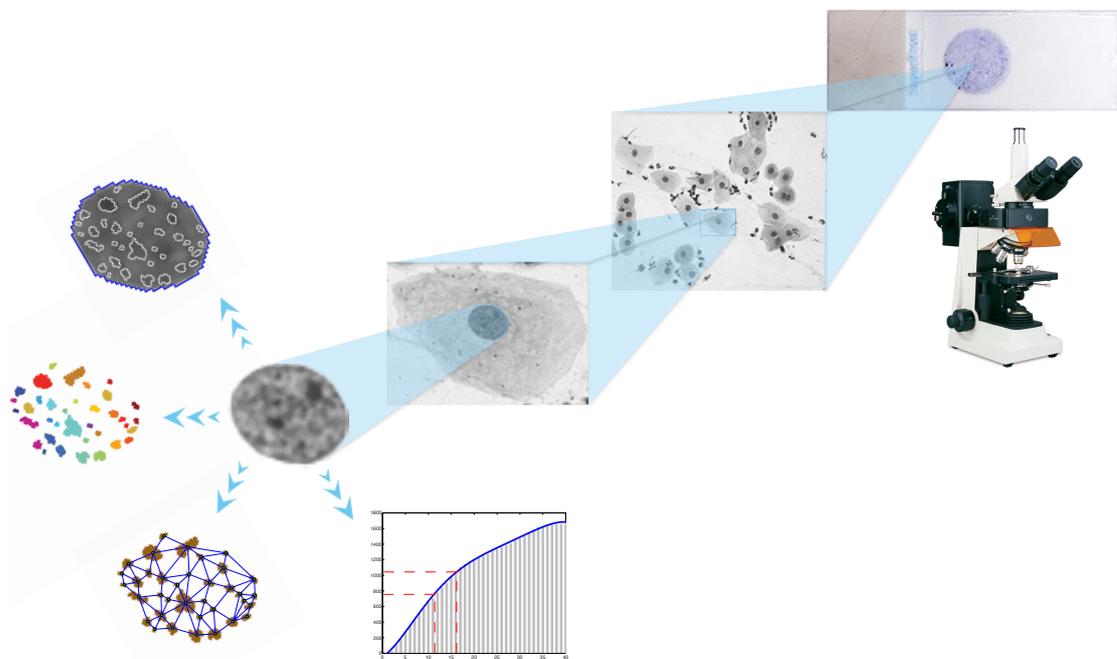


Chromatin pattern analysis of cell nuclei for improved cervical cancer screening

Master's Thesis in Biomedical Engineering



**Ramin
Moshavegh**

**Babak
Ehteshami Bejnordi**

Department of Signals & Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2013
Thesis number: EX030/2013

Chromatin pattern analysis of cell nuclei for improved cervical cancer screening

Master's Thesis in Biomedical Engineering

**Ramin
Moshavegh**

**Babak
Ehteshami Bejnordi**

Thesis supervisor and examiner:

**Assistant Professor
Andrew James Heinrich Mehnert**

Image analysis
Chalmers University of Technology
Department of Signals and Systems
SE-412 96 Gothenburg Sweden

Thesis number: EX030/2013
ISSN: 99-2747920-4

Abstract

The Papanicolaou (Pap) test or Pap smear is the primary screening test for cervical cancer. It involves the microscopic examination of cells sampled from the cervix. Two major factors affect the accuracy of the Pap test. The first is sampling error wherein no diagnostic cells make it on to the slide. The other is interpretation error for reasons including fatigue, inexperience, and habituation. Whilst computer-assisted interpretation can potentially address the issue of interpretation error it cannot address sampling error. However the malignancy-associated change (MAC) phenomenon may offer a solution. MACs are subtle sub-visual changes in the appearance of otherwise normal-looking cells from an abnormal Pap smear slide.

An essential first step in the development of an automated screener, based on MACs, is robust automatic segmentation of free-lying cell nuclei in digitized Pap smear images. This thesis presents and evaluates a fully automated algorithm for robustly detecting and segmenting free-lying cell nuclei in bright-field microscope images of Pap smears. The proposed novel segmentation algorithm makes use of grey-scale annular closings to identify free-lying nuclei-like objects together with marker-controlled watershed segmentation to accurately delineate the nuclear boundaries. The method was evaluated empirically using images digitised from Pap smears sourced from the Regional Cancer Centre in Thiruvananthapuram in India. The results show that the sensitivity and specificity of nucleus detection is 94.71% and 85.30% respectively, and that the accuracy of segmentation, measured using the Dice coefficient, of the detected nuclei is $97.30 \pm 1.3\%$.

This thesis also presents and evaluates a set of novel structural texture features for quantifying and classifying nuclear chromatin patterns in cells on a conventional Pap smear. The experimental results demonstrate the efficacy of the proposed structural approach and that a combination of the structural texture features and conventional features can be used to discriminate between normal and abnormal slides with high accuracy (0.954 ± 0.019 AUC \pm SE). They also demonstrate that it is possible to detect MACs in Papanicolaou stain (which is not stoichiometric). This in turn suggests the possibility of developing a fully automated Pap smear screener based on MACs.

