated Microscopic Image Analysis for Leukocytes Identification: A survey. *Micron*, 65:20–33, 2014.

[32] Erik R, Michael A, Bruce G, and Peter S. Colour Transfer Between Images. *IEEE Computer Graphics and Applications*, 21:34–41, 2001.

[33] Liu YC, Shih HH, Yang TH, Yang HB, Yang DS, and Sun YN. Quantitative Measurement for Pathological Change of Pulley Tissue From Microscopicimages Via Color-Based Segmentation. *In: Proceedings of the Asian conferenceon Intelligent Information and Database Systems*, 7198: 476–485, 2012.

[34] Magee D, Treanor D, Crellin D, Shires M, Mohee K, and Quirke P. Colournormalisation in Digital Histopathology images. *In: Proceedings of Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy (MICCAI Workshop)*, pages 100–111, 2009.

[35] Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, Schmitt C, and Thomas NE. A Method for Normalizing Histology

Slidesfor Quantitative Analysis. *In: Proceedings of IEEE International Symposium onBiomedical Imaging: From Nano to Macro*, pages 1107–1110, 2009.

[36] Adnan MK, Nasir R, Darren T, and Derek M. A Nonlinear Mapping Approach to Stain Normalization in Digital Histopathology Images Using Image-Specific Colour Deconvolution. *IEEE Transactions On Biomedical Engineering*, 61:1729–1738, 2014.

[37] Lin Y, Peter M, and David JF. Unsupervised Segmentation Based on Robust Estimation and Color Active Contour Models. *IEEE Transactions on Information Technology in Biomedicine*, 9:475–486, 2005.

[38] Can A, Bello M, Cline HE, Tao X, Ginly F, Sood A, Gerdes M, and Montalto M. Multi-modal Imaging of Histological Tissue Sections. *In: Proceedings of IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 288–291, 2008.

[39] Niethammer M, Borland D, Marron JS, Woosley J, and Thomas NE. Appearance Normalization of Histology Slides. *In: Proceedings of International conference on Machine Learning in Medical Imaging*, 6357:58–66, 2010.

[40] Saraswat M, Arya K, and Sharma H. Leukocyte Segmentation in Tissue Images Using Differential Evolution Algorithm. *Swarm and Evolutionary Computation*, 11:46–54, 2013.

[41] Erik M. Cell Segmentation: 50 Years Down the Road. *IEEE Signal Processing Magzine*, 29:140–145, 2012.

[42] Otsu N. A Threshold Selection Method From Gray-Level Histograms. *IEEE Transactions on Systems Man Cybernet SMC*, 9:62–66, 1979.

[43] Ridler TW and Calvard S. Picture Thresholding Using an Iterative Selection Method. *IEEE Transactions on Systems, Man and Cybernetics*, 8: 630–632, 1978.

[44] Pascal B and Brian L. Unsupervised Cell Nucleus Segmentation With Active Contours. *Signal Processing*, 71:203–213, 1998.

[45] Min H, Xijian P, and Yihong D. Automated Cell Nucleus Segmentation Using Improved Snake. In *International Conference on Image Processing (ICIP)*, pages 2737–2740, 2004.

[46] Cong G and Parvin B. Model-based Segmentation of Nuclei. *Pattern Recognition*, 33:1383–1393, 2000.

[47] Radu R, Giovanna B, Cristian S, and Vasile M. Segmentation of Cell Nuclei within Complex Configurations in Images with Colon Sections. In *Intelligent Computer Communication and Processing (ICCP)*, pages 243–246, 2010.

[48] Jingqi AO, Sunanda M, Rodney L, Brian N, and Sameer A. A Hybrid Watershed Method for Cell Image Segmentation. In *Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pages 29–32, 2012.

[49] Ondrej D, Pavel M, Carlos OS, Arrate MB, Martin M, and Michal K. Segmentation of Touching Cell Nuclei Using a Two-Stage Graph Cut Model. Lecture Notes in Computer Science. *Lecture Notes in Computer Science*, 5575:410–419, 2009.

