A Cell Level Automated Approach for Quantifying Antibody Staining in Immunohistochemistry Images

R. A. A. Khorshed

PhD

UNIVERSITY OF BRADFORD

2013

A Cell Level Automated Approach for Quantifying Antibody Staining in Immunohistochemistry Images

A Structural Approach for Quantifying Antibody Staining in Colonic Cancer Spheroid Images by Integrating Image Processing and Machine Learning towards the Implementation of Computer Aided Scoring of Cancer Markers.

R. A. A. Khorshed

Submitted for the degree of
Doctor of Philosophy

Department of Computing
School of Computing, Informatics and Media
University of Bradford

2013

Keyword: Image processing, Cancer, Immunohistological, Nuclei segmentation, Nuclei classification, Feature extraction.

Abstract

Immunohistological (IHC) stained images occupy a fundamental role in the pathologist's diagnosis and monitoring of cancer development. The manual process of monitoring such images is a subjective, time consuming process that typically relies on the visual ability and experience level of the pathologist.

A novel and comprehensive system for the automated quantification of antibody inside stained cell nuclei in immunohistochemistry images is proposed and demonstrated in this research. The system is based on a cellular level approach, where each nucleus is individually analyzed to observe the effects of protein antibodies inside the nuclei.

The system provides three main quantitative descriptions of stained nuclei. The first quantitative measurement automatically generates the total number of cell nuclei in an image. The second measure classifies the positive and negative stained nuclei based on the nuclei colour, morphological and textural features. Such features are extracted directly from each nucleus to provide discriminative characteristics of different stained nuclei. The output generated from the first and second quantitative measures are used collectively to calculate the percentage of positive nuclei (PS). The third measure proposes a novel automated method for determining the staining intensity level of positive nuclei or what is known as the intensity score (IS). The minor intensity features are observed and used to classify low, intermediate and high stained positive nuclei. Statistical methods were applied throughout the research to validate the system results against the ground truth pathology data. Experimental results demonstrate the effectiveness of the proposed approach and provide high accuracy when compared to the ground truth pathology data.

Declaration

I hereby declare that this thesis has been genuinely carried out by myself and has not been used in any previous application for a degree. The invaluable participation of others in this thesis has been acknowledged where appropriate.

Reema A Khorshed

Dedication

This thesis is dedicated to my beloved parents, whose love, help, support and prayers are the reason why I am where I am today. My wonderful sister Lamia, this thesis is for you.

Acknowledgment

All praise and thanks to Allah the All-Merciful the All-Beneficent, for granting me the ability and strength to complete this thesis.

I am very thankful to my supervisors Professor Jianmin Jiang, Dr. Roger Phillips and Dr. Robert Holton, whose encouragement, guidance and support enabled me to develop an understanding of the subject. It is an honour for me to have worked with them.

I am grateful to Dr. Philip A. Batman from the Bradford Royal Infirmary for providing the pathology validation data to support this thesis. I would also like to thank the Institute of Cancer Research ICT at the University of Bradford for providing me with the image datasets used to accomplish this thesis and for introducing me to the world of Life Sciences.

I would also like to thank my colleagues: Dr. Qudoos Yousuf for his encouragement and support throughout my Ph.D; Dr. Yousef Alqasrawi for his guidance in the image processing field; and Mr. Ahtasham Raza for his guidance in the tissue analysis field.

I am heartily thankful to my father and mother, for their love, support and prayers during my studies, thank you for believing always in me. My special thanks to all my sisters and brothers and to my amazing friend Shazia Shaik.

Publications and Contributions

Conference papers:

- Khorshed, R. A., Yousuf, Q. and Jiang, J. (2012), *An automated approach for Percent Stained Scoring of Antibodies in Immunohistochemistry Images, the 11th IEEE conference of Cybernetic Intelligent Systems CIS2012, , pp.118-122, Ireland.*

- Khorshed, R. A. ,Yousuf, Q. and Jiang, J. (2012), *A Cell Level Classification Approach for Measuring Antibody Staining in Immunohistochemistry Images Based on Neural Network, the 11th IEEE conference of Cybernetic Intelligent Systems CIS2012, pp.123-128, Ireland. (**selected for special issue journal publication)***

- Khorshed, R. A., Yousuf, Q. and Jiang, J. (2012), *An Intelligent Approach for the Automated Segmentation and Quantification of Immunohistologically Stained Nuclear, 6th Conference on Software Knowledge Information Management and Applications SKIMA, China.*

- *Q. Yousuf, R. A. Khorshed, R.L.Ward and J.Elliott, (2011) Optimization of ternary code with perfect periodic ACF for Location detection in RFID system, European Navigation Conference, London.*

- *Q.Yousuf, R. A. Khorshed. and J.Elliott, (2012),Optimization of RFID transceiver design for indoor location using DSSS" Royal Institute of Navigation (RIN) journal publication in progress.*

Journals:

- *Khorshed, R.A., Phillips, R., Jiang J,* Yousuf, Q,*and Philip B., 'A Cell Level Automated Approach for Quantifying and classifying Antibody Staining in Immunohistochemistry Images', Journal of Theoretical Biology. (submitted).*

- *Khorshed, R.A., Phillips, R., Jiang J, and* Yousuf, Q 'An automated evaluation of tumour heterogeneity in immunohistochemistry stained images', Journal of Applied Pattern Recognition. *(In progress, based on selected paper from the IEEE CIS2012 Conference).*

Abstracts and poster presentations:

- R. Khorshed, Structural image analysis towards automatic identification of colorectal cancer cell, international symposium on biomedical imaging, ISBI (2012).
- R. Khorshed Cancer cell detection and tracking using Nonlinear image filtering techniques, University of Bradford school of computing, School Research seminars (2011).

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D	Two Dimensions
3D	Three Dimensions
ATM	Automated Threshold Method
ANN	Artificial Neural Network
CT	Computer Tomography
DAB	Diaminobenzidine
DWT	Discrete Wavelet Transform
FSVM	Fuzzy Support Vector Machine
FN	False Negative
FP	False Positive
GA	Genetic Algorithm
GLCM	Grey Level Co-occurrence Matrix
HH	High High (filter)
HL	High Low (filter)
HSV	Hue Saturation Value
HSCORE	Histology score
IOD	Integrated Optical Density
IHC	ImmunoHistoChemistry
PS	Proportion Score (Percent Stained)
IS	Intensity Score
Lab	lightness and a and b color-opponent dimensions
LH	Low High (filter)
LLM	Local Linear Map
LL	Low Low (filter)
LoG	Laplacian of Gaussian

LI	Labelling Index
MDFS	Maximum Difference Features
MLP	Multilayer perceptrons
NN	Neural Network
PBS	Phosphate Buffered Saline
PCA	Principal Component Analysis
QDA	Quadratic Discriminant Analysis
RBF	Radial Basis Function
ROI	Region of Interest
RGB	Red Green Blue
SVM	Support Vector Machine
SVMC	Support Vector Machine Clustering
TP	True Positive
TN	True Negative

Chapter 1

Introduction

1.1 Background

Over the past two decades, an enormous range of cancer diagnostic techniques have been implemented. Imaging techniques have enabled practitioners to examine the entire human body, for example through magnetic resonance imaging, ultrasonography, computer tomography CT scans and microscopic images taken from biopsy samples. These methods of image examination have increased the chance of detecting abnormality at very early stages and helped physicians to understand cancer better and treat it accordingly (Tadashi et al., 2012; Sarvazyan et al., 1995).

Immunohistological stained images occupy a vital role in the pathologist's diagnosis and monitoring of cancer patients. Immunohistochemical staining is a procedure commonly used for tasks such as diagnosing abnormalities in cancer cells, drug development and biological research. This procedure is performed by applying different antibody stains to detect molecular markers on a tissue section. The molecular markers characterize certain cellular

events such as cell death (apoptosis) and proliferation (Khojasteh et al., 2012; Kostopoulos et al., 2009).

1.1.1 Visual scoring method of IHC slides

Stained cell nuclei allow pathologists to diagnose and grade tumour (Ruifrok et al., 2004). This procedure is conducted manually, which has proved to be very subjective and time consuming (Ficarra and Macii, 2006). The procedure is performed via a number of steps; once the stained sections are prepared, the pathologists view the samples using a microscope. This method enables quantitative and comparative studies of the samples. The score of cancer markers in stained nuclei is based on two main quantitative measures: the percentage of positively stained nuclei and the overall staining intensity of positively stained nuclei (Anderson et al., 1998).

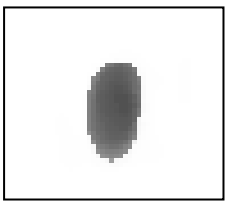
The traditional scoring method used by pathologists assigns a number to each staining intensity level, where (+) indicates weak staining, (++) indicates intermediate staining and (+++) indicates high staining. The percentage of positively stained nuclei is measured by manually counting the number of positive and negative stained nuclei objects. (Fang et al., 2003).

In this study, we use the visual scoring method as a benchmark to validate the results of the proposed automated method, the benchmark scoring results were provided by an experienced pathologist using the visual scoring method described above.

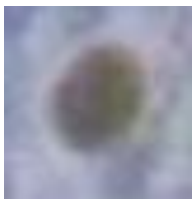
1.2 Overview of the thesis

A number of attempts to automate the manual procedure of scoring IHC images have been introduced, (Choudury et al., 2010; Kuo et al., 2010; Gavrielides, 2008; Hatanaka et al., 2003). However most of the previous automated methods have scored the images based on the average tissue analysis rather than providing an individual cell solution. Other cell based solutions meanwhile only provide a semi-quantitative or a semi-automated method rather than providing a fully quantitative automated method.

Our proposed method provides a comprehensive cell–level automated method for the quantification of antibody staining based on the individual cell analysis. This approach tackles each cell individually, which allows the direct extraction of information from each cell. Previous cell-level methods provide images which are segmented based on rectangular and square shaped windows. This means the segmented nuclei object will include some noise corresponding to the background. Our proposed system resolves this issue by providing a precise means of segmentation which will only segment nuclei objects, discarding the background surrounding pixels. This will allow features to be extracted directly from the nuclei object without the interference of the background.



A) Segmented nuclei using our method



B) segmented nuclei using other methods

Figure 2-1: segmentation method comparison

In order to achieve this, an efficient means for nuclei segmentation and quantification is essential, which will then be used to assist the automated scoring of percent stained or the Proportion Score (PS) and the Intensity Score (IS). An automated approach for nuclei segmentation and quantification is implemented, which automatically counts both positive and negative nuclei, (generating the total number of nuclei in an image). The second stage is the classification of the positive and negative nuclei.

The outcomes of these two stages yield the elements required to calculate the proportion score (PS). Finally, a novel approach for measuring the Intensity Score (IS) of heterogeneously stained positive nuclei based on a pixel level intensity classification is demonstrated.

The proposed method aims to provide an accurate automated scoring method for stained nuclei inside IHC images. This would ultimately reduce the human error and the time spent on analyzing stained biopsy samples from a few hours to a few seconds per slide. However, fully automated method for grading IHC images can only be designed to grade a specific type of IHC samples when implemented using fixed parameters. To increase the generality of such automated methods, Flexible parameters can be introduced in the implementation stage. In this case parameters can be adjusted to serve a wider range of IHC samples. In this study a number of parameters are defined at the implementation stage. The parameters were selected to work for all Ki67 and p27 images without any further adjustments by the user.

1.2.1 IHC samples

In order to develop this approach, the experimental model used is the multicell spheroid model of human colonic cancer. This is an in-vitro three dimensional model that mimics many of the features of solid tumour biology (Sutherland et al., 1986). This includes gradients of cell proliferation with dividing cells on the surface of the spheroid, where as those cells residing close to the Necrotic core are not proliferating. Dividing plus non-dividing cells can be visualised using antibodies Ki67 (Scholzen and Gerdes, 2000) and p27 ^{Kip1} (Karen et al., 2004) respectively as shown in figure 1-1.

Positive protein expressions (Brown nuclei) are identified by biotinylated-labeled antibody and indicate the positive response to the antibody raised against human Ki67 and p27 ^{Kip1} proteins. Whereas, negative protein expressions (blue nuclei) are counterstained with haematoxylin solution and they represent the negative response to the antibody raised against human Ki67 and p27 ^{Kip1} protein. The main focus in this study is the quantification of positive and negative nuclei objects. Necrotic cells are eliminated and only cells in the viable rim are targeted.

The IHC samples were provided by The Institute of Cancer Therapeutics, University of Bradford. Details regarding the IHC samples are provided in the appendix. The details provided in the appendix are part of a master dissertation by (H. Y. CHI, 2006).

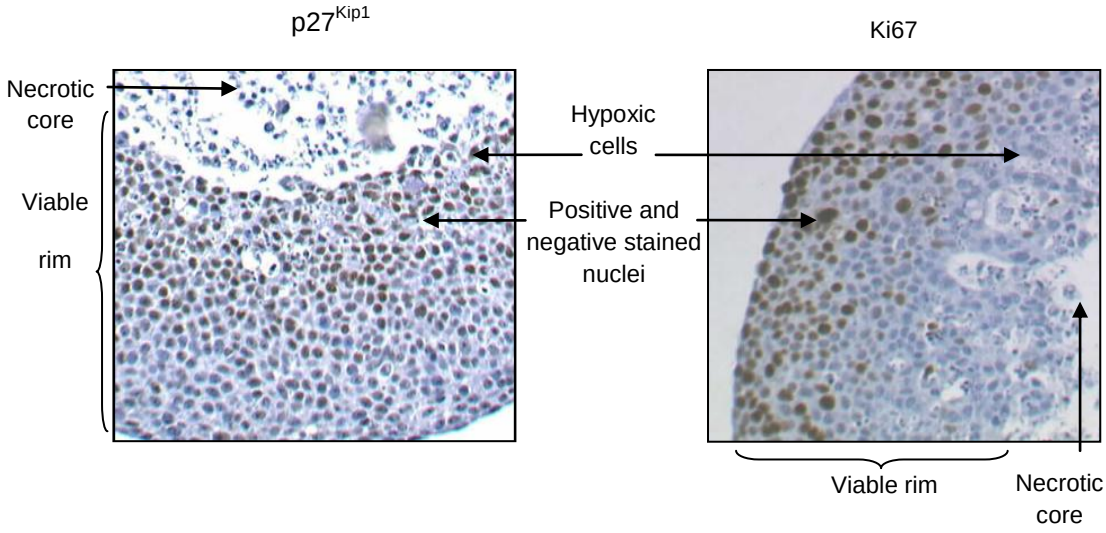


Figure 1-2: represents $p27^{Kip1}$ and Ki67 expressions in intact HT29 spheroids. Positive protein expression is identified by the brown stain indicting where secondary antibodies have bound to the primary. The brown colour is diaminobenzidine produced following the reaction with horseradish peroxidase conjugated to the secondary antibody. Basophilic structure of cell (blue) was counterstained with haematoxylin solution.

Such images were selected in this research as they represent the most common challenges faced in the automated quantification and analysis of IHC image tasks. Some of the main challenges faced in such analysis are presented below;

Despite active research and several attempts to automate the quantification process of cell nuclei, nuclei segmentation remains a challenging problem due to the variety and complexity of microscopy images across cell types and application frameworks, with mainly weak contrast level, touching nuclei, diffused background and changeable size and shape of cell nuclei all posing challenges to existing methods. Generality of existing methods is also a challenge. An illustration of these challenges is presented in figure 1-3.

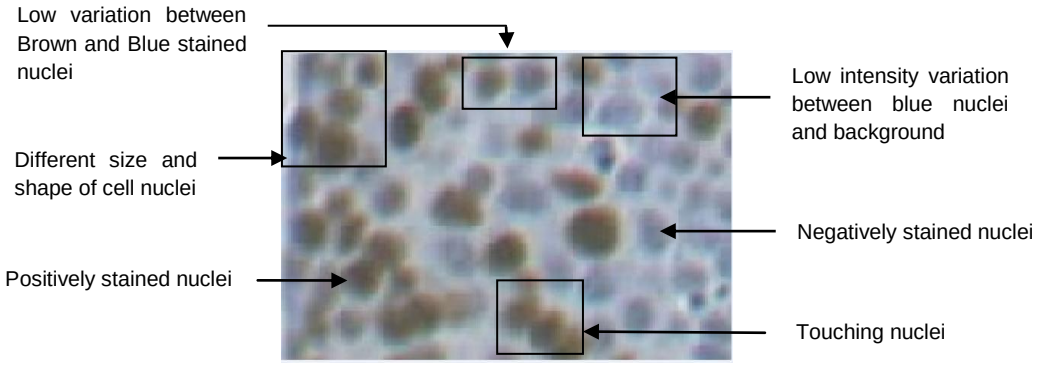


Figure 1-3 represents Ki67 expression in intact HT29 spheroids. Positive protein expression is identified by brown stain. Basophilic structure of cell (blue) was counterstained with haematoxylin solution.

A number of challenges are raised when attempting to analyse this type of image;

- Such kind of immunohistological images have low intensity variations between various substances within an image, and this makes differentiation among the objects (nuclei and background) a complex mission. This is particularly true of negatively stained nuclei objects (blue stained nuclei), which have a very low contrast variation against the background. This makes the process of extracting and segmenting this type of nuclei a challenging task.
- The morphological characteristics of the nuclei objects differ frequently, which makes the localization and quantification of nuclei, objects a challenging task.
- The quantity of brown stained nuclei compared to the quantity of blue stained nuclei provides a significant measuring tool for the determination of protein antibodies' activation inside cells. The computation of the Percent Stained (PS) is calculated as;

$$PS = \frac{The\ Number\ of\ Positive\ Nuclei}{The\ Total\ Number\ of\ Nuclei} * 100 \qquad (1\text{-}1)$$

This makes the distinction between positive brown stained nuclei and negative blue stained nuclei a crucial step in diagnosis and monitoring procedures. Following investigation, it was found that some pixel intensity levels of both positive and negative nuclei overlapped. This means that some positive nuclei intensity can possibly hold pixel values which represent negative nuclei intensity values. Consequently, this makes the process of distinguishing between positive and negative nuclei a complex task.

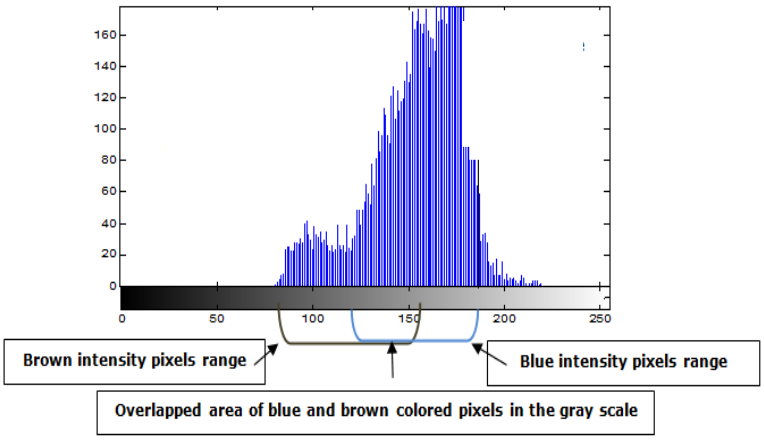


Figure 1-4: Overlapping of pixel intensity

- Intensity score (IS) is the means for measuring the staining intensity level of positively stained nuclei (brown nuclei). Heterogeneity of staining intensity amongst positively stained nuclei presents a challenge when attempting to set an automated measure. This is due to the absence of an existing threshold for the three intensity levels.

1.3 Aims and objectives

The main aim of this research is to provide a comprehensive cell-level automated approach for the quantification of antibody staining inside each nucleus cell in IHC images.

In this research, a cell-level analysis approach is targeted, rather than a tissue-level analysis. Using such an approach, nuclei objects are individually analyzed and the antibody staining is observed inside each cell. This will enable the extraction of features that are directly related to the nuclei object, unlike tissue-level analysis, where background and other cellular particles have an influence of the level of information retrieval and the final classification outcome.

To achieve the automated quantification of stained nuclei cells in IHC images a number of objectives would have to be met.

Nuclei segmentation and quantification: nuclei segmentation is the first task towards the cell-level analysis approach. This will mean segmenting the nuclei objects from the background and other cellular elements. Nuclei segmentation has remained a challenging task due to the variety and complexity of the nuclei characteristics in such images. In this regard, a number of image processing techniques are investigated to provide a structural approach for the nuclei segmentation task, tackling all the main challenges. Counting of positive and negative nuclei can be achieved upon successful nuclei segmentation.

Nuclei stain type classification: The second task is the classification of positive and negative nuclei in IHC images. This can be achieved based on a cell-level classification approach, where features are directly extracted to describe the characteristics of each nucleus. These features can then be used to classify both types of nuclei based on their discriminative characteristics. A huge number of feature extraction methods can be employed to extract features from nuclei objects. However, the investigation as to which of these features hold discriminative power is crucial as this determines the level of accuracy in the automated classification tasks. Upon successful classification, the PS can be calculated by combining the output of both nuclei classification and counting.

Positive nuclei heterogeneity classification: the third task towards the automated quantification of stained nuclei in IHC images is the measurement of the intensity score IS. Until now, the determination of the intensity level of positive stained nuclei has depended on giving a single threshold that described the intensity level of positive nuclei in an IHC image. However, a more sophisticated approach can be introduced by addressing the staining heterogeneity amongst positive nuclei in a single IHC image. To achieve this, a classification approach which classifies each positive nucleus based on its staining intensity needs to be introduced. This can be achieved by extracting discriminative features of each positive staining class (high, intermediate and low). However, highly responsive features will have to be extracted as the staining levels of positive nuclei are very minor and could prove difficult to define.

1.4 Contributions

The automated scoring method involves a number of novel algorithms which are exclusively designed to achieve the final goal of this research work.

Contributions can be outlined as follow;

- A segmentation and counting tool was developed for the automated quantification of stained nuclei objects in IHC images. A number of nuclei counting methods have emerged that provide an automated counting measure for stained nuclei objects. However, such tools have fixed parameters and only provide sufficient results for the set of images they are designed for. For this reason, developing our own general segmentation tool was necessary, as this allowed the extensive analysis of a diversity of IHC image datasets by adjusting the parameters as required. Designing our method has given us an insight into the challenges involved when developing such systems, whether due to the morphological complexity of nuclei or the efficiency of the technologies used for developing such tools. Understanding these challenges has enabled us to develop a suitable tool for this task, tackling the main challenges related to the characteristics of the IHC images and improving the performance of the existing technologies involved in the development of such systems.

- An image reconstruction method was introduced to resolve the pixel intensity overlapping between nuclei objects and the background. This

was achieved by extracting the global and regional maxima and minima from the IHC images directly, which were then used as thresholds to improve the contrast of differently stained nuclei objects and background.

- Watershed segmentations have been widely used for the task of segmenting overlapping and adjacent nuclei objects in microscopic image analysis (Cheng and Rajapakse, 2009; Cloppet and Boucher, 2011). However, the watershed performance is restricted by the quality of the images it is used for. For instance, when watershed transform was used on image datasets used in this study, it provided limited segmentation, such that overlapped and adjacent nuclei were not always segmented correctly. To improve the performance of watershed in such cases, we integrated the use of L*A*B colour space and morphological features which were defined through studying the nuclei shape characteristics closely. This has resulted in an improved segmentation performance.

- A nuclei extraction method was introduced, where each segmented nucleus was extracted and saved automatically in a separate image. This method was developed to create the segmented nuclei image datasets. This step is fundamental in achieving the cell-level image analysis of nuclei objects, by means of which features can be extracted directly from each nuclei object, and therefore classified accordingly.

- A cell-level classification approach was developed for classifying the positive and negative nuclei objects, based on colour, textural and morphological features. An extensive investigation was carried out to select only those features that provided discriminative power in differentiating between positive and negative nuclei. Another aim of this investigation was to study the influence of using integrated features extracted from segmented nuclei objects in providing efficient classification measures of the different types of nuclei cell objects. By using the output generated from the nuclei counting task and the positive and negative nuclei classification task, a new approach for measuring the PS was introduced.

- A cell-level classification approach was developed for classifying heterogeneously stained positive nuclei objects. Extensive experimental work was conducted in this task to study the influence of using textural, colour and morphological features in providing discriminative characteristics that describe the minor differences of each stain level (high, intermediate and low). Textural features based on the intensity histogram were developed to provide improved discriminative power compared to the baseline textural extraction methods. Finally, the IS was measured by providing cell-level details that presented the percentage of each positive stain type in each IHC image.

1.5 Thesis Organization

This thesis consists of six chapters which are organized as follows:

Chapter 1: this chapter includes introductory key points to the thesis, pointing out the significance of image processing techniques in the medical field, the importance of immunohistological stained images, a description of the manual process of measuring the antibodies in cancer microscopic images and how this process can be automated for an efficient grading tool. This chapter also includes an overview of the thesis and description of the IHC samples used in this thesis, it also includes the aims and objectives of the thesis, and finally the main contributions of the author.

Chapter 2: this chapter includes a comprehensive literature survey of the related research papers. The survey contains four main sections: nuclei segmentation in digital microscopic images, cancer cell feature extraction, classification approaches and finally IHC image scoring methods.

Chapter 3: a novel segmentation and counting method is presented in this chapter. The method is used to extract and segment nuclei objects form IHC images. This process allows the automated generation of the nuclei count, to yield the total number of stained nuclei in each input image, which then is used in the next chapter to calculate the PS. later in this chapter, an extraction method is introduced as a basis for creating the segmented nuclei image datasets. Results of the segmentation and counting method are presented, validated against ground truth pathologist data and discussed.

Chapter 4: this chapter demonstrates a novel approach for classifying positive and negative nuclei objects. Feature extraction and classification methods are applied on segmented nuclei images to achieve nuclei classification. The classification output of this process is used collectively with the output of Chapter 3 to generate the final PS of each input image. Two types of classifiers are compared to analyze their performance in nuclei classification tasks. Results are presented, validated against ground truth pathologist data and discussed.

Chapter 5: this chapter demonstrates a novel approach for measuring the intensity stain IS of positive nuclei based on a cell-level approach. A novel feature extraction method is introduced to measure the staining heterogeneity of positive nuclei. A classification method is then applied to make a decision regarding which intensity group a particular nucleus belongs to. Finally, the results are quantified to yield the final IS based on the percentage of nuclei belonging to each intensity group. The results are presented, validated against ground truth pathologist data, and discussed.

Chapter 6: this chapter includes the conclusion and future work. Suggestions as to how the current method could be improved further are also given.

Chapter 2

Literature Review

2.1 Nuclei segmentation in digital microscopic images

The term segmentation, when used in the image processing field means; separating the image into background and object of interest. The object of interest in this case is the nuclei. Nuclei segmentation plays a vital role in biological studies (Elowitz et al., 2002; Megason et al., 2007). Efficient segmentation of nuclei from microscopy images is an essential step in single cell measurement studies (Rosenfield et al., 2005). Segmentation methods usually involve a study of the object's shape, size and intensity level, and then use of these details to detect and identify the object of interest (Carpenter et al., 2006; Li et al., 2008; Wang et al., 2008; Wang et al. 2009).

A number of approaches for nuclei segmentation have been implemented. The approaches differ based on the purpose of use and nature of images.

Details about different segmentation approaches and how they have been utilised for this purpose are reviewed below;

2.1.1 Global and local thresholding methods

Global adaptive thresholding is an uncomplicated method, which can be used for conditions where nuclei objects are well separated from the background, and the intensity variation between nuclei objects and background is relatively high. (Otsu, 1979; Wu et al., 1995; Wu et al., 2000).

Mouroutis et al. (Mouroutis et al., 1998) used histogram thresholding to extract the region of interest (ROI). The threshold was set to the peaks of the histogram which correspond to the cytoplasm surrounding the nuclei objects. This type of segmentation sets a global threshold for the full image, subsequently dividing the image into background and ROI based on the intensity variation between the two. This means that the ROI will be successfully segmented from background only when the variation levels between the two regions are relatively high.

Nandy et al. (Nandy et al., 2012) presented an automatic segmentation of nuclei cells in 2D images. Automated thresholding was used to identify boundary regions. This thresholding method works based on the gradient magnitude differences of both nuclei objects and the background. As a result the boundary regions of the nuclei are extracted, yet again, segmentation efficiency is dependent on the level of intensity variation between the background and ROI.

Coelho et al. (Coelho et al., 2009) considered using three different thresholding methods for segmenting nuclear objects in microscope cell images. The first thresholding method used was Ridler-Calvard (Ridler and

Calvard, 1978), the second method used was Otsu (Otsu, 1979), and finally the mean pixel value thresholding method was used. The three threshold adjacent regions are considered nuclei objects. “Otsu and Ridler-Calvard” thresholding presented less accurate results when compared to the mean thresholding results. This is due to the presence of especially bright nuclei objects, leading the threshold methods to set a threshold between the bright nuclei objects and the less bright nuclei objects, instead of setting the threshold between the nuclei objects and the background. The mean thresholding presented better thresholding results in this case, where grey-scale images with high intensity variance between nuclei and background are used.

Hu et al. (Hu et al., 2004) calculated dual thresholds by iteration. The two thresholds T1 and T2, are firstly initialized to divide the whole grey image into three sections; R1, R2 and R3. The average intensities of the three sections of each threshold are then calculated to form the final threshold. The results present the nucleus regions in black, cytoplasm regions in grey and background regions in white. This kind of threshold presents improved thresholding compared to the global thresholding method in cases where an image is required to be divided into three regions.

Phansalkart et al. (Phansalkart et al., 2011) implemented an improved local adaptive thresholding method for segmentation and detection of nuclei objects in stained cytology images. The local mean and standard deviation become adaptive to the value of the local threshold according to the contrast in the local neighbourhood of the pixel. The value of the threshold goes

under the mean when the contrast in the local neighborhood is fairly low. This results in efficient elimination of fairly dark regions of the background, and therefore leads to better identification and segmentation of nuclei objects. This type of thresholding method provides improved segmentation results compared to the global thresholding method, and particularly in images with considerable background noise, low variation between ROI and background, and variation in contrast within the ROI.

2.1.2 Watershed segmentation method

Watershed segmentation methods are parameter free methods which are applied in complex image processing problems (Vincent and Soille, 1991; Meyer, 1994; Cloppet and Boucher, 2011). Watershed is implemented in open source software as well as commercial software. It is available in image processing packages such as Matlab® and ImageJ (Abramoff et al., 2004). The method has proved to be effective for separating overlapped and adjacent objects in images. However, over-segmentation is a common problem in this method, where a number of unnecessary segmentations are made between objects. Marker-controlled and hierarchical watershed methods are then often employed to resolve the over-segmentation problem (Beuchar, 1992; Beuchar, 1994; Najman and Schmitt, 1996). Other approaches use active contours, such as the snakes approach, in which objects are signified as smooth contours (Kass et al., 1988; Chan and Vese, 2001).

Cheng and Rajapakse (Cheng and Rajapakse, 2009) present a method to divide clustered nuclei objects from fluorescence microscopy cellular images, using a marker-controlled watershed algorithm. This algorithm demonstrates how the Euclidean distance can be improved using an outer distance transform to improve the segmentation of adjacent nuclei. The method produced smoother watershed lines separating the clustered nuclei objects when compared to the traditional watershed transform method. Cloppet and Boucher (Cloppet and Boucher, 2011) present a method to segment the overlapping nuclei cells. This method is based on the watershed segmentation algorithm. However, the authors demonstrate how some prior information about the nuclei morphological characteristics could help in optimizing the correct set of markers, to represent the starting points from which the watershed flooding will be initiated.

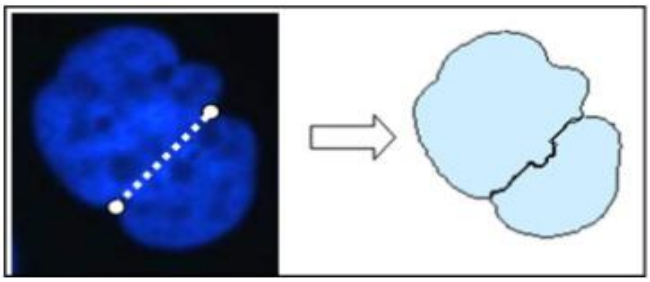


Figure 2-1: Segmentation results of the overlapping nuclei based on the watershed algorithm (Cloppet and Boucher, 2011).

From the literature reviewed above, it is evident that the traditional watershed transform algorithm failed to provide efficient separation of nuclei in a number of cases. To improve the performance of watershed, integration of other algorithms such as the nuclei cell morphology and improved Euclidean distance could provide enhanced watershed segmentation.

2.1.3 Morphological operations

Morphological operations are non-linear functions which are used to describe the shape and structure of objects in images (Meyer and Beucher, 1990). Morphological operations have always been a valuable tool in the processing of microscopic digital images: they are used to extract information concerning the image object by describing its geometrical structures in a prescribed way (Soille, 1999; Beucher and Meyer, 1993).

Nuclei cell segmentation is a challenging problem, due to the complex shape and structure of the objects. Morphological operations can be used in this case to regulate the shape of the cell objects. Some of the most commonly used morphological operations are; erosion, dilation, opening and closing operations, which are used to regulate the nuclei object (Comer and Delp, 1999; Goutsias et al., 1995; Louverdis et al., 2002). Dorini et al. (Dorini et al., 2007) used simple morphological operations along with scale-space properties to improve white blood cell segmentation. Cheng and Rajapakse (2009) used the active contour method to segment the clustered nuclei cell objects. Following this, a number of morphological operations were performed to refine the irregular nuclei shapes. Firstly, the holes inside the segmented nuclei are filled using the dilate operation to prevent the incorrect separation of clusters, when the subsequent distance transforms is calculated. Small objects which are unlikely to be fragments of the nuclei objects are eliminated using the morphological opening. Although simple morphological operations could refine irregular nuclei shapes, leading to improved segmentation, these do not provide the means for defining the

border around the nuclei objects. Such an operation is essential to defining nuclei objects, as it works as the first step towards the automated counting of nuclei objects.

2.1.4 Edge detection methods

Edge detection is a signal processing algorithm common in artificial intelligence and image recognition programmes. Tabor et al. (Tabor et al., 2009) built a genetically encoded edge detection algorithm which programmed an isogenic community of E. coli to sense an image of light and communicate to differentiate and identify the dark and light edges: the results of the calculation were then presented visually. In order to implement this algorithm, the authors used multiple genetic circuits in which they differentiate between the light and dark areas by using an engineered light sensor. In the dark, a diffusible chemical signal that diffuses into light regions is produced through the cells. Their idea was that only cells that sense light and the diffusible signal should produce a positive output, and to achieve this they used “genetic logic gates” The performance of the whole program is predicted using a mathematical model built from the first values and parameterized measurements of the component circuits.They proposed an accurate model which could assist in the engineering of more complex biological behaviours.

Aarnink et al. (Aarnink et al., 1994) proposed an automated method for identifying the contours of the prostate in ultrasonographic images. In this

proposal, a specific kind of edge detection based on nonlinear Laplace filtering is used. Edges are found at zero-crossings of the second derivative of the image.

The gradient value reflects the strength of the edges at that position. The intensity of the edge from the initial ultrasonographic image is calculated based on a combination of information taken from the location and strength of the edge. Edge enhancement is also used to identify and enhance the edges.The edges which represent boundaries are linked using a technique called “interpolation” which is used to fill the gaps between identified boundary edges.

Kothari et al. (Kothari et al., 2009) present a semi-automatic method for cell cluster segmentation and cell counting of digital tissue image samples. The method used consisted of three key steps. The first step was a pre-processing stage on the RGB images, which was performed to detect the boundary of the nuclei objects using a binary mask. The second step was performed to detect the points where nuclei objects overlapped. Finally, the third step was performed to segment the connected nuclei objects using an ellipse-fitting technique. Once the connected nuclei objects are segmented successfully, the separated nuclei objects are counted to give the final nuclei count. Conventional methods such as edge detection segmentation (e.g. Canny, Sobel and Prewitt; Kanopoulos et al., 1998; Ali and Clausi, 2001) and active contours were previously used for edge detection and localization (Bamford and Lovell, 2001). Further improved methods have also emerged, that have proved to produce enhanced results when applied to

Immunohistological images. One of these methods demonstrated the selection of tumour areas and segmentation of nuclear membranes in tissue using a morphological approach rather than the traditional active contour method. This approach has been demonstrated to improve results when applied to immunohistochemistry IHC of lung cancer tissue cells. The approach seemed to accomplish a promising nuclear segmentation. Nevertheless, it did not provide a quantitative description which gave automatic counting of the differently stained nuclei present within the same image (Di Cataldo et al., 2010). On the other hand, another investigation concentrated on the development of an automatic algorithm for counting positive and negative cancer cells collectively in immunohistological stained slides from breast cancer tissue. The results demonstrated a method for counting positive and negative cells using neural networks and morphological operations. However, an automated separation and counting method for the different colour stained cells individually was not achieved through this method (Phukpattaranont and Boonyaphiphat, 2007).

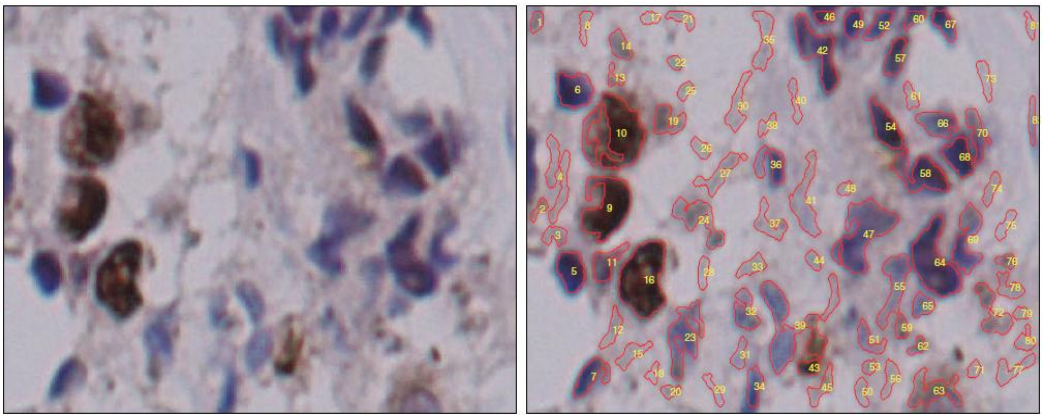


Figure 2-2: Results of the nuclei edge detection based on local adaptive thresholding (Phukpattaranont and Boonyaphiphat, 2007).