Acknowledgements

First and foremost we would like to express our deepest appreciation to our supervisor Assistant Professor Andrew Mehnert, whose guidance, stimulating suggestions and encouragement, helped us to coordinate the thesis and write this report.

We also extend a big thank you to our co-supervisors Professor Ewert Bengtsson and Patrik Malm, CBA Uppsala, for their guidance and support and also for providing the data needed.

We express our gratitude to MedTech West and its board for providing an inspiring research environment.

Furthermore we would also like to acknowledge, with much appreciation, our fellow researchers at MedTech West, Yazdan Shirvany, Qaiser Mahmmod and Mohammad Alipoor for their friendship and support.

Finally, we thank our families, and especially our parents, for their unflagging love and support throughout our life; this thesis would have been impossible without them.

The Authors, Gothenburg, Sweden 2013/03/20

Contents

1	Introduction	1
1.1	The Pap smear test	2
1.1.1	Shortcomings of the Pap smear test	3
1.1.2	Automated screening	3
1.2	Malignancy associated changes (MACs)	4
1.3	Aim and objectives	4
1.4	Scope	5
1.5	Structure of the thesis	7
2	Literature review	9
2.1	Review of existing methods for segmenting cells and nuclei in Pap smear images	9
2.2	Review of existing features devised to quantitatively characterise chromatin distribution	12
2.2.1	Review of chromatin texture features	12
3	Novel algorithm for segmenting free-lying cell nuclei	15
3.1	High-level description of the proposed algorithm	15
3.2	Extraction of inner markers (step 1)	16
3.3	Marker-controlled watershed segmentation of the detected nucleus-like objects (steps 2-3)	17
3.4	Artefact Rejection (steps 4-5)	20
3.4.1	Size criterion	20
3.4.2	Shape criterion	21
3.4.3	Texture criterion	22
3.5	Empirical evaluation of the proposed segmentation algorithm	28
3.5.1	Image data	28
3.5.2	Method and Experiment	31
3.5.3	Results	31
3.5.4	Discussion	32

3.6	Summary	34
4	Texture features characterising chromatin distribution	35
4.1	Chromatin segmentation	35
4.2	Proposed structural texture features	37
4.2.1	Graph-based features	37
4.2.2	Margination features	39
4.2.3	Clustering features	39
4.2.4	Contextual chromatin particle features	41
4.2.5	Statistics of chromatin particle morphometric features	41
4.3	Summary	43
5	Evaluation of the proposed texture features for screening Pap smear slides	45
5.1	Pap smear images and ground truth	45
5.2	Feature selection methods considered in this study	46
5.2.1	The curse of dimensionality and the need for feature selection	46
5.2.2	MSVM-RFE	46
5.2.3	Guided Regularized Random Forest (GRRF)	47
5.2.4	L1-regularization procedure with generalised linear model	48
5.3	Experiment I: Determination of the most discriminatory subset of features for the classification of Pap smears	48
5.3.1	Cytology slides for experiment I	49
5.3.2	Feature extraction	49
5.3.3	Feature selection and classification	50
5.3.4	Results of experiment I	51
5.4	Experiment II: MAC-based classifiers for Pap smear screening	52
5.4.1	Cytology slides for experiment II	53
5.4.2	Classification performance evaluation of the best subset of features derived in experiment I	53
5.4.3	Results of experiment II	54
5.5	Summary	55
6	Summary and Conclusion	57
6.1	Thesis summary	57
6.2	Key contributions and findings	58
6.3	Conclusion	59
6.4	Limitations	60
6.5	Opportunities for further research	61
	Appendices	63

A	Fundamental operations from Mathematical Morphology	64
A.1	Mathematical morphology for binary images	64
A.2	Mathematical morphology for grey-scale images	65
B	Segmentation methods underlying methods in Chapters 2 and 3	66
B.1	Thresholding	66
B.2	Active contours and Deformable models	67
B.3	Seeded Region Growing	72
B.4	Watershed Transform	72
C	Cytology features	75
D	Candidate cell nuclei marker extraction methods	76
D.1	Morphological reconstruction	76
D.1.1	Grey-scale reconstruction:	76
D.1.2	Regional minima:	77
D.1.3	H-extrema:	78
D.1.4	Morphological Top-hat Transform	79
D.1.5	Improved Morphological Top-hat Transform	80
E	Shape Criteria	82
E.1	Danielsson's G shape factor	82
E.2	Dice similarity coefficient	83
E.3	Eccentricity	83
F	Classifiers used in the experiment II	85
F.1	SVM classifier with a linear kernel	85
F.1.1	Theory of linearly separable binary classification	85
F.2	Logistic regression	87
G	Publications	91
	Bibliography	109