[50] Ling Z, Hui K, Chien TC, Shaoxiong L, Zhi C, Tianfu W, and Siping C. Segmentation of Cytoplasm and Nuclei of Abnormal Cells in Cervical Cytology Using Global and Local Graph Cuts. *Computerized Medical Imaging and Graphics*, in press, 2014.

[51] Marina EP and Christophoros N. Overlapping Cell Nuclei Segmentation Using a Spatially Adaptive Active Physical Model. *IEEE Transactions on Image Processing*, 21:4568–4580, 2012.

[52] Stephan W, Daniel H, Kai S, Albrecht S, Michael B, Peter H, Manfred D, Carsten D, and Frederick K. Detection and Segmentation of Cell Nuclei in Virtual Microscopy Images: A Minimum-Model Approach. *Scientific Reports*, 2, 2012. doi: 10.1038/srep00503.

[53] Dalle JR, Li H, Huang CH, Leow WK, Racoceanu D, and Putti TC. Nuclear Pleomorphism Scoring by Selective Cell Nuclei Detection, 2009.

[54] Francesco R, Marios AG, Georgia K, Kleoniki L, and Ioannis P. Automated Evaluation of Her-2/neu Status in Breast Tissue From Fluorescent In Situ Hybridization Images. *IEEE Trans Image Process*, 14:1288–1299, 2005.

[55] Veta M, Huisman A, Viergever MA, van Diest PJ, and Pluim JPW. Marker-Controlled Watershed Segmentation of Nuclei in H&E Stained Breast Cancer Biopsy Images. In *International Symposium on Biomedical Imaging (ISBI)*, pages 618–621, 2011.

[56] Muhimmah I, Kurniawan R, and Indrayanti I. Automated Cervical Cell Nuclei Segmentation Using Morphological Operation and Watershed Transformation. In *IEEE International Conference on Computational Intelligence and Cybernetics (CyberneticsCom)*, pages 163–167, 2012.

[57] Beucher S and Lantuejoul C. Use of Watersheds in Contour Detection. In *Int. Workshop on Image Processing, Real-Time edge and motion detection/estimation*, 1979.

[58] Beucher S. The Watershed Transformation Applied to Image Segmentation. *Scanning Microsc Suppl*, 6:299–314, 1992.

[59] Jesus A. Nucleus modelling and segmentation in cell clusters. In *Mathematics in industry*, pages 217–222, 2010.

[60] Kim KB, Lee SG, and Kim GH. Nucleus Extraction of Uterine Cervical Pap-Smears Using Marker Information and Watershed Algorithm. *Journal of Medical Imaging and Health Informatics*, 4:1–7, 2014.

[61] Malpica N, Solorzano CO, Vaquero JJ, Santos A, Vallcorba I, Garcia-Sagredo JM, and Pozo F. Appliying Watershed Algorithms to the Segmentation of Clustered Nuclei. *Wiley-Liss, Inc, Cytometry*, 28:289–297, 1997.

[62] Jierong C and Jagath CR. Segmentation of Cluster Nuclei With Shape Markers and Marking Function. *IEEE Transactions on Biomedical Engineerring*, 56:741–748, 2009.

[63] Xiaodong Y, Huoqiang L, and Xiaobo Z. Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift, and Kalman Filter in Time-Lapse Microscopy. *IEEE Transactions on Information Technology in Biomedicine*, 9:2405–2414, 2006.

[64] Chao Z, Changming S, and Tuan DP. Clustered Nuclei Splitting Using Curvature Information. In *Digital Image Computing Techniques and Applications (DICTA)*, pages 352–357, 2011.

[65] Yunyong S and Pornchai P. Overlapping Nuclei Segmentation Using Direction-Based Flow Tracking. *IEEJ Transactions on Electrical and Electronic Engineering*, 8:387–394, 2013.