2.2 Feature extraction

The performance of automated analysis of cell level microscopic images depends profoundly on the features extracted to represent the characteristics of the nuclei. The foremost features that signify the nuclei objects are; textural, colour and morphological features. These features have been explored through various studies in an attempt to investigate and determine the optimal nuclei features. Diverse methods have been employed to extract nuclei features (Liu, 2011) and use a set of features to classify cell nuclei in fluorescent microscopy images. A variety of features such as morphological features, Zernike moments, (Boland and Murphy, 2001), Haralick texture features (Doyle et al., 2007) and wavelet features (Doyle et al., 2008) were used to extract features of the cell. Plissiti et al. (Plissiti et al., 2011) used a combination of shape, texture and intensity features for cell nuclei extraction in Pap smear images. To define shape features, the authors used a number of cell properties such as circularity, perimeter, eccentricity, and the major and the minor axis length. Statistical properties of the intensity histogram were extracted to define the textural features. Finally, intensity features were extracted based on the average intensity value of the pixel. This approach had the advantage of integrating a variety of features concerning the morphological characteristics of the nuclei as well as the textural and intensity features. However, colour features were not considered in this study.

Plissiti et al. (Plissiti et al., 2011) proposed an approach for extracting the colour and textural features in the spectral domain to recognize centroblasts

in histopathological images. The variation of the power spectrum along the frequency scale was used to extract textural features of the centroblasts. Principal component analysis (PCA) was used to extract colour texture features in the spectral domain. Finally, a quadratic discriminant analysis (QDA) classifier was used to classify the centroblasts and the non-centroblasts. Although colour features were considered in this study, PCA was used to extract the colour features, which was shown to produce ineffectual results when compared to Discrete Wavelet Transform (DWT) for colour features extraction. The use of DWT of different types in image feature extraction is reviewed below;

Discrete Wavelet Transform (DWT) is a multi-resolution technique used for various image processing purposes such as image compression, texture and colour analysis (Daubechies, 1990; Kokare et al., 2007; Sun and Ozawa, 2003). It has also been utilized in the medical imaging field. Gabor Wavelet was used to extract features from mammograms (Buciu and Gacsadi, 2009). The features were used to discriminate benign and malign tumour types in mammogram. PCA was used to reduce dimensionality. Finally, SVM was selected for classification. The author concluded that wavelet features seem to produce better discriminative power than features extracted by using PCA. Niwas et al. (Niwas et al., 2010) shed light on how the wavelet transform has become a valuable tool for numerous biomedical image and signal application tasks (Ma and Manjunath, 1995; Mabrouk et al., 2005; Selesnick et al., 2005). The author used the complex wavelet transform on nuclei objects of breast cancer images to extract textural features. Statistical

textural features were then used to differentiate malignant samples from benign in cytological breast cancer images. The study demonstrated how DWT could be used to extract textural features. However, DWT could be used to extract colour characteristics of the nuclei under appropriate colour spaces.

Wavelet transform has the capability to decompose the signal to a number of coefficients, each represented by a vector. The vectors hold information that describes the coefficients at different levels. The high level coefficient holds global features, while the low level coefficient holds local features (Kokare et al., 2007, Serrano et al., 2004). Gupta et al. (Gupta et al., 2007) utilised the wavelets functionality to extract features from nuclear receptors. The features were extracted based on wavelet variance over the different levels of coefficient, which describes the seven important physicochemical properties of amino acids. In this study, wavelet transform has been used to analyse both grey scale and colour images. Chen et al. (Chen et al., 2001) used wavelet transform to extract grey-scale and grey-scale gradient features from grey scale cancer cell images, to aid in the diagnosis of lung cancer. The features are then fed into a neural network classifier to classify the samples. Ma (Ma, 2009) represented an approach for colour image retrieval based on HSV colour space and wavelet to extract features. The method works by building a dimensional feature vector that represents the colour features. Textural features based on wavelet transform were also extracted. Experimentation disclosed that texture features based on wavelet provided more effective results. This, and previously mentioned studies,

demonstrates the discriminative power which wavelet features can provide. This has made DWT one of the most influential tools in providing discriminative features in nuclei cell classification tasks. However, the use of the appropriate decomposition levels play a vital role in the selection of the DWT coefficients, as these levels determine the quality of discriminative features provided by the DWT.

Another factor that could determine the performance of the extracted features is the type of classifier used. The use of Support Vector Machines (SVM) and artificial neural networks (ANN) is presented in the following section. These two types of classifier are reviewed in more detail as they are the primary classifiers used in the conducting of experiments in this thesis. Other classifiers are briefly mentioned in the next section.

2.3 Classification methods

Support Vector Machines (SVM) are a well established algorithm which recently provided efficient output when used to structure machine learning models based on large feature vectors, (Cristianini and Shawe-Tayor, 2000; Hearst, 1998; Mjolsness and DeDoste, 2001).

Machine learning algorithms have also become an important tool in medical image processing applications. Such algorithms have been used in various implementations to enhance and improve the classification, detection and tracking of cancer cells.

SVMs have provided a high classification efficiency in numerous biomedical image processing tasks such as; blood cell and tissue cell analysis (Wang, 2003), monitoring cell density and viability on the basis of dark field microscopy (Wei, 2007), cell cycle phase identification (Wang, 2008), and classification of sections of brain tumours (Glotsos, 2004). When compared to other similar approaches, SVM supplies improved results, particularly when used in real world applications (Cristianini and Shawe-Tayor, 2000; El-Naqa et al., 2002; Nattkemper et al., 2003; Han et al., 2007; Wei et al., 2011).

Wei et al. (Wei et al., 2011) presented a method for blood cell counting using a multi-class support vector machine. Segmentation of blood cells was performed based on morphological and thresholding methods. Once the blood cells were successfully segmented, features were extracted from segmented regions to identify the blood cells: features such as circularity, area and cytoplasm ratio were selected as features. Finally, SVM classification was applied to classify positive and negative cells. In previous work, Wei et al. (Wei et al., 2008) also presented a classification method using two SVMs to classify dead cells from live cells and separate cells from background. In both studies, the authors used a regional base feature extraction rather than using object feature extraction approach. The use of regional based feature extraction methods describes features of the ROI along with background and other regions included in the region, which could affect the classification accuracy of the objects. The authors also relied mainly on extracting morphological characteristics of the nuclei, disregarding

the valuable details that could be extracted from the colour and textural characteristics of the ROI.

Kumar et al. (Kumar et al., 2011) presented a support vector machine based method for the detection and removal of artefacts in cervical cytology images. The pre-processing stage was performed to remove background noise and enhance the image in the RGB colour space. A Gaussian loG edge detector was used to detect the edges by defining the zero crossing points subsequent to filtering the image with a laplacian of Gaussian filter. For feature extraction, a combination of ranked features were used, such as; the densitometric feature of Integrated Optical Density IOD (Bengtsson, 1999), the Fourier coefficient, and the morphological characteristics of nuclei regions. SVM was then applied to classify the epithelial cells and artefacts. The features selected in this study targeted features describing the density and morphological features of nuclei objects and cellular artefacts. Fourier coefficients were selected in the top feature ranking. However, Fourier coefficients have been shown to provide less discriminative power when compared to DWT coefficients in object recognition and classification tasks (Apatean et al., 2007).

Glotsos et al. (Glotsos et al., 2004) applied support vector machine clustering (SVMC) for the classification of nucleus and non-nucleus objects based on a pixel level classification. The method was applied to hematoxylin and eosin stained microscopic images of urinary bladder cancer. Active contours were used for edge detection subsequent to the classification procedure. However, analyses of images were performed at the grey scale rather than

taking benefit of the colour information of hematoxylin and eosin stains which could provide discriminative power for the classification of both stain types. Integrating such features can improve the level of classification accuracy, as will be demonstrated later in this thesis.

Kong et al. (Kong et al., 2011) utilized a method focusing primarily on colour transformations for the segmentation of microscopic cells in histopathological images. In this method, the training data is created by defining learning patches to train the algorithm of which particles represent the nuclei and which particles represent the background and other cellular regions. The procedure is performed by manually pointing 150 positions of the microscopic image, which include both nuclei and other extra-cellular particles. Each of the pointed locations is then cropped using an 11×11 local neighbourhood window, and used as a training patch. However, the training data set is performed using the manual procedure of a cropped window image which represents the nuclei as well as the extra- cellular particles. This again allows the classification of objects based on the collection of regional features extracted from the images.

Another mechanism that has been widely used for classification is artificial neural networks (ANN). Artificial neural networks work based on a parallel structure, which gives them the ability to learn from experience. The learning experience is acquired via a connected form of weights. The weights are the functions connecting the neurons, which then make a decision to classify the input (Al-Timemy et al., 2009). ANN could provide a good level of accuracy when used on classification tasks. However, classification performance is

heavily dependent on the training sets provided by the user (Al-Timemy et al., 2008). ANNs have been used in tasks such as disease risk estimation (Chester, 1993) and tumour growth models (Kazmi et al., 2012).

A review off ANN applications in the medical field was carried out by Papik et al. (1998). Certain radiologists have employed artificial neural networks to identify benign or malignant breast tumours. This kind of ANN has been developed by the American College of Radiology. The network was trained and tested via 206 cases, in which 77 were malignant and 133 were benign (Baker et al., 1996). Most neural network approaches for nuclei detection focus on investigating the pixel intensities of image blocks. This approach works by determining which blocks of pixels contain cell nuclei and which blocks represent the background. An example of this approach was presented by Phukpattaranont and Boonyaphiphat (Phukpattaranont and Boonyaphiphat, 2007), where an algorithm for segmenting nuclei from background in immunohistologically stained slides from breast cancer was conducted. The procedure for the approach consists of colour categorization using a neural network based on pixel colour contents and mathematical morphology.

A different approach to using neural networks was used by Lee and Street (Lee and Street, 2003). The learning procedure of the neural network was conducted using cluster shapes of nuclei rather than using the pixel intensities of image blocks. The hidden layer of such neural networks is increased every time a new cluster is detected. Each hidden node represents a cluster, which is then used to classify the nuclei objects. This

approach provides better segmentation results than previously mentioned neural network segmentation approaches. However, testing the method on larger data sets with a high number of cells could make this approach computationally expensive, due to the increased number of nuclei clusters in the hidden layer of the neural network.

Nattkemper et al. (Nattkemper et al, 2001) used an artificial neural net of local linear map (LLM) for the classification of fluorescent lymphocytes in tissue sections quantization. The algorithm works by manually selecting a set of cell patches. These patches are then used to train the neural network. The fluorescent cells are then evaluated based on their position, number and phenotype. The system provides a high level of cell detection. However, the system involves manual interaction with the user for the selection of cell patches.

Other classification methods such as Fuzzy C-means (Wang and Kumar, 2004), K-means (Raghavendra et al., 2010), Genetic Algorithms (GA) (Lukac et al., 2004; Huang and Hung, 2012) and Fuzzy Support Vector Machines (FSVM) (Zhang et al., 2012) have also been applied in the imaging classification field.

Fuzzy measures work by identifying the relationship between the various features that illustrate the properties of the images. Plissiti et al. (Plissiti et al., 2011) provided an automated method for nuclei detection in Pap smear images. The classification stage was performed using an unsupervised fuzzy C-means and supervised SVM classifiers. When using fuzzy C-means for

classification, the method did not require the implementation of a training set. On the other hand, SVM classification requires a training set, which in this case was performed using 34 images randomly selected form the dataset. Comparative analysis showed that the classification performance of SVM provided improved results when compared to the fuzzy C-means classification results.

Genetic algorithms can be used for a number of tasks such as feature classification and feature selection. Wu et al. (Wu et al., 2005) present a multispectral system for the early detection of cervical cancer in microscopic smear images. A genetic algorithm GA was selected for the feature extraction and SVM was selected for the classification stage. The GA provided refinement of unrelated noisy features. In conclusion, various studies have shown that conventional classifier systems such as SVM provided improved performance when compared to other fuzzy and genetic algorithms. Fuzzy and Genetic Algorithms showed improved performance when used alongside SVM and neural networks, (Seetha et al., 2008).

2.4 IHC image scoring methods

Inspection of Immunohistochemistry IHC images and standard scoring methods is purely based on visual inspection in clinical practice, which makes the procedure time consuming and inaccurate (Cregger et al., 2006; Wang et al., 2001; Ciampa et al., 2006). Pathologists use a scoring method

based on visual assessment of IHC stained images, which mainly rely on two main characteristics;

1) The Percent Stained or Proportion Score (PS), which is the percentage of positively stained tissue that is identified by the pattern of stained cells.
2) The Intensity Score (IS), which is the overall staining intensity of the positively stained tissue in the IHC image.

The traditional scoring method used by pathologists assigns a number to each staining intensity level, where 0 indicates no staining, 1 indicates weak staining, 2 indicates intermediate staining and 3 indicates high staining. (Fang et al., 2003).

Allred et al. (Allred et al., 1998) developed an IHC scoring system. The system works by combining two characteristics; stain intensity score (IS) and staining proportion score (PS). Each of the characteristics is then assigned a single value. Staining intensity is assigned a number between 0 and 3, where 0 indicates no staining and 3 indicates high intensity staining. The proportion score is given a geometric value where 0 = no stained cells, ≤1/100 cells stained =1, 1/10 cells stained = 2, ≤1/3 cells stained = 3, ≤2/3 cells stained = 4, and all cells stained =5. The staining intensity score and proportion score are then added up using the following equation; PS + IS, and the score outcome is a value ranging from 0 to 8. The Allred scoring system has an advantage over traditional manual scoring methods, as it is able to get a single scoring value describing both intensity and proportion score. However,

the proposed scoring method remains manual and fully reliant on the visual assessment of the pathologist, this is time consuming process and it can also lead to inaccurate scoring when there is low positive staining, due to the compression part of the scoring system.

Hatanaka et al. (Hatanaka et al., 2003) performed a semi-quantitative analysis for the examination of nuclear staining intensity and correlated the results with the HSCORE method. 1000 cells were randomly selected from five areas taken from an IHC breast cancer image. HSCORE was used to calculate the percentage of positively stained nuclei using the following equation; HSCORE $=\sum(i * Pi)$, where Pi is the percentage of positively stained nuclei and i is their staining intensity. i is assigned a single value ranging between 1 and 3 to define the intensity of staining, and Pi is assigned a value ranging between 0 and 100% to define the percentage of positively stained cells. The score outcome is a value ranging from 0 to 300. Commercially available software called WinROOF® image analysis (WinROOF, N.D.) was then used to count the stained positive and negative cells**.** Labelling index (LI) represents the percentage of positively brown stained nuclei in the image. This was performed by comparing the positively brown stained nuclei against the total number of nuclei in the image (brown positive and blue negative nuclei combined). The authors proposed two threshold methods to distinguish the positively stained brown nuclei. The first threshold extracts the nuclei presenting different staining intensities ranging from 1 to 3 (strong, intermediate and weak intensity), while the second threshold extracts only the nuclei showing strong staining (staining intensity

=3). The first threshold method seems to provide better results in determining hormone receptor status. The authors provided a semi-automated method based on cell level analysis. However the authors used existing commercially available software for counting the nuclei in the IHC images. This makes the final output of the semi-automated system fully dependent on the output provided from the commercial software (WinROOF).

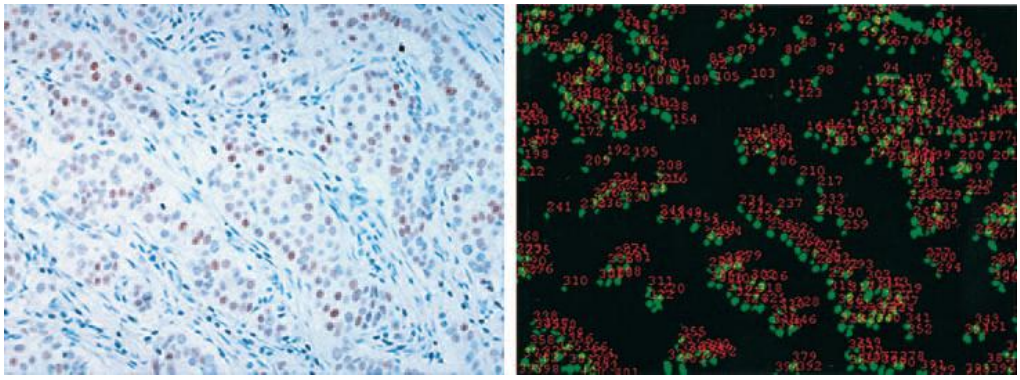


Figure 2-3: Results of counting the nuclei in the IHC image using the commercial software (WinROOF) (Hatanaka et al., 2003)

Choudury et al. (Choudury et al., 2010) presented an automated threshold method (ATM), based on the mean value of brown pixel intensity in the neoplastic area. The ATM is a single quantifying measure that includes both the proportion and intensity scores. This method has been developed mainly to allow for large-scale processing: however, it is appropriate in situations in which the analysis is prepared based on tissue area rather than measuring based on individual cells. This means that areas in the image above a given level of intensity, called the threshold (t), are regarded as stained, and the rest of the image is regarded as unstained. The percentage of staining in the image is then calculated by measuring the total area stained in comparison

to the rest of the image (the unstained area). Although this method is computationally inexpensive, it does not provide a solution based on individual cell analysis, and this could lead to less accurate results in the final scoring. The authors implemented a method based on a tissue level measurement rather than a cell level quantification. The method is based on an average thresholding technique, which could provide inaccurate results in cases where the targeted nuclei have low intensity variation with that of the background. In such cases, this method would result in sections of the background being considered as nuclei containing tissue.

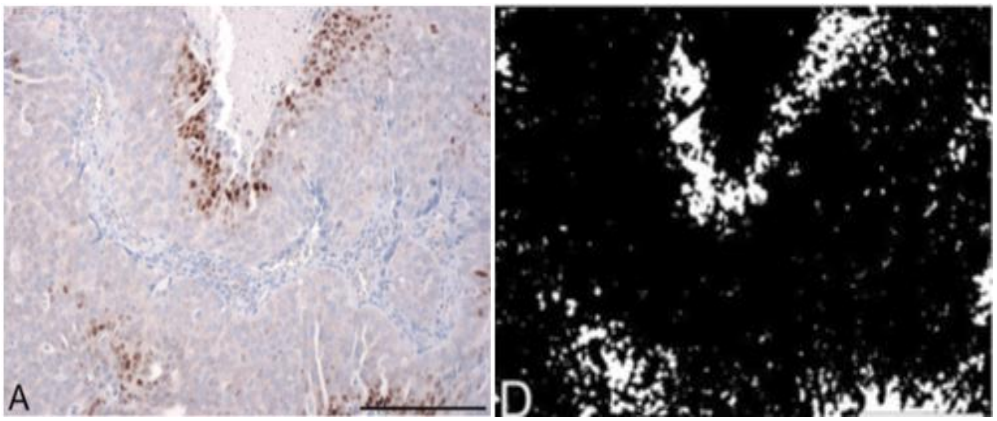


Figure 2-4: Results of the automated threshold method (ATM), based on the mean value of brown pixel intensity in the neoplastic area (Choudury et al., 2010).

This chapter has introduced an overview of the work related to cell segmentation, cell feature extraction, image classification and HIC scoring methods. The literature has provided a wide range of techniques and ideas which are associated with the work presented in this thesis. The majority of the reviewed literature focused on providing a solution to a minor section of the overall problem, rather than providing a complete solution for quantifying stained nuclei in IHC image quantification.

The proposed automated method in this thesis is the integration of discrete elements which are built up to form a comprehensive solution for quantifying stained nuclei in IHC images, tackling most if not all the challenges faced when building such systems. A limited number of studies have provided a complete automated method for measuring antibodies in IHC images (Choudury et al., 2010; Hatanaka et al., 2003). Such methods allow theoretical comparison with our proposed method. However, tissue-level measurement rather than cell-level quantification was proposed in previous methods. Our proposed method provides a cell-level based measurement, for the analysis and quantification of each individual cell within IHC images. This solution ultimately provides a high level of accuracy for the quantification of IHC images, as will be demonstrated in the following chapters. As part of the proposed method, a nuclei segmentation and counting method is proposed, designed and implemented. This method works as a flexible alternative to commercially used software which could restrict the research prospects due to its fixed parameters, and lack of literature resources and method descriptions. Our proposed system uses a novel extraction method: the segmented nuclei are used to form the datasets, which are then used in the classification step. This is a vital step as it determines the level of efficiency in the classification stage. Our proposed segmentation method also providing a precise means of segmentation by only segmenting nuclei objects, discarding the background surrounding pixels. This will allow features to be extracted directly from the nuclei object without the interference of the background.

Chapter 3

Nuclei Segmentation and Counting

3.1 Introduction

In this chapter, an automated approach for the segmentation and quantification of nuclei objects in IHC images is presented. The main aim of this chapter is to create an efficient automated method for segmenting nuclei objects (ROI) from the background and other cellular partials. The nuclei segmentation will serve two main purposes. First, it will allow the automated counting of each nucleus object in the IHC images. Secondly, it will allow the cell-level analysis of each segmented stained nucleus. Upon successful nuclei segmentation and counting the nuclei extraction method is presented. This method will be used to extract all the segmented nuclei from the IHC images, and nuclei objects are then automatically saved in a database to generate the segmented nuclei image datasets. The experimental work of this chapter is carried out using ImageJ. ImageJ is open source Java-based image processing software used for the analysis of medical images. To build up the methods described in this chapter, two plug-ins were written by the author using the functions available in the ImageJ Fiji library. The first plug-in was designed to perform the nuclei segmentation and counting procedure.

The second plug-in was designed to extract the segmented nuclei and form the segmented nuclei image datasets. The algorithms used to build up the plugins are described in details in the following sections.

3.2 Nuclei segmentation and counting

Nuclei segmentation is the process of separating the nuclei objects from the background and other cellular noise. This process will aid the automated counting of positive and negative nuclei collectively. The nuclei segmentation and counting method consists of a number of structural algorithms that work sequentially to achieve the final segmentation results. The structural segmentation model can be presented in the following flowchart: see Figure 3-1.

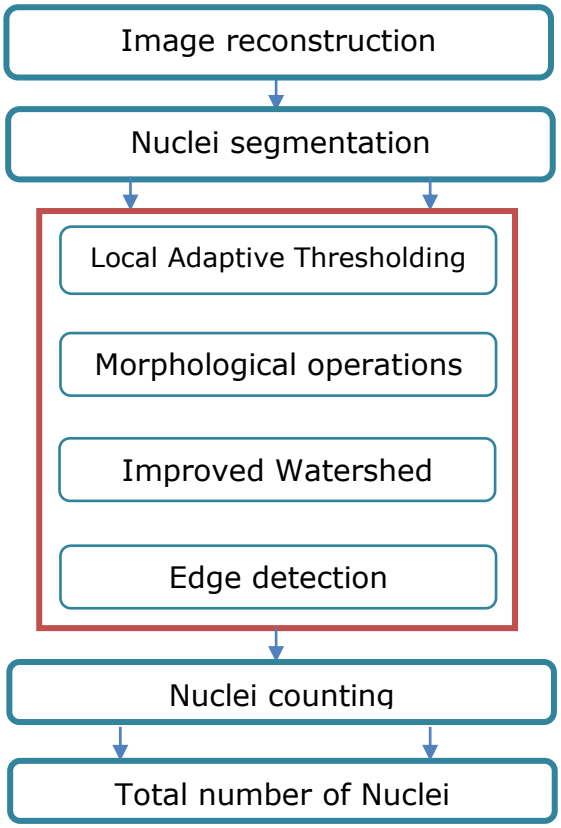


Figure 3-1: Flowchart of the structural segmentation model

The major concern at this stage is the automated counting of nuclei objects in an image. To achieve this, nuclei are detected comprehensively using a segmentation method to separate the nuclei from the other particles. The three main characteristics of our images are; positive protein expression (brown nuclei), negative protein expression (blue nuclei) and finally the background particles, which must be considered as noise to minimise false negative response. The process consists of two major steps: firstly, to detect the nuclei and separate it from the other particles in the image; and secondly, to count positive brown stained nuclei and negative blue stained nuclei collectively. This will aid in the automatic generation of the total number of nuclei in an image.

In this work we investigate images of colonic cancer spheroid slides stained with diaminobenzidine and haematoxylin, which were retrieved using a Leica light microscope fitted with QICAM digital camera and photographed at 20x magnification. The images are then saved as JPEG files into a computer for the image analysis. Three sets of colonic cancer spheroid images are used: Ki67 day 19, p27 day 19 and Ki67 day 22.

3.2.1 Image reconstruction

First we start with reconstructing the image using a novel method. This method is performed to increase the contrast variation level between the nuclei and the other particles within the image. This stage is vital in analysing images with low intensity variation between the nuclei objects and the

background. Since the intensity variation between the negative protein expression (blue nuclei) and background is very low, a false segmentation response could occur when some average values of pixels belonging to the background matches some average values of pixels belonging to the negative protein expression (blue nuclei).

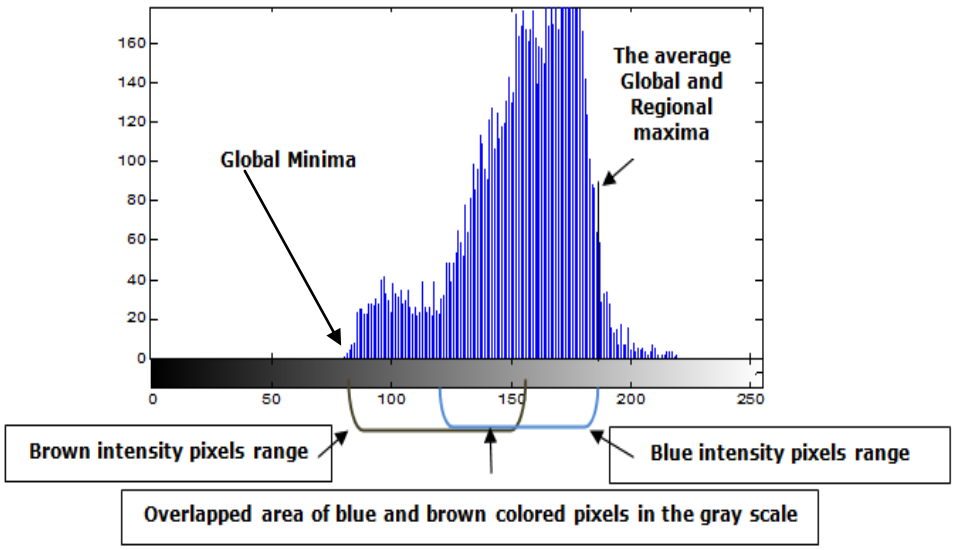


Figure: 3-2 Selections of Maxima and Minima

This would result in some parts of the background being identified as negative blue nuclei-like regions which would subsequently increase the level of segmentation noise. To avoid this occurrence we employ an exclusive contrast enhancer method based on the minima and maxima values of the brown and blue nuclei intensity in the grey scale. This is achieved by selecting the global minimum, which represents the darkest shade of the brown pixels' intensity and the average of global and regional maxima, which represents the average pixel intensity value belonging to the lightest shade of blue nuclei-like regions: see Figure 3-2. Improved intensity

variation levels between the background and blue nuclei are visible after employing the image reconstruction method: see Figure 3-3, 3-4 (A) and (B).

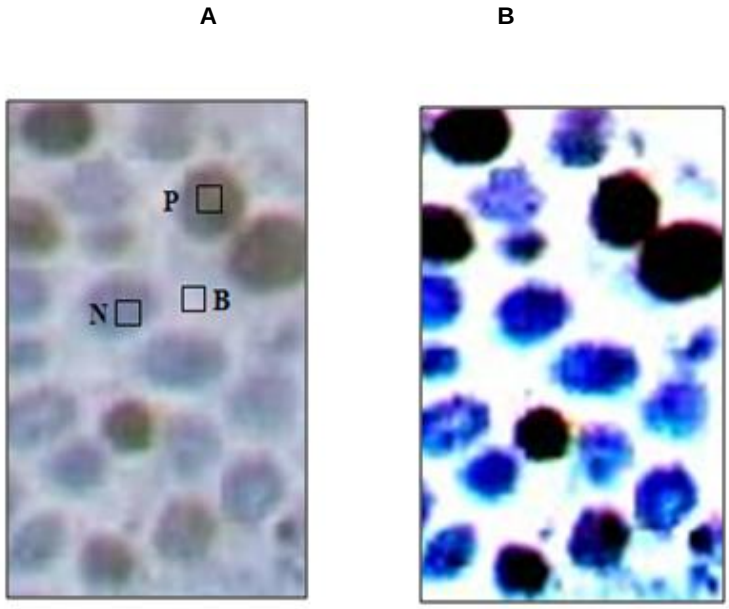


Figure 3-3: **(A)** the original IHC image in RGB colour space. **P** represents the positive brown stained nuclei-like regions, **N** represents the negative blue stained nuclei-like regions and **B** represents the background. Figure 3-3: **(B)** the image after increasing the contrast variation level between the nuclei and the other particles within the image using the image reconstruction method.

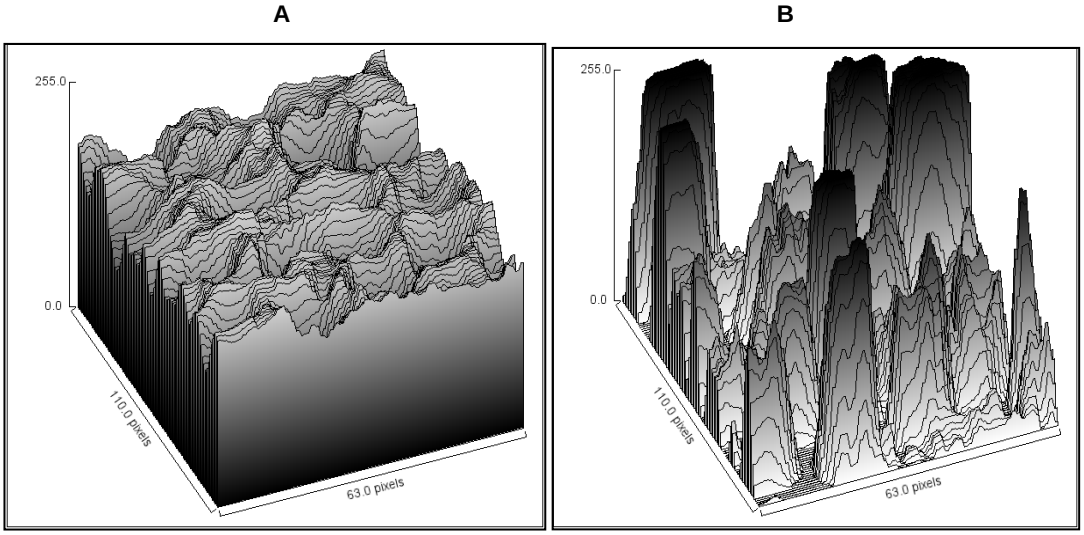


Figure 3-4: is a three dimensional representation of the IHC image surface presented in figure 3-3, prior to and after applying the image reconstruction method. Improved intensity variation levels between the background and nuclei objects are visible.

3.2.2 Local adaptive threshold

Thresholding is a method that is used to separate the ROI from the background. When using a thresholding method, pixels are divided into two categories; background, with a pixel value equivalent to “0”, and Regions of Interest, with a pixel value equivalent to “1”. A binary image is then formed with black and white pixel label (Zhao and Ong, 1998). Local adaptive threshold has revealed better segmentation performance than global thresholding (Mehmet and Sankur, 2004), since the pixels are labeled based on local statistics such as pixel variance and contrast values rather than the general thresholding of pixel values, which works based on the average intensity level of the whole image. Local thresholding works by calculating the average of the grey pixel values in some neighbourhoods based on a specific window size. If the pixel value is considerably darker than the average of the neighbouring pixels, the pixel is considered an object: otherwise, it is considered background. The preceding method is used to increase the contrast between nuclei and background particles. Nevertheless, complete elimination of background noise is not achieved, due to some pixel values of the background corresponding to some average values of blue stained nuclei pixels, occurring from inhomogeneous illumination and inconsistent staining. Consequently, local adaptive thresholding is used, which works based on the local distribution of intensities. The size of neighbourhood is particularly selected based on an 8x8 window, taking into consideration features of the blue stained nuclei-like

regions, such as contrast and intensity of pixels. Results are demonstrated in Figure 3-9 (A).

3.2.3 Morphological operations

Following effective separation of nuclei regions form background, morphological operations such as filling holes and opening are applied to refine the irregular nuclei shapes and eliminate noise (Plaza and Plaza, 2008). Morphological opening is an operation applied on the greyscale image after thresholding. The morphological opening operation consists of erosion followed by dilation (Droogenbroeck and Buckley, 2005). The morphological opening operation selected is based on a two-dimensional connectivity. Pixel connectivity is determined based on pixel neighbourhoods of 4-connected pixels: see Figure 3-5. This function is applied to remove small particles from the images. This is required to prevent noise objects being considered as nuclei regions. The opening operations depend on the size of the image and level of magnification. In our experimental work it is set to 40 connective pixels of a 4-connected neighbourhood. The parameter was selected based on the analysis of 5000 nuclei cells taken from various images at 20x magnification level. A hole-filling operation (Kwak et al., 2011) was finally performed to fill the gaps inside the nuclei objects: see Figure 3-9: (B).

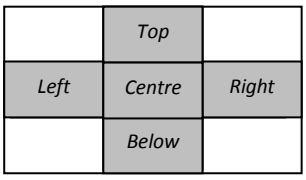


Figure 3-5: Pixel connectivity based on pixel neighbourhoods of 4-connected pixels.

3.2.4 Improved watershed transform

At this stage, watershed is performed to segment the overlapped nuclei. The watershed transform works based on a flooding mechanism, which is interpreted as water flooding the surface of an image until it reaches a local minimum. Watershed identifies the objects by calculating the Euclidian distance of each object within the image (the distance between the edge of an object and the edge of another object). Then it looks for the local minima between the overlapping objects to segment the objects at that point. However, watershed transformation does not provide a complete separation to all the overlapping nuclei (Malpica et al., 1997). This is due to the weak contrast between adjacent nuclei objects, and in particular brown positive nuclei. This makes it difficult to locate the local minima between the overlapping objects. Consequently, in such cases the watershed fails to separate these objects. To improve the performance of the watershed, RGB images are converted to the *L*a*b** colour space. The Euclidean distance of the overlapped nuclei seems to increase when using the *L*a*b** colour space, which then provides improved segmentation as the local minima between overlapping objects are easily located . Improved segmentation results are reported in Figure 3-6.

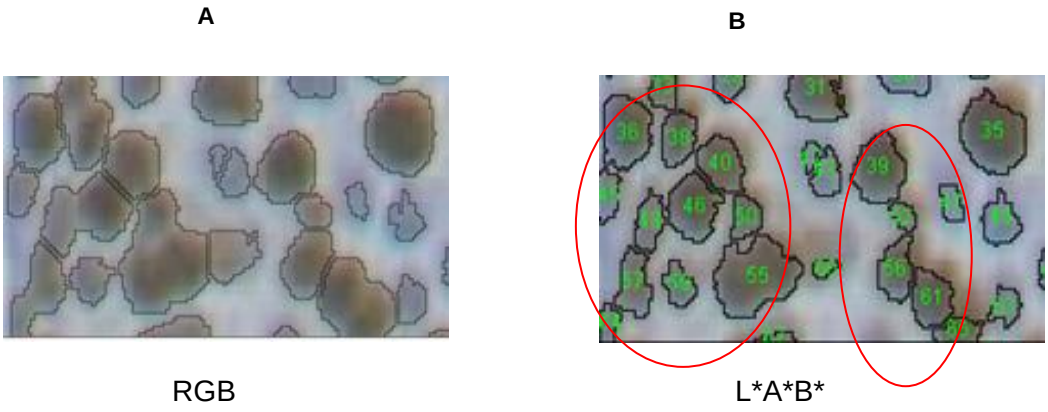


Figure 3-6: (A) represents the results of applying the watershed transformation, (B) represents the same image after applying the L*a*b* colour space prior to watershed. The results show improved segmentation and therefore improved nuclei counting.

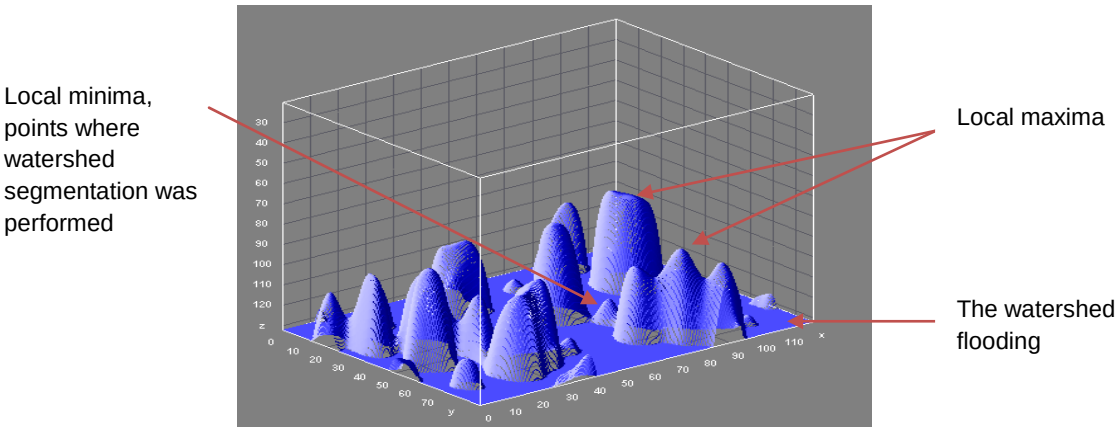


Figure 3-7: a 3D representation of the watershed transform operation on the RGB image.