List of Figures

1.1	Data acquisition and nucleus segmentation in an automated screening system	6
3.1	Extracting an inner marker for a free-lying cell nucleus.	17
3.2	Segmentation of the detected nuclei-like objects.	19
3.3	Cells of the squamous epithelium	21
3.4	Ellipse fitting to an arbitrary shape	22
3.5	Computing averages at different scales	23
3.6	Measure of Tamura coarseness for 6 different cell nuclei.	24
3.7	Artefacts rejected by the Tamura coarseness feature	24
3.8	Two sample FOVs from a Pap smear slide	25
3.9	GUI for Ground truth generation	29
3.10	GUI for Ground truth generation	30
3.11	DSC scores for the evaluation of the proposed automatic segmentation . .	32
3.12	Examples of false positive detection by the algorithm	33
3.13	missed nuclei due to presence of background noise and watershed failure .	34
3.14	Missed nuclei due to unsuitable choice of parameter values	34
4.1	Representation of a nucleus as a topographic relief (grey-level represents height)	36
4.2	Chromatin segmentation	37
4.3	Illustration of how the area-Voronoi diagram and generalised Delaunay graph are computed.	38
4.4	Graphs related to the Delaunay graph	38
4.5	Illustration of how the features $M1$ and $M2$ are computed.	39
4.6	Illustration of how the features $C1$ and $C2$ are computed.	40
4.7	Illustration of how the features $C3$ and $C4$ are computed.	41
5.1	Pie chart showing the contribution of the four main object feature categories in this study.	50

5.2	Illustration of 100 times iterated 5-fold cross validation feature selection	51
5.3	Boxplots of the features F1, F2, F3, F4, F5, F6 for experiment 1.	52
5.4	An example of a digitalized Pap smear FOV	54
5.5	Boxplots of the features $F1, F2, F5, F6$ for experiment II.	55
6.1	Focus stacking for extended depth of field in optical microscopy.	60
B.1	Global Thresholding	67
B.2	Parametric representation of a contour	69
B.3	Active contour segmentation	70
B.4	Cell nuclei segmentation by GVF snake.	71
B.5	Simulating the flooding of a topographic surface	73
B.6	Computing the magnitude of gradient image: Watershed transform considers the gradient image as a topographic surface.	74
D.1	Illustration of regional minimum at level 2 in an image	78
D.2	H-maxima transformation on a 1-D signal and extracting the regional maxima by computing $f - R_f^\delta(f - h)$	78
D.3	H-minima transformation on a 1-D signal and extracting the regional minima by computing $R_f^\varepsilon(f + h) - f$	79
D.4	Morphological Top-hat Transform on an image	80
E.1	G Shape-factor calculation	82
E.2	Parametric description of an ellipse	83
F.1	SVM for two-class data classification by a hyperplane	86
F.2	Plot of a logistic function	88
F.3	Two-class data classification by a linear logistic regression decision boundary	89

List of Tables

2.1	Comparison between cell and cell nucleus segmentation approaches found in the literature	11
3.1	Proposed algorithm to detect and segment free-lying intermediate cell nuclei	16
3.2	Shape and granularity features for 6 different cell nuclei.	26
3.3	Shape and granularity features for 6 non-nuclei artefacts	27
4.1	Summary of the extracted nucleus features	42
5.1	Frequency table for top 5 ranking features in experiment I.	51
5.2	Classification results for experiment I.	52
5.3	Classification results for experiment II.	55

1

Introduction

CERVICAL CANCER is the third most diagnosed cancer and the fourth leading cause of cancer death among women worldwide accounting for 9% of all malignancies among females in 2008 [1]. More than 85% of the cervical cancer cases occur in developing countries where public health infrastructure does not support Papanicolaou testing [1]. In countries with developed healthcare systems, widespread cervical screening programmes, aimed at detecting precancerous changes that can then be treated to prevent invasive cancer, have significantly reduced the number of deaths from the disease. The Papanicolaou (Pap) smear, is the primary screening test for cervical cancer. It has been largely responsible for diagnosing cancerous and precancerous lesions in many developed countries [2]. There is compelling evidence showing the efficacy of organised screening programmes based on performing the Pap smear test every 3-5 years. In some of the Nordic countries the incidence of invasive cervical cancer has declined by 80% since the introduction of organised screening programmes [3]. In Sweden the incidence of cervical cancer has decreased by 65% during the last 40 years but the incidence of invasive cervical cancer and cervical cancer mortality figures have been rather stable over the last decade [4].

Fortunately, invasive cervical cancer takes years to develop from slowly progressing precancerous lesions, thus enabling early detection through screening and diagnostic tests. Currently screening by conventional cervical cytology (Pap smear test) remains the principal global strategy to prevent invasive cervical cancer. However, the Pap smear test has several shortcomings including: subjective nature (dependent on individual interpretation), low sensitivity (i.e. ability to detect abnormal changes) and the need for frequent retesting [5].

The remainder of the present chapter is organised as follows: Section 1.1 describes the Pap smear test and discusses its shortcomings, and then presents the rationale for automated screening of Pap smear slides and describes the problems inherent in current automated screening systems. Section 1.2 then discusses a phenomenon known as ma-

lignancy associated changes (MACs), which can offer a solution to the automated Pap smear screening problem. The aim and objectives of the thesis are presented in section 1.3 and section 1.4 outlines the scope of the research. Finally, section 1.5 outlines the structure of this thesis.