[66] Alessandro I, Robert PW D, and Marco L. Supervised Localization of Cell Nuclei on TMA Images. In *Mathematical Methods in Biomedical Image Analysis (MMBIA)*, pages 4321–4327, 2012.

[67] Anna M, Tobias M, Sandro H, Nadine V, Benjamin D, and Jurgen P. Automated Seeding-based Nuclei Segmentation in Nonlinear Optical Microscopy. *Applied Optics*, 52:6979–6994, 2013.

[68] Peter A, Yousef AK, Richard S, Michael D, and Gerardo F. Integrated Segmentation of Cellular Structures. *SPIE*, 7962, 2011. doi: 10.1117/12. 876722.

[69] Johannes S, X Jens C. Otte, Andrei K, Andreas B, Ariel G, G. Ulrich N, Uwe S, and Ralf M. Fast Segmentation of Stained Nuclei in Terabyte-Scale, Time Resolved 3D Microscopy Image Stacks. *PLOS one*, 2014.

[70] Hatice CA, Hui K, Camille E, Jessica H, Barrie M, Jin M, Elizabeth P, Rachel R, Mitchell W, Rebecca Z, Gerard L, and Metin NG. Automated detection of cells from immunohistochemically-stained tissues: application to Ki-67 nuclei staining. *SPIE*, 8315, 2012. doi: 10.1117/12.911314.

[71] Xiaobo Z, Fuhai L, Jun Y, and Stephen TCW. A Novel Cell Segmentation Method and Cell Phase Identification Using Markov Model. *IEEE Transactions on Information Technology in Biomedicine*, 13:152–157, 2009.

[72] Umesh A, Ravikanth M, Rodrigo FG, and Carlos O de S. High-Throughput Analysis of Multispectral Images of Breast Cancer Tissue. *IEEE Transactions on Image processing*, 15:2259–2268, 2006.

[73] Cataldo SD, Ficarra E, Acquaviva A, and Macii E. Achieving the Way for Automated Segmentation of Nuclei in Cancer Tissue Images Through Morphology-based Approach: A Quantitative Evaluation. *IEEE Transactions on Information Technology in Biomedicine*, 12:453–461, 2009.

[74] Bjorn N and Anders H. Segmentation of Complex Cell Clusters in Microscopic Images: Application to Bone Marrow Samples. *Cytometry Part A*, 66:24–31, 2005.

[75] Kaustav N, Prabhakar RG, Ryan A, Karen JM, Tom M, and Stephen JL. Automatic Segmentation and Supervised Learning-Based Selection of Nuclei in Cancer Tissue Images. *Cytometry Part A*, 81:743–754, 2012.

[76] Oleh D, Wiggert A.van C, Jeroen E, Wiro JN, and Erik M. Advanced Level-Set based Multiple-Cell Segmentation and Tracking in Time-Lapse Fluorescence Microscopy Images. In *IEEE International Symposium on Biomdical Imaging: From Nano to Macro*, pages 185–188, 2008.

[77] Huang PW and Lai YH. Effective Segmentation and Classification for HCC Biopsy Images. *Pattern Recognition*, 43:1550–1563, 2014.

[78] Zimmer C and Olivo-Marin JC. Coupled Parametric Active Contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27: 1838–1842, 2005.

[79] Aditya PP, Rushi L, and Malik LG. Segmentation of Nuclei in Cytological Images of Breast FNAC Sample: Case Study. *International Journal of Computer Science and Mobile Computing (IJCSMC)*, 3:226–232, 2014.

[80] Korde VR, Bartels H, Barton J, and Ranger-Moore J. Automatic Segmentation of Cell Nuclei in Bladder and Skin Tissue for Karyometric Analysis. *Analytical and Quantitative Cytology and Histology*, 31:83–89, 2009.

[81] Jung C, Kim C, Chae SW, and Oh S. Unsupervised Segmentation of Overlapped Nuclei Using Bayesian Classification. *BioMedical Engineering*, 55:2825–2832, 2010.