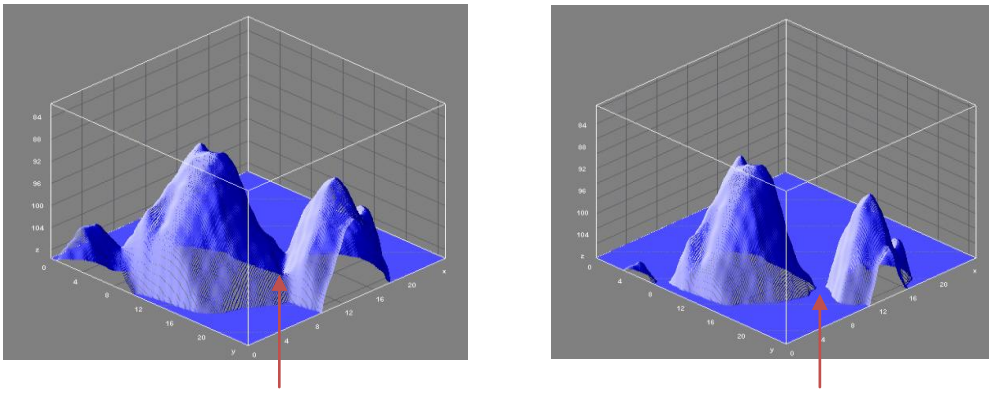


Figure 3-8: 3D representation of the watershed transform operation. (A): The watershed failed to segment due to the absence of local minima. (B) Successful watershed segmentation after applying the L*a*b* colour space. Increased Euclidian distance between the objects resulting in visible local minima.

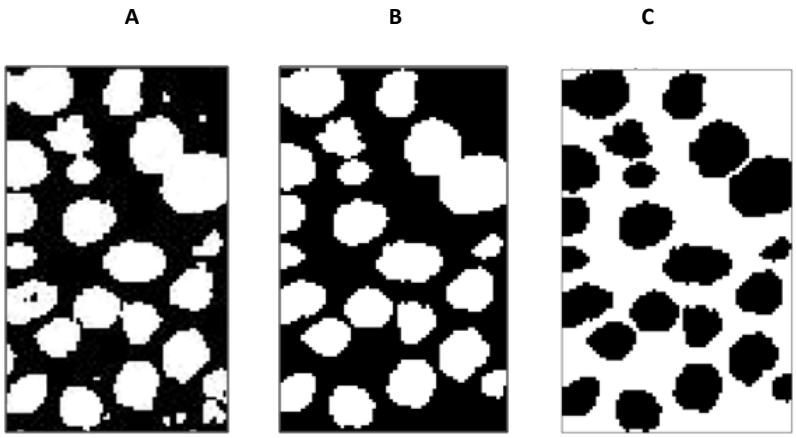


Figure 3-9: (A) Binary image of segmented nuclei after applying the local adaptive threshold. (B) Filling holes and opening morphological operations. (C) Watershed segmentation of overlapping nuclei.

3.2.5 Nuclei border detection and counting

Circularity analysis is applied at this stage in sequence to detect nuclei edges. Circularity grade is evaluated from 0 to 1 as shown in Equation 1 (Kuznetsov et al., 2011).

$$Circularity\ [0\text{-}1] = 4\pi \frac{A}{P^2} \qquad (3\text{-}1)$$

where P (Perimeter) is the distance round the boundary of each nuclei and A (Area) is the amount of space within a closed object representing the nuclei regions. A perfect circle has a circularity of “1”, which is unlikely in nuclei shaped regions, while a very irregular nucleus has a circularity value closer to “0”. Perimeter and area measurements are extracted from each nuclei object to allow the calculation of circularity for each object.

Table 3:1: represents the nuclei objects and the elements required to calculate the circularity for each nucleus.

No. nuclei	Area	Perim.	Circ.
1	266	62.527	0.855
2	157	48.284	0.846
3	116	45.355	0.709
4	239	56.77	0.932
5	130	43.799	0.852
6	294	63.698	0.911
7	57	28.627	0.874
8	82	34.728	0.854
9	183	51.113	0.88
10	47	27.213	0.798
11	213	58.77	0.775
12	47	25.556	0.904
13	156	46.284	0.915
14	148	48.284	0.798
15	152	47.579	0.843
16	132	41.432	0.965
17	129	42.027	0.917
18	174	56.954	0.673
19	174	60.77	0.591
20	49	28.867	0.738
21	112	41.23	0.827
22	139	46.442	0.809

When computing the circularity formula, a circularity parameter was not required at this stage. The parameter works as a filter to eliminate small circular non-nuclei objects. Since the elimination of small objects has been implemented using the opening morphological operation, the circularity was set to default (00.0-00.1, see Figure 4(A)). After applying the circularity measure, edges of nuclei objects are finally defined, and the closed objects within the perimeters are now counted individually and labelled to conclude the number of negative in addition to positive stained nuclei. (Fang et al., 2003): see Figure 3-10 (A-C).

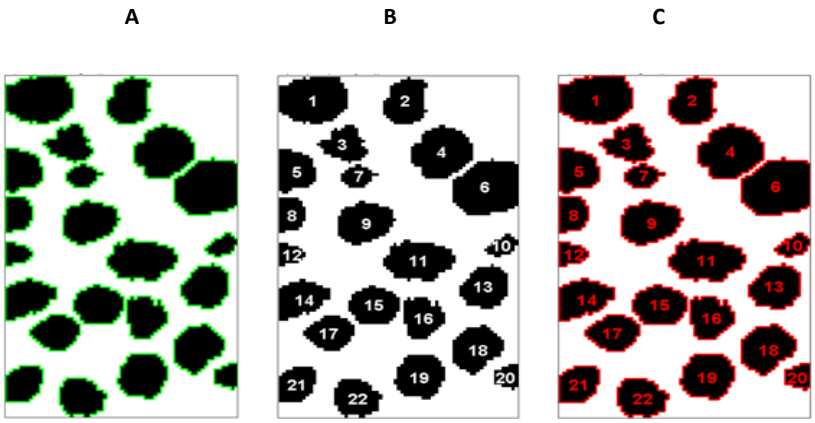


Figure 3-10 (A) Border detection of nuclei. (B) Total number of stained nuclei. (C) Representation of border and total number of stained nuclei.

The nuclei edges are now converted into a mask, the original image is then recalled and the mask is placed as a window on the original image. This step is performed to define the edges of the positive and negative nuclei on the original IHC image. The mask also includes the nuclei number labels. The following image represents the result of the automated segmentation and counting method.

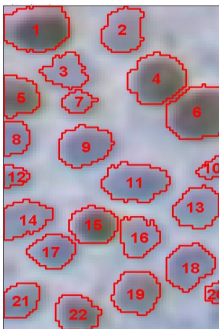


Figure 3-11: total number of positive brown stained nuclei and negative blue stained nuclei represented on the IHC image.

Result presented in figure 3-11, provides clear evidence of successful nuclei detection in addition to segmentation of nuclei regions from the background, which includes other cell particles such as cytoplasm and membrane. Segmentation at this stage provides the morphological characteristic representation (nuclei shape and dimension) of both positive protein

expression (brown nuclei) and negative protein expression (blue nuclei), devoid of any distinction in the monochromatic characteristic representation (nuclei colour). The product of this stage facilitates the automated quantification of both blue and brown stained nuclear collectively. This stage successfully provides the automated counting of nuclei cells in an IHC image.

3.3 Nuclei extraction and datasets generation

At this stage each nucleus needs to be extracted and saved individually as an image. This is an essential step in the analysis of the individual cell. This is performed using an exclusive extraction method. This method is computed as follows: Each labelled nuclei is saved as a ROI (R1,R2,.....Rn), and a command is then given to save each ROI as an individual image. Each nucleus object is then saved automatically to form the segmented nuclei image datasets. The same is performed on all the images used. A dataset including all the extracted segmented nuclei images is now generated.

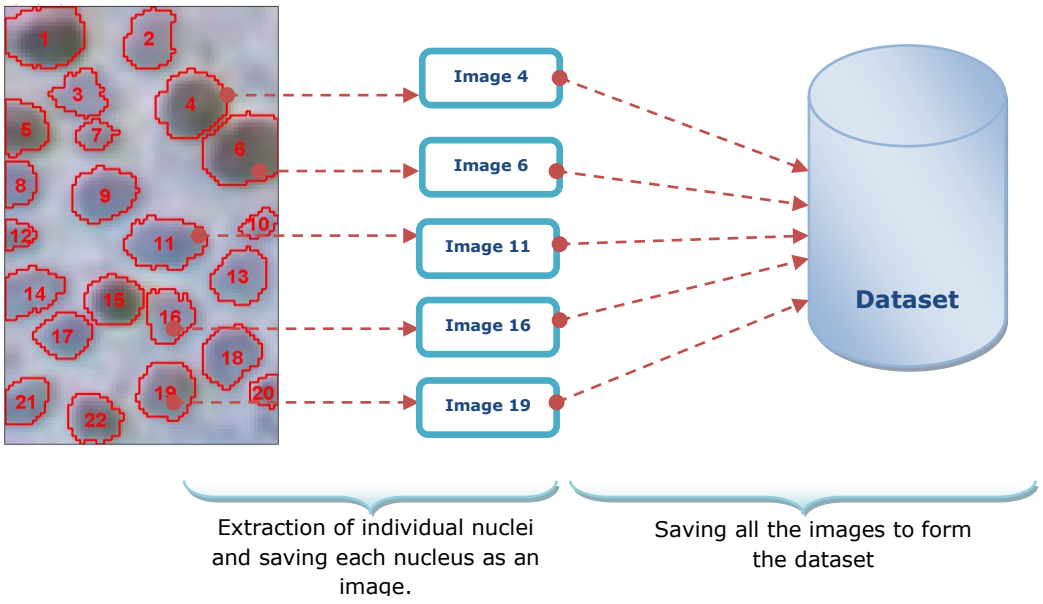


Figure 3-12: the nuclei extraction model

This process is the foundation for the individual analysis of nuclei cells. The image datasets generated will be used in various tasks such as nuclei feature extraction, and positive and negative nuclei classification, as well as the analysis of staining heterogeneity amongst the positive nuclei objects. This will be discussed in detail in the following chapters (Chapters 4 & 5).

3.4 Results and discussion

Image segmentation and counting was performed using the approach described earlier in this chapter. The method was applied to three different datasets of spheroid images; *Image dataset 1*: Ki67 Day 19, this set consists of 120 images (69 high staining images, 23 intermediate staining images and 26 low staining images). *Image dataset 2*: p27 Day 19, this set consists of 110 images (23 high staining images, 46 intermediate staining images and 39 low staining images). *Image dataset 3*: Ki67 Day 22, this set consists of 136 images (87 high staining images, 39 intermediate staining images and 8 low staining images). Images with different staining intensity levels from every dataset were used for the evaluation of the automated counting method. This is essential to assess the performance and generality of the automated method. 5 images from each set are presented in this section to demonstrate results of the automated counting method. Finally, generic data that include the results of all images from all datasets are presented to calculate the overall performance of the automated method. Results of nuclei segmentation and counting of Ki67 Day 19, p27 Day 19 and Ki67 Day 22 image sets are provided below in Figure 3-13 (A-C);

Image 1 (Ki67 Day 19 high staining)

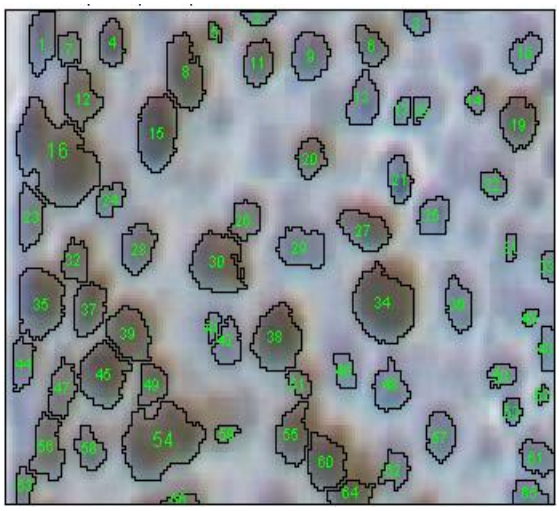


Image 2 (Ki67 Day 19 high staining)

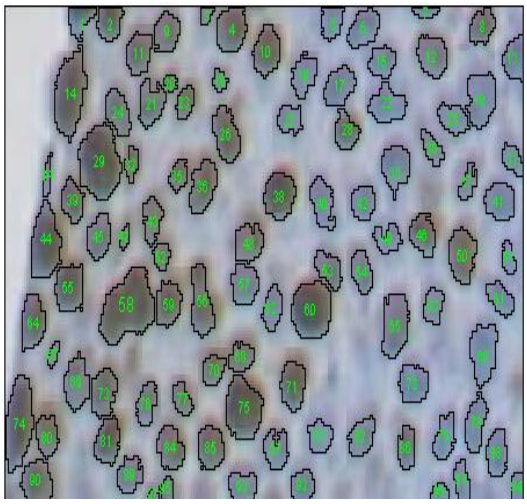


Image 3 (Ki67 Day 19 high staining)

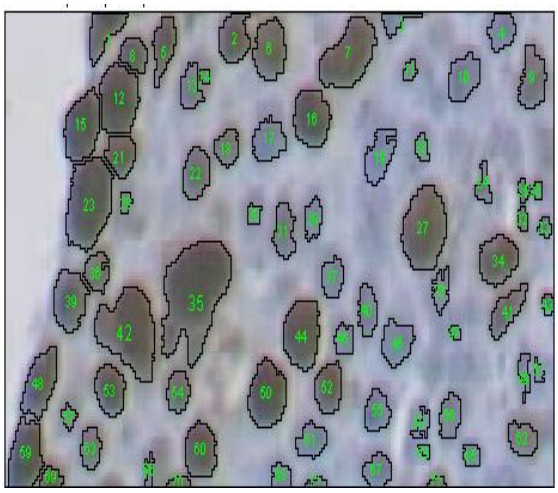


Image 4 (Ki67 Day 19 intermediate staining)

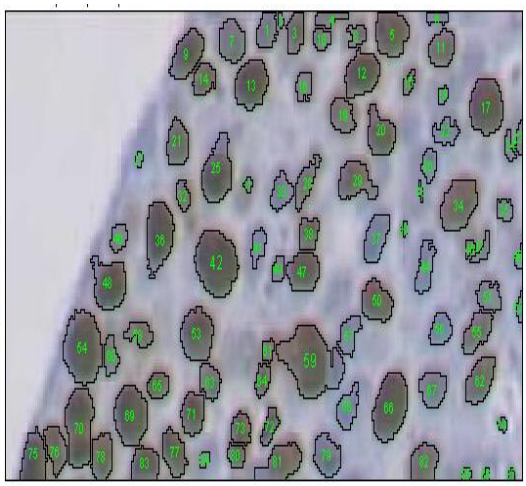


Image 5 (Ki67 Day 19 low staining)

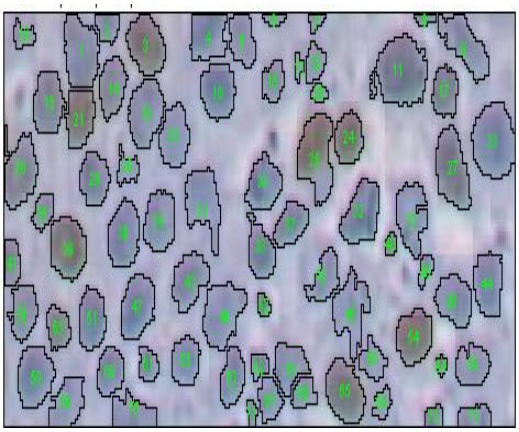


Figure 3-13: (A) represents the segmentation and counting results for the Ki67 Day 19 image dataset

Image 1 (p27 Day 19 high staining)

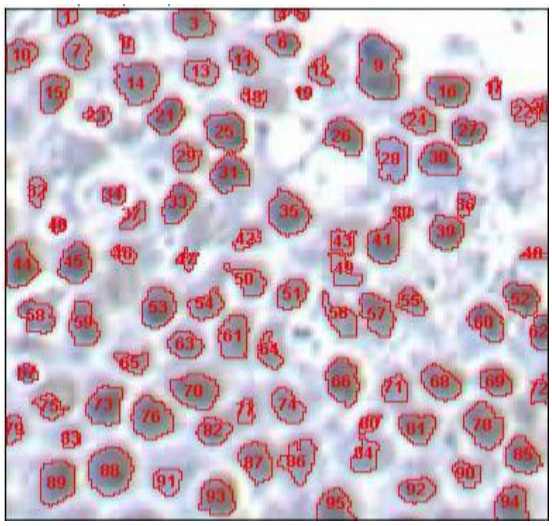


Image 2 (p27 Day 19 intermediate staining)

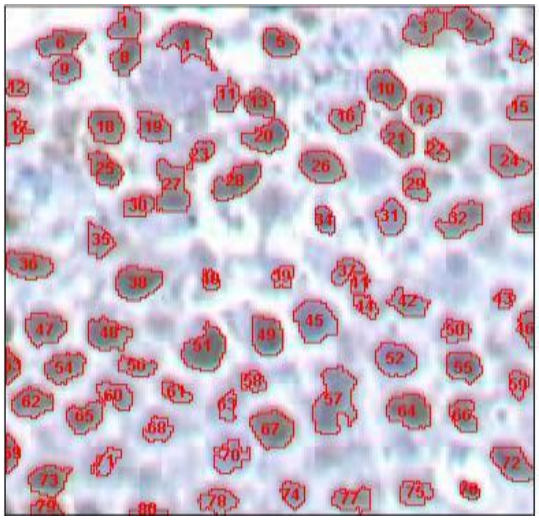


Image 3 (p27 Day 19 intermediate staining)

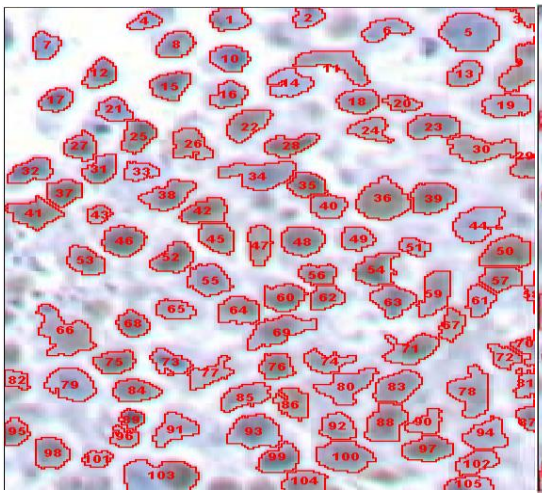


Image 4 (p27 Day 19 low staining)

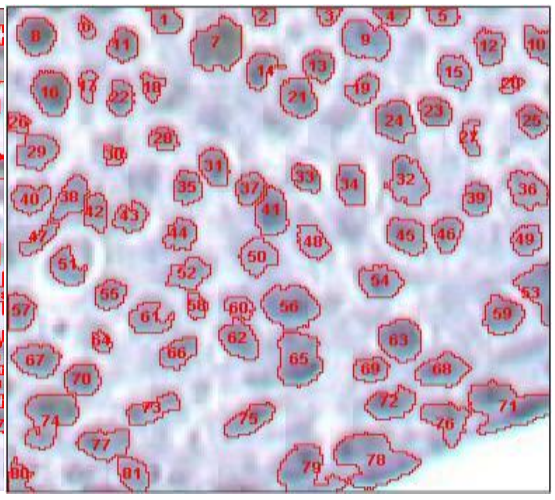


Image 5 (p27 Day 19 low staining)

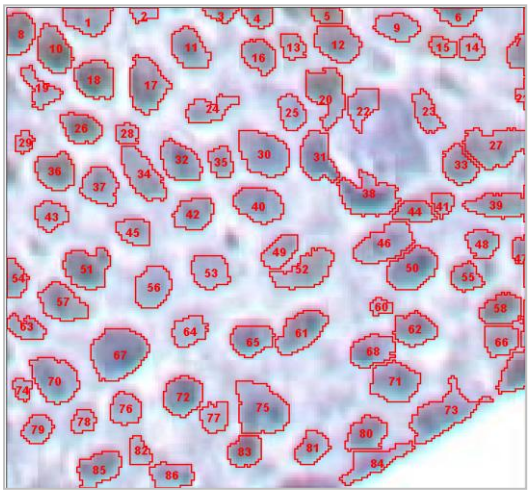


Figure 3-14: (B) represents the segmentation and counting results for the p27 Day 19 image dataset

Image 1 (Ki67 Day 22 high staining)

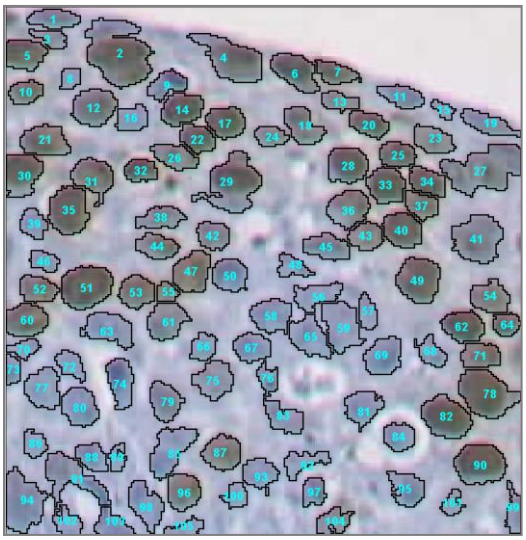


Image 2(Ki67 Day 22 intermediate staining)

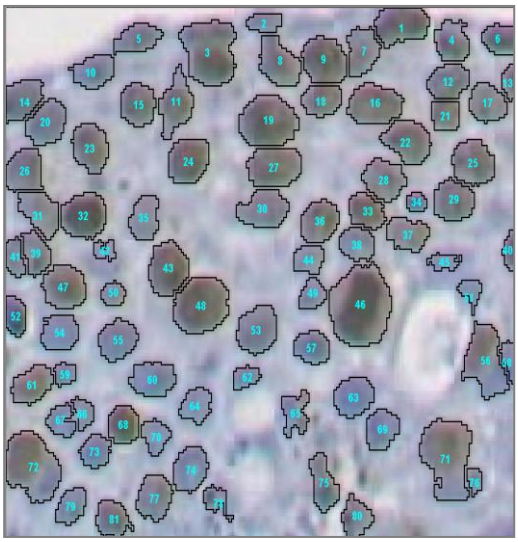


Image 3 (Ki67 Day 22 intermediate staining)

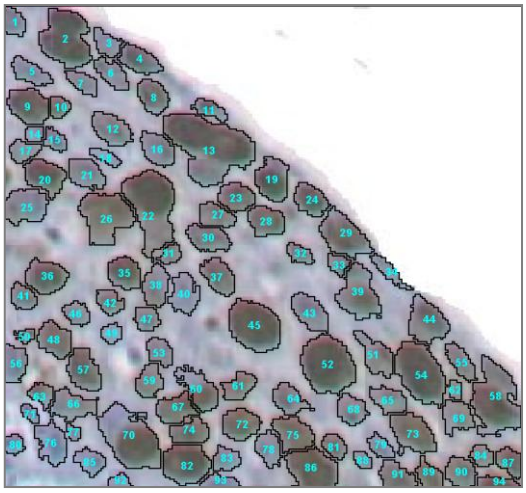


Image 4 (Ki67 Day 22 intermediate staining)

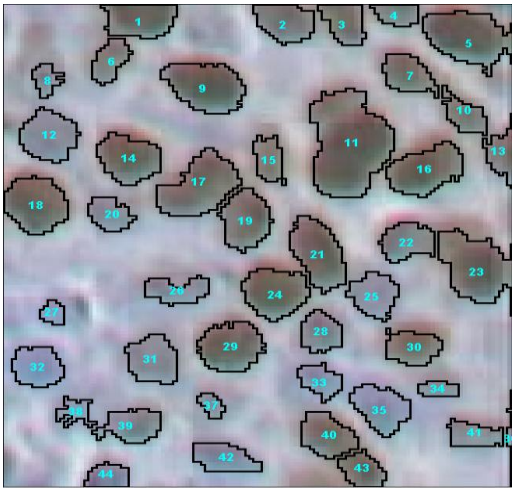


Image 5 (Ki67 Day 22 intermediate staining)

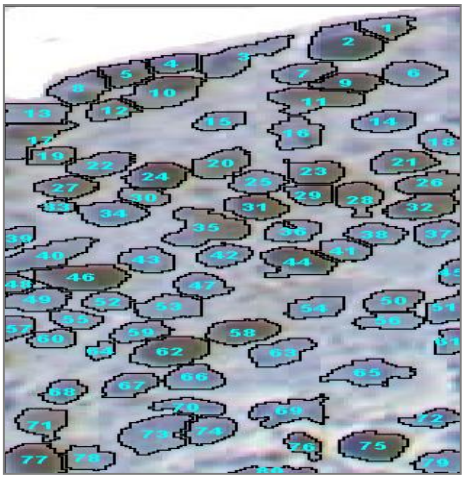


Figure 3-13 (C): represent the segmentation and counting results for the Ki67 Day 22 image dataset

To increase the accuracy of nuclei counting, further refinement to our approach is added: larger nuclei objects exceeding a total number of 240 pixels are divided into two individual cells. The parameter was selected based on the analysis of 5000 nuclei cells taken from various images at 20x magnification level, where objects exceeding 240 pixels are observed to be two nuclei objects. The command is computed as follows;

If Area $\geq$ *240 pixels*

Then

Count as 2 cells

Else

Count as 1 cell

Using this approach, overlapping nuclei objects which failed to be divided by the watershed segmentation are separated, e.g. (nucleus number 16 from image “1) Ki67 Day 19”, nucleus number 58 from image “2) Ki67 Day 19”, nucleus number 35 from image “3) Ki67 Day 19”, nuclei number 72 and 46 from image “2) Ki67 Day 22” . A number of divisions are made when applicable to refine the number of counted nuclei. When a nucleus is divided into two nuclei objects, the segmented nucleus is duplicated into the image dataset.

Table 3-2: (A-C) represents the results of the automated segmentation approach for the 15 randomly selected IHC images from the three image datasets.

A

Images of Ki67 Day 19	Pathologist counting	Automated counting	Pathologist staining intensity score
	Total No. Of nuclei	Total No. Of nuclei	
1) Ki67 Day 19 high staining	65	68	+++
2) Ki67 Day 19 high staining	103	100	+++
3) Ki67 Day 19 high staining	78	77	+++
4) Ki67 Day 19 intermediate staining	100	93	++
5) Ki67 Day 19 low staining	67	73	+

B

Images of p27 Day 19	Pathologist counting	Automated counting	Pathologist staining intensity score
	Total No. Of nuclei	Total No. Of nuclei	
1) p27 Day 19 high staining	93	95	+++
2) p27 Day 19 intermediate staining	74	80	++
3) p27 Day 19 intermediate staining	95	105	++
4) p27 Day 19 low staining	69	81	+
5) p27 Day 19 low staining	76	86	+

C

Images of Ki67 Day 22	Pathologist counting	Automated counting	Pathologist staining intensity score
	Total No. Of nuclei	Total No. Of nuclei	
1) Ki67 Day 22 high staining	100	105	+++
2) Ki67 Day 22 intermediate staining	74	83	++
3) Ki67 Day 22 intermediate staining	94	97	++
4) Ki67 Day 22 intermediate staining	49	47	++
5) Ki67 Day 22 intermediate staining	73	80	++

3.5 Statistical evaluation of the automated counting method

To evaluate the efficiency of our automated counting method, we compare our results to results which have been conducted manually by an experienced pathologist. Four measures were used in order to evaluate the performance of the automated counting method: precision (positive predictive value), recall (sensitivity), accuracy and error (Salton, 1968; Powers, 2007; Taylor, 1999; Golub et al., 1996).

In the context of this chapter, precision is defined as the ratio of the nuclei which have been *correctly* detected by the automated method to the *total* number of nuclei detected by the automated method. This indicates the probability that the detection of a nucleus by the automated method is in fact associated with a nucleus counted by the pathologist. Precision is given as:

$$Precession = \frac{TP1}{TP1+FP1} \qquad (3\text{-}2)$$

Where *TP1* (True Positive) are the number of nuclei which have been counted by the pathologist and also detected by the automated system. *FP1* (False Positive) are the number of nuclei which have not been counted by the pathologist but are detected by the automated system.

In contrast, recall is defined as the ratio of the nuclei number which has been correctly detected by the automated method to the *total* number of nuclei detected by the *pathologist*. This indicates the probability that where a nucleus is counted by the pathologist, it will be detected by the automated method. Recall is given as:

$$Recall = \frac{\boldsymbol{TP1}}{\boldsymbol{TP1+FN1}} \qquad (3\text{-}3)$$

Where *FN1* (False Negative) are the number of nuclei which have been counted by the pathologist but are not detected by the automated system.

Both precision and recall describe different qualities of the automated counting system results. In this case, high precision indicates that most of the nuclei detected by the automated counting method are counted by the pathologist, and high recall indicates that most of the nuclei counted by the pathologist are detected by the automated method. The recall is one if all the nuclei counted by the pathologist are detected by the automated system.

Accuracy is calculated to evaluate the overall performance of the automated method in comparison to the pathologist results. An accuracy of 1 means the automated method results correlated exactly with the pathologist results.

Accuracy is given as;

$$Automated\ Nuclei\ Counting\ Accuracy = \frac{\boldsymbol{TP1}}{\boldsymbol{TP1+FP1+FN1}} \qquad (3\text{-}4)$$

TN is one of the variables used to calculate accuracy in cases where two types of classes are analysed. The performance analysis at this stage only refers to the number of nuclei, without any classification based on the type of stain (positive or negative). The variable *TN* is set to null, and discarded in the calculation of accuracy at this stage. *TN* will be used in the next chapter to indicate the correct detection of negatively stained nuclei by the automated classification method.

Table 3-3: (A-C) represent the Precision, Recall and Accuracy average of the automated counting approach for the 15 randomly selected IHC images from the three image datasets.

A

Images of set Ki67 Day 19	Precision	Recall	Accuracy
1) Ki67 Day 19 high staining	0.95	1.0	0.955
2) Ki67 Day 19 high staining	1.0	0.97	0.970
3) Ki67 Day 19 high staining	1.0	0.98	0.987
4) Ki67 Day 19 intermediate staining	1.0	0.93	0.93
5) Ki67 Day 19 low staining	0.91	1.0	0.917
Average	**0.972**	**0.976**	**0.9518**

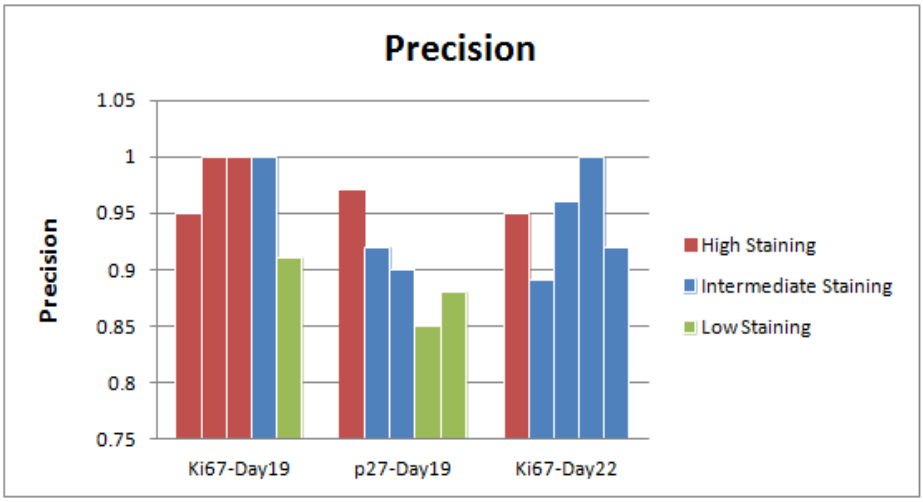
B

Images of set p27 Day 19	Precision	Recall	Accuracy
1) p27 Day 19 high staining	0.97	1.0	0.97
2) p27 Day 19 intermediate staining	0.92	1.0	0.92
3) p27 Day 19 intermediate staining	0.90	1.0	0.90
4) p27 Day 19 low staining	0.85	1.0	0.85
5) p27 Day 19 low staining	0.88	1.0	0.88
Average	**0.904**	**1.0**	**0.904**

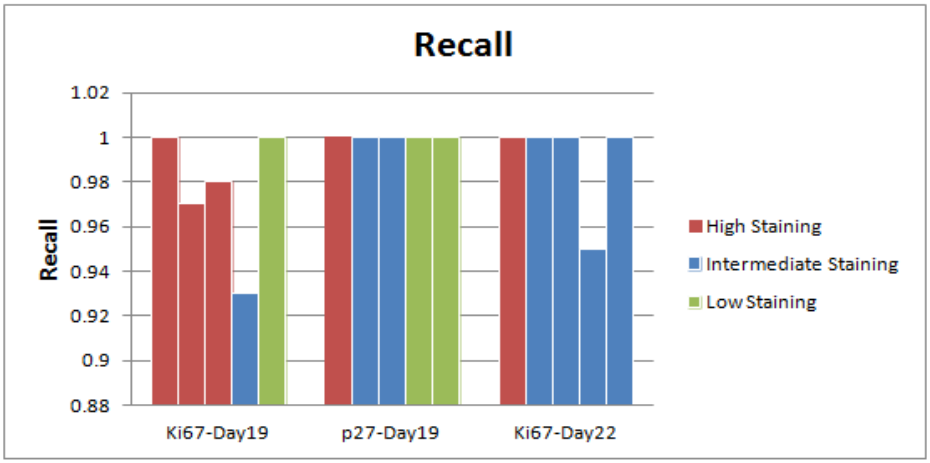
C

Images of set Ki67 Day 22	Precision	Recall	Accuracy
1) Ki67 Day 22 high staining	0.95	1.0	0.95
2) Ki67 Day 22 intermediate staining	0.89	1.0	0.89
3) Ki67 Day 22 intermediate staining	0.96	1.0	0.96
4) Ki67 Day 22 intermediate staining	1.0	0.95	0.95
5) Ki67 Day 22 intermediate staining	0.92	1.0	0.92
Average	**0.944**	**0.99**	**0.934**

A



B



C

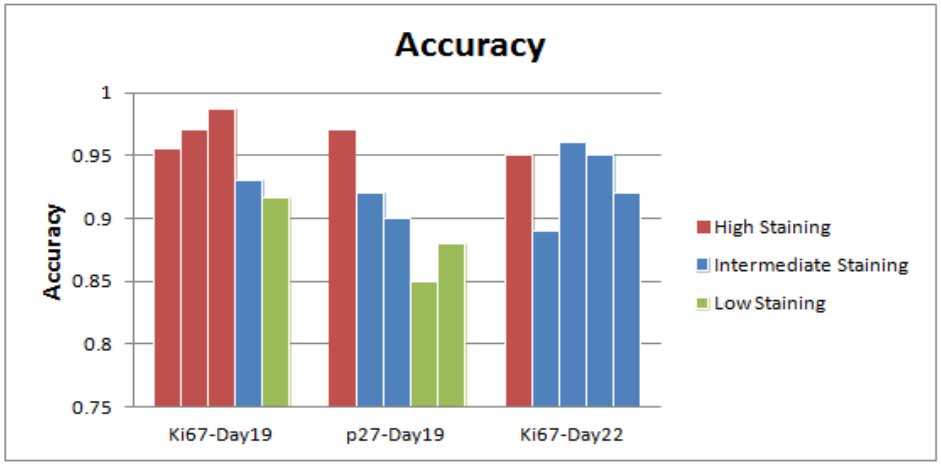


Figure 3-14: (A-C) represent the represent the Precision, Recall and Accuracy of the automated counting approach for each of the 15 randomly selected IHC images from all image datasets. Each bar in these graphs represents an image of a certain staining level.

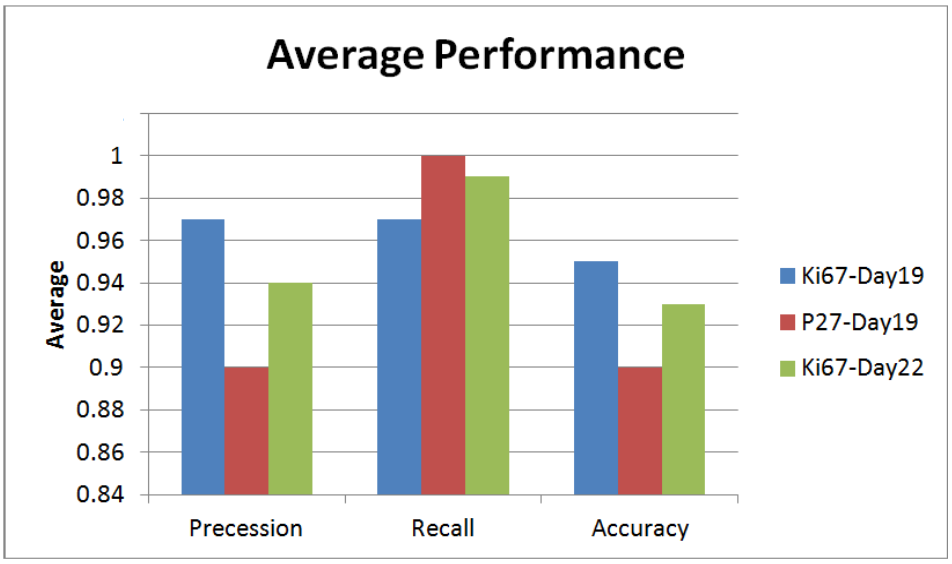


Figure 3-15: represents the Precision, Recall and Accuracy average of the automated counting approach for the 15 randomly selected IHC images from all image datasets.

Table 3-4: (A-C) represent the Precision, Recall and Accuracy average for all IHC images from the three image datasets.