1.1 The Pap smear test

The Papanicolaou (Pap smear) test is the most cost-effective cancer prevention and detection program ever invented [6]. It was devised by George N. Papanicolaou. In 1928 he first presented his findings that malignant cells from the cervix can be detected in a small sample of cells collected from the cervix (vaginal smears) [7].

Over a decade passed before collaboration between Dr. Papanicolaou and Dr. Herbert Traut, a gynaecologist and pathologist, provided scientific evidence of the potential of vaginal smears for the identification of cervical cancer and precancerous changes. Traut provided Papanicolaou with a large number of clinical samples from female patients at Cornell's Hospital. Papanicolaou published a detailed description of pre-invasive cervical lesions in his treatise, "Diagnosis of uterine cancer by the vaginal smear" [8] and in a major paper [9].

The conventional Pap test is a simple procedure comprising the following steps [6]:

- A speculum is inserted into the vagina to widen the opening so that the cervix can be viewed;
- Cells are sampled from inside and around the cervix using a swab, brush, or spatula;
- Cells are pressed on a glass slide and a fixative (preservative) is applied to preserve the sample;
- The samples are stained to ameliorate the contrast in the specimen and highlight the structural patterns to be analysed with a light microscope in a cytology laboratory [10, 11].

Liquid-based cytology (LBC) is a new method for cellular sample preparation for cytological tests. The main difference between LBC and the conventional Pap smear is related to the underlying preparation technique [12]. The sample is collected in a similar way to the Pap smear, but rather than smearing the cells onto a glass slide, the cellular material is immediately rinsed into a preservative liquid solution. The sample is then sent to the laboratory where special filtering techniques are used to remove non-diagnostic materials such as mucus, pus and blood cells. In the next step, a thin layer of cells is deposited onto a slide. Finally, the slide is examined under the microscope by a cytologist in the same way as in the conventional smear test [12].

1.1.1 Shortcomings of the Pap smear test

It usually takes several days or weeks to prepare the final results of a Pap smear test. Doctors and nurses sample the cervix and send the specimen to a pathology laboratory for visual evaluation under a microscope [13]. The microscopic examination itself is laborious and time-consuming involving the review of possibly hundreds of thousands of cells for signs of cancer or precancer. It is not surprising therefore that 1 in every 10 to 20 positive cases is missed in routine screening [5].

Two major factors affect the accuracy of the Pap smear test. The first is sampling error wherein no diagnostic cells make it onto the slide. This occurs when health care providers fail to adequately sample the cervix (failing to sample precancerous/cancerous cells when they are present). It also occurs when the precancerous/cancerous cells on the collecting device do not make it onto the glass slide. The second factor affecting the accuracy of the Pap smear test is interpretation error by the laboratory specialist (for any of a number of reasons including fatigue, inexperience, and habituation). These shortcomings have motivated the research and development of automated screening systems [5, 13].

1.1.2 Automated screening

Automated screening machines can analyse Pap smear slides in a short time without fatigue, providing consistent and objective classification results. The rationale for automated screening is to improve the limitations of the conventional Pap smear test in the following ways [14]:

- Increase the sensitivity¹ and specificity² of the Pap smear test;
- Decrease the workload of technicians and pathologists;
- Reduce the cost for cervical cancer screening programmes; and
- Lower the probability of incidence of cervical cancer and the mortality rate from the disease.

At present there are two FDA-approved cervical cancer screening systems: BD FocalPoint GS Imaging System (formerly known as TriPath AutoPap system) and HOLOGIC ThinPrep Imaging System (formerly known as Cytoc Thin Prep Imaging System) [15].

The FocalPoint GS Imaging System works on LBC slides and any Pap stained conventional smear. The machine scores and ranks the slides based on the likelihood of abnormality and categorizes them into four groups: review, no further review (NFR), process review and quality control (QC) review [15]. Among the analysed slides the NFR group, which comprises up to 25% of the total qualified slides, can be archived without human review. The ThinPrep AutoPap system, by contrast, can only work on ThinPrep

¹The sensitivity measures the proportion of actual positives that are correctly classified.

²The specificity measures the proportion of negatives which are correctly classified.

LBC slides with a special stain which is nearly stoichiometric¹ for DNA content. This system automatically selects 22 field of view (FOV) images that are of diagnostic interest for cytopathologists, from among a total of 120 FOVs acquired on each slide. These 22 FOVs must then be fully reviewed by a cytopathologist. The system provides no scoring or ranking of the slides [15].

Whilst automated screening systems can reduce false negatives attributed to interpretation errors, they cannot reduce false negatives due to sampling errors. Research [16] suggests that a phenomenon known as malignancy-associated changes (MACs) may offer a solution.

1.2 Malignancy associated changes (MACs)

The expression malignancy associated changes (MACs) was coined by Nieburgs et al. (1959) [17]. They reported subtle changes in nuclei of apparently normal looking cells “adjacent to or distant from malignant tumours” [18]. Research in the 1980s identified sub-visual alterations in intermediate cells from cervical atypical smears [19].

The development of an automated screener based on detecting MACs can potentially overcome the problem of sampling error because it is not necessary to perform an exhaustive review of all of the cellular material to identify diagnostic cells but rather to look for subtle nuclear texture changes in a sub-population of cells sampled from the slide.