[82] Raman S, Parvin B, Maxwell C, and Barcellos-Ho MH. Geometric Approach To Segmentation And Protein Localization In Cell Cultured Assays. *Journal of Microscopy*, 225:22–30, 2007.

[83] Quan W, Hang C, and Parvin B. A Delaunay Triangulation Approach for Segmenting Clumps of Nuclei. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 9–12, 2009.

[84] Xiangzhi B, Changming S, and Fugen Z. Splitting Touching Cells Based on Concave Points and Ellipse Fitting. *Pattern Recognition*, 42:2434–2446, 2009.

[85] Filipczuk P, Fevens T, Krzyzak A, and Monczak R. Computer-Aided Breast Cancer Diagnosis Based on the Analysis of Cytological Images of Fine Needle Biopsies. *IEEE Transactions on Medical Imaging*, 32:2169–2178, 2013.

[86] Wu HS, Xu R, Harpaz N, Burstein D, and Gil J. Segmentation of intestinal gland images with iterative region growing. *Journal of Microscopy*, 220:190–204, 2005.

[87] Naik S, Doyle S, Michael F, John T, and Madabhushi A. Gland segmentation and computerized Gleason grading of prostate histology by integrating low-, high-level and domain specific information. In *In 2nd MICCAI Workshop Microscopic Image Analysis with Appl. in Biology*, 2007.

[88] Maqlin P, Robinson T, Joy JM, and Atulya KN. Automatic Detection of Tubules in Breast Histopathological Images, Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications. *Advances in Intelligent Systems and Computing*, 202:311–321, 2013.

[89] Xu J, Sparks R, Janowcyzk A, Tomaszewski JE, Feldman MD, and Madabhushi A. High-Throughput Prostate Cancer Gland Detection , Segmentation , and Classification from Digitized Needle Core Biopsies. *Cancer Imaging*, pages 77–88, 2010.

[90] Kien N, Anil KJ, and Ronald LA. Automated Gland Segmentation and Classification for Gleason Grading of Prostate Tissue Images. In *International Conference on Pattern Recognition (ICPR)*, pages 1497–1500, 2010.

[91] Nguyen K, Sabata B, and Jain A. Prostate cancer grading: Gland segmentation and structural features. *Pattern Recognition Letters*, 33:951–961, 2010.

[92] Hao F, Guoping Q, Ilyas M, and Jie S. A Novel Polar Space Random Field Model for Glandular Biological Structure Detection. In *British Machine Vision Conference (BMVC)*, pages 42.1–42.12, 2012.

[93] Jones MJ and Rehg JM. Statistical Color Models with Application to Skin Detection. In *IEEE International Conference on Computer Vision and Pattern Recognition*, 1999.

[94] Phung SL, Bouzerdoum A, and Chai D. Skin Segmentation Using Color Pixel Classification: Analysis and Comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:148–154, 2005.

[95] Rudra PK P, Jian JZ, David L, and Hammadi NC. Skin Color Detection Using Region-Based Approach. *International Journal of Image Processing (IJIP)*, 7:385–394, 2013.

[96] Vincent L. Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms. In *IEEE Transactions on Image Processing*, pages 176–201, 2002.

[97] Dalal N and Triggs B. Histograms of Oriented Gradients for Human Detection. *IEEE Conf. Computer Vision and Pattern Recognition*, 2:886–893, 2005.

[98] Bosch A, Zisserman A, and Munoz X. Representing Shape With a Spatial Pyramid Kernel. In *ACM International Conference on Image and Video Retrieval*, 2007.

[99] Matos LL, Trufelli D, Graciela M, and Matos LD. Immunohistochemistry as an Important Tool in Biomarkers Detection and Clinical Practice. *Biomarker Insights*, 9:9–20, 2010.

[100] Landini. Plugin of colour deconvolution. URL http://www.dentistry.bham.ac.uk/landinig/software/cdeconv/cdeconv.html.