A

Images of set Ki67 Day 19	Pathologist counting	Auto. counting	Precision	Recall	Accuracy
Images (1-69) Ki67 Day 19 high staining	82	84	0.97	1.0	0.97
Images (70-93) Ki67 Day 19 intermediate staining	77	81	0.95	1.0	0.95
Images (94-120) Ki67 Day 19 low staining	91	98	0.92	1.0	0.92
Average	**83.3**	**87.6**	**0.94**	**1.0**	**0.94**

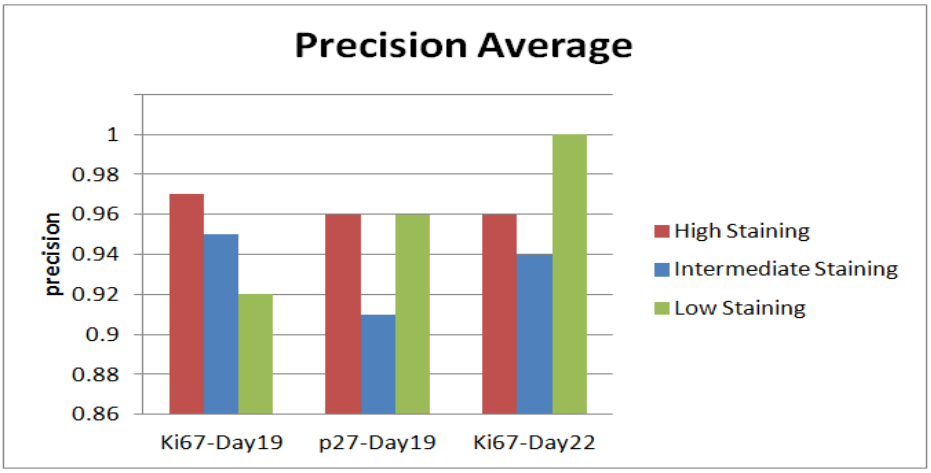
B

Images of set p27 Day 19	Pathologist counting	Auto. counting	Precision	Recall	Accuracy
Images (1-23) p27 Day 19 high staining	101	105	0.96	1.0	0.96
Images (24-70) p27 Day 19 intermediate staining	81	89	0.91	1.0	0.91
Images (71-110) p27 Day 19 low staining	56	63	0.96	1.0	0.88
Average	**79.3**	**85.6**	**0.94**	**1.0**	**0.91**

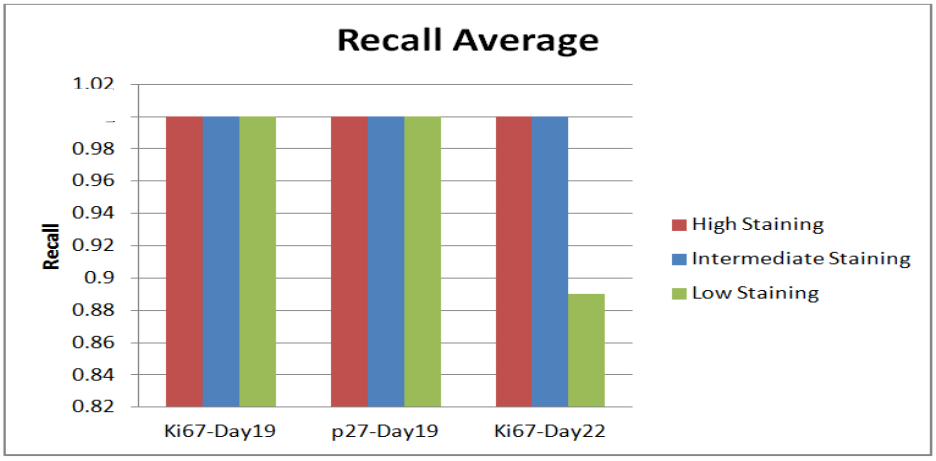
C

Images of set Ki67 Day 22	Pathologist counting	Auto. counting	Precision	Recall	Accuracy
Images (1-87) Ki67 Day 22 high staining	86	88	0.96	1.0	0.97
Images (88-127) Ki67 Day 22 intermediate staining	67	71	0.94	1.0	0.94
Images (128-136) Ki67 Day 22 low staining	48	43	1.0	0.89	0.89
Average	**67**	**67.3**	**0.96**	**0.96**	**0.93**

A



B



C

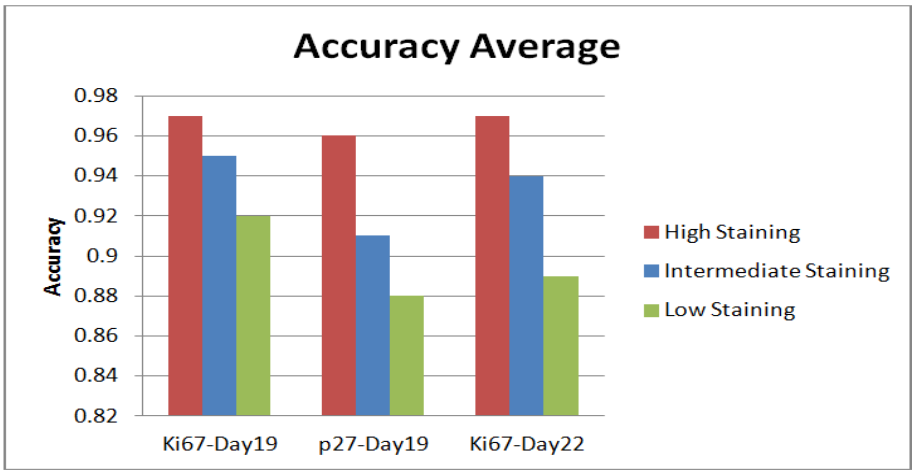


Figure 3-16: (A-C) represent the Precision, Recall and Accuracy average for all IHC images from the three image datasets.

Figure 3-17: represents the overall Precision, Recall and Accuracy average of the automated method for all IHC image datasets combined.

- Precision average represents the percentage of nuclei correctly detected by the automated system.
- Recall average represents the percentage of nuclei counted by the pathologist, which has been detected by the automated system.
- Accuracy average represents the overall performance of the automated method in comparison to the pathologist results.

Since three different IHC image datasets have been selected (Ki67-Day19, Ki67-Day22 & p27-Day19), the error is calculated to compare the efficiency of the automated counting method when dealing with different stained images at different staining intensity levels. This analysis also investigates those cases where the automated method and the pathologist differed. The error is calculated as;

$$\textbf{Error} = \frac{|Experimental - Theoretical|}{Theoretical} \qquad (3\text{-}5)$$

where *Experimental* indicates the automated method results and *Theoretical* indicates the pathologist results. Results are demonstrated in Figure 3-18.

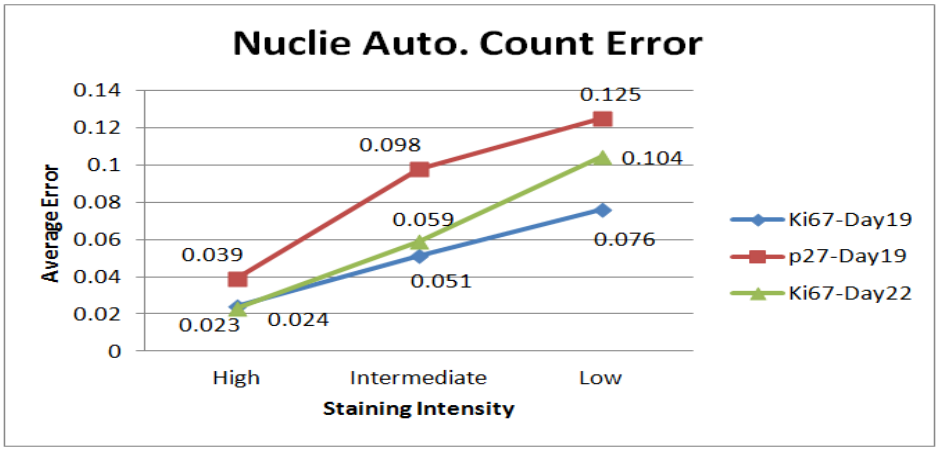


Figure 3-18: Comparisons of Nuclei counting error between Ki67-Day19, Ki67-Day19 & p27-Day19 at different staining intensity levels.

3.6 Discussion

The automated counting method seems to produce a higher accuracy rate when counting the high and intermediate stained nuclei. The accuracy rate of counting high stained images of sets Ki67-day19 and day 22 are relatively high: this is due to the morphological characteristics and the position of such nuclei. High stained nuclei are usually situated on the viable rim, and such nuclei tend to have regular circular shapes, which makes their detection easier for the automated system.

A slight reduction in the accuracy rate is reported when counting intermediate stained images of sets Ki67-day19 and day 22. This is due to some of the intermediate stained nuclei being located near the necrotic core. The necrotic core contains mostly lifeless cells which tend to have irregular shrunken nuclei shapes (necrosis cells). This type of nuclei cells is not counted by pathologists since it does not provide any valuable details in measuring the PS and IS. In this study, necrosis cells are automatically eliminated using the morphological opening operation described in Section (2.1.3). However, in some cases where such cells have relatively large shapes (larger than 30 connected pixels), these could be confused by the automated segmentation and counting system with viable nuclei and therefore counted by the automated system, which subsequently increases the number of false positives. However, this kind of false response is limited, as the dead cells would have to be larger than the smallest viable nuclei in an image for them to be counted by the automated system.

A clear reduction in the accuracy rate is reported when counting low stained images of sets (Ki67-day19 and day 22). This is again due to the position of the low stained nuclei, which tend to be mostly located near the necrotic core, which will consequently result in an increase in the number of false positives compared to the rate of false positives accruing from high and intermediate stained images.

A slight difference between the error rate of Ki67-day19 and day 22 is reported when counting the low stained nuclei. Ki67-day22 images seem to have a higher error rate when compared to the Ki67-day19 images. This is

due to a failure to segment the adjacent nuclei of low stained positive nuclei and negative nuclei, which occurs due to the low intensity variation between such nuclei, causing them to be counted as one nucleus by the system instead of three separate nuclei. This would eventually increase the number of false negatives.

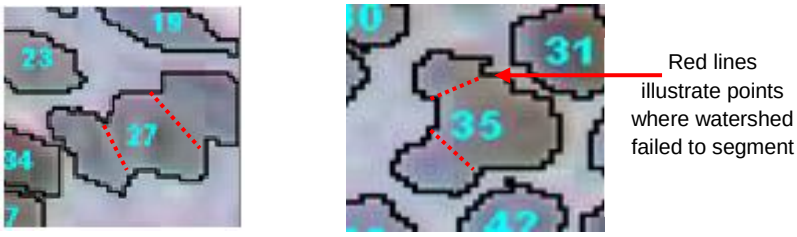


Figure 3-19: illustrates the watershed under segmentation

Figure 3-19 illustrates the failure in segmenting adjacent cells using watershed. Applying the L*a*b colour space to images prior to watershed segmentation improves the performance of watershed to an extent, as demonstrated earlier in Section (3.2.4) of this chapter. However, it does not provide a complete solution to the problem. This does not apply only to the Ki67-Day22 set, as it could occur with any image where adjacent nuclei that have exceedingly low intensity variation. To minimise the error rate further, a division algorithm is added as explained in Section (3.4). If a nuclei object exceeds the maximum value, it is then divided into two nuclei. This method provides a sufficient solution in cases where two adjacent nuclei have been detected as one nucleus. However, in cases where three nuclei have been detected as one nucleus, the algorithm will only divide the detected object into two nuclei.

Further adjustments to the algorithm to divide large objects into three nuclei was avoided, as the system is more likely to detect two adjacent nuclei as

one object than it is to detect three objects as one object. If the division algorithm was extended to divide objects above a certain size into three objects, this could result in some of the two adjacent nuclei being divided as three objects rather than two, consequently increasing the number of false positives generated by the proposed counting system.

P27-Day19 images dataset seem to have the highest percentage of error and the lowest accuracy rate when compared to the other datasets used. This is due to the nature of the p27 images, which tend to have a substantially low intensity variation between the negative protein expression (blue nuclei) and background. This aspect makes the detection of the true positive nuclei a complex task and causes the automated system to consider some of the diffused background parts as nuclei objects. As a result, the number of false negatives is increased when counting the nuclei of such images. This would affect the accuracy rate of counting nuclei objects of all staining levels. The increase in error reported between the high, medium and low stained nuclei images of the p27 dataset is due to the same reasons mentioned for the Ki67 datasets.

Another factor which caused an increase in false positives was the over-segmentation problem, where a nucleus object is divided by the watershed into two individual nuclei. This incident was relatively rare, and therefore further refinements to the watershed performance were avoided in this regard.

Chapter 4

Positive and Negative Nuclei Classification

4.1 Introduction

In this chapter, a cell-level classification approach is presented. The approach is based on an object feature extraction, where descriptive features are directly extracted for each segmented nuclei image. DWT, GLCM and morphological features are used to extract colour, texture and shape features respectively. Features from each category are optimized through a series of extensive experiments, to select only those features that hold the most discriminative power. These features are then fed into a machine learning to classify the positive and negative segmented nuclei images accordingly. Two types of machine learning algorithms are used in this chapter; ANN and SVM. The results of each classification method are presented based upon a number of experiments. Finally, the results from Chapter 3 are used collectively with the results output from this chapter, to measure the PS of the IHC images. The experimental work of this chapter is carried out using Matlab. The flow chart of the proposed classification approach is presented in Figure 4-1.

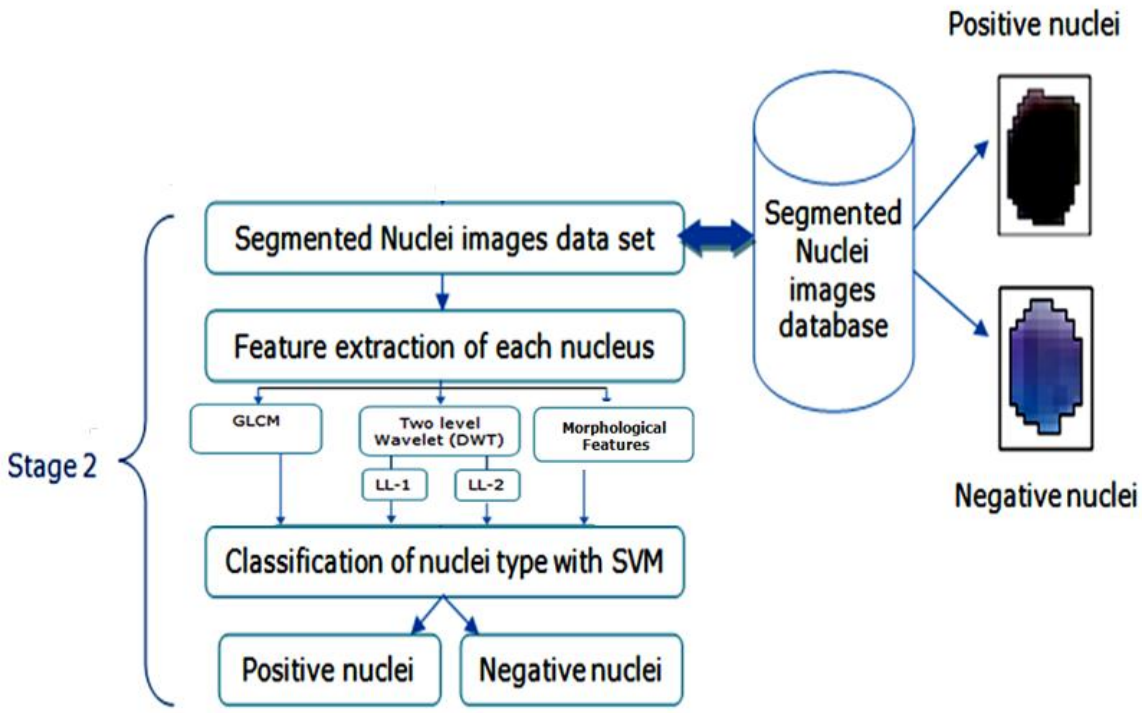


Figure 4-1: the proposed approach for the automated classification of segmented positive and negative nuclei images.

Feature extraction is the key to differentiating between positive and negative nuclei images. The features extracted from the positive and negative nuclei images hold discriminative characteristics which can be used to differentiate between these two types of stained nuclei. These features will then be used to classify segmented nuclei images into two classes; positive stained nuclei and negative stained nuclei.

In image processing, feature extraction can be performed at two levels. At image level: this level of feature extraction describes the global features of an image. This means that the visual contents of an image are described

collectively, certain regions of the image cannot be defined individually using this level. Colour histogram is one of the tools to describe such features.

At region level: this type divides the image into rectangular blocks or regions using a fixed size grid. The grid window is placed on an image to perform the region segmentation. Colour and texture features are then extracted from the blocks, to define their characteristics. Each feature can then be stored in a vector and fed into a classifier for categorization purposes.

Region level segmentation is efficient in tasks where a certain object or certain regions of an image require analysis. However, in images that contain a vast amount of minor details and objects closely adjacent to each other, object segmentation cannot be achieved using this method. Therefore, in this chapter an object feature extraction approach is introduced, this approach will target extracting features specifically from the ROIs, which in this case are the segmented nuclei objects. The extraction method described in Section (3.3) ensures that only the object of interest is segmented, therefore allowing the features which directly describe the object of interest to be extracted.

In this chapter, three types of feature extraction methods are selected to describe the colour, textural, and morphological characteristics of segmented nuclei images. The features extracted will then be used to discriminate between the positive and negative nuclei.

4.2 Feature extraction

4.2.1 Discrete wavelet transform for colour features

Discrete wavelet transform DWT has become a valuable tool in signal and image processing applications, e.g. image compression, image classification and feature extraction (Wang et al., 2001; Kokare et al., 2007; Serrano et al., 2004), where it has shown a high level of efficiency in colour and textural feature extraction tasks. In image processing, DWT works by decomposing the image into four sub-bands; (LL), (LH), (HL) and (HH). The low frequency sub-band (LL) represents the horizontal and vertical low frequency components of an image, and is known as the approximation coefficients. The low high frequency sub-band (LH) represents the horizontal low and vertical high frequency components of an image. The high low frequency sub-band (HL) represents the horizontal high and vertical low frequency components of an image. Finally, the high frequency sub-band (HH) represents the horizontal and vertical high frequency components of an image.

In this work, a 'two-level' wavelet decomposition using two dimensional DWT functions available in Matlab is used to extract the colour features from the segmented nuclei images, where the (LL) approximation coefficients of the first level are decomposed further into further four sub-bands: see Figure 4-2. The Haar wavelet type is used in this work due to its validity in representing images colour and texture.

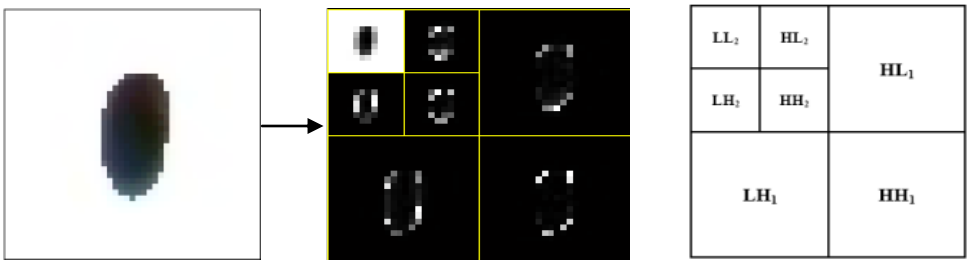


Figure 4-2: nuclei feature extraction using a 'two-level' wavelet decomposition.

In order to extract colour features from images, a colour space needs to be selected first. The colour space is what represents the colour information in an image. A number of colour spaces have been used to represent colour, one of which is HSV. HSV has been shown to provide more efficient colour representation than other colour spaces. This is due to its uniformity in representing colour and ability to represent colour in the same way that the human eye observes colour. This makes HSV a preferable colour space in image analysis and feature extraction tasks. For these reasons, HSV colour space is selected in this work to extract colour information from the images.

To extract the colour features (energy measures) from the wavelet coefficients, nuclei images are first converted into HSV. The two-level DWT is then applied on each of the three colour components; the hue component of the nuclei image (H), the saturation component (S) and the intensity value (V) for the LH, HL and HH sub-bands of each wavelet decomposition level (level one and two). In result, a feature vector of length 18 (3 components x 6 energy measures) is produced: see Figure 4-3.

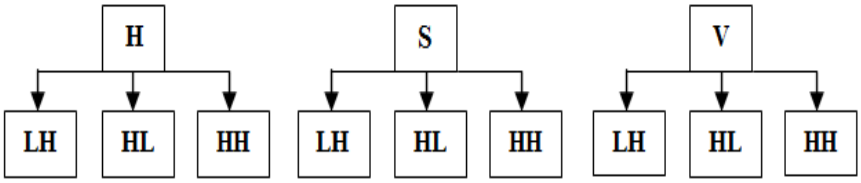


Figure 4-3: Feature vector of 'one-level' wavelet decomposition

The energy of a wavelet sub-band is computed as;

$$Energy = \frac{1}{ab} \sum_{i=1}^{a} \sum_{j=1}^{b} |C_{ij}| \qquad (4\text{-}1)$$

where ab is the size of the wavelet sub-band and C is the wavelet coefficient.

4.2.2 Grey level co-occurrence matrix for texture features

A number of textural feature extraction approaches have been presented and implemented in recent research, such as statistical moments and the intensity histogram of an image or area. This kind of method measures textural features based on the distribution of intensities. However, information regarding the position of pixels in respect to each other cannot be extracted by using these methods.

Statistical approaches such as the grey level co-occurrence matrix provide textural features which describe the distribution of grey level intensity pixels in relation to their position with the neighbouring pixels in an image.

A co-occurrence matrix is a two-dimensional array (Chaddad et al., 2011; Materka and Strzelecki, 1998) where an image is presented by a number of row and column values. GLCM is defined as;

$$G_{v,r}[c,d] = n_{cd} \qquad (4\text{-}2)$$

A GLCM *G* is defined by specifying a displacement vector ***v****=(vc,vd).* Where *v* is the space between the two pixels holding the values *c* and *d, r is the direction of pixels.* n_{cd} is the number of occurrences of the pixel holding the values *c* and *d.*

A number of features were defined by (*Haralick et al., 1973)* to represent an image. In this chapter, we use and compute four of Haralick's features to analyse the texture of positive and negative segmented nuclei images. The features used are; variance *(σ),* contrast *(Cont),* entropy *(Ent) and* correlation *(Corr).* These texture features are calculated as weighted averages of the co-occurrence matrix contents. The features were selected based on the discriminative values they provide (vectors), GLCM features which provided a high difference in feature values between positive and negative nuclei were selected to differentiate between the features of negative and positive nuclei images. GLCM features which did not provide discrimination in the vector values of positive and negative nuclei were discarded. The experimental results are presented later in this chapter. The four selected GLCM texture features are described in the following section.

4.2.2.1 GLCM Variance

The variance feature relies on the mean parameter (***μ***): the mean is measured by calculating the occurrence frequency of a reference pixel value in combination with a certain neighbouring pixel value. The performance of the GLCM variance is similar to the common descriptive statistic known as variance. It is calculated based on the distribution of grey-level differences at a certain distance, v, from the mean value. The GLCM variance is given as;

$$\mu_c = \sum cG_v[c,d], \quad \sigma_c^2 = \sum c^2 G_v[c,d] - \mu_c^2 \qquad (4\text{-}3)$$

4.2.2.2 GLCM Contrast

The contrast feature is calculated based on the image intensity contrast variation of a pixel value and its neighbouring pixel value in an image. This feature describes the fineness of a texture in an image. High variation describes rough textures, while low variation values between pixels describe smooth textures. The GLCM contrast is given as;

$$Cont = \sum_{c=1}^{N} \sum_{d=1}^{N} (1-c)^2 \, . \, G_{v,r}[c.d] \qquad (4\text{-}4)$$

where *N* is the number of grey levels in the image.

4.2.2.3 GLCM Entropy

The entropy feature measures the randomness of the distribution of pixel intensities in an image. Randomness corresponds to an image where there is no preferred intensity level for a combination of two pixels with a distance

v, Low entropy values indicate an irregular GLCM. Entropy attains its peak value when all elements of the GLCM are equal. Irregular entropy values hold descriptive features of the nuclei objects. The GLCM entropy is given as;

$$Ent = -\sum_{c=1}^{N}\sum_{d=1}^{N} G_{vr}[c.d].\ln\left(G_{v,r}[c.d]\right) \quad (4\text{-}5)$$

4.2.2.4 GLCM Correlation

The correlation feature measures the linear relationship between grey level pixel values in an image. A single object tends to have a higher correlation value between its pixels. The correlation values between nuclei object pixels will be used to differentiate between negative and positive nuclei images. The GLCM correlation is given as;

$$Corr = \sum_{c=1}^{N}\sum_{d=1}^{N}(c-\mu x)(d-\mu y).\frac{G_{v,r}[c.d]}{\sigma x.\sigma y} \quad (4\text{-}6)$$

4.2.3 Morphological feature extraction

For the morphological description of nuclei objects, features which have been directly extracted from the segmented nuclei images are used. Three morphological features have been selected, which are; circularity, area, and perimeter. Area can be defined as the number of pixels contained inside the border of the segmented nuclei object (surface of the ROI): this feature describes the size of the nuclei object. Circularity describes the level of roundness of the nuclei shapes and is calculated using equation (3-1).

Finally, perimeter can be defined as the length of the nuclei outside the boundary: in this case it is calculated by the total number of connected pixels on the segmented nuclei boundary. These features were extracted using the proposed segmentation method explained in Section (3.3). A vector of length 3 is generated for each nuclei image from this stage. The aim of extracting such features is to use the morphological characteristics of positive and negative nuclei in providing discriminative power, which can then be used in classifying both types of nuclei.

4.3 Machine learning for classification task

Image classification is the process of classifying or categorising images into a number of classes. Classification has two main approaches; discriminative and generative. The discriminative approach works by estimating boundaries or driven probabilities to discriminate between the different classes. The aim of this approach is to differentiate between the different classes by finding classification boundaries. Support vector machine and neural networks are examples of discriminative classifiers which have been widely used in image classification and pattern recognition analysis. Support vector machine is used in this thesis as the main classifier to classify positive and negative segmented nuclei images. Neural network was used as a secondary classifier to compare the classification result output with the SVM classification results. Although support vector machines and neural networks are different in concept, they are both types of machine learning algorithms capable of providing data classification. In this chapter, the performance

analysis of both classifiers is provided to evaluate the efficiency of both classifiers in stained nuclei categorisation tasks. The features described in the previous section are used as feature vectors to train the classifiers and then used to predict the classification results of the test images. A feature vector of length 25 (18 DWT feature + 4 GLCM feature + 3 morphological features) is generated for each individual nucleus to be used as an input feature in both classifiers.

4.3.1 Support Vector Machines

Support vector machine (SVM) is a supervised learning technique which has become a valid tool in image classification since it was introduced by (Vapnik et al., 1996). SVM is an excellent tool for binary and multiclass classification, and works based on statistical theory, seeking out the optimal separating hyperplane which presents the maximum space or margin that separates two or more data points. In other words, SVM takes the data points represented by *p*-dimensional vectors, then finds hyperplanes to separate the two points into two classes based on the maximum distance between the nearest two points that represent each class. This learning procedure is performed by the SVM during the training process, where SVM finds a number of hyperplanes that separate the *p*-dimensional vectors, and then only selects the hyperplane that represents the maximum margin classifier. The SVM presented by (Vapnik et al. 1996) was a linear classifier: this type of classifier does not provided efficient classifications in tasks where a non-linear decision is required. SVM kernel functions are therefore introduced to

provide a solution for classification tasks where this is the case. A number of kernels such as polynomial, sigmoid and radial basis function (RBF) can be used in such cases. SVM kernel functions are defined by (Giuliodori, 2011) as:

$$f(x) = \sum_{i=1}^{N_s} a_i y_i K(s_i, x) + b \qquad (4\text{-}7)$$

where N_s is the number of points located closest to the hyperplane on each side, or what is known as the support vectors. a_i is the weight of the point s_i learnt by the SVM during the training process. y_i is the class label of s_i which can be either -1 or +1 in cases of binary classification tasks. $K(s_i, x)$ is the value of the kernel function for the training sample s_i and the testing sample x. x belongs to class -1 if the value f for test point x is negative and belongs to class +1 if the value f for test point x is positive.

In this chapter SVM is used as the main classifier for categorising the segmented nuclei images into two classes; -1 for negative nuclei and +1 for positive nuclei. The SVM method is performed using the publicly available LIBSVM software. The discrimination between positive and negative nuclei cannot be performed using a linear classifier due to the complexity of the nuclei characteristics, and therefore the Gaussian radial basis function RBF is selected for its efficiency in binary none-linear classification tasks. The RBF kernel is defined as;

$$K(x_i, x_j) = exp\left(-\frac{\|x_i - y_i\|^2}{2\sigma^2}\right) \qquad (4\text{-}8)$$

where x_i is the support vector and x_j is the testing data point. σ is set to the

median of the Euclidean distances from each positive training set point to the nearest negative point (Zhu et al., 2011).

Support vector machines are used in this chapter to evaluate the discriminative power of the different feature extraction sets, to select only those features that provide the highest level of classification accuracy.

Experiments using the SVM classifier were performed using a 10-fold cross validation method. Cross validation is a model evaluation method used at the training process of the SVM. It measures how well the classifier is learning the new data. The image dataset included both negative and positive nuclei image sets. Each set was randomly divided into 2 subsets. The first subset included 90% of the nuclei images which were selected arbitrarily for the learning procedure of the SVM. The second subset included the remaining 10% of nuclei images, which were left for the testing procedure. This procedure was repeated 10 times (10-fold) to allow the testing of all nuclei images by the SVM classifier. Parameters were selected based on 10-fold cross validation. The average of the results over the 10 splits was then calculated to yield the overall classification accuracy. The number of images holding the label 1 are counted to yield the total number of positive nuclei images classified by the SVM. The number of images holding the label -1 are counted to yield the total number of negative nuclei images classified by the SVM. 10-fold cross validation was selected to attain more accurate performance estimation of the SVM classification results.

4.3.2 Artificial neural networks

A three-layer multilayer perceptrons (MLP) neural network was used for the purpose of classifying positive and negative stained nuclei. This neural network was selected in our study to investigate the possibilities of using such ANNs to classify stained nuclei images and compare the classification results to the output results from the SVM classification. In the input layer, the number of neurons is equal to the length of the 25 feature vectors. In the hidden layer 2N nodes were used, where 'N' is the input feature vector elements. Sigmoid was used in this layer as the activation function due to its efficiency in nonlinear classification tasks. The output layer has only one neuron: this neuron represents the decision of the classification process (either positive or negative nuclei). Purelin was used in this layer as the activation function. The training process uses a function based on Levenberg-Marquardt optimization. This function was selected due to its efficiency when used with non-linear functions such as neural networks training tasks. 90% of positive and negative nuclei images are arbitrarily selected for the training purposes and the remaining 10% is used for testing purposes. This procedure is performed 10 times to ensure the testing of all nuclei images. The average of the results over the 10 splits is then calculated to yield the overall classification accuracy. The number of images holding the label 1 are counted to yield the total number of positive nuclei images classified by the ANN. The number of images holding the label -1 are counted to yield the total number of negative nuclei images classified by the ANN. Parameters of the ANN are listed in Table 4-1.

Table 4-1: Parameters of the ANN

Number of neurons in the input layer	25 feature vectors
Number of neurons in the hidden layer	25x 2
Transfer function in the hidden layer	Sigmoid
Number of neurons in the output layer	"1"
Transfer function in the output layer	Purelin
Training function	Trainlm
Gradient	0.0001.
Number of Epochs	20000

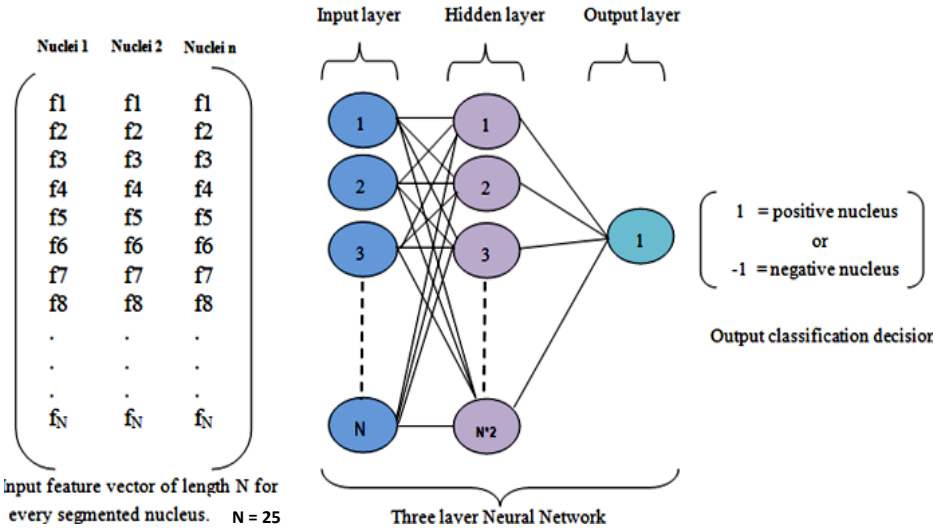


Figure 4-4: the structure of the ANN classifier

4.4 Feature selection

4.4.1 Texture features selection

Feature selection and optimization play a vital role in determining the level of classification accuracy. In this section, textural features selection is targeted. The total number of GLCM features defined by (Haralick et al., 1973) are 14. The goal of this section is to select features that provide discriminative characteristics between positive and negative nuclei. Using all the GLCM features without optimization can result in a high dimensioned vector with insufficient discriminative features, which degrades the classification

accuracy. This is due to the pattern of positive and negative nuclei which exhibit similar textural characteristics. To select only those features which provide discriminative information, a feature selection method is applied. A key point for the selection of relevant features is to eliminate similar features between positive and negative nuclei. The Maximum Difference Features (MDFS) algorithm introduced by (Nithya and Santhi, 2011) is used to select only those features which exhibit discriminative characteristics of the positive and negative nuclei. The four features that provide the highest level of difference between the positive and negative nuclei are selected as the optimal textural features. The algorithm is calculated using a number of steps; first, the sum of features extracted from all positive nuclei images using the first feature of GLMC, e.g. *contrast,* is calculated using the following formula;

$$F1 = \sum_{i=1}^{n} P_i \qquad (4\text{-}9)$$

where *n* is the number of nuclei images and *P* is the features extracted from positive nuclei images. The same is calculated for all negative nuclei images using the following formula;

$$F2 = \sum_{i=1}^{n} N_i \qquad (4\text{-}10)$$

N is the features extracted from negative nuclei images. The difference *D* of *F1* and *F2* are then calculated using the following algorithm;

If F1>F2

D = (F1-F2) / (F1+F2).

Else

D = (F2-F1) / (F1+F2).

The procedure is repeated for all 14 GLCM features. The four top features with the highest difference *D* are selected to be the optimal texture features.

Following calculations, it was found that the four features with the highest difference values were variance *(s),* contrast *(Cont),* entropy *(Ent), and* correlation (Corr). To evaluate the results of this algorithm, the 14 GLCM features are divided into 3 sets based on the *D* value in descending order. The first set includes the 5 GLCM features which were found to have the lowest *D* value. The second set includes the 5 GLCM features with intermediate *D* values: and finally the third set includes the 4 GLCM features with the highest *D* values. The 3 GLCM sets are presented in Table 4-2 below;

Table 4-2: GLCM high, intermediate and low features difference value sets.

GLCM features with high difference values	*Contrast, Entropy, Variance and Correlation.*
GLCM features with intermediate difference values	*Difference entropy, Homogeneity, Sum average, Sum variance, and Maximum correlation coefficient.*
GLCM features with low difference values	*Energy, Sum entropy, Difference variance, Information measure of correlation1, Information measure of correlation2.*

A vector of length 4 is generated for each nuclei image from this stage. Results of the classification accuracy using the different sets of GLCM texture features are demonstrated in Section (4.5).

4.4.2 Wavelet level selection

The selection of the wavelet level of decomposition is an essential step since the coefficients produced at every level are the key components in defining the discriminative characteristics of positive and negative nuclei images. In this section, wavelet decomposition levels 1-5 are examined to select the optimal level of wavelet decomposition.

4.5 Experiments

4.5.1 First set of experiments

In this chapter, four sets of experiments are presented. Experiments were conducted using the SVM classifier. A 10-fold cross validation method was used, where 90% of the nuclei images of each class were selected arbitrarily for the learning procedure of the SVM, while the remaining 10% were left for the testing procedure. The two classes are the positive and negative nuclei classes. The first set of experiments was performed to optimize the selection of features that best describe the discriminative characteristics between positive and negative segmented nuclei images.

For this set of experimental work, three image datasets were used;

Nuclei image dataset 1: 1000 segmented nuclei images from Ki67 Day 19 image dataset were randomly selected. The 1000 were selected to include two classes; 500 images are positive nuclei and the remaining 500 are negative nuclei images. *Nuclei image dataset 2*: 1000 segmented nuclei

images from p27 Day 19 image dataset were randomly selected. The 1000 were selected to include two classes; 500 images are positive nuclei and the remaining 500 are negative nuclei images. *Nuclei image dataset 3*: 1000 segmented nuclei images from Ki67 Day 22 image dataset were randomly selected. The 1000 were selected to include two classes; 500 images were positive nuclei and the remaining 500 were negative nuclei images.

All images of segmented positive and negative nuclei were extracted and saved using the extraction method described in Section (3.3). For this set of experiments only the TP segmented positive and negative nuclei images were used to create the image datasets.