Based on the available literature, the most effective MAC features seem to be textural features [20]. To date, most of the approaches for defining nucleus texture have been based on stochastic approaches. However, the work of Mehnert [20] suggests that an alternative structural approach to defining texture, which corresponds well to what cytopathologists perceive from cell nuclei, is more effective to describe the chromatin² distribution inside the nuclei.

1.3 Aim and objectives

The aim of this research was to fully explore the structural approach to chromatin pattern description and to evaluate the efficacy of the features derived from it for discriminating between normal and abnormal Pap slides. The research had the following objectives:

1. To develop a robust algorithm for detecting and segmenting cell nuclei in digitized Pap smear images obtained using bright-field microscopy;
2. To develop structural texture features that quantitatively characterise the pattern (arrangement, size, shape, etc.) of the nuclear chromatin;

¹Stains for which the amount of stain uptake in the nucleus is proportional to the amount of DNA are Stoichiometric stains.

²Chromatin is the combination of DNA and proteins that make up the contents of the nucleus of a cell.

3. To determine the most discriminatory subset of features for discriminating between normal and abnormal slides using real clinical data; and
4. To evaluate the performance of a classifier(s), based on the selected features, using real clinical data.

1.4 Scope

As noted in section 1.3, MACs are subtle sub-visual alterations in the appearance of normal looking cells from an abnormal Pap smear. The features which appear to have the most discriminatory power are nuclear texture features [21]. These features reflect chromatin structure and characterise the distribution of chromatin inside the nucleus. The MAC approach to analysing Pap smear slides is conceptually straight forward (see Figure 1.1). The process involves automatically capturing digital images of individual FOVs from a Pap smear slide, identifying the location of nucleus like objects (scene segmentation), segmenting the nuclei like objects in the image (nucleus segmentation), extracting quantitative texture and other features for each nucleus and finally classifying the slides as either normal or abnormal based on these features.

The proposed MAC-based cervical screening approach has the following steps:

1. Scanning the Pap-stained slide using a light microscope coupled with a CCD camera with multiple objectives.
2. Capturing multiple images at different focal planes from interesting fields of view (FOVs) on the slide.
3. Generating extended depth-of-field (EDF) images for each FOV. This involves combining multiple focal planes for each FOV to obtain a single image where each object is all in focus.
4. Locating and segmenting the free-lying cell nuclei in each EDF image, and performing artefact rejection to make sure only nuclei-like objects are retained;
5. Segmenting the nuclear chromatin inside each nucleus into texture primitives (blobs);
6. Extracting features from this structural model to quantitatively characterise the chromatin pattern;
7. Deriving slide-based features from the features in 6 in order to classify the slide as normal or abnormal.

This thesis does not consider the slide scanning and FOV acquisition steps (steps 1 and 2).