[101] Plugin of cmyk. URL http://imagej.nih.gov/ij/plugins/cmyk/index.html.

[102] Ruifrok AC, Katz RL, and Johnston DA. Comparison of Quantification of Histochemical Staining by Hue-Saturation-Intensity (HSI) Transformation and Color-Deconvolution. *Appl Immunohistochem Mol Morphol*, 11:85–91, 2003.

[103] Rittscher J. Characterization of Biological Processes Through Automated Image Analysis. *Annual Review of Biomedical Engineering*, 12:315–344, 2010.

[104] Vincenza C, Andrew KB, Richard S, Pinel O M, Federica G, Gioacchino L, Sergio M, Maria P, Ilias X, Gian PG, David P, Dominic Y, James OB, and Amar PD. Computer-Assisted Image Analysis of Liver Collagen:Relationship to Ishak Scoring and Hepatic Venous Pressure Gradient. *Hepatology, Liver Failure/Cirrhosis/Portal Hypertension*, pages 1236–1244, 2009.

[105] Chaudhuri D and Samal A. A Simple Method for Fitting of Bounding Rectangle to Closed Regions. *Pattern Recognition*, 40:1981–1989, 2007.

[106] Soille P. Morphological Image Analysis: Principles and Applications. *Springer-Verlag*, pages 170–171, 1999.

[107] Log nuclei segmentation from farsight. URL http://www.farsight-toolkit.org/wiki/Main_Page.

[108] Ranefall P, Egevad L, Nordin B, and Bengtsson E. A new method for segmentation of colour images applied to immunohistochemically stained cell nuclei. *Analytical Cellular Pathology*, 15:145–156, 1997.

[109] Smochina C, Manta V, and Kropatsch W. Sampling Step Importance in Hierarchicalsemantic Segmentation of Microscopic Images. In *International Conference on System Theory, Control, and Computing (ICSTCC)*, pages 1–6, 2011.

[110] Chang CC and Lin CJ. LIBSVM : A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.

[111] Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, and et al. Histological Grading and Staging of Chronic Hepatitis. *Journal of Hepatology*, 22:696–699, 1995.

[112] Fontana RJ, Goodman ZD, Dienstag JL, Bonkovsky HL, Naishadham D, Sterling RK, and et al. Relationship of Serum Fibrosis Markers with Liver Fibrosis Stage and Collagen Content in Patients with Advanced Chronic Hepatitis C. *Hepatology*, 47:1–10, 2008.

[113] Manousou P, Dhillon AP, Isgro G, Calvaruso V, Luong TV, Tsochatzis E, Xirouchakis E, Kalambokis G, Cross TJ, Rolando N, O'Beirne J, Patch D, Thornburn D, and Burroughs AK. Digital Image Analysis of Liver Collagen Predicts Clinical Outcome of Recurrent Hepatitis C Virus 1 Year After Liver Transplantation. *Liver Transplantation*, 17:178–188, 2011.

[114] Sethasine S, Jain D, Groszmann RJ, and Garcia-Tsao G. Quantitative Histological-Hemodynamic Correlations in Cirrhosis. *Hepatology*, 55: 1146–1153, 2012.

[115] Thomas JF and Joachim MB. Computational Pathology: Challenges and Promises for Tissue Analysis. *Computerized medical imaging and graphics*, 35:515–530, 2011.

[116] Anderson JE, Hansen LL, Mooren FC, Post M, Huh H, Zuse A, and Los M. Methods and Biomarkers for the Diagnosis and Prognosis of Cancer and Other Diseases: Towards Personalized Medicine. *Drug Resist Updates*, 9:198–210, 2006.

[117] Epstein JI, Allsbrook WC, Amin JR MB, and Egevad L. Update on the Gleason Grading System for Prostate Cancer: Results of an International Consensus Conference of Urologic Pathologists. *Advances in Anatomic Pathology*, 13:57–59, 2006.