13 experiments of features and their combinations were performed as follows:

Three types of feature extraction methods were used; DWT was used to extract colour features, GLCM features were used to extract the textural features, and morphological features were used to describe the shape of the segmented nuclei. To evaluate the performance of each feature set, features were used separately to report their discriminative power. In case of DWT the coefficients extracted from every level were used as feature vectors to classify the positive and negative nuclei using the SVM classifier. The first and second level of wavelet decomposition presented the optimal features in classifying the positive and negative nuclei images. Integration of wavelet decomposition coefficients of level one and two reported higher classification accuracy compared to the rest of the wavelet decomposition levels. A

feature vector of length 18 is produced from level one and two wavelet coefficients as described in Section (4.2.1) of this chapter. Classification performances of the DWT are reported in Figures 4-5 to 4-8.

The three groups of GLCM features described in Section (4.2.2) of this chapter are evaluated using the SVM classifier. The feature vectors of the high difference GLCM are integrated together to form a feature vector of length 4. The high difference GLCM features reported the highest level of classification accuracy compared to the intermediate and low difference GLCM features.

Classification performances of the GLCM are reported in Figures 4-5 to 4-8. The morphological features are then evaluated using the SVM classifier and the performance is reported in Figures 4-5 to 4-8. To optimize the feature selection further, integration of the uppermost discriminative features is performed. Integrating DWT (level one and two) and morphological features gained a good classification performance compared to the use of DWT and morphological features individually. Although morphological features did not obtain the best classification results, they can be used to improve the performance of the DWT. Similarly, with the high difference GLCM features, integration with the DWT improves classification accuracy further. However, they do not provide the highest classification performance. Integration of the two-dimensional DWT (level one and two), high difference GLCM features and morphological features provided the highest classification performance. Classification performances of the uppermost features are reported in Figures 4-5 to 4-8.

Finally, the three image datasets described earlier in this section were integrated collectively to form one dataset that includes the 3000 segmented nuclei images. This experiment was conducted to evaluate the performance of the classification approach when used on a larger dataset with different categories of stained nuclei images. Classification performances are reported in Figure 4-8.

Table 4-3: represents the SVM classification accuracies resulted from using the 13 experiments of the different feature sets and their combinations.

Features	Classification accuracy% of Ki67 Day 19	Classification accuracy% of P27 day 19	Classification accuracy% of Ki67 Day 19	Classification accuracy% of Ki67 Day 19 + Ki67 Day 22 + P27 day 22
Wavelet-level-1	74	71	75	72
Wavelet-level-2	69	68	71	66
Wavelet-level-3	48	53	51	53
Wavelet-level-4	43	41	44	44
Wavelet-level-5	31	29	34	31
Wavelet-level (1 and 2)	79	74	81	75
High Difference GLCM	71	68	74	69
Intermediate Difference GLCM	57	55	54	59
Low Difference GLCM	46	31	38	44
Morphological	58	52	60	50
Wavelet level (1 and 2) + High difference GLCM	87	85	89	84
Wavelet level (1 and 2) + Morphological	82	79	84	83
Wavelet level (1 and 2) + High difference GLCM + Morphological	95	93	96	93

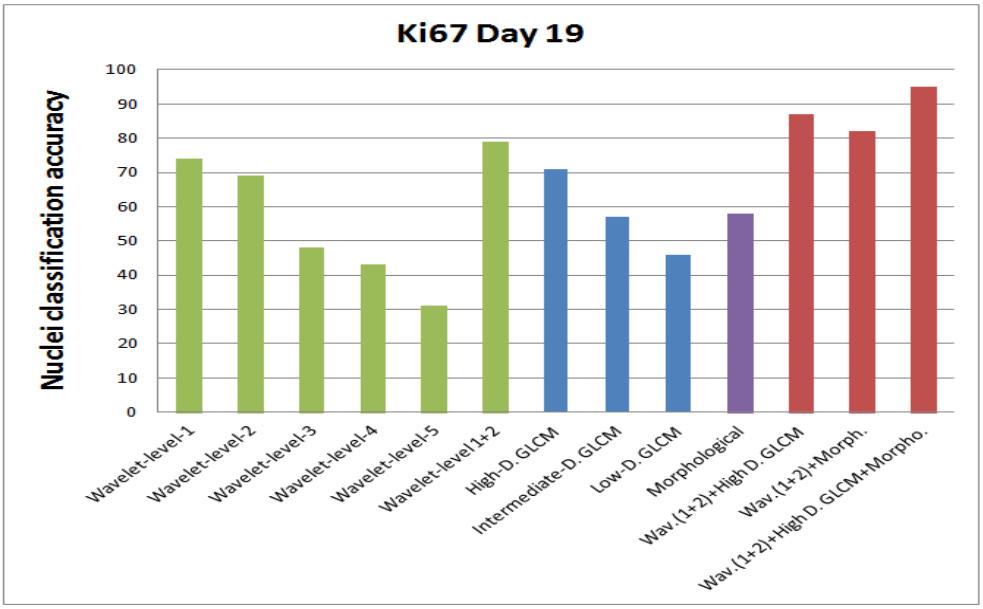


Figure 4-5: represents the SVM accuracies of classifying segmented negative and positive nuclei images from the Ki67 Day 19 image dataset as results of the 13 experiments of the different features and their combination.

Bars in green show the performance of nuclei image classification when using the different levels and combinations in DWT. Bars in blue show the performance of nuclei image classification when using the GLCM features. The bar in purple shows the performance of nuclei images when using morphological features. Bars in red show the performance of nuclei image classification when using the combination of the uppermost discriminative features.

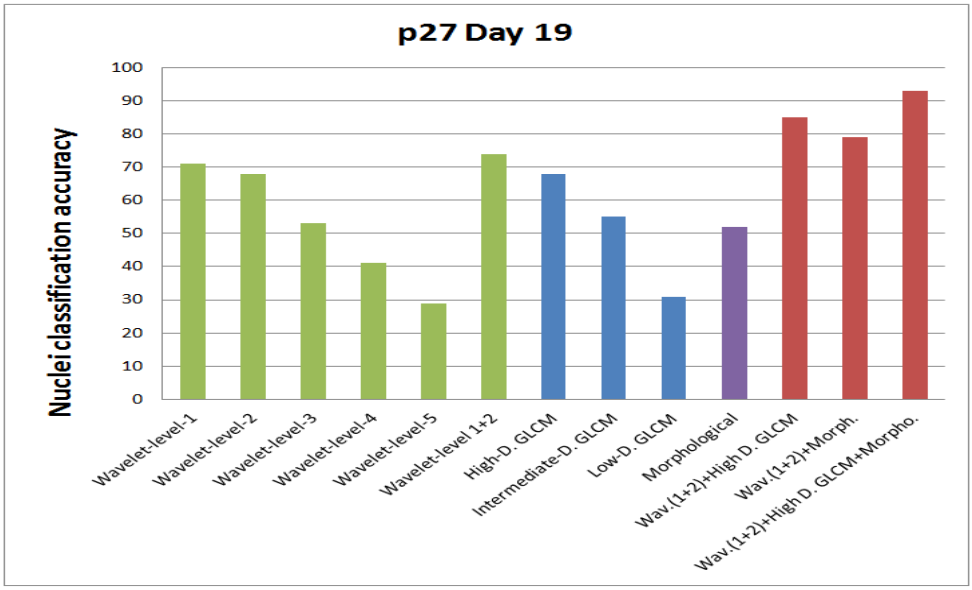


Figure 4-6: represents the SVM accuracies of classifying segmented negative and positive nuclei images from the p27 Day 19 image dataset as results of the 13 experiments of the different features and their combination

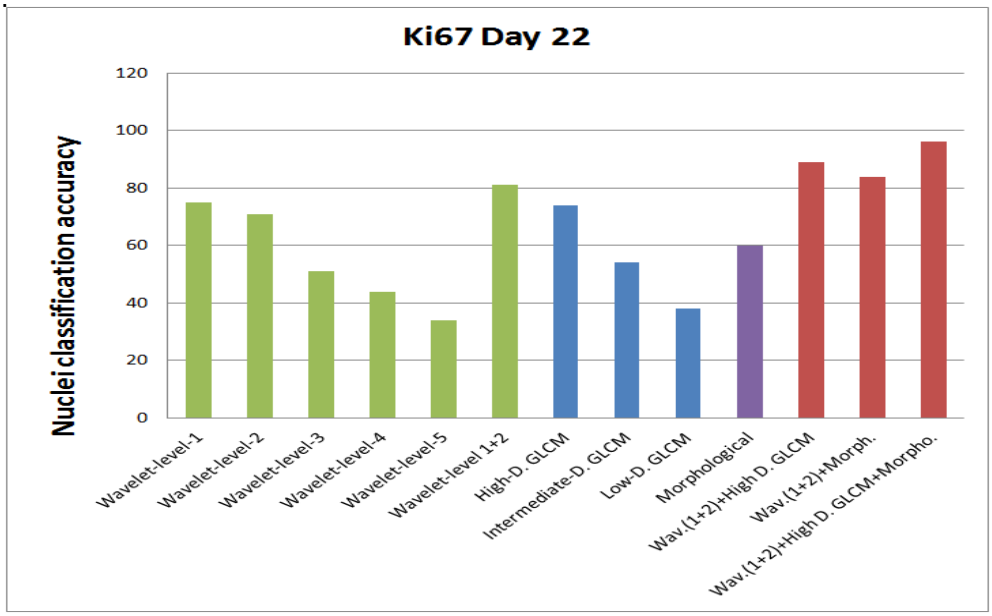


Figure 4-7: represents the SVM accuracies of classifying segmented negative and positive nuclei images from the Ki67 Day 22 image dataset as results of the 13 experiments of the different features and their combination.

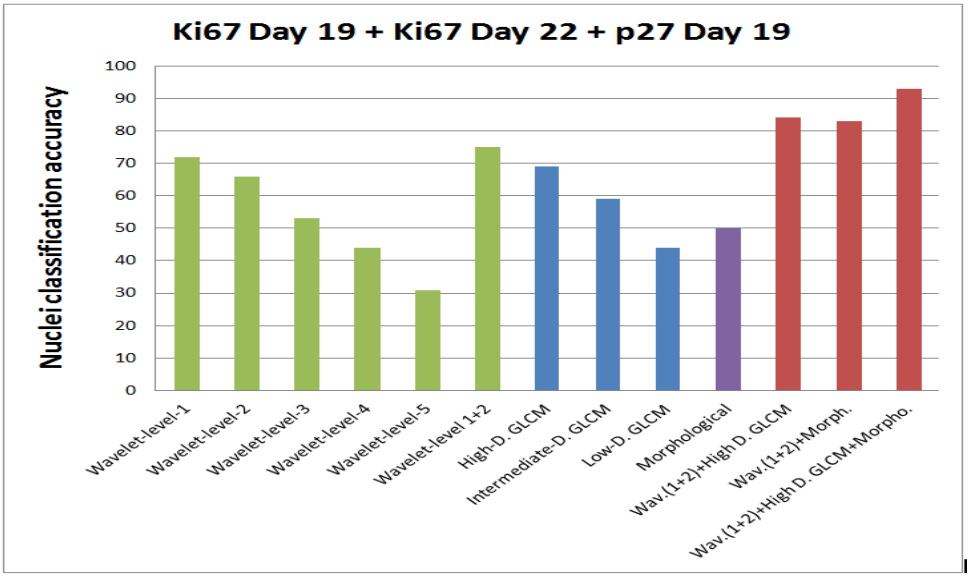


Figure 4-8: represents the SVM accuracies of classifying segmented negative and positive nuclei images from the Ki67 Day 19, p27 Day 19 and Ki67 Day 22 image dataset collectively as results of the 13 experiments of the different features and their combination.

4.5.2 Second set of experiments

In the second set of experiments, the classification performance of ANN is investigated. This experiment is conducted to evaluate the performance of both SVM and ANN classifiers using the same features and image dataset. The image dataset used in this experiment is the integration of the three image datasets described in Section (4.5.1): it includes the 3000 positive and negative nuclei images. The use of ANN is described in detail in Section (4.3.2). Although both classifiers work based on different concepts, it was interesting to evaluate the classification performance of the two classifiers when used on the same image datasets and same feature selection for the task of classifying positive and negative segmented nuclei images. Since a 10-fold cross validation method was used in the SVM

classifier, where 90% of images are used for training and the remaining 10% were used for testing, a similar concept was used for the ANN classification procedure. The ANN training and testing procedure is divided in the same manner (90% for training and 10% for testing). This procedure is repeated 10 times to allow the testing of all nuclei images: a different set of images is used for testing at each time. 13 experiments of features and their combinations are performed in this experiment, set up in the same manner as was used in the previous section (4.5.1). It is obvious from the experiments that SVM outperforms ANN in all types of features and their combinations: see Figure 4-9.

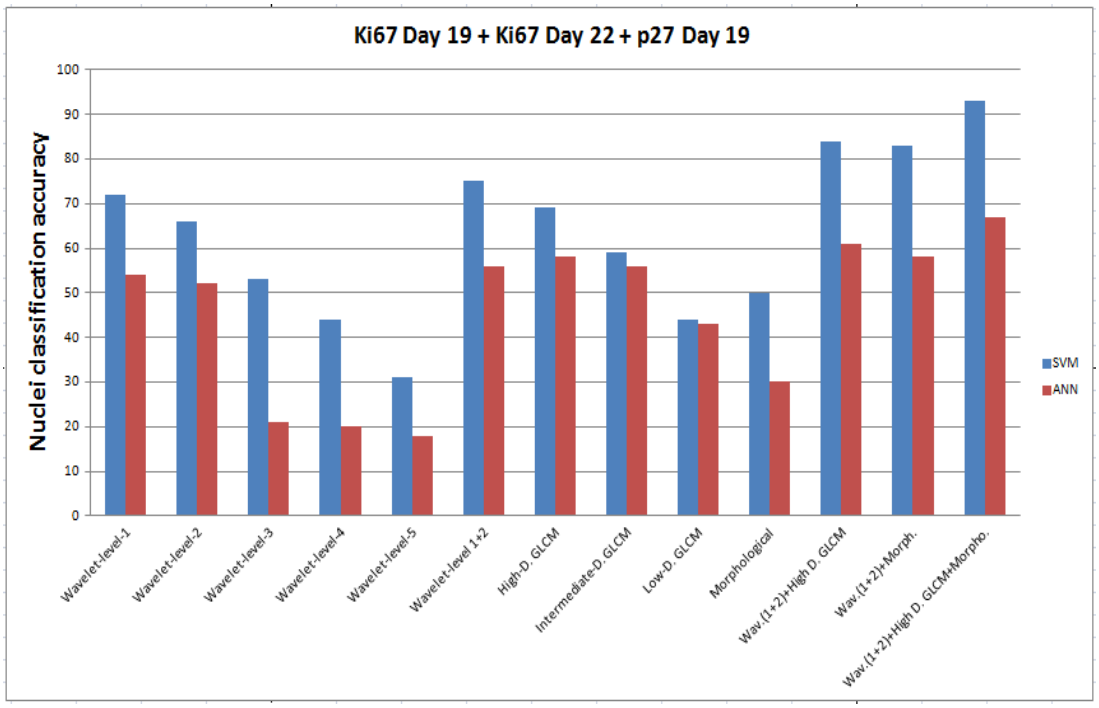


Figure 4-9: The classification performance for each of the features tested, resulting from the 13 experiments using SVM and ANN classifiers. Blue bars show the performance of nuclei image classification using SVM classifier. Bars in red show the performance of nuclei image classification using ANN classifier.

4.5.3 Third set of experiments

Experiment 3 of this chapter presents the statistical analysis that is used to calculate the performance accuracy of the proposed SVM classification approach in terms of classifying both positive and negative nuclei using the following formula;

$$Classification\ Accuracy\ \% = \frac{correctly\ predicted\ data}{total\ testing\ data} * 100 \qquad (4\text{-}11)$$

SVM classification accuracy was measured at 93.6% when tested on 3000 segmented nuclei images as described in Section (4.5.1). Since half of the image dataset contains positive nuclei and the other half contains negative nuclei, it was interesting to measure the performance of the proposed classification approach in classifying positive and negative nuclei independently. To calculate this, two statistical measures were used; sensitivity and specificity (Minnen et al., 2006).

In the context of this chapter, sensitivity is defined as the ratio of the number of positive nuclei which have been *correctly* classified by the proposed approach to the *total* number of positive nuclei in the ground truth image dataset. This will indicate the classification accuracy of the proposed approach in classifying segmented positive nuclei images. Sensitivity is given as:

$$Sensitivity = \frac{\boldsymbol{TP2}}{\boldsymbol{TP2+FN2}} \qquad (4\text{-}12)$$

Where *TP2* (True Positive) is the number of positive nuclei which have been classified correctly by the automated approach. *FN2* (False Negative) are the number of positive nuclei which have been classified as negative nuclei by the proposed approach.

In contrast, specificity is defined as the ratio of the number of negative nuclei which have been *correctly* classified by the proposed approach to the *total* number of negative nuclei in the ground truth image dataset. This will indicate the classification accuracy of the proposed approach in classifying segmented negative nuclei images. Specificity is given as:

$$Specificity = \frac{TN2}{TN2+FP2} \qquad (4\text{-}13)$$

Where *TN2* (True Negative) is the number of negative nuclei which have been classified correctly by the automated approach. *FP2* (False Positive) is the number of negative nuclei which have been classified as positive nuclei by the proposed approach. Each label of the positive nuclei class (1) is counted to yield the total number of correctly classified positive nuclei. (-1) are also counted to yield the total number of correctly classified negative nuclei.

Table 4-4: represents the classification accuracy results summary.

	Positive and negative nuclei	Positive nuclei	Negative nuclei
Ground truth nuclei image dataset	3000	1500	1500
Total number of nuclei images classified correctly by the proposed approach	2808	1441	1367
Accuracy %	93.6	96	91.1

4.5.4 Fourth set of experiments

In this section, a demonstration of how the automated nuclei counting output described in Chapter 3 and the proposed classification approach output described in this chapter can be used collectively to calculate the percentage of positive staining in IHC images. The percentage of positive staining PS is defined as the ratio of the positively stained nuclei number to the total number of nuclei. To demonstrate how the PS can be calculated, the 15 images presented in Section (3.4) Figure 3-13: (A-C) are used. This will include the total number of nuclei generated from each of the 15 IHC images. The total number of nuclei objects generated from each image will be used as the first element to calculate the PS. Each nucleus is then extracted and saved as an image using the extraction method described in Section (3.3). The classification approach is then applied to the segmented nuclei images from each IHC image individually as follows; the uppermost feature (two-level-wavelet + the high difference GLCM features + morphology features) are used to extract the discriminative features from both positive and negative nuclei. For each IHC image, 90% of the segmented nuclei images of each class were selected arbitrarily for the learning procedure of the SVM, while the remaining 10% segmented nuclei images of each class were left for the testing procedure. A 10-fold cross validation was used to ensure the testing of each segmented nucleus image. The positive nuclei that had been correctly classified from all 10-fold were then counted to yield the second element required to calculate the PS (total number of positive nuclei). In cases where the automated counting of nuclei exceeded the number of

nuclei counted by the pathologists the (FN) background objects that had been extracted as blue cells, were considered as blue nuclei in this study. Classification results for each image used are presented in Table 4-5 (A-C). The last two columns represent the positive and negative nuclei images which have been correctly classified by the proposed classification approach respectively. The positive and negative nuclei datasets are the ground truth data which were used to generate the training and testing set.

Table 4-5:

A

Images of set Ki67 Day 19	Patho. Coun.	Path. Positive Nuclei	Patho. negative nuclei	Auto. coun.	positive nuclei dataset	Negative nuclei dataset	Classifica. of positive nuclei	Classifica. of negative nuclei
1) Ki67 Day 19 high staining	65	44	21	68	44	24	43	23
2) Ki67 Day 19 high staining	103	54	49	100	54	46	53	43
3) Ki67 Day 19 high staining	78	37	41	77	38	39	36	37
4) Ki67 Day 19 intermediate staining	100	60	40	93	59	34	57	31
5) Ki67 Day 19 low staining	67	15	51	73	14	59	14	55

B

Images of set p27 Day 19	Patho. Coun.	Path. Positive Nuclei	Patho. negative nuclei	Auto. coun.	positive nuclei dataset	Negative nuclei dataset	Classifica. of positive nuclei	Classifica. of negative nuclei
1) p27 Day 19 high staining	93	60	33	95	58	37	56	34
2) p27 Day 19 intermediate staining	74	46	28	80	43	37	42	36
3) p27 Day 19 intermediate staining	95	44	51	105	43	62	43	58
4) p27 Day 19 low staining	69	18	51	81	18	63	17	58
5) p27 Day 19 low staining	76	18	58	86	17	69	17	63

C

Images of set Ki67 Day 22	Patho. Coun.	Path. Positive Nuclei	Patho. negative nuclei	Auto. coun.	positive nuclei dataset	Negative nuclei dataset	Classifica. of positive nuclei	Classifica. of negative nuclei
1) Ki67 Day 22 high staining	100	51	49	105	51	54	49	52
2) Ki67 Day 22 intermediate staining	74	51	23	83	51	33	50	30
3) Ki67 Day 22 intermediate staining	94	66	28	97	69	28	66	24
4) Ki67 Day 22 intermediate staining	49	28	21	47	25	22	24	21
5) Ki67 Day 22 intermediate staining	73	32	41	80	31	51	31	46

Table 4-6: (A-C) represent the PS score given by the pathologist, the PS score generated using the proposed method and the PS accuracy of the proposed method.

A

Images of set Ki67 Day 19	PS pathology	PS Auto.	PS accuracy %
1) Ki67 Day 19 high staining	67	63.2	93.3
2) Ki67 Day 19 high staining	52.4	53	98
3) Ki67 Day 19 high staining	47.4	46.7	98.6
4) Ki67 Day 19 intermediate staining	60	61.2	98
5) Ki67 Day 19 low staining	22.3	19.1	85.6

B

Images of set p27 Day 19	PS pathology	PS Auto.	PS accuracy %
1) p27 Day 19 high staining	64.5	58.9	94.7
2) p27 Day 19 intermediate staining	62.2	52.5	84.4
3) p27 Day 19 intermediate staining	46.3	40.9	88.4
4) p27 Day 19 low staining	26.1	20.9	80.4
5) p27 Day 19 low staining	23.7	19.7	83.1

C

Images of set Ki67 Day 22	Ps pathology	Ps Auto.	Ps accuracy %
1) Ki67 Day 22 high staining	51	46.6	91.5
2) Ki67 Day 22 intermediate staining	71.8	60.2	83.9
3) Ki67 Day 22 intermediate staining	70.2	68	96.8
4) Ki67 Day 22 intermediate staining	57.1	51	89.3
5) Ki67 Day 22 intermediate staining	43.8	38.7	88.3

It is worth mentioning that PS score accuracy was reduced in cases where the number of nuclei objects counted by the automated method exceeded the total number of nuclei objects found by the pathologist. The average PS accuracy achieved for each image dataset is; 94% for the Ki67-Day-19 dataset, 86% for the p27-Day-19 dataset and 90% for the Ki67-Day-22 dataset. Overall PS accuracy for all datasets was 90% for all the above tested IHC images based on 1270 segmented nuclei images.

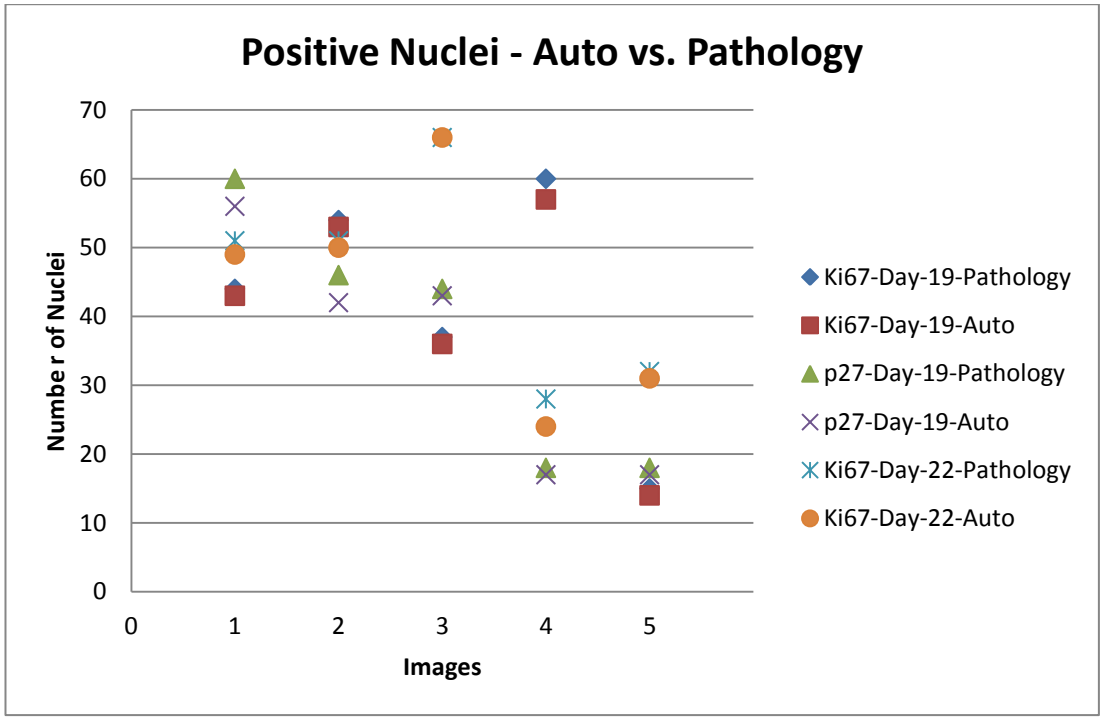


Figure 4-10: scatter plot representing the correlation between the pathologist results and the proposed method results in classifying the positively stained nuclei.

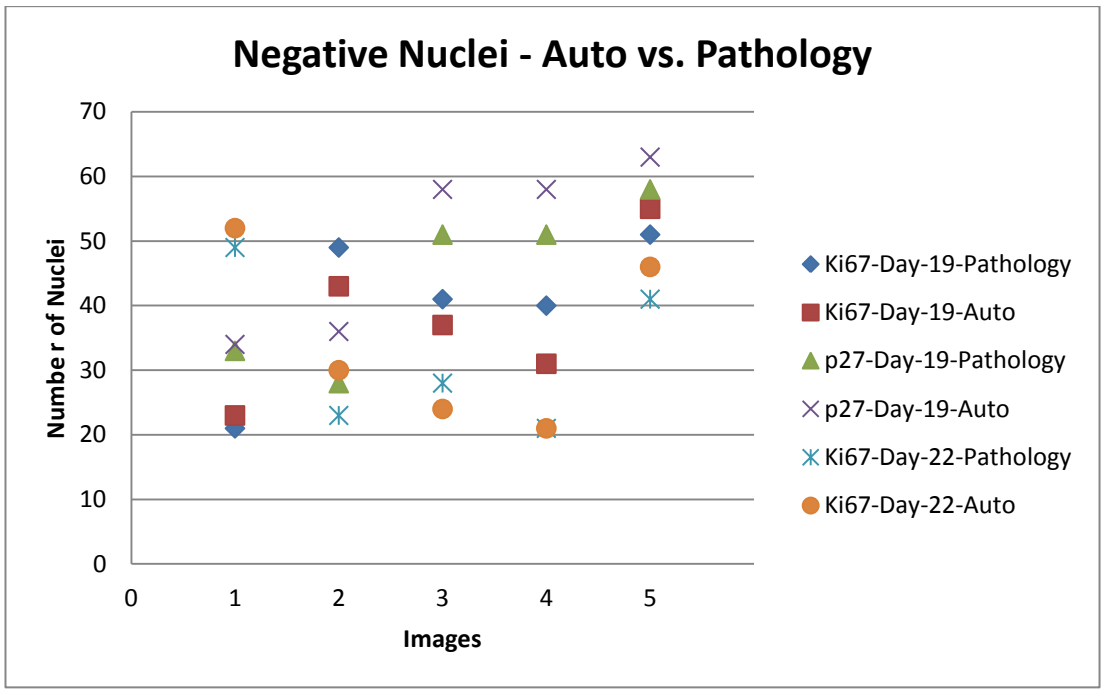


Figure 4-11: scatter plot representing the correlation between the pathologist results and the proposed method results in classifying the negatively stained nuclei.

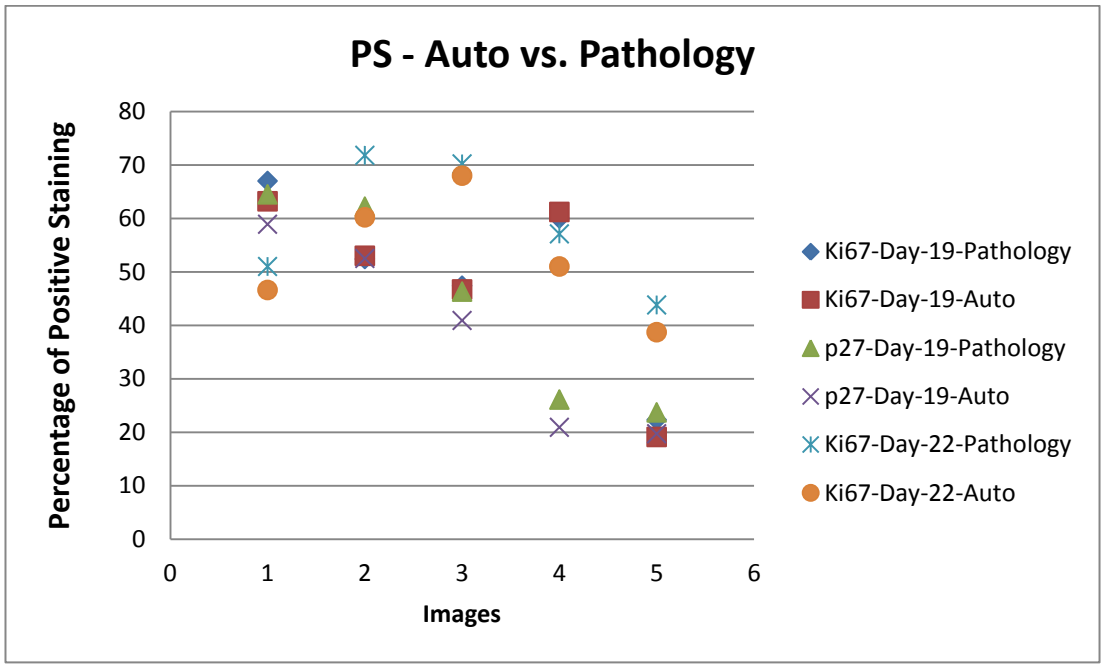


Figure 4-12: scatter plot representing the correlation between the pathologist results and the proposed method results in determining the PS score.

Chapter 5

Positive Nuclei Stain Intensity Classification

5.1 Introduction

Antibody staining plays a vital role in the quantification of protein expression in IHC images. However, immunohistochemical staining is often heterogeneous, which results in variable staining levels within a tissue sample. Pathologist score the staining intensity of positively stained nuclei (IS) based on their experience: each IHC image is given a score of low (+), intermediate (++) or high (+++), regardless of the staining heterogeneity amongst the positively stained nuclei in one image or slide.

In this chapter we propose a cell-level heterogeneity automated method that scores each positive nuclei based on the intensity level of the positive stain. Each nucleus will be given a value of (+1, +2, or +3), depending on the intensity level of the positive stain inside each nuclei. Finally, the percentage of each positive staining intensity level is calculated. This quantitative measurement of the cell-level heterogeneity in IHC images aims to evaluate tumours and identify underlying issues in histopathology which could remain unidentified if relying upon the visual abilities of pathologists.

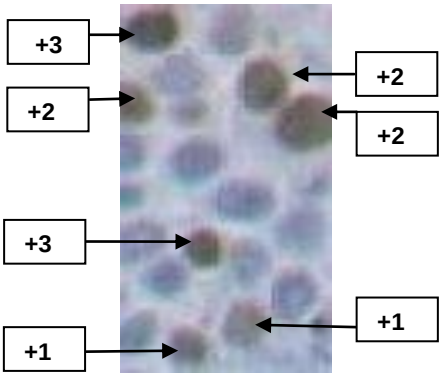
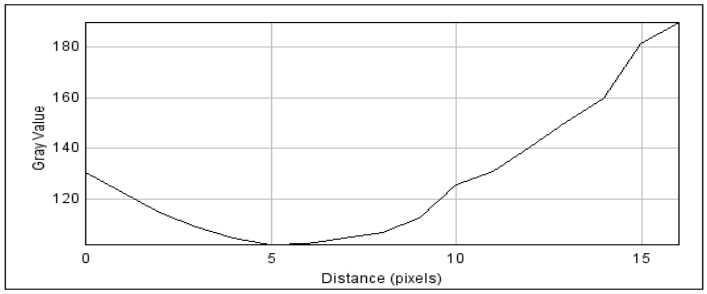
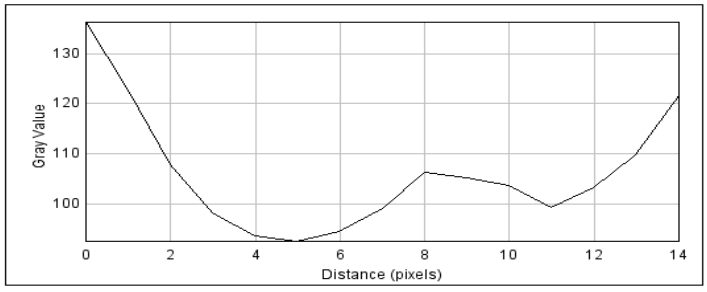


Figure 5-1: represents the staining heterogeneity of positive nuclei. (+1) Denoting low staining, (+2) Denoting intermediate staining and (+3) Denoting high staining.

(A) Low staining (+1)



(B) Intermediate staining (+2)



(C) High staining (+3)

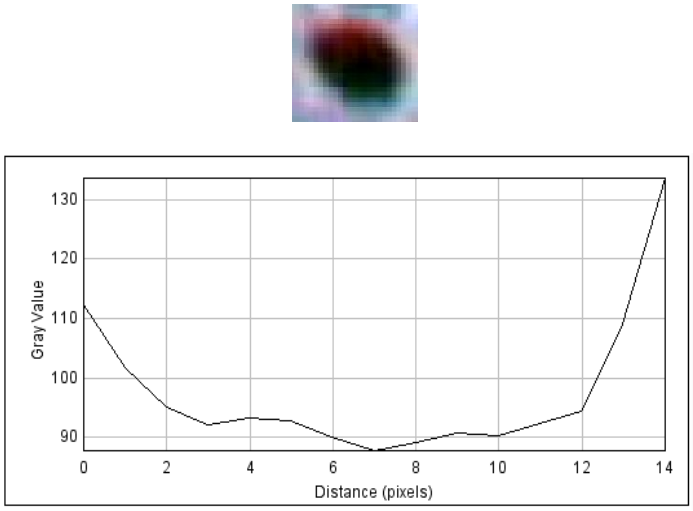


Figure 5-2: (A-C) demonstrates the staining heterogeneity of the three staining levels of positive nuclei. The line plots represent the staining intensity differences between the three staining levels. The one-dimensional line plots were made from the centre row of pixels of the nuclei objects. This was performed to extract their intensity values against their positions.

In this chapter a novel method for quantifying and classifying the staining intensity of positive nuclei is presented. As discussed in the previous chapter, feature extraction is the key point for classifying different objects. In this chapter, since the main objective is to classify the staining intensity of positive nuclei, textural features are investigated to provide the discriminative features required to classify the positive nuclei based on their staining intensity level. One of the most commonly used textural features is the Intensity histogram or pixel intensity features. This chapter demonstrates how other histogram mechanisms such as histogram equalization and k-means clustering could provide features with improved discriminative values compared to the intensity histogram features extracted from the baseline

intensity histogram. This therefore improves the classification and identification of the different staining intensity of positive nuclei, based on the minor textural differences of each intensity level. Image thresholding intensity features are also investigated and analysed in this chapter.

5.2 Textural feature methods

5.2.1 Intensity histogram

The histogram of intensity levels is a summarizing and uncomplicated method, used to generate and observe the statistical information enclosed in an image. It is also known as the first-order statistical information, since it involves the gray level value of each pixel. The probability mass of occurrence is the function used to generate the intensity histogram of an image. This function is given as;

$$p(i) = \frac{h(i)}{MN}, i = 0,1, \dots, G-1 \qquad (5.1)$$

Where $h(i)$ is a certain intensity value, MN is the image dimensions, i is the discrete intensity values in an image and G is the total number of intensity levels in an image.

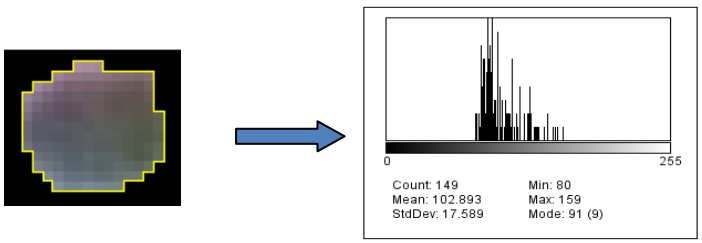


Figure 5-3: represents the intensity histogram of a high stained positive nucleus

The shape of the histogram provides many indications as to the characteristics of the image. This means image features can be described and extracted from the histogram to quantitatively illustrate the first-order statistical features of an image. The main features that can be extracted from the intensity histogram are: the maximum histogram intensity value or (max), minimum histogram intensity value or (min), mean, standard deviation, skewness, and kurtosis. The min and max value represent the lowest intensity value (0) and the highest intensity value (255) in a histogram respectively. The mean describes the average level of intensity in the region of interest, which in this case is the segmented positive nuclei object, whereas standard deviation describes the variance from the mean gray value within the region of interest. The skewness is a measure of the asymmetry of the distribution of the gray values in a histogram around the mean within the region of interest. Kurtosis is a measure of the peakedness or the flatness of histogram gray values around the mean. When kurtosis scores a higher value, this indicates that the variance is due to infrequent excessive deviations (Hafizah et al., 2012).

Table 5-1: represents the intensity textural features extracted from the image intensity histogram.