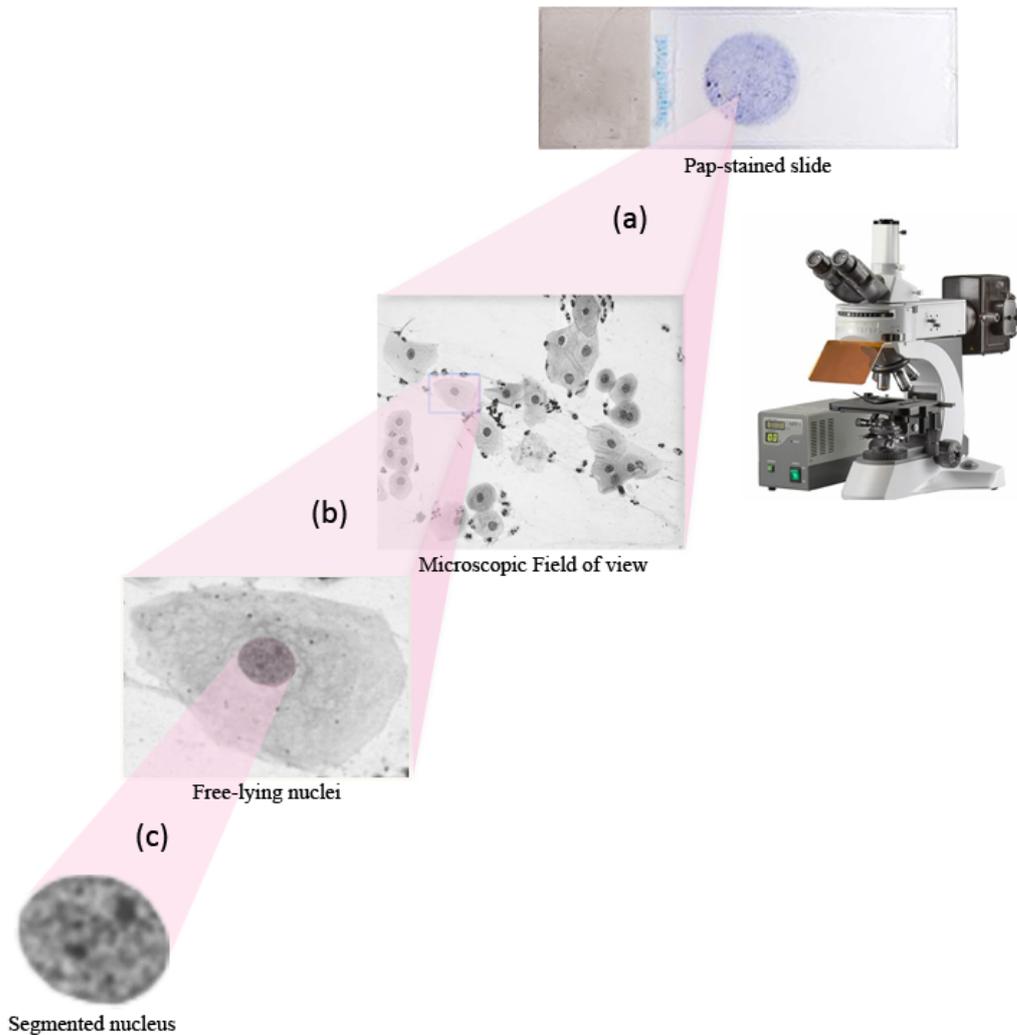


Figure 1.1: Data acquisition and nucleus segmentation in an automated screening system. (a) The cytometer¹ scans the Pap-stained slide and captures scenes from the deposition area on the slide in a predefined way. (b) The location of the cervical epithelial nuclei are identified inside each microscope field of view. (c) Each nucleus-like object is then segmented by defining its boundaries.

¹The cytometer utilises a CCD camera mounted on a light microscope (fitted with a 40× objective lens) to acquire 8-bit images of Papanicolaou-stained cells.

1.5 Structure of the thesis

This chapter has:

- Provided an overview of the Pap smear test, including its shortcomings, and discussed the rationale for automated screening of Pap smear slides.
- Described the malignancy associated changes (MACs) phenomenon and its relevance to automated screening, in particular in addressing the problem of sampling error.
- Defined the aim, objectives, and scope of this research.

The remainder of the thesis is organised as follows:

Chapter 2 This chapter presents two literature reviews pertinent to chapter 3 and chapter 4. The first is a review of the cell and cell nucleus segmentation methods in Pap smear images published in the literature. The second is a review of the texture features available in the literature devised to quantify chromatin texture/distribution.

Chapter 3 This chapter deals specifically with the problem of accurately and robustly segmenting the cervical cell nuclei in digitized light microscopy images of Pap smears. A novel algorithm is developed to detect and segment free-lying intermediate cell nuclei. The algorithm includes three main steps of locating the free-lying nuclei, delineating nuclei and rejecting the artefacts. This chapter also presents an empirical evaluation of the proposed segmentation algorithm for both detection and delineation of free-lying nuclei in Pap smear images.

Chapter 4 This chapter deals with the problem of quantitative characterisation of chromatin texture and presents a set of novel structural texture features to describe nuclear chromatin patterns in cells on a conventional Pap smear. These features are derived from a segmentation of the chromatin into blob-like primitives. The proposed set of features are, in particular, derived from statistics of morphometric features and contextual features computed for these blobs.

Chapter 5 This chapter presents an evaluation of the performance of the proposed structural chromatin texture features. In particular, it presents an investigation of the most discriminatory subset of features, from among the proposed features and a wide range of features drawn from the literature, for discriminating between normal and abnormal Pap smears using the MAC approach. The chapter presents the details of the two experiments carried out in this study. The first is a feature selection experiment performed to obtain the most discriminatory subset of features. The second experiment is to evaluate the performance of a variety of classifiers built using the feature subset obtained in the first experiment to discriminate between the normal and abnormal slides.

Chapter 6 This chapter reviews the work that is presented in this thesis and summarises the major contributions and findings. In addition, it outlines the limitations of the research undertaken, and proposes an avenue of future research.

The material presented in chapters 3, 4 and 5 have been published in the Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2012) [22] and the Proceedings of the 2013 SPIE Medical Imaging Conference [23].

2

Literature review

THIS chapter presents two literature reviews pertinent to chapter 3 and chapter 4. The first is presented in Section 2.1. It is a review of the cell and cell nucleus segmentation methods in Pap smear images published in the literature. The conclusions motivate the method proposed in chapter 3. The second is presented in section 2.2. It is a review of the standard texture features available in the literature devised to quantify chromatin texture/distribution.

2.1 Review of existing methods for segmenting cells and nuclei in Pap smear images

This section provides a review of the existing literature on the topic of automated detection and segmentation of cells and cell nuclei in Pap smear images. The following scientific databases were searched: IEEEXplore¹, Inspec², Medline³, ScienceDirect⁴. The main keywords chosen were: Segmentation, cell segmentation, nuclei segmentation, Pap smears. Image segmentation is the process of partitioning an image into sub-images corresponding to the objects of interest and background. In general, automated segmentation is one of the most difficult tasks in image processing. Numerous algorithms have been published in the literature for segmenting cells and cell nuclei in microscopy images. They can be categorized according to the primary underlying segmentation methodology used: global and adaptive thresholding [24], watershed transform [24, 25], boundary detection algorithms and deformable models [26, 27, 28], and edge enhancement based techniques [29, 30]. These methods are summarized in table 2.1.