Histogram intensity textural features	
Maximum histogram intensity	(Max)
Minimum histogram intensity	(Min)
Mean	$\mu = \sum_{i=0}^{G-1} i \times p(i)$ (5.2)
Standard Deviation	$\sigma = \sqrt{\sum_{i=0}^{G-1} (i - \mu)^2 \; p(i)}$ (5.3)

Skewness	$\mu_3 = \sigma^{-3} \sum_{i=0}^{G-1}(i - \mu)^3 p(i)$	(5.4)
Kurtosis	$K = \sigma^{-4} \sum_{i=0}^{G-1}(i - \mu)^4 p(i) - 3$	(5.5)

These textural features are selected in this chapter to extract discriminative characteristics, which will then be used to classify positively stained nuclei into three classes based on their intensity level (+1, +2 or +3).

5.2.2 Histogram equalization

Histogram equalization is a technique used to enhance the contrast of an image by increasing the intensity histogram uniformity of an image. This is usually performed using the normalized cumulative histogram function to remap the histogram intensities into more equalized distribution along the intensities gray scale. The normalized cumulative histogram function can be defined as;

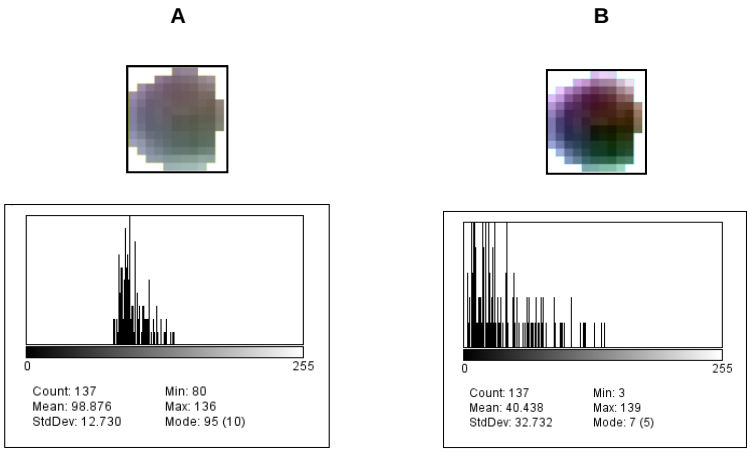
$$cdf(pi) = \sum_{j=0}^{i} p(j) \qquad (5.6)$$

where p is an image histogram and i is a certain intensity level within the histogram (Krutsch and Tenorio, 2011; Rubner et al., 1998).

In this chapter, the histogram equalization method was used to improve the contrast of the nuclei image with different staining intensities (+1, +2 and +3) by stretching out the intensities of a nuclei image into more equalised arrangements over the intensity scale. The aim of this step is to be able to

extract more identifying and discriminative features form the equalized image histogram, which will then be used to classify the three different staining intensities of positively stained nuclei. This will be performed using a number of steps; first, the intensities of the segmented nuclei image histogram are adjusted by limiting the range to a certain maxima and minima. The minima and maxima are adjusted to include what is thought to be a pixel intensity belonging to a positive nuclei object. This step is used to eliminate noise which could accrue from pixels on the border of the nuclei object that could belong to the background. This can be achieved by adjusting the minima to a certain intensity level which is thought to be the lowest intensity level of positive nuclei. In this study, the lowest intensity level is thought to be 180. The maxima is set to the highest intensity level of positive nuclei objects and is believed to be 80. The aim of performing this step is to limit the intensity range to only those pixels that identify a positive nuclei object. The maxima and minima in this case were selected based on the analysis of 1000 positive nuclei objects: see Figure 5-4: (A). The second step of the histogram equalization is to normalize the adjusted cumulative histogram from 0 to 255 using the normalize histogram cumulative remapping function. This will stretch the range of the selected intensities over the total intensity scale: see Figure 5-4: (B).

Figure 5-4: (A) illustrates the selection of minima and maxima in the intensity histogram of a positive nuclei object. (B) illustrates the equalised histogram after the remapping function.



Intensity histogram Equalization histogram

Figure 5-5: (A) & (B) illustrate the results of the equalised histogram of a positive nuclei object: improved contrast is visible as result of the equalised histogram.

5.2.3 K-means clustering

The *K*-means is one of the uncomplicated unsupervised learning algorithms which is used to cluster pixels into regions of pixels that belong together. Pixel clustering using *K*-means divides data into a number of groups known as *k* clusters. The number of clusters is selected to suit the task and targeted output. *K*-means uses *k* centroids by assigning one centroid for each cluster: the centroids are placed at different points, so that they are distant from each other. The maximum distance between each centroid point is selected. Once the centroid points are determined, each pixel or data point is assigned to the

nearest centroid. The next step is to take the mean of the pixel values inside each cluster to generate new centroid points, and this will cause the centroids to move location. The same procedure is repeated until the same pixels are assigned to each cluster in every round at the same location. The main aim of this procedure is to minimize square error (Witten and Frank, 2005; Maurya et al., 2011).

K-means clustering can be summarized in four main points indicating each step of the clustering procedure;

1. A number of *K* points are selected based on the maximum distance between each *K* point: these points represent the initial centroids.
2. Each pixel or data point is allocated to the nearest centroid.
3. The mean of data points for each cluster is recalculating and used to generate new centroids.
4. The procedure is repeated until the locations of all centroids are stationary.

In this chapter, *k*-means clustering is used to cluster the intensities of pixels into three clusters: each cluster will represent a certain intensity of the brown stain within the positive nuclei (see Figure 5-6). The aim of using these clusters is to be able to extract more distinctive features from the clustered data, which will then be used to classify the high, intermediate and low stained positive nuclei.

A B

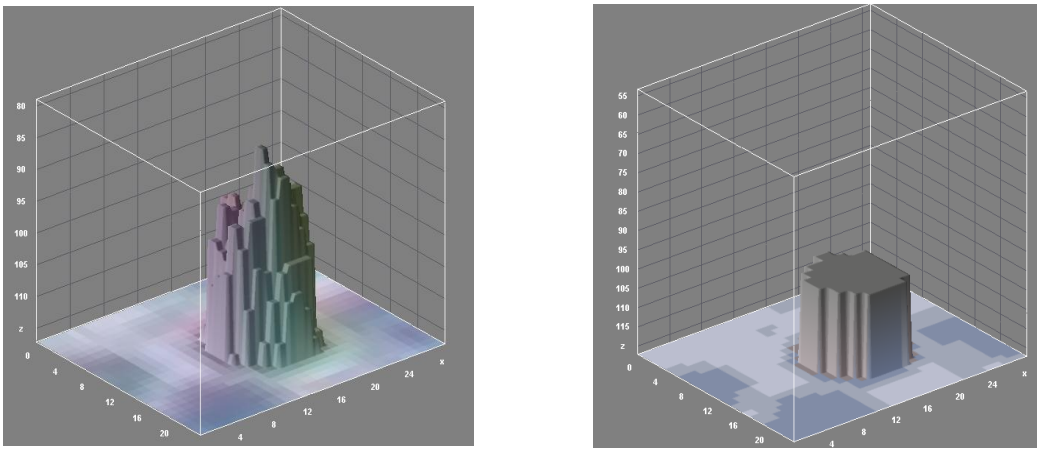


Figure 5-6: (A) represent the pixel intensities of a positive nucleus before Appling the K-means clustering technique. (B) Represents the clustered nucleus object using 3 clusters, the graph represents the quantity of pixels belonging to a certain cluster in the positive nucleus.

5.3 Textural feature analysis

In this section, the use of three methods on the different staining intensities of positive nuclei is demonstrated. The purpose of using these methods in this chapter is to compare the feature vectors of these methods, and analyse their validity in providing powerful discriminative textural features of heterogeneously stained positive nuclei.

The first method used is thresholding or binarization, in which the segmented nuclei image is binarized using a thresholding method (Choudury et al., 2010). Since the thresholding method is directly applied to the segmented nuclei image, the pixel values inside each nucleus are divided into two categories, 0 and 1. In this case, pixels with a value of 0 represent pixels with

low intensity values and pixels which have been assigned a value of 1 represent pixels with high intensity values. From Figure 5-7, it is concluded that this method provides weak variation between high and intermediate stained nuclei, while improved distinction of low stained nuclei is reported.

The second method is histogram equalization: this method is used on the segmented nuclei images. At this stage, histogram equalization is used in its default form, in which the maximum and minimum range is set to the standard form (0-255) intensity levels. The results are demonstrated in the actual nuclei image as well as the image histogram: see Figure 5-7.

It is visible from the nuclei images that the contrast of the segmented nuclei is significantly improved. However, background noise around the edges of the nuclei objects could not be eliminated, and this has resulted in an unnecessary stretch in the equalised histogram intensities, which could influence the values of the textural features. The third method examined in this section is the *K*-means clustering; the *K*-means clustering is applied to the equalized image. The purpose of using this method is to cluster those equalized intensities into three clusters, each representing a certain intensity category of the positive nuclei. The results of using this method are demonstrated in Figure 5-7.

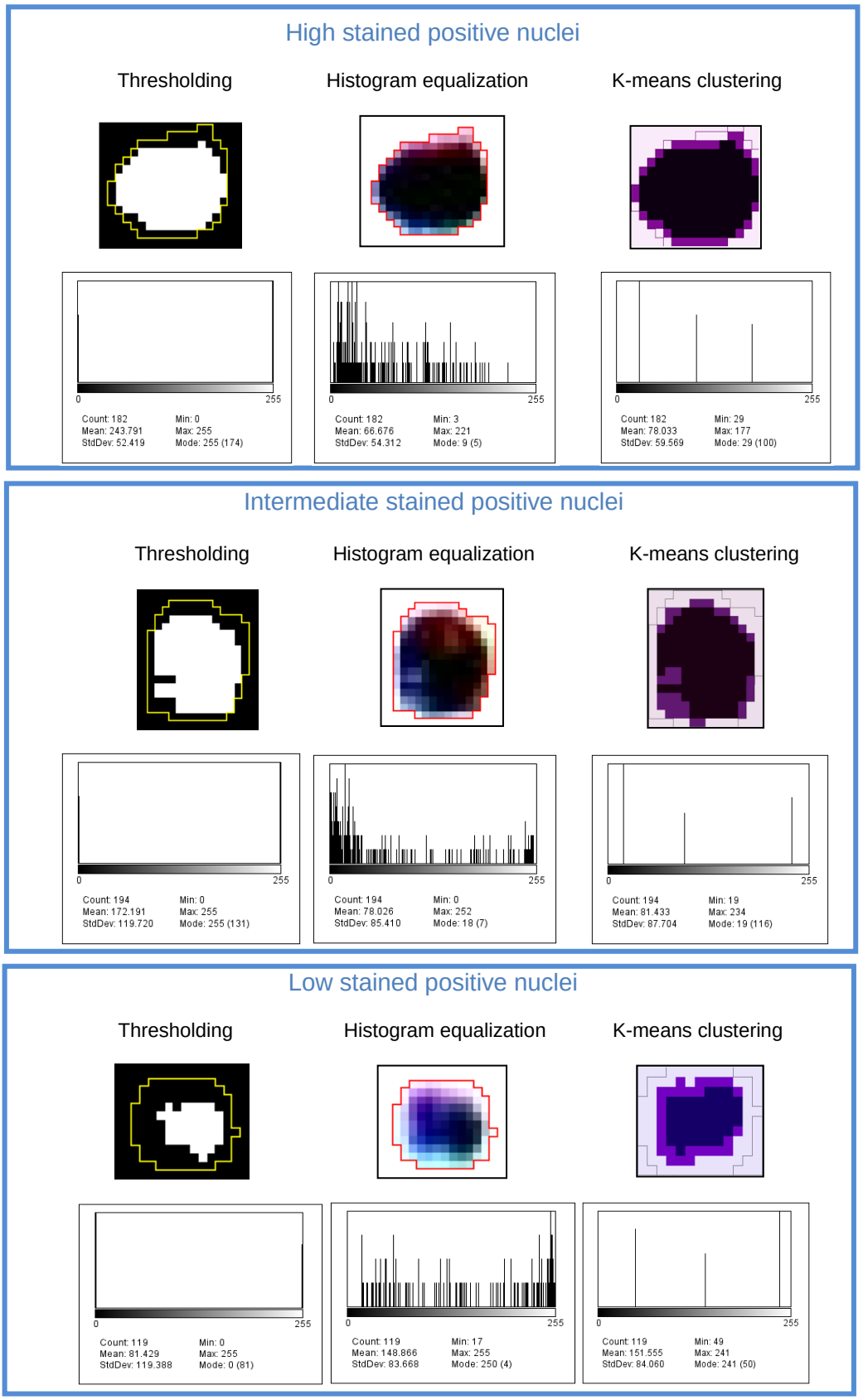


Figure 5-7: represents the effects on the nuclei object at different staining intensities after Applying; thresholding, histogram equalization and *K*-means clustering.

5.4. Discussion

From the previous section, it is evident that histogram equalization achieved an improved nuclei image contrast. This could ultimately improve the quality of textural features extracted from the equalized image. However, a certain level of background noise accrued around the edges of the nuclei objects. To minimise this noise, the minima and maxima is adjusted in the manner discussed in Section (5.2.2). This approach eliminates the low pixel intensity values around the edges of the nuclei by remapping the intensities of those pixels into higher intensities within the selected range.

After the nuclei image is successfully equalised, the K-means clustering based on three clusters is applied. It is interesting to analyse the results of applying the k-means clustering on the image prior to and following the histogram equalisation. Such results are presented in Figures 5-8 to 5-10. The results demonstrate that clustered images which had been previously equalised provided more intensity details in the form of three intensity clusters. Clustered images which had not been formerly equalised showed a reduction in the number of clusters, e.g. two clusters as presented in Figures 5-8 to 5-10 (D). This provides that the use of equalised images prior to *K*-means clustered images provided more intensity details which are represented in the form of intensity clusters. Results are provided in Figures 5-8 to 5-10 for low intermediate and low stained positive nuclei respectively.

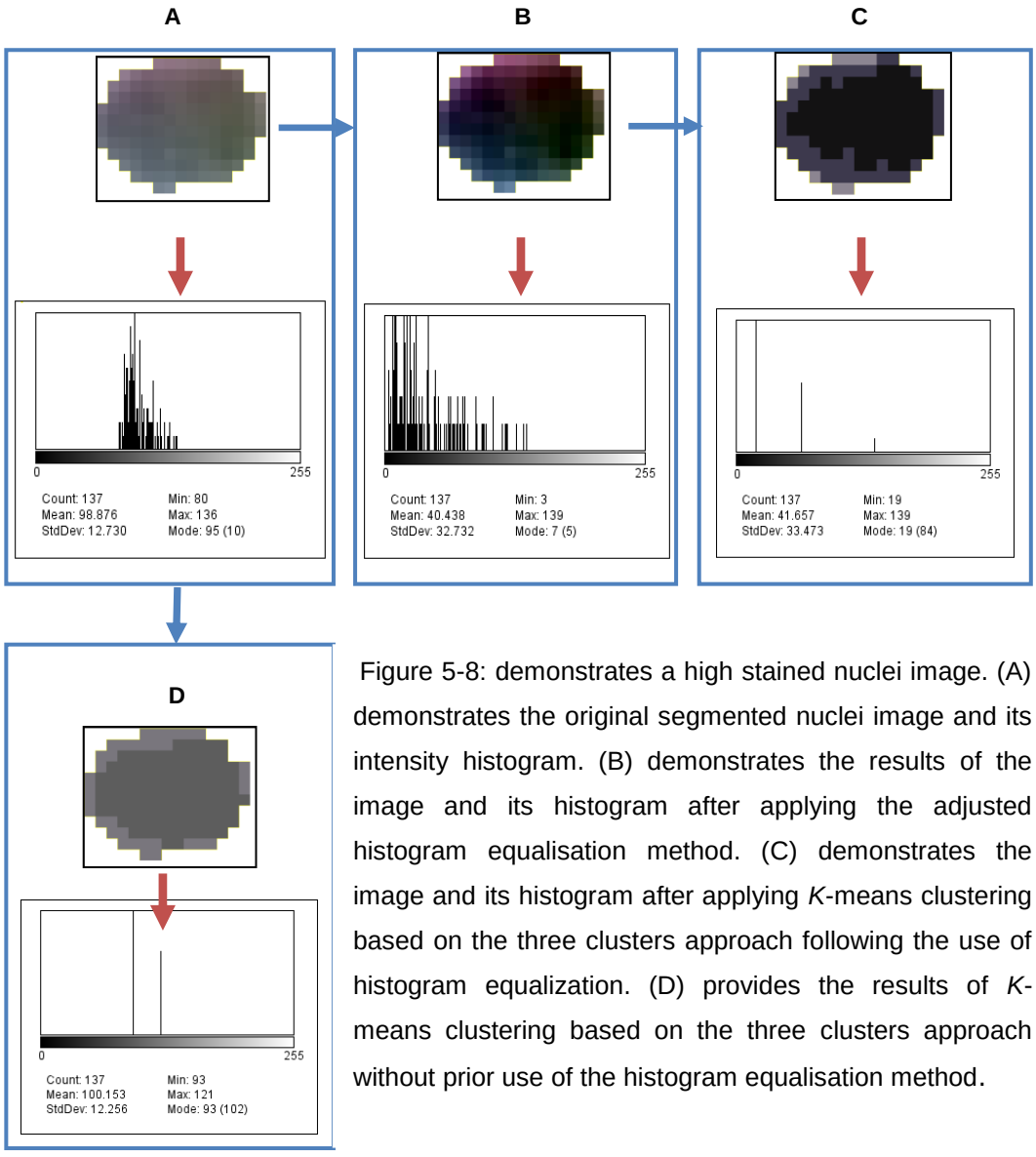


Figure 5-8: demonstrates a high stained nuclei image. (A) demonstrates the original segmented nuclei image and its intensity histogram. (B) demonstrates the results of the image and its histogram after applying the adjusted histogram equalisation method. (C) demonstrates the image and its histogram after applying *K*-means clustering based on the three clusters approach following the use of histogram equalization. (D) provides the results of *K*-means clustering based on the three clusters approach without prior use of the histogram equalisation method.

This figure provides the result of applying the histogram equalisation in its adjusted form followed up by the use of *K*-means clustering. it is noticeable that the equalised histogram in Figure 5-8 (B) has a higher density of data at the higher intensity levels (closer to 0), while the density seem to gradually reduce as in moves down along the intensity scale, and the last third of the

intensity scale is empty, indicating that pixels belonging to this range of low intensities do not exist in the high stained positive nuclei objects.

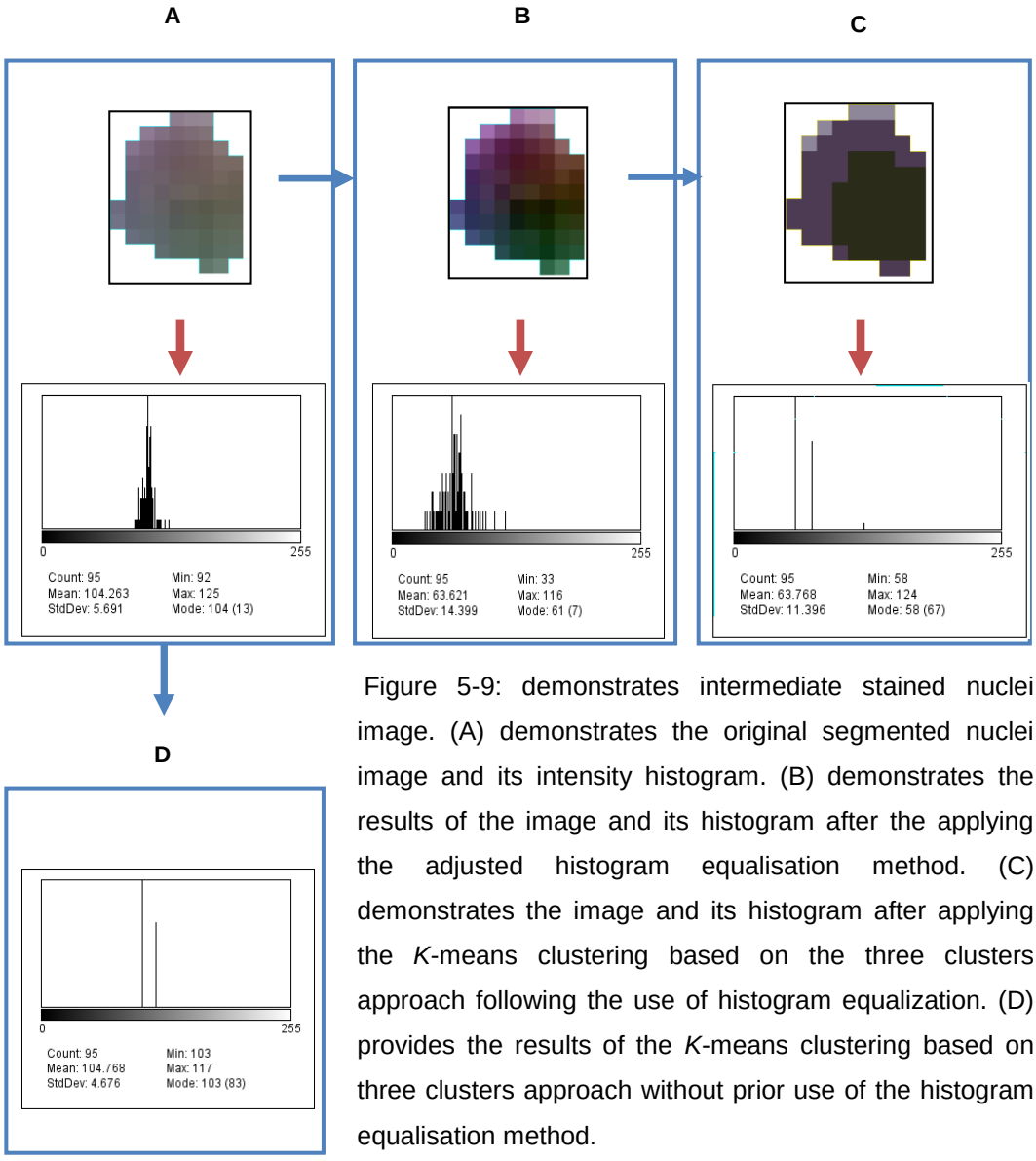


Figure 5-9: demonstrates intermediate stained nuclei image. (A) demonstrates the original segmented nuclei image and its intensity histogram. (B) demonstrates the results of the image and its histogram after the applying the adjusted histogram equalisation method. (C) demonstrates the image and its histogram after applying the *K*-means clustering based on the three clusters approach following the use of histogram equalization. (D) provides the results of the *K*-means clustering based on three clusters approach without prior use of the histogram equalisation method.

From Figure 5-9 (B), it is apparent that the intensity scale values from 0 to 33 are empty, and this indicates that pixels belonging to this value range do not exist inside the intermediate stained positive nuclei. It also provides that

underpopulated intensities seem to exist between the peak of the intensities and the middle point of the scale, until these disappear totally before the midpoint of the scale.

These indications can be used to discriminate between this type of stained nuclei and other types as it provides different characteristics compared to the high and low positive stained nuclei.

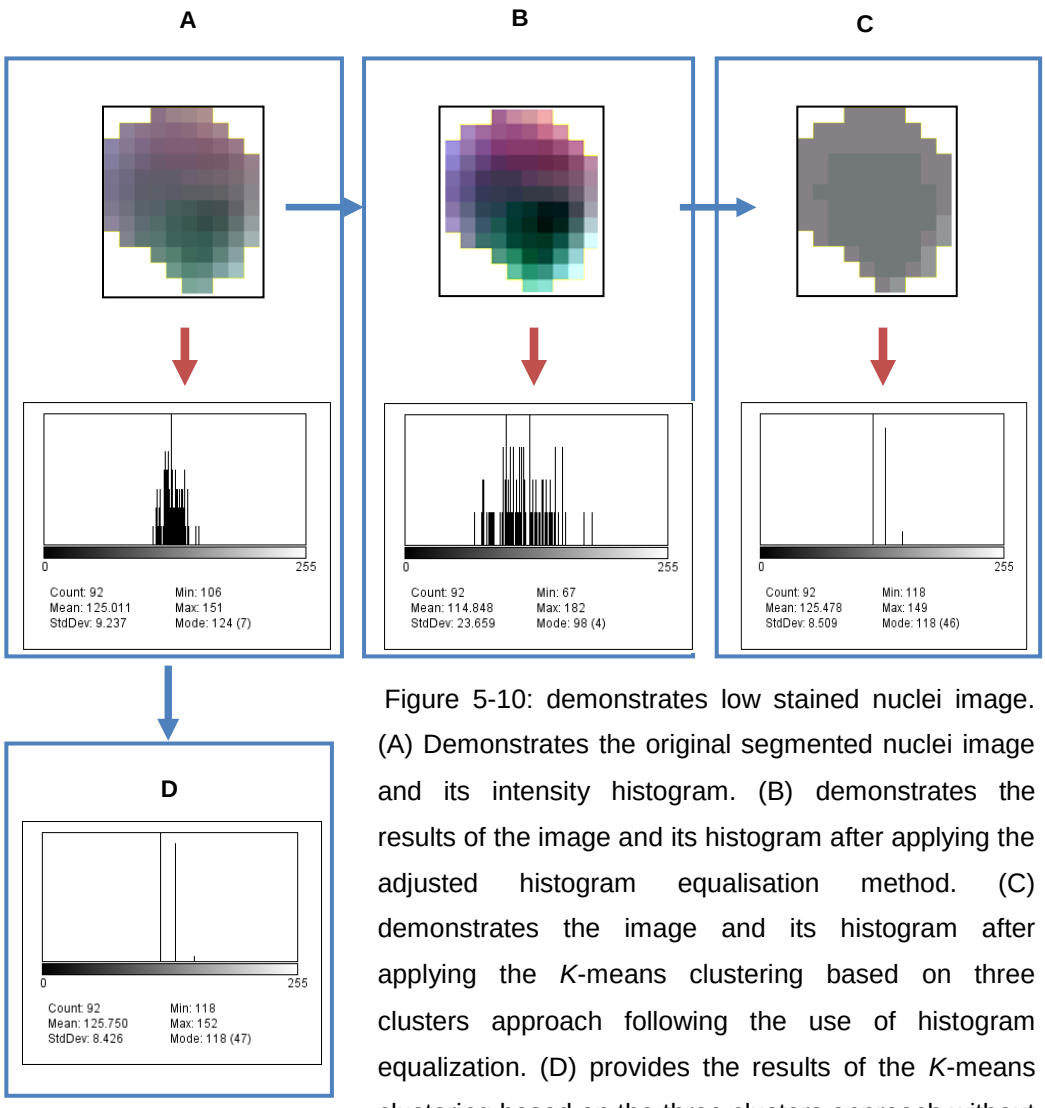


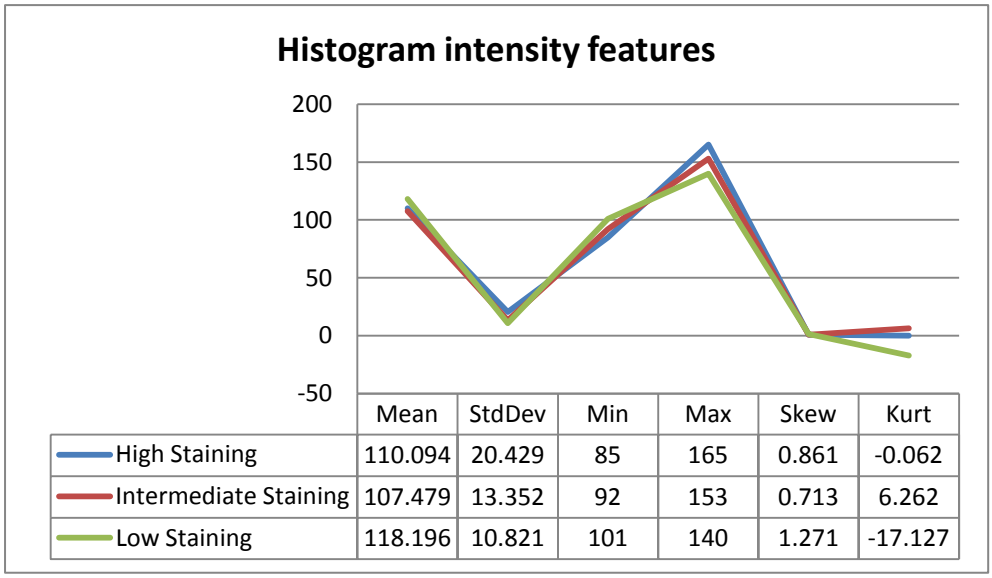
Figure 5-10: demonstrates low stained nuclei image. (A) Demonstrates the original segmented nuclei image and its intensity histogram. (B) demonstrates the results of the image and its histogram after applying the adjusted histogram equalisation method. (C) demonstrates the image and its histogram after applying the *K*-means clustering based on three clusters approach following the use of histogram equalization. (D) provides the results of the *K*-means clustering based on the three clusters approach without prior use of histogram equalisation method.

Figure 5-10 (B), the equalised histogram, shows that the intensity levels are intermediately dense around the middle part of the scale. Pixels with relatively high and low intensities do not exist as noticeable from this particular histogram. However, the data in the middle part of the scale seem to be scattered, which could be an indication of the low stained nuclei objects which seem to have medium to low intensity pixels. From Figures 5-8 to 5-10, it can be concluded that the equalised histogram provided evident discriminative characteristics between high, intermediate and low stained nuclei. On the other hand, discrimination between the different staining intensities of positive nuclei using original image intensity histograms appears to be insufficient due to the lack of discriminative characteristics.

To evaluate the performance of the methods mentioned (intensity histogram, threshold, histogram equalization and *k*-means clustering), textural features are extracted from 1200 positive nuclei (400 high stain, 400 intermediate stain and 400 low stain). The average of each staining category is calculated and the validity of their use in providing discriminative characteristics of high intermediate and low stained positive nuclei is presented through line graphs. Finally the use of the proposed improved histogram equalisation followed up by the *K*-means clustering method is demonstrated. The results are demonstrated below;

Histogram intensity features provided inefficient discriminative characteristics between the high intermediate and low stained positive nuclei. The mean, standard deviation, minima, and skewness features almost overlapped for the three staining categories. Maxima and kurtosis features provided

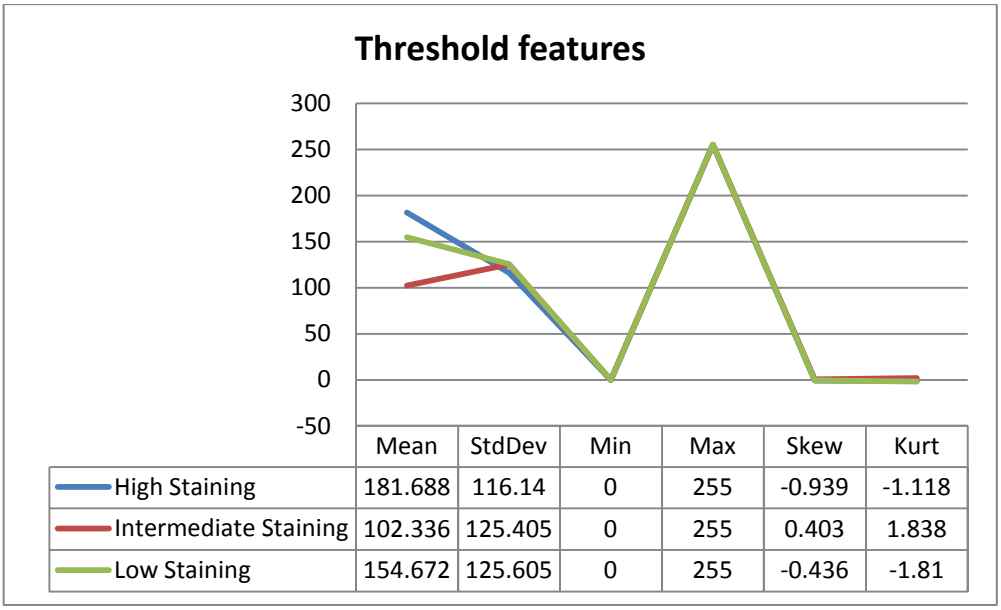
relatively improved discriminative values over the three staining categories. The results are demonstrated in Figure 5-11.



	Mean	StdDev	Min	Max	Skew	Kurt
High Staining	110.094	20.429	85	165	0.861	-0.062
Intermediate Staining	107.479	13.352	92	153	0.713	6.262
Low Staining	118.196	10.821	101	140	1.271	-17.127

Figure 5-11: represents the intensity histogram features for the three levels of nuclei positive staining.

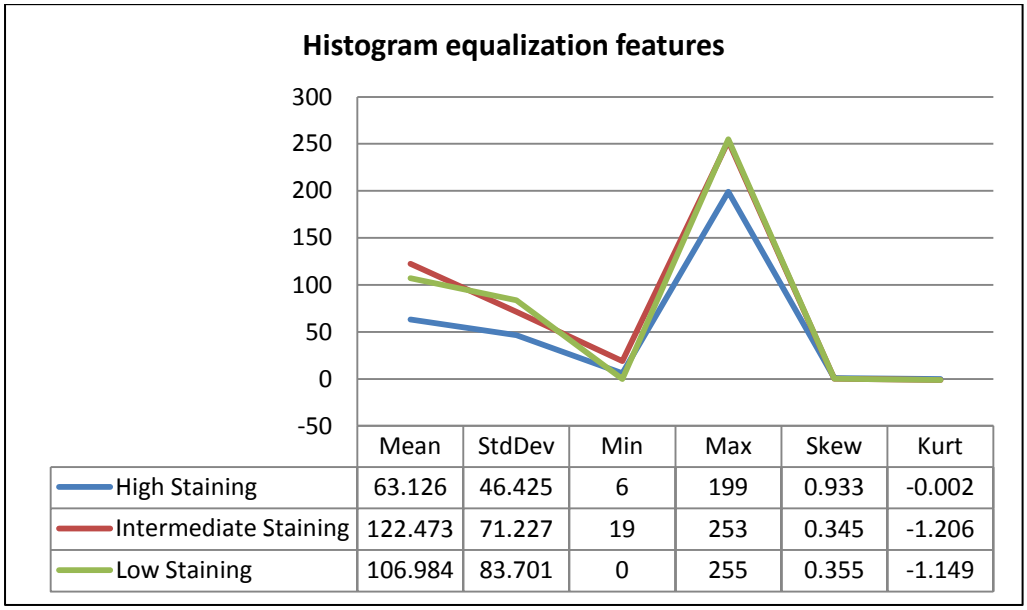
Thresholding also provided inefficient discriminative characteristics between high intermediate and low stained positive nuclei. The standard deviation, min, skewness, max and kurtosis features almost overlapped for all the three staining categories. The mean feature provided relatively improved discriminative value over the three staining categories. The results are demonstrated in Figure 5-12.



	Mean	StdDev	Min	Max	Skew	Kurt
High Staining	181.688	116.14	0	255	-0.939	-1.118
Intermediate Staining	102.336	125.405	0	255	0.403	1.838
Low Staining	154.672	125.605	0	255	-0.436	-1.81

Figure 5-12: represents the threshold features for the three levels of nuclei positive staining.

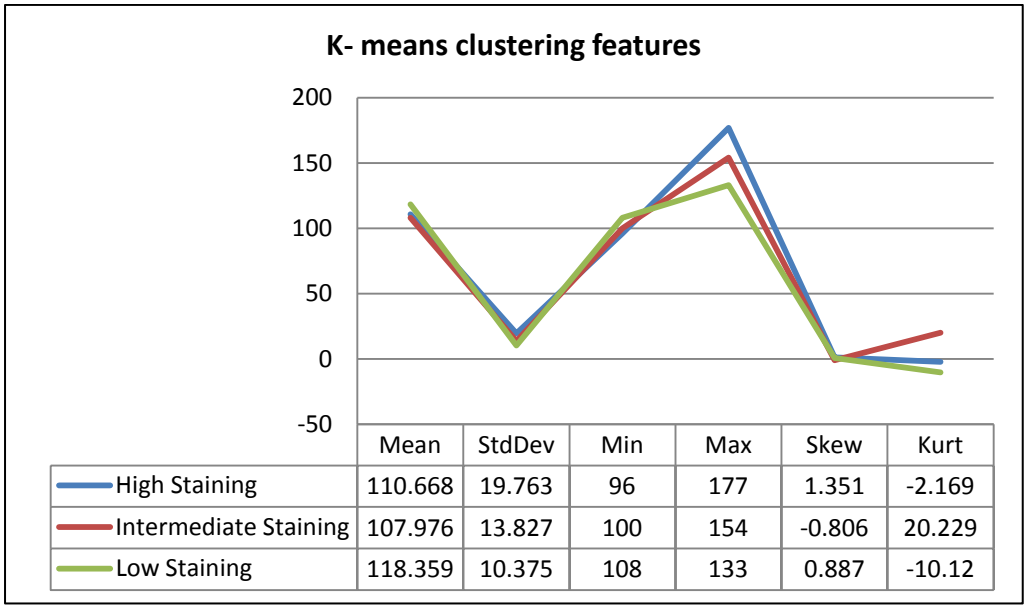
Histogram equalization features provided improved discriminative characteristics between the three categories compared to the two previously demonstrated methods. The mean, standard deviation, min and max features provided a reasonable level of discrimination between the three staining categories. However, overlapping between the high and low stained nuclei at the min feature and between intermediate and low at the max feature were apparent. Skewness and kurtosis features more or less overlapped over all staining categories.



	Mean	StdDev	Min	Max	Skew	Kurt
High Staining	63.126	46.425	6	199	0.933	-0.002
Intermediate Staining	122.473	71.227	19	253	0.345	-1.206
Low Staining	106.984	83.701	0	255	0.355	-1.149

Figure 5-13: represents the equalized histogram features for the three levels of nuclei positive staining.

The *K*-means feature, when used on its own, provided a reasonable level of discrimination between the three staining categories at the max and kurtosis features. Low discriminative characteristics were reported when using the mean, standard deviation, min and skewness.