¹<http://ieeexplore.ieee.org>

²<http://www.engineeringvillage.com>

³<http://www.ncbi.nlm.nih.gov/pubmed>

⁴<http://www.sciencedirect.com>

The basic methods underlying the existing cell and cell nucleus segmentation algorithms mentioned in table 2.1 are presented in Appendix B. These methods employ fundamental concepts and elements of mathematical morphology (see Appendix A).

An objective comparison between the performances of these methods is difficult, either because no quantitative analysis of performance is provided or because widely different evaluation methodologies (including data and methods) have been used. The lack of information about how some parameters are derived further complicates the assessment of these methods. Table 2.1 shows a detailed comparison of the cell and cell nuclei segmentation methods in Pap smear images available in the literature.

The watershed approach to segmentation has proved to be a powerful and fast segmentation technique for both object boundary delineation and region-based segmentation. Simplicity, speed and complete division of the image are the properties that make the watershed transform a popular method for many different image segmentation applications. The method has been frequently applied to biological images and has produced good results. Watershed-based methods have also been applied for segmentation of clustered cell nuclei [25, 31]. However, these methods usually lead to over-segmentation. In order to alleviate this issue, heuristic rules are defined (as a post-processing step) to merge the over-segmented regions to produce the final segmented image. The alternative approach is to use marker-controlled watershed, which effectively handles the problem of over segmentation [32]. The method requires that the extracted markers represent true cell nuclei. In the following chapter we present a novel algorithm for extracting markers of candidate free-lying nuclei-like objects for subsequent marker-controlled watershed segmentation to obtain the nucleus boundaries.

Table 2.1: Comparison between cell and cell nucleus segmentation approaches found in the literature

Author	Segmentation Method	Images	Image size	Number of cells	Quantitative measure of performance
Bamford et al. [26]	Active contours for nucleus segmentation	20130	128×128	20130 (Single cell images)	Visual evaluation: 99.6% of the Pap stained cell images were correctly segmented.
Wu et al. [28]	Template fitting for nucleus segmentation	1	80×100	1	Misclassification rate of the parametric fitting approach for the synthetic cell image $\leq 5\%$
Garrido et al. [27]	Template fitting for nucleus segmentation	3	Unknown	Unknown	No Quantitative measures
Lezoray et al. [25]	Watershed transform for nucleus segmentation	10	Unknown	209	Vinet measure*: 2.24 and 3.41 in RGB and HSL (Hue, Saturation, Luminance) color space. Mean difference from the groundtruth: 2.8% (RGB) - 0.47% (HSL).
Lassouaoui et al. [33]	Genetic algorithm for cell segmentation	2	256×256	Unknown	No Quantitative measures
Yang-Mao et al. [29]	Edge detector for nucleus and cytoplasm segmentation	Unknown	64×64	124 (single cell images)	The average segmentation errors (MHD, EMM, RAE, ME)** of 0.1523 and 0.0775 for nucleus and cytoplasm segmentation, respectively.
Linnet al. [30]	Edge detector for nucleus and cytoplasm segmentation	10	Unknown	10 (single cell images)	Average segmentation error (the average of ME, RAE, and MHD) of 0.1323 for nucleus segmentation.
Plissitiet al. [34]	Edge detector for nucleus segmentation	38	1536×2048	5617	Sensitivity: 90.57% fuzzy C-means (FCM), 69.86% support vector machine (SVM) Specificity: 75.28% (FCM), 92.02% (SVM)

* Vinet measure [35] is used to quantify the distance between two segmentations and corresponds to the correct classification rate.

** Misclassification error (ME), edge mismatch (EMM), relative foreground area error (RAE), and modified Hausdorff distance (MHD), are often used as segmentation performance measures varying from 0 for an absolutely correct segmentation to 1 for a totally erroneous case [36].

2.2 Review of existing features devised to quantitatively characterise chromatin distribution

This section presents a review of methods proposed for quantitative characterisation of the distribution of chromatin; i.e chromatin texture.

2.2.1 Review of chromatin texture features

A taxonomy of features for cytometry (cell measurement) on microscope images can be found in Appendix C. One class of features in this taxonomy are texture features. In principle such features can be applied to any image object including a cell nucleus.

Two main approaches exist for describing the chromatin arrangement/texture in the cell nucleus. The first approach, called the statistical approach, assumes that texture is a realization of a stochastic process governed by a set of parameters [37] and characterises the chromatin distribution by second or higher order statistics. In the other, the structural approach, the chromatin distribution is assumed to be composed of primitives that are arranged according to certain placement rules.

Traditionally MAC features have been based on a statistical approach to defining texture. However such features do not correspond well to the terms used by cytopathologists to describe chromatin texture such as heterogeneity, granularity, margination, condensation, compaction, clumping, diffuse, blobs and particles [20]. Another difficulty is that these features are sensitive to changes in, or non-uniformity of, illumination and staining. This motivates interest in a structural approach to chromatin texture analysis.