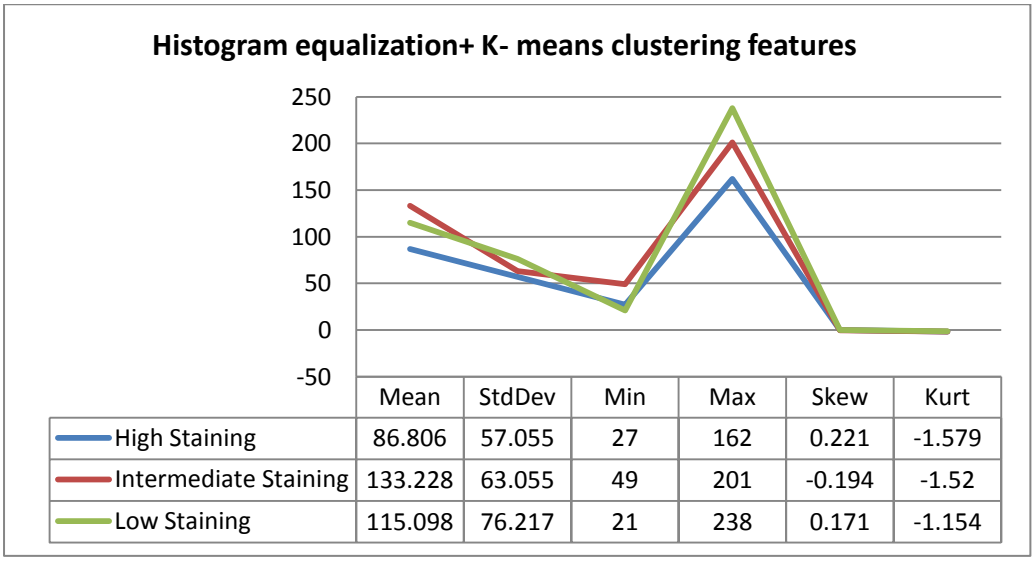


	Mean	StdDev	Min	Max	Skew	Kurt
High Staining	110.668	19.763	96	177	1.351	-2.169
Intermediate Staining	107.976	13.827	100	154	-0.806	20.229
Low Staining	118.359	10.375	108	133	0.887	-10.12

Figure 5-14: represents the K-means clustering features for the three levels of nuclei positive staining.

Finally the proposed method of using the improved histogram equalization followed by *K*-means clustering outperformed all the previous mentioned methods. The mean, standard deviation, max features provided high discriminative characteristics between the three staining categories, with max features being the highest. The Min feature provided a reasonable level of discrimination. Skewness and kurtosis features more or less overlapped over all staining categories.

From Figure 5-15, it is evident that textural features extracted using the proposed method provided an improved level of discrimination over the three categories of positive nuclei staining. Therefore the proposed method is used as the optimal feature extraction method in this study. To optimize the method further, the features providing high discriminative value are selected to be the optimal textural features. Skewness and kurtosis are excluded to reduce the classification error.



	Mean	StdDev	Min	Max	Skew	Kurt
High Staining	86.806	57.055	27	162	0.221	-1.579
Intermediate Staining	133.228	63.055	49	201	-0.194	-1.52
Low Staining	115.098	76.217	21	238	0.171	-1.154

Figure 5-15: represents the intensity histogram features for the three levels of nuclei positive staining.

5.5 Experimental work

5.5.1 First Set of Experiments

In this chapter, three sets of experiments are presented. Experiments were conducted using the SVM classifier. A 10-fold cross validation method was used, where 90% of the nuclei images of each image dataset were selected arbitrarily for the learning procedure of the SVM, while the remaining 10% of nuclei images of each image dataset were left for the testing procedure. The first set of experiments was performed to analyse the performance of the histogram intensity textural features (mentioned in the previous section) in terms of their efficiency in classifying the heterogeneously stained positive nuclei. The six histogram intensity textural features extracted from the original image, thresholded image, the equalized image, the *K*-means clustered image and finally the output image of the proposed method mentioned in Section (5.4) Figure (5-15) that used the improved equalized image followed by the clustering algorithm is used to extract the last set of histogram intensity textural features. The goal of this stage is to select only those textural features that best describes the discriminative characteristics of the high, intermediate and low stained positive nuclei.

For this set of experimental work, three image datasets are used;

Positive nuclei image dataset 1: 1200 segmented nuclei images from the Ki67 Day 19 image dataset are randomly selected. The nuclei selection include three classes; 400 images are high stained positive nuclei, 400 images are intermediate stained positive nuclei and the remaining 400

images are low stained positive nuclei.

Nuclei image dataset 2: 1000 segmented nuclei images from the p27 Day 19 image dataset were randomly selected. The nuclei selection include three classes; 400 images are high stained positive nuclei, 400 images are intermediate stained positive nuclei and the remaining 400 images are low stained positive nuclei.

Nuclei image dataset 3: 1000 segmented nuclei images from Ki67 Day 22 image dataset were randomly selected. Nuclei selection includes three classes; 400 images are high stained positive nuclei, 400 images are intermediate stained positive nuclei and the remaining 400 images are low stained positive nuclei.

All images of segmented positive and negative nuclei were extracted and saved using the extraction method described in Section (3.3). For this set of experiments only the TP segmented positive nuclei images are used to create this image datasets. Following the performance analysis of the histogram intensity textural features in classifying the heterogeneously stained positive nuclei, the features mentioned in the previous chapter section (5.4) are used. The aim of this step is to integrate the wavelet colour features, high difference GLCM features and morphological features with the textural features proposed in this chapter to improve the classification accuracy of the three different types of stained positive nuclei.

19 experiments of features and their combinations were performed as follows:

Four types of feature extraction methods were used; histogram intensity GLCM features were used to extract the textural features, DWT was used to extract colour features and morphological features were used to describe the shape of the segmented positive nuclei.

To evaluate the performance of each feature set: features were used separately to report their discriminative power in classifying the three types of positive stained nuclei. The features that provided the highest level of classification accuracy were then integrated together. The histogram intensity textural features extracted from the original image, thresholded image, equalized image, *K*-means clustered image and the proposed intensity features extraction method were used to classify the three types of stained positive nuclei.

It was evident that the proposed method of feature extraction (equalized + *K*-means clustering) outperformed the rest of the intensity histogram features extracted from the standard extraction methods. To improve the classification performance further, the intensity textural features that provided the most discriminative power were selected (The mean, standard deviation, max and min features) as described in Section (5.4) Figure (5-15) of this chapter.

The DWT features GLCM and Morphological features were used in the same manner described in Section (4.5.1). The coefficient extracted from DWT level one and two collectively provided the highest level of classification

accuracy compared to the rest of the DWT level features. The high discriminative GLCM features provided the highest level of classification accuracy compared to the rest of the GLCM features. To improve the classification accuracy further, features that provided the highest discriminative power from each feature extraction collection were integrated together to form a feature vector of length 29 (4 optimized textural features+ 18 DWT feature + 4 GLCM feature + 3 morphological features) for each individual nucleus. It was evident that the proposed intensity textural features improved the classification accuracy and provided the highest discriminative power when integrated with the 2-level DWT, high difference GLCM and morphological features. The results of the classification performances are reported in Figures 5-16 to 5-19.

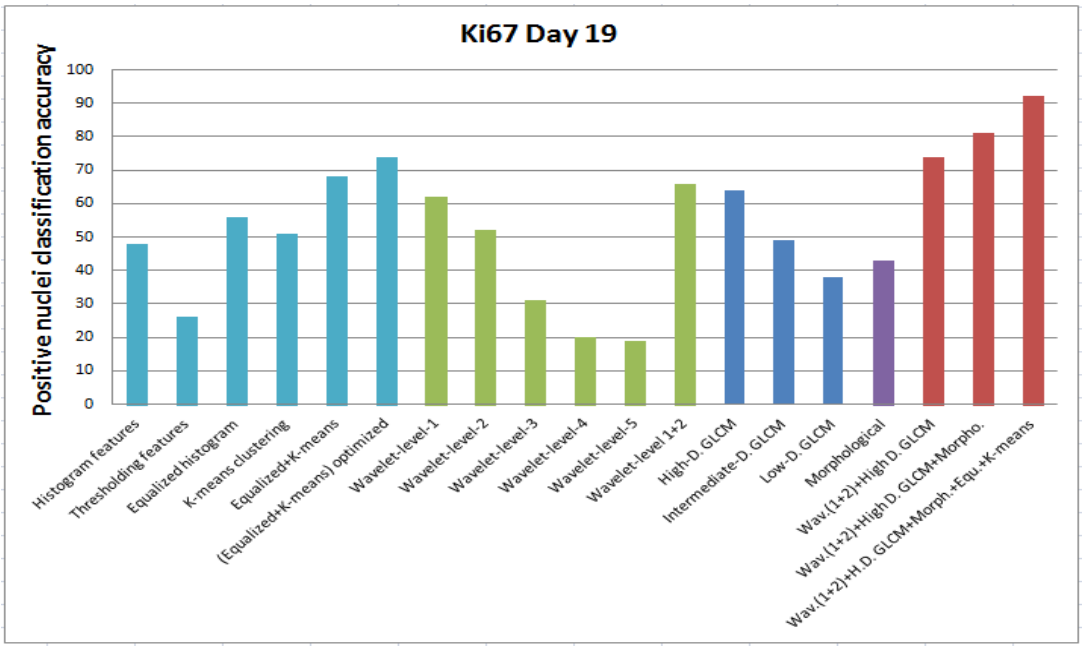


Figure 5-16: accuracies of classifying segmented (high, intermediate and high) positive nuclei images from the Ki67 Day 19 image dataset.

Figure 5-16 represents the classification performance resulting from the 19 experiments of the different features and their combination when used on the Ki67 Day 19 image dataset using the SVM classifier. Bars in cyan represent the performance of histogram intensity textural features, while Bars in green show the performance of positive nuclei classification when using the different levels and combination of DWT. Bars in blue show the performance of positive nuclei classification when using the GLCM features. The bar in purple shows the performance of positive nuclei dataset when using the morphological features. Bars in red show the performance of nuclei image classification when using the combination of the uppermost discriminative features.

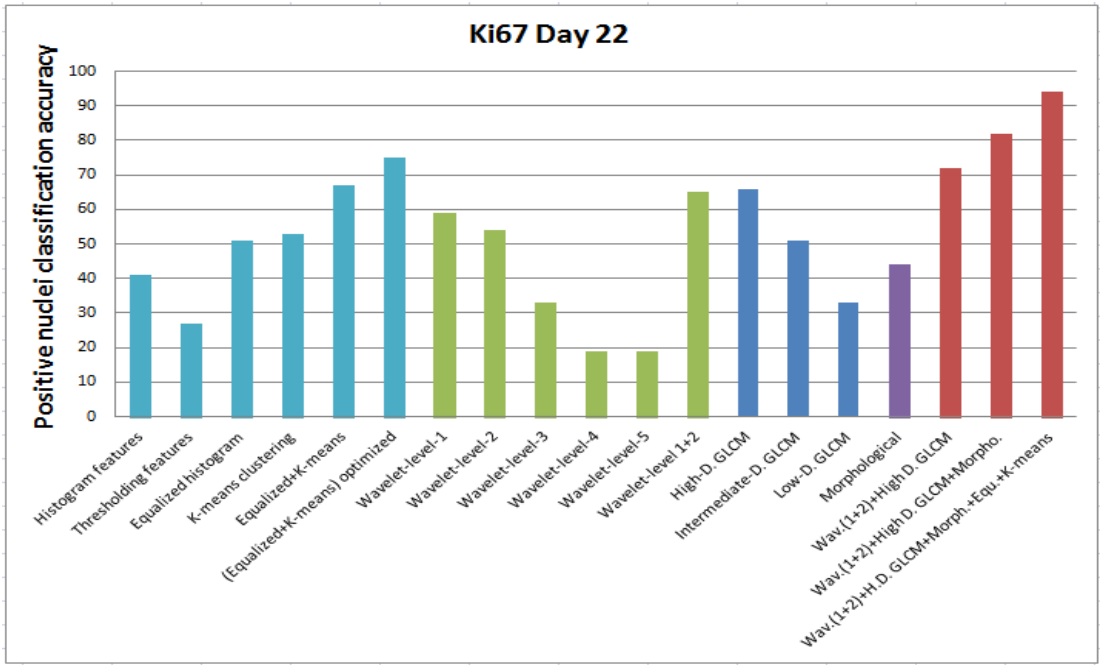


Figure 5-17: represents the SVM accuracies of classifying segmented (high, intermediate and high) positive nuclei images from the Ki67 Day 22 image dataset as results of the 19 experiments of the different features and their combinations.

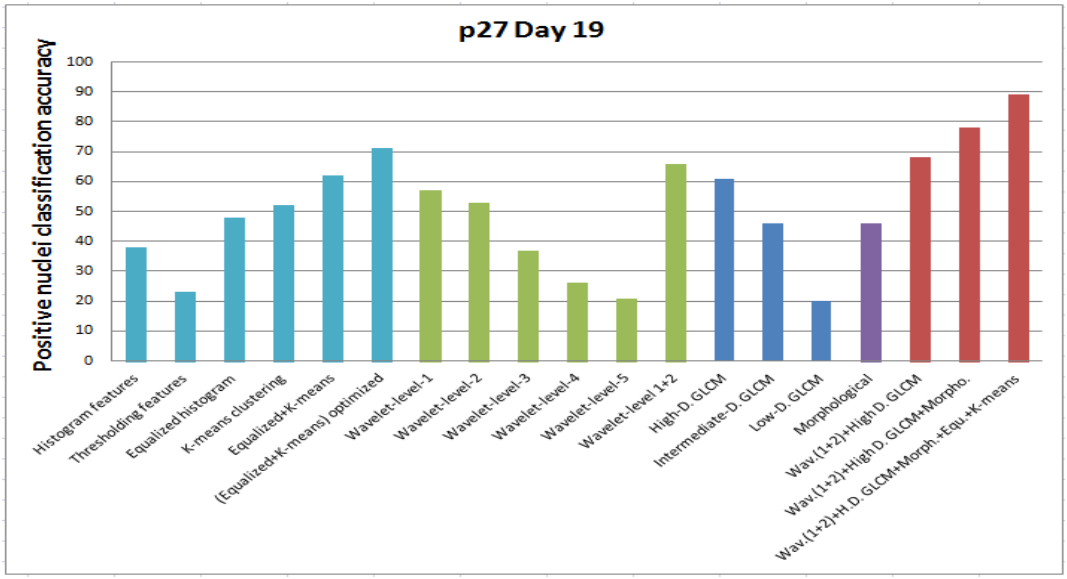


Figure 5-18: represents the SVM accuracies of classifying segmented (high, intermediate and high) positive nuclei images from the p27 Day 19 image dataset as results of the 19 experiments of the different features and their combinations.

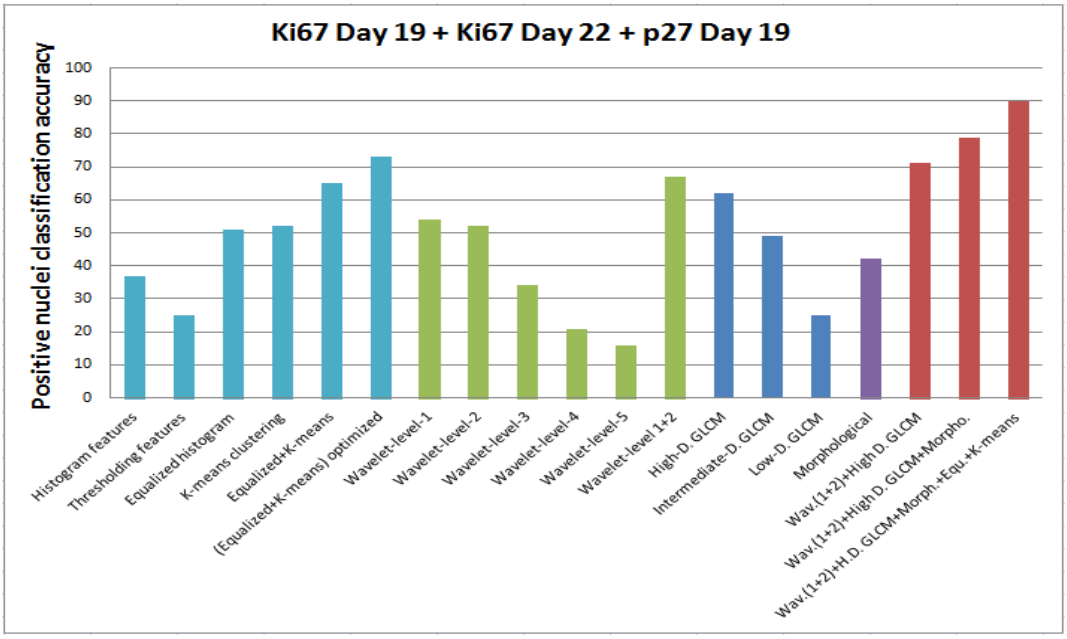


Figure 5-19: represents the SVM accuracies of classifying segmented (high, intermediate and high) positive nuclei images from Ki67 Day 19, Ki67 Day 22 and p27 Day 19 collectively as results of the 19 experiments of the different features and their combinations.

Table 5-2: represents the classification accuracies resulting from the 19 experiments of the differently used features for the segmented positive nuclei images datasets using SVM.

Features	Classification accuracy% of Ki67 Day 19	Classification accuracy% of Ki67 Day 22	Classification accuracy% of P27 day 19	Classification accuracy% of Ki67 Day 19 + Ki67 Day 22 + P27 day 22
Histogram features	48	41	38	37
Thresholding features	26	27	23	25
Equalized histogram features	56	51	48	51
K-means clustring features	51	53	52	52
Equalized+K-means features	68	67	62	65
(Equalized+K-means) optimized features	74	75	71	73
Wavelet-level-1	62	59	57	54
Wavelet-level-2	52	54	53	52
Wavelet-level-3	31	33	37	34
Wavelet-level-4	20	19	26	21
Wavelet-level-5	19	19	21	16
Wavelet-level (1 and 2)	66	65	66	67
High Difference GLCM	64	66	61	62
Intermediate Difference GLCM	49	51	46	49
Low Difference GLCM	38	33	20	25
Morphological	43	44	46	42
Wavelet level (1 and 2) + High difference GLCM	74	72	68	71
Wavelet level (1 and 2) + High difference GLCM + Morphological	81	82	78	79
Wavelet level (1 and 2) + High difference GLCM + Morphological_(equalization +K-means clustering)	92	94	89	90

5.5.2 Second set of experiments

This experiment presented the statistical analysis that is used to calculate the performance accuracy of the proposed classification approach in terms of classifying the high intermediate and low stained nuclei. SVM classification accuracy was measured at 90% when tested on 3600 segmented positive nuclei images of different staining intensities as described in Section (5.5.1). Since the image dataset contains positive nuclei images of three different staining intensities, it was interesting to measure the performance of the proposed classification approach in classifying high, intermediate and low stained positive nuclei independently. The measure selected for this purpose in this section is accuracy. Labels presenting each class are counted to yield the total number of correctly classified high stained positive nuclei, intermediate stained positive nuclei and low stained positive nuclei respectively. In the context of this section, accuracy is used in three forms, each of which represents the classification accuracy for a certain staining intensity. Accuracy is defined as follows;

$$Classification\ Accuracy\ of\ high\ stained\ \ nuclei\ \%\ = \frac{\boldsymbol{correctly\ classified\ high\ stained\ nuclei}}{\boldsymbol{total\ high\ stained\ nuclei}} * 100 \qquad (5.7)$$

$$Classification\ Accuracy\ of\ intermediate\ \ stained\ nuclei\ \%\ = \frac{\boldsymbol{correctly\ classified\ intermediate\ stained\ nuclei}}{\boldsymbol{total\ intermediate\ stained\ nuclei}} * 100 \qquad (5.8)$$

$$Classification\ Accuracy\ of\ Low\ stained\ nuclei\ \%\ = \frac{\boldsymbol{correctly\ classified\ low\ stained\ nuclei}}{\boldsymbol{total\ low\ stained\ nuclei}} * 100 \qquad (5.9)$$

Table 5-3: summarises the classification accuracies for 3600 segmented positive nuclei images.

	All positive nuclei	High stained Positive nuclei	Intermediate stained positive nuclei	Low stained positive nuclei
Ground truth positive nuclei image dataset	3600	1200	1200	1200
Total number of segmented positive nuclei images classified correctly by the proposed approach	3252	1113	1062	1077
Accuracy %	90.3	92.7	88.5	89.7

From the results, it is concluded that the proposed feature extraction and classification approach provided a high level of accuracy in classifying the heterogeneously stained positive nuclei. It is evident that the approach provided the highest level of classification accuracy when used to classify high stained positive nuclei. The lowest classification accuracy reported was when the approach was used to classify intermediate stained nuclei. This could be due to the variable stain characteristics, which could lead the system to incorrectly classify intermediate stained nuclei as low or high stained nuclei. Meanwhile, in cases of the low stained nuclei, accuracy was slightly lower than that reported for high stained nuclei. However, the accuracy for this nuclei type was improved compared to that reported for the intermediate stained nuclei.

5.5.3 Third set of experiments

This section demonstrates how the approach described in this chapter can be used to measure the intensity score IS of individual IHC images. The segmented positive nuclei image datasets that resulted from Section (4.5.4) is used in this section, such that, only the correctly classified positive nuclei

are used. The feature extraction and the classification approach described in this chapter is applied to each segmented positive nuclei individually as follows; the optimal feature (optimized intensity features, two-level-wavelet + high difference GLCM features + morphology features) are used to extract the discriminative features from each class (high, intermediate and low stained positive nuclei).

A 10-fold cross validation method was used to ensure the testing of each nucleus image. 90% of the nuclei images of each class were selected arbitrarily for the learning procedure of the SVM, the remaining 10% nuclei images of each class were left for the testing procedure. The positive nuclei that had been correctly classified were then counted for all three classes to yield the total number of nuclei belonging to each class.

Classification results for each IHC image used are presented in Table 5-4: (A-C). The first column of each table represents the total number of positive nuclei which has been correctly classified using the automated classification method described in Section (4.5.4). The second column represents the total number of correctly classified high, low and intermediate positive nuclei collectively. The remaining three columns represent the total number of correctly classified positive nuclei for each positive stain class.

Table 5-4

A

Images of set Ki67 Day 19	positive nuclei dataset	Correctly classified heterogeneously stained nuclei	No. of high stained nuclei	No. of intermediate stained nuclei	No. of low stained nuclei
1) Ki67 Day 19 high staining	43	41	18	14	9
2) Ki67 Day 19 high staining	53	50	25	14	11
3) Ki67 Day 19 high staining	36	34	18	9	7
4) Ki67 Day 19 intermediate staining	57	53	15	24	14
4) Ki67 Day 19 low staining	14	10	0	6	4

B

Images of set Ki67 Day 22	positive nuclei dataset	Correctly classified heterogeneously stained nuclei	No. of high stained nuclei	No. of intermediate stained nuclei	No. of low stained nuclei
1) p27 Day 19 high staining	56	52	21	22	9
2) p27 Day 19 intermediate staining	42	41	8	23	10
3) p27 Day 19 intermediate staining	43	38	6	19	13
4) p27 Day 19 low staining	17	14	2	8	4
5) p27 Day 19 low staining	17	15	0	6	9

C

Images of set Ki67 Day 22	positive nuclei dataset	Correctly classified heterogeneously stained nuclei	No. of high stained nuclei	No. ff intermediate stained nuclei	No. of low stained nuclei
1) Ki67 Day 22 high staining	49	46	22	13	11
2) Ki67 Day 22 intermediate staining	50	45	6	25	14
3) Ki67 Day 22 intermediate staining	66	61	12	29	20
4) Ki67 Day 22 intermediate staining	24	22	5	11	6
5) Ki67 Day 22 intermediate staining	31	29	6	16	7

Results of the automated intensity score for each IHC image used are presented in Table 5-5 (A-C). The first column represents the IS score provided by the pathologies where +++ denotes high positive staining, ++ denotes intermediate positive staining and + denotes low positive staining. The remaining three columns represents the detailed IS measurements, where +3 denotes the percentage of positive nuclei with high staining intensity for each IHC image, +2 denotes the percentage of positive nuclei

with intermediate staining intensity for each IHC image, and +1 denotes the percentage of positive nuclei with low staining intensity for each IHC image .

Table 5-5

A

	pathology	IS Automated		
Images of set Ki67 Day 19	**IS**	**% of (+3)**	**% of (+2)**	**% of (+1)**
1) Ki67 Day 19 high staining	+++	43.9	34.1	22
2) Ki67 Day 19 high staining	+++	50	28	22
3) Ki67 Day 19 high staining	+++	52.9	26.5	20.6
4) Ki67 Day 19 intermediate staining	++	28.3	45.3	26.4
5) Ki67 Day 19 low staining	+	0	60	40

B

	pathology	Automated		
Images of set p27 Day 19	**IS**	**% of (+1)**	**% of (+2)**	**% of (+3)**
1) p27 Day 19 high staining	+++	40.4	42.3	17.3
2) p27 Day 19 intermediate staining	++	19.5	56.1	24.4
3) p27 Day 19 intermediate staining	++	15.8	50	34.2
4) p27 Day 19 low staining	+	14.3	57.1	28.6
5) p27 Day 19 low staining	+	0	40	60

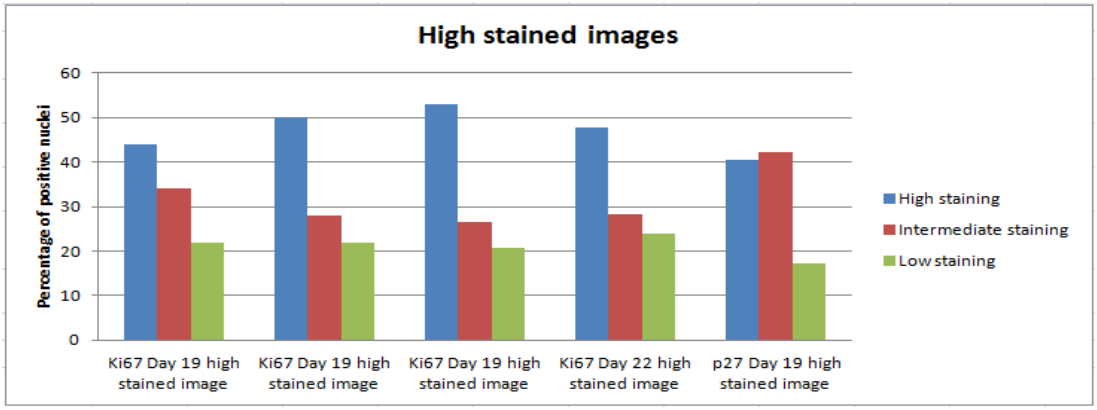
C

	pathology	Automated		
Images of set Ki67 Day 22	**IS**	**% of (+1)**	**% of (+2)**	**% of (+3)**
1) Ki67 Day 22 high staining	+++	47.8	28.2	24
2) Ki67 Day 22 intermediate staining	++	13.4	55.5	31.1
3) Ki67 Day 22 intermediate staining	++	19.7	47.5	32.8
4) Ki67 Day 22 intermediate staining	++	22.7	50	27.3
5) Ki67 Day 22 intermediate staining	++	20.7	55.2	24.1

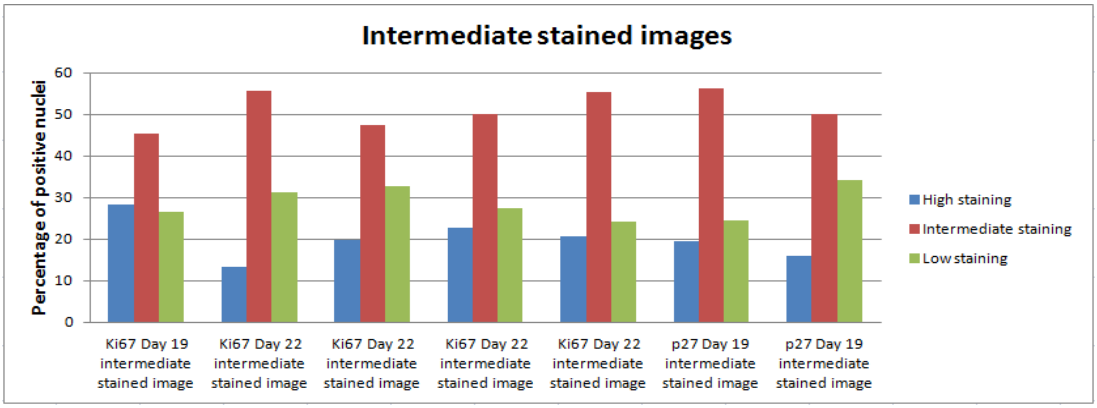
From Table: 5-5 (A-C), it is concluded that the percentage of each stain intensity could be indicative of the IS for each IHC image. For example, in the case of high stained images, the percentage of high stained nuclei in each image was the highest compared to the percentage of both intermediate and low stained nuclei present in the same image. The same theory could apply to the intermediate stained IHC images. However, in the case of low stained images the percentage of low stained nuclei did not always score the highest percentage when compared to intermediate and high stained nuclei. In the tested low stained IHC images, the percentage of high stained nuclei scored the lowest value, while intermediate stained nuclei scored the highest percentage in both cases of Ki67 and p27 day 19 images.

When analysing further images of the same type (low stained), it was found that the PS of such images was below 30%, which could have an influence on the way the intermediate stained nuclei are judged by the eye, leading the observer to recognise the intermediate stained nuclei as low stained amongst the higher number of negative nuclei in that image. Therefore, the automated approach proposed in this chapter for measuring the IS of the heterogeneously stained positive nuclei, could provide pathologists with higher level of information, by providing the number and percentage of each positive staining level for each IHC image.

A



B



C

Figure 5-20: (A-C) represent the percentage of positive nuclei with high, intermediate and low stained images respectively.

Chapter 6

Conclusion and Future Work

6.1 Summary of contributions and Conclusions

This thesis proposed a number of methods for the automated quantification of antibodies in immunohistochemistry stained image. These methods have been applied to three challenging problems in the automated quantification of such images; nuclei cell segmentation and counting, positive and negative nuclei classification and positive nuclei stain intensity classification.

The performance of these tasks is usually influenced by the techniques and approaches utilised to retrieve information. Despite many efforts during the past few years investigating the robustness of automated quantification methods, this task is still an open and highly challenging task. This is due to the variability of the image types and the unpredictable characteristics of nuclei cells.

This thesis investigated different approaches introduced by the author for quantifying stained nuclei cells based on a cell-level approach. The proposed approaches have addressed the individual analysis of each nuclei cell, to extract powerful discriminative features. These features are then used to

provide high accuracy performance in classifying the different stained nuclei cells. The contributions are evaluated on tumour multicell spheroid images for three different quantification measures, as follows;

A novel approach for the automated segmentation and quantification of nuclei objects in IHC images was presented. The automated method consists of a number of image reconstruction, thresholding, segmentation and edge detection algorithms that work consecutively to achieve the segmentation and quantification of positive and negative nuclei collectively. Experimental results were demonstrated on Ki67 and p27 ^{Kip1} antibodies in intact HT29 colonic cancer spheroid images. Images with different staining intensity levels from every dataset were used to assess the performance and generality of the automated method. Upon extensive experiments, it was found that the position of the nuclei plays a fundamental role in the automated method performance accuracy. Nuclei which are located closer to the viable rim in the Ki67 image dataset, achieved a higher accuracy when compared to the intermediate and low stained nuclei. This is due to the location of such cells which tend to be located closer to the necrotic core. Necrotic core contains hypoxic and necrosis cells, this type of cells can be detected by the automated method and falsely identified as viable cell. This ultimately reduces the accuracy and increase the error rate of the automated method performance.

The Watershed segmentation method imposed a number of limitations to the automated approach. Watershed segmentation is a parameter

free algorithm, and therefore adjustment to the segmentation level was a challenging task. To improve the performance of the watershed segmentation method, a number of morphological and colour map methods were suggested by the author, improved results were reported after the additional adjustments. However, a complete elimination of the under segmentation problem was not achieved. This has resulted in an increase in the error rate of the automated counting and segmentation method.

Despite the limitations imposed over the performance of the automated counting and segmentation method, the overall accuracy rate was calculated at 92.6% for all examined images of all image datasets. This provides that the automated proposed method achieved a high level of accuracy when compared to the benchmark data provided by the pathologist.

The second automated method provided by pathologists was the cell-level approach for classifying positive and negative segmented stained nuclei images. Colour, textural and morphological features are extracted from each individual positive and negative nuclei to observe the discriminative characteristics of each nuclei class. Features were then optimized to select only those features that provide high discriminative power between positive and negative nuclei images. Upon extensive experiments of feature extraction and selection, a feature vector of length 25 was generated for each nucleus image and demonstrated to provide the uppermost discriminative power in classifying the nuclei classes. The 25 proposed features consisted of 18 DWT features to describe the colour characteristics

of the stained nuclei objects, 4 GLCM features to describe the textural features, and 3 morphological features.

Upon extensive experiments, it was found that DWT coefficients of level one and two provided the highest level of classification accuracy using both SVM and ANN classifiers when compared to the rest of the proposed features presented in this thesis. This provides that, DWT can be used as a powerful tool not only to extract textural information from images but also to extract powerful colour discriminative features, which can then be used to differentiate between the diverse types of cells based on their stain colour information.

It was also found that the levels of wavelet compression have a direct impact on the quality of information retrieved from the wavelet coefficient. Wavelet works based on a compression mechanism which means that the colour information of the image are compressed further as the level of compression increases. This will result in the reduction of the colour information that provides the discriminative power. As a result; coefficient of further levels (beyond level one and two) provided reduced classification accuracy.

Two types of machine learning algorithms were investigated in this chapter; SVM and ANN, for the purpose of classifying positive and negative nuclei. 13 experiments of feature collections and their combinations were conducted on different image datasets (Ki67-day19, P27-day19 and Ki67-day22). From the experiment outcomes it was evident that the SVM classifier outperforms the ANN in all types of features and their combinations.

It was also evident that the proposed 25 discriminative features provided the highest level of classification accuracy, at 93% for all image datasets, using the SVM classifier. This again shows that the proposed classification automated method achieved a high level of accuracy when compared to the benchmark data provided by the pathologist.

Finally the output from Chapter 3 which provided the total number of nuclei objects in an image, was combined with the classification output of Chapter 4, which provided the number of positive nuclei. The output of both chapters was used to measure the PS of individual IHC images. Results demonstrated that the proposed automated PS achieved an overall accuracy of 90% for all tested IHC images when compared to the benchmark data of the experienced pathologist.

In Chapter 5, a novel method for quantifying and classifying heterogeneously stained positive nuclei was proposed. The cell-level automated method scores each positive nucleus based on the intensity level of the positive stain. High, intermediate and low stained nuclei are classified using a number of discriminative features. Histogram intensity textural features are extracted from heterogeneously stained positive nuclei, to provide discriminative features for the three positive staining intensity levels. To enhance the discriminative power of the baseline histogram intensity textural features, image thresholding, histogram equalization and *K*-means clustering techniques are applied. From extensive experiments, it was found that the image thresholding technique provided weak discriminative power when used on segmented positive nuclei images. Histogram equalization and *K*-

means clustering provided improved discriminative power. An improved form of histogram equalization was proposed and provided higher discriminative power when followed by the use of the *K*-means clustering method. This provides that clustering algorithms such as *K*-means and equalized histogram features improves the discriminative power of the baseline intensity textural features, and can be used prior to the extraction of textural features to enhance the discriminative power of such features.

Uppermost DWT, GLCM and morphological features were integrated with the proposed intensity textural features to form 29 discriminative features. The proposed features achieve an overall classification accuracy of 90% using the SVM classifier. The accuracy was based on the testing of images from all datasets (Ki67-day19, P27-day19 and Ki67-day22).

From the above, it is concluded that the cell-level automated approach provided high level of accuracy on all tested data, and achieved accuracies in the range of 90% to 93% when compared to the benchmark pathology manual data. This means; the automated method can potentially be used as a second reader by pathologists in the clinical practice or by researchers in the medical imaging field, who seek to use automated time efficient methods for quantifying large image datasets,

Although the proposed method is implemented based on an automated approach which does not involve the manual input of a user, the addition of flexible parameters can broaden the use of the proposed method to include the analysis of a variety of image datasets and stain types. Such parameters

can be included to adjust the, threshold, colour balance and morphological characteristics to suite a variety of image datasets. The proposed method can also potentially work on various image magnification levels. However, the user would have to apply some adjustments to the parameters that concern the size of the nuclei objects.

6.2 Future work

In this section, a number of suggestions for the future work are given for the continuation of the work presented in this thesis.

- In this thesis, a number of IHC image data sets were used to demonstrate the validity of an automated system in scoring cancer markers. It would be interesting to apply the proposed automated method to a wider range of IHC images addressing different cancer types and antibody stains.

- The automated feature extraction approach presented in Chapters 4 and 5 provided a number of unique features that described the colour, textural and morphological characteristics of stained nuclei objects. It would be interesting to interpret these features in describing further biological genetic characteristics and behaviour change dynamics.

- It would be interesting to analyse other cellular particles from tumour microenvironment images such as membranes and other cell types. A

dataset can be generated to include different classes, each representing a cell type or a particular cellular object. Cellular object classes can be fed into a classifier to determine the class an object belongs to (membrane, positive nuclei etc.). Classification results could then be used to quantify different aspects of the microscopic tumour images.

- The automated quantitative descriptions provided in this thesis can be used to track and monitor the progression of cancer throughout its various stages. This could include drug delivery optimization tasks and determining the output of a particular cancer drug. It could also be used for assessment in tasks such as quantifying image patterns, extracting phenotypic and molecular features and discovering new targeted cancer drugs.

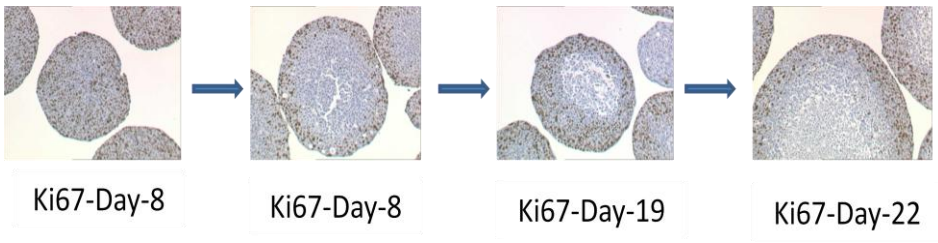


Figure 6-1: progression of cancer throughout its various stages.

- A tumour's response to treatment can be monitored via a number of approaches; including the cell level approach, as represented in this thesis, or a tissue level approach. A tissue level approach can be integrated with cell-level analysis to provide generic characteristics concerning the tumour tissue pattern.

- The automated quantitative approach could be extended to provide an optimization approach for treatment regimes to suit the genetic structure of the individual patient and type of tumour. This could be achieved by Integrating imaging modalities with genomics, which could ultimately aid in the understanding of cancer genome heterogeneity and predict individualised clinical outcomes of cancer patients.

- It will also be interesting to integrate the automated functionalities of the proposed method to be used directly with the microscope. As known until today, most of the microscopes have limited functionalities in the way they acquire images and display them to the user using complex yet limited functionality software. The integration of more custom designed software such as the automated proposed method can improve the quality and the level of information acquired by the microscope. This would give the used the ability to control the level of information retrieved from the microscope and provide a clearer output image that best suit the task required. This could be achieved by adjusting some of the acquisition hardware of the microscope or simply by integrating the software to work alongside the microscope software.

References

Aarnink, R. G., R. J. B. Giesen, A. L. Huynen, J. J. M. C. H. De la Rosette, F. M. J. Debruyne and H. Wijkstra;, “A practical clinical method for contour determination in ultrasonographic prostate images”, Ultrasound Med Biol.,; 20(8):705-17, 1994.

Abramoff, M.D., Magelhaes, P.J. & Ram, S.J., “Image Processing with ImageJ”, Biophotonics International 11(7), 36-42, 2004.

Ali, M.; Clausi, D.; , "Using the Canny edge detector for feature extraction and enhancement of remote sensing images," *Geoscience and Remote Sensing Symposium, IGARSS '01, IEEE International* , vol.5, no., pp.2298-2300 vol.5, 2001.

Allred, D., J. Harvey, M. Berardo, G. Clark, “Prognosticand predictive factors in breast cancer by immunohistochemicalanalysis”, Mod Pathol 11:155–168, 1998.

Anderson E, Clarke RB, Howell A. Estrogen responsiveness and control of normal human breast proliferation. J Mammary Gland Biol Neoplasia 1998; 3: 23-35.

Apatean Anca (discant), Alexandrina Rogozan, Simina Emerich, Abdelaziz Bensrhair, Wavelets as features for objects recognition, Journal of acta technica napocensis, 2007.

Al-Timemy, A. H., F. M. Al-Naima and S. Mahdi, “Data Acquisition System for Myocardial Infarction Classification based on Wavelets and Neural Networks,” *Proc. of the Fifth International Multi-Conference on Systems, Signals and Devices (IEEE SSD'08),* Amman, Jordan, 20-23 July 2008.

Al-Timemy, A.H.; Al-Naima, F.M.; Qaeeb, N.H., "Probabilistic Neural Network for Breast Biopsy Classification," *Developments in eSystems Engineering (DESE), 2009 Second International Conference on* , vol., no., pp.101-106, 14-16 Dec 2009.

Amara Graps, “An Introduction to wavelets”, IEEE Computational Science and Engineering, vol. 2 no. 2, 1995.

Belkacem-Boussaid, K.; Sertel, O.; Lozanski, G.; Shana'aah, A.; Gurcan, M.; , "Extraction of color features in the spectral domain to recognize centroblasts in histopathology," *Annual International Conference of the IEEE* , vol., no., pp.3685-3688, 3-6 Sept 2009.

Baker, A., P. Kornguth, J. Lo, et al., “ Artificial neural network: improving the quality of breast biopsy recommendations. Radiology, vol. 198, pp. 131–135, 1996.

Battari Assou el, Jose Luis, Jean-Michel Martin, Jacques Fantini, Jean-Marc Muller,Jacques Marvaldi and Jacques Pichon, The vasoactive intestinal peptide receptor on intact human colonic adenocarcinoma cells (HT29-D4)Evidence for its glycoprotein nature, Biochemestry Journal, 242, 185-19, 1987.

Beuchar, S., “The watershed transformation applied to image segmentation”, presented at 10th Pfefferkorn Conf. on Signal and Image Processing in Microscopy and Microanalysis, 1992.

Beuchar, S., “Watershed, hierarchical segmentation and waterfall algorithm”, In J. Sera and P. Soille, editors, Mathematical Morphology and its applications on image and signal processing. 69-76. Kluwer Academic Press, 1994.

Bengtsson E. “Fifty years of attempts to automate screening forcervical cancer”, Medical Imaging Technology, 17:203-210, 1999.

Boland, M. and R. F. Murphy, “A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of hela cells,” *Bioinformatics*, vol. 17, no. 12, pp. 1213–1223, 2001.

Bamford P. and Lovell B., “Method for accurate unsupervised cell nucleus segmentation”, Proc of IEEE Engineering in Medicine and Biology, Vol.3, , pp.2704-2708, 2001.

Buciu Joan and Gacsadi A., "Gabor Wavelet Based Features for Medical Image Analysis and Classification", 2009.

Chaddad Ahmad, Camel Tanougast, Abbas Dandache, Ahmed Bouridane, Extraction of Haralick Features from Segmented Texture Multispectral Bio-Images for Detection of Colon Cancer Cells, First International Conference on Informatics and Computational Intelligence, 2011.

Choudury, B. Kingshuk Roy, Kevin J. Yagle, Paul E. Swanson, Kenneth A. Krohn, and G. Rajendran, “A robust Automated measure of average antibody staining in immunohistochemistry images”, vol. 58(2): pp. 95-107, 2010.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C. Kang, I.H.,Friman, “CellProfiler: image analysis software for identifying and quantifying cell phenotypes”, Genome Biology 7:R100, 2006.

Cheng Jierong; Rajapakse, J.C.; , "Segmentation of Clustered Nuclei With Shape Markers and Marking Function," *Biomedical Engineering, IEEE Transactions on* vol.56, no.3, pp.741-748, March 2009.

Cregger, M., A. J. Berger, and D. L. Rimm, “Immunohistochemistry and quantitative analysis of protein expression,” Arch. Pathol. Lab. Med., vol. 130, pp. 1026–1030, Jul 2006.

Ciampa, B. Xu, G. Ayata, D. Baiyee, J. Wallace, M. Wertheimer, K. Edmiston, and A. Khan, “HER-2 status in breast cancer, correlation of gene amplification by fish with immunohistochemistry expression using advanced cellular imaging system,” Appl. Immunohistochemistry Molecular Morphol., vol. 14, pp. 132–137, Jun 2006.

Cloppet, F. and Boucher, A., “Segmentation of complex nucleus configurations in biological images”, Pattern Recognition Letters 31 755-761, 2011.

Chan, T. & Vese, L., “Active contours without edges”, IEEE Trans. on Image Processing, 2, 266-277, 2001.

Chester M., “*Neural Networks: A Tutorial”*, Englewood Cliffs, NJ: Prentice Hall, ch.2, 1993.

Coelho, L.P.; Shariff, A.; Murphy, R.F.; , "Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms," *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on* , vol., no., pp.518-521, June 28 2009-July 1 2009.

Comer, Mary L. and Delp, Edward J., Morphological operations for color image processing, Journal of Electronic Imaging, vol. 8, no.3, pp. 279-289, 1999.

Cristianini, N., Shawe-Tayor, J.; “An Introduction to Support Vector Machines”, The Press Syndicate of The University of Cambridge, Cambridge, 2000.

Daubechies, I.;, “The wavelet transform time-frequency localization and signal analysis", *IEEE Transactions on Information Theory,* 36, 961-1005, 1990.

Di Cataldo, S., E. Ficarra, A. Acquaviva, E. Macii, “Achieving the way for automated segmentation of nuclei in cancer tissue images through morphology-based approach: A quantitative evaluation”, Computerized Medical Imaging and Graphics 34-453–461, 2010.

Droogenbroeck, M., and Buckley, M.J., Morphological Erosions and Openings: Fast Algorithms Based on Anchors,Journal of Mathematical Imaging and Vision, vol 22, issue 2-3, 2005.

Doyle, S., M. Hwang, K. Shah, A. Madabhushi, and et al., “Automated grading of prostate cancer using architectural and textural image features,” in *In Biomedical Imaging: From Nano to Macro. ISBI 2008. 5th IEEE International Symposium*, 2007.

Doyle, S., S. Agner, A. Madabhushi, M. Feldman, and et al., “Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features,” *Biomedical Imaging: From Nano to Macro. 5th IEEE International Symposium*, 2008.

Dorini, L.B.; Minetto, R.; Leite, N.J.; , "White blood cell segmentation using morphological operators and scale-space analysis," *Computer Graphics and Image Processing, 2007. SIBGRAPI 2007. XX Brazilian Symposium on* , vol., no.,pp.294-304,7-10Oct.2007

El-Naqa, I., Yang, Y.Y., Wernick, M.N., Galatsanos, N.P., Nishikawa, R.: “Support vector machine learning for detection of microcalcifications in mammograms”, In IEEE Transactions on Medical Imaging, vol. 21, pp. 1552–1563, 2002.

Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S., “Stochastic gene expression in a single cell”, Science 297, 1183-1186, 2002.

Fang, B., W. Hsu, and M. L. Lee, “On the accurate counting of tumor cells," IEEE Trans. Nanobio ci., vol. 2, no. 2, pp. 94-103,2003.

Fawcett, T., An introduction to ROC analysis. Pattern Recognition Letters, 27(8): pp. 861-874, 2006.

Ficarra, E., and Macii, E., Computer-aided evaluation of protein expression in pathological tissue images,Proc IEEE Int Symp on Computer-Based Med Systems, pp.413-418, 2006.

Fuhua Chen; Jun Xie; Hong Zhang; Deshen Xia; , "A technique based on wavelet and morphology transform to recognize the cancer cell in pleural effusion," *Medical Imaging and Augmented Reality Proceedings. International Workshop on* , vol., no., pp.199-203, 2001.

Goutsias, J., Heijmans, H.J.A.M. and Sivakumar, K.. "Morphological Operators for Image Sequences". Computer Vision and Image Understanding 62. 3 :326-346, 1995.

Gavrielides, M.A.; Masmoudi, H.; Petrick, N.; Myers, K.J.; Hewitt, S.M.; , "Automated evaluation of HER-2/neu immunohistochemical expression in breast cancer using digital microscopy," Biomedical Imaging: From Nano to Macro, 5th IEEE International Symposium , pp.808-811, 14-17, May 2008.

Glotsos, D, Spyridonos P, Cavouras D, Ravazoula P, Dadioti PA, Nikiforidis G., “Automated segmentation of routinely hematoxylin-eosin-stained microscopic images by combining support vector machine clustering and active contour models”, Anal Quant Cytol Histol. 26(6):331-40, 2004.

Golub, Gene; Charles F. Van Loan, Matrix Computations – Third Edition. Baltimore: The Johns Hopkins University Press. pp. 53. ISBN 0-8018-5413-X, 1996.

Gupta, R.; Mittal, A.; Singh, K.; Sharma, P.; Sharma, A.; , "Identification of Nuclear Receptors and Its Family Using Wavelet Variance Features," *Information Technology (ICIT), 10th International Conference on* , vol., no., pp.68-73, 17-20 Dec 2007.

Glotsos, D., Spyridonos, P., Petalas, P., Cavouras, D., Ravazoula, P., Dadioti, P., Lekka, I., Nikiforidis G.;, “Computer-based malignancy grading of astrocytomas employing a support vector

machine classifier, the who grading system and the regular hematoxylineosin diagnostic staining procedure", Anal. Quant. Cytol. Histol., 26, 2004.

Giuliodori, M. A., Statistical classification of images, PhD theses, Universidad Carlos III de Madrid, 118, 2011.

Hafizah Wan Mahani, Eko Supriyanto, Jasmy Yunus Feature Extraction of Kidney Ultrasound Images based on Intensity Histogram and Gray Level Co-occurrence Matrix, Sixth Asia Modelling Symposium, 2012.

Han, J.W., Lane, P.C.R., Davey, N., Sun, Y.;, "Comparing the performance of single-layer and two-layer support vector machines on face detection", In Proceedings of The Seventh UK Workshop on Computational Intelligence, 2007.

Hatanaka A.Yutaka, K. Hashimzume, K. Nitta, T. Kato, I. Itoh and Y. Tani, "Cytometrical image analysis for immunohistochemical hormone receptor status inbreast carcinomas", vol. 53 pp. 693-699, 2003.

Haralick Robert M., Shanmugam K. and Dinstein Itshak, Textural feature for image classification, IEEE transaction on systems man and cybernetics, vol 3 no. 6 pp 610-621, November, 1973.

Herold, J., Friedenberger,M., Bode,M., Rajpoot,N., Schubert,W., Nattkemper, T.W.;, "Flexible synapse detection in fluorescence micrographs by modeling human expert grading", In Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1347–1350, 2008.

Hearst, M.A.;, "Support vector machines", IEEE Intell. Syst., 13, 18–28,1998.

Hiremath, P.S and Iranna Y., H.; , "Automated Cell Nuclei Segmentation and Classification of Squamous Cell Carcinoma from Microscopic Images of Esophagus Tissue," *Advanced Computing and Communications, 2006. ADCOM 2006. International Conference on* , vol., no., pp.211-216, 20-23 Dec. 2006.

Hua Zhang; Wenzhong Shi; Kimfung Liu; , "Fuzzy-Topology-Integrated Support Vector Machine for Remotely Sensed Image Classification," *Geoscience and Remote Sensing, IEEE Transactions on* , vol.50, no.3, pp.850-862, March 2012.

Huang Der-Chen; Kun-Ding Hung; ,"Leukocyte nucleus segmentation and recognition in color blood-smear images," *Instrumentation and Measurement Technology Conference (I2MTC), 2012 IEEE International* , vol., no., pp.171-176, 13-16 May 2012.

H. Y. CHI, *Is tumor necrosis really death*, MSc dissertation, Department of Biomedical Science, University of Bradford, 2006, 71 pages.

Khojasteh, M.; Ward, R.; MacAulay, C., "Quantification of membrane IHC stains through multi-spectral imaging," *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on* , vol., no., pp.752-755, 2-5 May 2012

Kokare, M., Biswas, P. K. & Chatterji, B. N., Texture image retrieval using rotated wavelet filters. *Pattern Recognition Letters,* 28, 1240-1249, 2007.

Kumar, R.R.; Kumar, V.A.; Sharath Kumar, P.N.; Sudhamony, S.; Ravindrakumar, R.; , "Detection and removal of artifacts in cervical Cytology images using Support Vector Machine," *IT in Medicine and Education (ITME), International Symposium*, vol.1, no., pp.717-721, 9-11 Dec 2011.

Kong, H., Gurcan M, Belkacem-Boussaid K., “Partitioning Histopathological Images: An Integrated Framework for Supervised Color-Texture Segmentation and Cell Splitting”, IEEE Trans Med. Imaging, 30(1):1661-1677, 2011.

Karen E. A. LaRue, Mona Khalil and James P. Freyer, Microenvironmental Regulation of Proliferation in Multicellular Spheroids is Mediated through Differential Expression of Cyclin-Dependent Kinase Inhibitors Cancer Research;64:1621-163, 2004.

Kazmi, N., Hossain, M.A. and Phillips, R., “Cellular and subcellular automaton model for tumour growth dynamics using artificial neural network”, 6th International Conference on Practical Applications of Computational Biology and Bioinformatics PACBB, 2012.

Kass, M., Witkin, A. & Terzopoulos, “Snakes, active contour model”, International Journal of Computer Vision, 321-331,1988.

Kostopoulos, S.; Cavouras, D.; Daskalakis, A.; Bougioukos, P.; Georgiadis, P.; Kagadis, G.C.; Kalatzis, I.; Ravazoula, P.; Nikiforidis, G.; , "Colour-Texture based image analysis method for assessing the Hormone Receptors status in Breast tissue sections," *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* , vol., no., pp.4985-4988, 22-26 Aug. 2007.

Kanopoulos, N.; Vasanthavada, N.; Baker, R.L.; , "Design of an image edge detection filter using the Sobel operator ," *Solid-State Circuits, IEEE Journal* , vol.23, no.2, pp.358-367, Apr 1988.

Kothari, S.; Chaudry, Q.; Wang, M.D.; , "Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques," *Biomedical Imaging: From Nano to Macro, ISBI '09. IEEE International Symposium*, pp.795-798, June 28 2009-July 1 2009.

Kokare, M., Biswas, P. K. and Chatterji, B. N.;, “Texture image retrieval using rotated wavelet filters”, *Pattern Recognition Letters,* 28, 1240-1249, 2007.

Krutsch Robert and Tenorio David Histogram Equalization application note for Microcontroller Solutions Group Guadalajara, june 2011.

Kuznetsov, D. Ślęzak, DH. Hepting,BG. Mirkin, Rough Sets, Fuzzy Sets, Data Mining and Granular, RSFDGrC, 1st Edition., XIII, 370 p, 2011.

Kuo, Yung-Lung, Chien-Chuan Ko, Jen-Yuan Lai; , "Automated assessment in HER-2/neu immunohistochemical expression of breast cancer," Computer Communication Control and Automation (3CA), 2010 International Symposium on , vol.2, no., pp.585-588, 5-7, May 2010.

Kwak Jin Tae, Hewitt Stephen M, Sinha Saurabh and Bhargava Rohit, Multimodal microscopy for automated histologic analysis of prostate cancer, BMC cancer, 2011.

Lee, K. young-Mi and W. Nick Street, “ An adaptive resource- allocating network for automated detection, segmentation and classification of breast cancer nuclei”, IEEE Transactions on Neural Networks, vol. 14, no. 3, 2003.

Liu Song, Mundra, P.A.; Rajapakse, J.C., "Features for cells and nuclei classification," *Engineering in Medicine and Biology Society,EMBC, Annual International Conference of the IEEE* , vol., no., pp.6601-6604, Aug. 30 2011-Sept 3 2011.

Louverdis, G., M.I. Vardavoulia, I. Andreadis, Ph. Tsalides, “A new approach to morphological color image processing”, Pattern Recognition, Volume 35, Issue 8, Pages 1733-1741, ISSN 0031-3203, 10.1016/S0031-3203(01)00166-2, , August 2002.
(http://www.sciencedirect.com/science/article/pii/S0031320301001662)

Li, G., Liu, T., Nie, J., Guo, L., Chen,J., Zhu,J., Xia,W., Mara, A., Holley, S. & Wong, S.T.C., “Segmentation of touching cell nuclei using gradient flow tracking”, Journal of Microscopy 231, 47-58, 2008.

Lukac, R.; Plataniotis, K.N.; Smolka, B.; Venetsanopoulos, A.N.; , "Color image filtering and enhancement based genetic algorithms," *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on* , vol.3, no., pp. III- 913-16 Vol.3, 23-26 May 2004.

Malpica, C. Solo´rzano, J. Jose´ Vaquero, “Applying watershed algorithms to the segmentation of clustered nuclei”, Cytometry, Vol.28, pp.289-297, 1997.

Maurya Rohit, Dr. Shalini Singh, Dr. P.R Gupta, Manish Kumar Sharma Road extraction using K-means clustering and morphological operations International journal of advanced engineering sciences and technologies Vol No. 5, Issue No. 2, 290 – 295, 2011.

Materka. A, M. Strzelecki, *Texture Analysis Methods – A Review*, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels 1998.

Meyer, F., “Topographical distance and watershed lines Signal Processing”, 38(1), 113-125, 2004.

Mehmet Sezgin and Bulent Sankur Survey over image thresholding techniques and quantitative performance evaluation, Journal of Electronic Imaging 13(1), 146–165, January 2004.

Masmoudi, H., Hewitt, S.M., Petrick, N., Myers, K.J., Gavrielides, M.A.; , "Automated Quantitative Assessment of HER-2/neu Immunohistochemical Expression in Breast Cancer," Medical Imaging, IEEE Transactions on , vol.28, no.6, pp.916-925, June 2009.

Min Hu, Xijian Ping, Yihong Ding , "Automated cell nucleus segmentation using improved snake," *Image Processing, 2004. ICIP '04. 2004 International Conference on* , vol.4, no., pp. 2737- 2740 Vol. 4, 24-27 Oct. 2004.

Minnen David , Westeyn Tracy , Starner Thad, Performance metrics and evaluation issues for continuous activity recognition,” Performance Metrics for Intelligent Systems, 2006.

Megason, S.G. and Fraser, S. E., “Imaging in systems biology”, Cell 130, 784-795, 2007.

Mjolsness, E., DeDoste, D.; “Machine learning for science: state ofthe art and future prospects”, Science 293, 2051–2055, 2001.

Ma Ji-quan; , "Content-Based Image Retrieval with HSV Color Space and Texture Features," *Web Information Systems and Mining (WISM), International Conference*, vol., no., pp.61-63, 7-8 Nov. 2009.

Mabrouk, S., J. Malek and R. Tourki, “Cytological image wavelettexture- based feature extraction”, Proc. of 3rd International Conference on Systems, Signals & Devises, pp. 341-346, 2005.

Meyer,F., S. Beucher, Morphological segmentation, Journal of Visual Communication and Image Representation, Volume 1, Issue 1, September 1990, Pages 21-46, ISSN 1047-3203, 10.1016/1047-3203(90)90014-M.

Ma, Y. and B. Manjunath, “A comparison of wavelet transform features for texture image annotation”, ICIP, 1995.

Niwas, S.I.; Palanisamy, P.; Sujathan, K.; , "Complex wavelet based texture features of cancer cytology images," International Conference on *Industrial and Information Systems (ICIIS),* vol., no., pp.348-353, July 29 2010-Aug. 1 2010.

Nattkemper, T.W., Twellmann, T., Schubert,W., Ritter, H.: “Human vs. machine: evaluation of fluorescence micrographs”, Comput. Biol. Med. 33 2003.

Nattkemper, T.W., Ritter, H.J., Schubert, W., “A neural classifier enabling high-throughput topological analysis of lymphocytes in tissue sections”, IEEE Trans. Inf. Technol. Biomed. 5, 138–149, 2001.

Nithya.R, Santhi.B, Mammogram classification using maximum difference feature selection method, Journal of Theoretical and Applied Information Technology, 2011.

Nandy, K., Gudla, P. R., Amundsen, R., Meaburn, K. J., Misteli, T. and Lockett, S. J., Automatic segmentation and supervised learning-based selection of nuclei in cancer tissue images, Cytometry, 81A: 743–754, (2012) doi: 10.1002/cyto.a.22097

Najman, L. & Malik, J., “Geodesic Saliency of Watershed Contours and Hierarchical Segmentation”, IEEE Trans. On Pattern Analysis and Machine Intelligence 18(12), 1163-1173, 1996.

Otsu, N., “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, January 1979.

Plissiti M. E., Christophoros Nikou and Antonia Charchanti, “Combining shape, texture and intensity features for cell nuclei extraction in pap smear images”, Pattern recognition letters, Vol.32, No. 6, pp. 838-853, 2011.

Plissiti, M.E., Nikou, C., Charchanti, A., "Automated Detection of Cell Nuclei in Pap Smear Images Using Morphological Reconstruction and Clustering," *Information Technology in Biomedicine, IEEE Transactions on* , vol.15, no.2, pp.233-241, March 2011.

Papik, K., B. Molnar, R. Schaefer, Z. Dombovari, “Application of neural networks in medicine- a review”, Medical Sciences, 1998.

Phansalkar, N.; More, S.; Sabale, A.; Joshi, M.; , "Adaptive local thresholding for detection of nuclei in diversity stained cytology images," *Communications and Signal Processing (ICCSP), 2011 International Conference on* , vol., no., pp.218-220,10-12Feb.2011

Phukpattaranont, P. and P. Boonyaphiphat, “Color based segmentation of nuclear stained breast cancer cell images”, Transactions on Electrical Eng., Electronics and Communications, vol.5, no.2, 2007.

Powers, David M W, Evaluation From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation, Journal of Machine Learning Technologies 2 (1): 37–63, 2007/2011. http://www.bioinfo.in/uploadfiles/13031311552_1_1_JMLT

Plaza A., Plaza J.: Parallel Morphological Classification of Hyperspectral Imagery Using Extended Opening and Closing by Reconstruction Operations, 58-61, 2008.

Pratt W. K., Digital Image Processing. New York: Wiley Interscience, 1991.

Qiongshui W., Zeng L. , Ke H. , Zheng H. , Xijian Gao , Diancheng Wang, “A multispectral imaging analysis system for early detection of cervical cancer”, Proc. SPIE 5745, Medical Imaging: Physics of Medical Imaging, 801,2005.

Raghavendra, P.S.; Chowdhury, S.R.; Kameswari, S.V.; , "Comparative study of neural networks and k-means classification in web usage mining," *Internet Technology and Secured Transactions (ICITST), 2010 International Conference for* , vol., no., pp.1-7, 8-11 Nov 2010.

Ridler, T. and S. Calvard, “Picture thresholding using an iterative selection method,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, no. 8, pp. 630–632, Aug. 1978.

Rehan Ali, Gooding, Mark, Szilágyi, Tünde, Vojnovic, Borivoj, Christlieb, Martin and Brady, Michael, "Automatic segmentation of adherent biological cell boundaries and nuclei from brightfield microscopy images", Machine Vision and Applications,Springer Berlin / Heidelberg (2012) Issn: 0932-8092, Vol. 3 issue.4

Rosenfield, N., Young, J.W.,Alon, U., Swain, P.S. & Elowitz, M.B., “Gene regulation at the single-cell level”, Science 307,1962-1965, 2005.

Rubner Yossi, Carlo Tomasi, and Leonidas J. Guibas, A Metric for Distributions with Applications to Image Databases, Proceedings of the IEEE International Conference on Computer Vision, Bombay, India, 1998.

Ruifrok, A.C., RL. Katz and DA. Johnston, Comparison of quantification of histochemical staining by Hue-Saturation- Intensity (HSI) transformation and color deconvolution. Appl Immunohistchem Mol Morphol, 2004, 11(1):85–91.

Seetha, M., I.V.Muralikrishna, B.L. Deekshatulu, B.L.Malleswari, Nagaratna, P.Hegde, “Artificial Neural Networks and other methods of Image Classification”, Journal of Theoretical and Applied Information Technology, vol 4, no. 11,2008.

Selvi, S.T.; Malmathanraj, R.;,"Segmentation and SVM Classification of Mammograms," *Industrial Technology, 2006. ICIT 2006. IEEE International Conference on* , vol., no., pp.905-910, 15-17 Dec 2006.

Sarvazyan, A.P.; Skovoroda, A.R.; Pyt'ev, Y.P.; , "Mechanical introscopy-a new modality of medical imaging for detection of breast and prostate cancer," *Computer-Based Medical Systems, 1995., Proceedings of the Eighth IEEE Symposium on* , vol., no., pp.4-5, 9-10 Jun 1995.

Scholzen Thomas and Gerdes Johannes, The Ki-67 Protein: From the Known and the Unknown Journal of cellular physiology, 182:311–322, 2000.

Serrano, N., Savakis, A. E. and Luo, J., Improved scene classification using efficient low-level features and semantic cues. *Pattern Recognition,* 37, 1773-1784, 2004.

Soille Pierre, Morphological Image Analysis: Principles and Applications, edition 2, Springer-Verlag, New York, USA, 2003.

Selesnick, I.W., R.G. Baraniuk and N. Kingsbury, “The dual-tree complex wavelet transform”, IEEE Signal Processing Magazine, 22 (6), pp. 123-151, 2005.

Sun, Y. and Ozawa, S.;, “Semantic-meaningful content-based image retrieval in wavelet domain”, Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval, 122-129, 2003.

Sutherland, R. M, B. Sordat, J. Bamat, Gabbert, H. Bourrat, and W. Muellerklieser, Oxygenation and Differentiation in Multicellular Spheroids of Human-Colon Carcinoma. Cancer Research, 46: 5320-5329, 1986.

Taylor John Robert, An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, University Science Books. pp. 128–129. ISBN 0-935702-75-X,1999. http://books.google.com/books?id=giFQcZub80oC&pg=PA128.

Tirtajaya, A.; Santika, D.D.;,"Classification of Microcalcification Using Dual-Tree Complex Wavelet Transform and Support Vector Machine," *Advances in Computing, Control and*

Telecommunication Technologies (ACT), Second International Conference on , vol., no., pp.164-166, 2-3 Dec 2010.

Tabor Jeffrey J., Howard M. Salis, Zachary Booth Simpson, Aaron A. Chevalier, Anselm Levskaya, Edward M. Marcotte, Christopher A. Voigt, Andrew D. Ellington", A Synthetic Genetic Edge Detection Program Cell - Vol. 137, Issue 7, pp. 1272-1281, 26 June 2009.

Tadashi K.; Junji U.; Shoichiro T.; , "Medical image diagnosis of lung cancer by hybrid multi-layered GMDH-type neural network using knowledge base," *Complex Medical Engineering (CME), 2012 ICME International Conference on* , vol., no., pp.663-668, 1-4 July 2012

Theodoros Mouroutis, Stephan J. Roberts and Anil, A. Bharath, "Robust cell nuclei segmentation using statistical modelling", Bio imaging, vol.6, No. 2, pp. 79-91,1998.

Vincent, L. & Soille, P., "Watersheds in digital spaces: an efficient algorithm based on immersion simulations", IEEE Trans. Pattern Anal. Mach. Intell. 13, 583-598, 1991.

Vapnik, V., Golowich, S. E., Smola, A., Support vector method for function approximation, regression estimation, and signal processing. Advances in Neural Information Processing Systems, NIPS, 281-287, 1996.

Wang, H., Zheng, C., Li, Y., Zhu, H., Yan, X.: "Application of support vector machines to classification of blood cells", Sheng Wu Yi Xue Gong Cheng Xue Za Zhi 20, 2003.

Wdward Dougherty, Mathematical morphology in image processing, Marcel Dekker, New York, 1993.

Wei-Liang Tai; Rouh-Mei Hu; Hsiao, H.C.W.; Rong-Ming Chen; Tsai, J.J.P.; , "Blood Cell Image Classification Based on Hierarchical SVM," *Multimedia (ISM),IEEE International Symposium on* , vol., no., pp.129-136, 5-7 Dec 2011.

Wei, N., You, J., Friehs, K., Flaschel, E., Nattkemper, T.W.: "An in situ probe for on-line monitoring of cell density and viability on the basis of dark field microscopy in conjunction with image processing and supervised machine learning". Biotechnol. Bioeng 97, 2007.

Wei, N., Flaschel, E., Friehs, K., Nattkemper, T.W.: "A machine vision system for automated non-invasive assessment of cell viability via dark field microscopy, wavelet feature selection and classification", BMC Bioinform. 9, 449, 2008.

Wang, M., Zhou, X., Li, F., Huckins, J., King,R.W. & Wong, S.T.C. , "Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy", Bioinformatics 24:1, 94-101, 2008.

Wang, Y.-P.; Ashok Kumar Dandpat; , "Classification of M-FISH images using fuzzy C-means clustering algorithm and normalization approaches," *Signals, Systems and Computers,Conference Record of the Thirty-Eighth Asilomar Conference on* , vol.1, no., pp. 41- 44 Vol.1, 7-10 Nov 2004.

Wang, Q., Neimi, J., Tan,C., You, L. & West, M., "Image Segmentation and dynamic lineage analysis in single-cell fluorescence microscopy", Journal of Synth. Biol, 2009.

Wang, S., H. Saboorian, E. P. Frenkel, B. B. Haley, M. T. Siddiqui, S. Gokaslan, F. H.Wians, L. Hynan, and R. Ashfaq, "Automated cellular imaging system (ACIS)-Assisted quantitation of immunohistochemical assay achieves high accuracy in comaprison with fluorescence in situ hybridization assay as the standard," Anatomic Pathol., vol. 116, pp. 495–503, 2001.

WinROOF®, [internet], N.D., available from: http://www.deltaunited.cn/t/Winroof%20t.htm [Accessed 10-10-2010].

Witten, I. H. & Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques*, second edition, Elsevier: San Francisco, ISBN 0-12-088407-0, 2005.

Wu, H.S., Barba, J. & Gill, J., “Iterative thresholding for segmentation of cell images”, J. Microsc. 97, 296-304, 2000.

Wu, K., Gauthier, D. & Levine, M.D., “Live cell image segmentation”, IEEE Trans. Biomed. Eng. 42, 1-12, 1995.

Zhao X. and S. H. Ong, Adaptive local thresh-olding with fuzzy-validity-guided spatial partitioning," in International Conference on Pattern Recognition, vol. 2, pp. 988-990., 1998.

Zhu Jun, Ning Chen, Eric P. Xing Infinite SVM: A Dirichlet Process Mixture of Large-margin Kernel Machines, Proceedings of the 28th International Conference on Machine Learning, Bellevue, WA, USA, 2011.

Appendix

A. Cell line and monolayer cell culture

HT29 cell line, which is a human adenocarcinoma of the colon cell line were maintained in Roswell Park Memorial Institute (RPMI) 1640 medium (Sigma) supplemented with 10% fetal bovine serum (FBS) (Costar), 1% of penicillin (10,000 IU), 1% of streptomycin (10,000 mg/ml), 1 mM glutamine (Sigma) and 2mM of sodium pyruvate (Sigma). Cells were kept at 37 ℃ in a 5% COS_2 humidified atmosphere.

For routine maintenance, cells were lifted using 5 ml 0.25% trypsin/ ethlyenediaminetetracetic acid (EDTA) solution for 5 mins at 37 ℃ incubation and resuspended in 10ml fresh medium. A 10µl sample of cell suspension was then placed into a haemcytometer chamber. Cell counts were taken from 5 grids of the haemocytometer chamber and calculated as the mean of 5 counts. Cell numbers were expressed as (mean cell count) x 10^4/ ml medium. 8×10^4 cells in T-75 flasks were re-seeded into fresh medium in T75 flasks.

B. Spheroids cell culture, diameter measurement

Three T-75 flasks were base-coated with 1% agarose in dH_2O to prevent adhesion of cells to the plastic. Following complete trypsinisation of a confluent HT29 monolayer, 5×10^5 cells were seeded into each flask. After

incubation at 37℃ for 3 days, aggregates were selected by sedimentation in a 20ml tube. Immature spheroids were transferred into 500 ml spinner flasks containing 250 ml of culture medium. The flasks were placed at 37 ℃ under stirring at 50 rpm. Culture medium was changed every 2–3 days. The diameters of 20 spheroids were measured by light microscope using an eyepiece graticule (calibrated using stage micrometer at the same magnification) every 2 days. Spheroids growth curves were performed in triplicate and the results were expressed as mean spheroids diameter ± standard deviation. Meanwhile the images of spheroids were taken under light microscopy.

C. Paraffin embedding of grown spheroids and sections

Spheroids were processed manually after being allowed to form a pellet by sedimentation in a universal tube. After being fixed in 10% phosphate buffered formalin (pH7.5) for 2 hours they were dehydrated by soaking in 70% ethanol for 1 hour, 90% ethanol for 30mins, 100% ethanol for 3 × 30 mins and cleared by 100% xylene for 3 × 30 mins. They were then washed in paraffin wax for 3 × 30 mins before being embedded in paraffin wax. After processing into wax blocks, the spheroid blocks were sectioned to 5µm thickness using Leica RM2155 microtome. Sections were transferred onto 3-aminopropyltriethoxysilane (APES) coated slides prior to immunohistochemical staining.

D. Haematoxylin and eosin staining of spheroid sections

Spheroid sections were de-waxed and rehydrated by incubating in 100% xylene for 20 mins, 50% xylene/ ethanol for 5 mins, absolute ethanol for 10 mins , 90% ethanol for 5mins, 70% ethanol for 5 mins and distilled water for 5 mins successively. The spheroids were then stained in Harris's haematoxylin solution (Sigma) for 10 mins. After being washed with tap water, they were soaked in alcohol/acid solution (0.5% HCl in 70% ethanol) for 5 seconds, washed in tap water again, "blued" in Scott' tap water for 2 mins and counterstained with 1% aqueous eosin for 1 min. Finally, the sections were dehydrated by soaking in 100% ethanol (2 × 2 mins), cleared in 50% xylene/ethanol and 100% xylene (2 × 5 mins) and coverslip mounted with DPX.

E. Immunohistochemistical staining of spheroid sections

Spheroid sections were de-waxed and rehydrated as describe above. Slides were washed in distilled water for 5 mins and treated with 1% H_2O_2 for 30 mins to remove endogenous peroxidase. After a 10 min phosphate buffered saline (PBS) wash, the slides were covered with antigen retrieval solution (0.01mol/L citrate buffer, pH 6), microwaved on high power for 25 mins and then left to cool for 30mins. Following another PBS wash, the sections on the slides were circled with a glue pen and then incubating with 1:200 horse blocking serum in PBS for 20 mins, they were then incubated in primary antibody overnight at 4℃. Ki67 protein was stained using a 1:200 dilution of mouse IgG antibody raised against human Ki67 protein (Santa Cruz

Biotechnology, Inc) in PBS. p27 ^{Kip1} protein was stained using a 1:100 dilution of mouse IgG antibody raised against human p27 ^{Kip1} (Dako, Denmark) in PBS. After a PBS wash, the sections were incubated with 1:200 dilution of secondary antibody (a biotinylated anti-mouse IgG (H+L) made in horse supplied by Vector, U.K) for 45 mins, washed with PBS again and incubated in avidin and biotinylated horseradish complex solution (Vectastain® ABC kits supplied by Vector, UK) for another 45 mins. Sections were then treated with 3,3'-diaminobenzidine (DAB substrate kit for proxidase Vector) for 5mins and counterstained with haematoxylin for 20 seconds before being dehydrated, cleared and mounted as described above. Each experiment was performed with a negative control (no primary antibody).