Several methods based on structural texture analysis have been proposed to detect structural alterations of the nuclear chromatin. Beil et al. [38] proposed a dual approach to structural texture analysis for microscopic cell images by region or by lines. In particular they describe the texture in terms of:

- the properties and the arrangement of regions; and
- the properties and the arrangement of lines.

Beil et al. [39] proposed a set of region texture features that corresponds to human vision, such as: number of regions, number of large regions (size>threshold), number of small regions (size<=threshold), number of regions at the boundary of the analysed area, average size of region, etc. They also proposed features for line textures including directionality and fractal dimension of line structures. Albrechtsen et al. [40] presented a structural texture analysis approach that “for each pixel in the image uses concepts from adaptive filtering to find the neighbouring pixels that belong to the local texel”. Thereafter, moment based features are extracted to characterise the grey level texel as an object. Walker and Jackway [41] used features based on the statistics of geometrical (SGF) attributes of connected regions developed by Chen et al. (statistical geometrical features for texture analysis) to quantitatively characterise the chromatin in the nuclei of Papanicolaou-stained cervical cells [41]. Mehnert [20] proffered a method, called the

adjacency graph attribute co-occurrence matrix (AGACM) that combines both structural and statistical/stochastic aspects of texture for characterising both blob-like and mosaic patterns (texture) in the plane.

The Cyto-Savant imaging system computes 116 nucleus features, which can be grouped into 3 general categories: nuclear morphological features, chromatin texture, and DNA content [42]. The measured texture features can be divided into statistical and structural groups. Chromatin distribution is statistically described by Markovian and non-Markovian texture features, fractal texture and run length features. Discrete texture features, however, reflect the structural aspect of chromatin distribution and are computed by first segmenting the nucleus into regions of low, medium, and high optical density. These regions are defined by two global thresholds. Discrete texture features characterise the segmented regions by computing their size, shape, optical density, and spatial distribution. Details of algorithms are described elsewhere [42].

3

Novel algorithm for segmenting free-lying cell nuclei

THE present chapter deals specifically with the problem of accurately and robustly segmenting the cervical cell nuclei in digitized light microscopy images of Pap smears. Given that the aim of the segmentation is to support MAC analysis, the goal is not to segment every nucleus but rather to segment free-lying nuclei. These are easier to segment and the nuclear texture is less likely to be affected by overlapping. Based on the literature review in Section 2.1 we opted to develop our own novel algorithm based on the marker controlled watershed transform for detecting and segmenting free-lying intermediate cell nuclei.

The remainder of this chapter is organised as follows. Section 3.1 presents the high-level description of the proposed algorithm. Section 3.2 presents the proposed marker extraction algorithm corresponding to the first step of the segmentation algorithm. Section 3.3 describes the implementation of the marker-controlled watershed algorithm for segmenting and delineating the detected nuclei. Section 3.4 presents the artefact rejection strategy developed to discard segmented objects that are not nuclei. Section 3.5 describes the choice of features for the artefact rejection step. An empirical evaluation of the performance of the proposed segmentation algorithm is presented in section 3.6. Finally, section 3.7 summarises the chapter.

3.1 High-level description of the proposed algorithm

The proposed segmentation approach is conceptually a 3 step process (see Table 3.1):

1. Detecting or locating the objects of interest;
2. Delineating and labelling those objects; and

3. Artefact rejection to ensure only desired objects are retained in the segmentation results.

Table 3.1: The proposed algorithm to detect and segment free-lying intermediate cell nuclei

- Input: Grey-scale image containing a field-of-view (FOV) from a Pap smear slide.
- Output: Binary image containing connected components, each corresponding to a free-lying intermediate cell nucleus.

Steps:

1. Extract inner markers for free-lying nuclei-like objects (these locate the interiors of candidate objects).
2. Apply the marker-controlled watershed transform on the FOV image with respect to the inner markers (this yields an outer marker that lies between the candidate objects).
3. Apply the marker-controlled watershed transform on the gradient image with respect to the inner and outer markers (this yields the object boundaries/masks).
4. Compute the area and quantitative measures of shape and texture granularity for each segmented nuclei-like object.
5. Reject objects that are too small or large to be intermediate cells, that do not have an elliptical shape, and that do not have a granular texture.

3.2 Extraction of inner markers (step 1)

In Pap-stained cervical images, the cell nucleus appears darker than the rest of the cellular material. Approaches to finding a marker (connected component) within each nucleus exploit this fact. Several candidate methods for extracting marker for nuclei were considered including: tophat transform [43], Jackway tophat [44], h-minima [45]. Appendix D presents the basic methods underlying the existing nuclear marker extraction algorithms. However, we found that none of these methods robustly detects free-lying nuclei. Therefore, we developed a new nuclear marker extraction algorithm for this purpose.

The algorithm is based on the tophat transform defined in terms of an annular closing. The grey-scale annular closing operator is defined

$$\Psi_{\text{ancl}}(f, B) = (f \ominus B) \vee f \quad (3.1)$$

where f is a grey-scale image and B is a symmetric structuring element that does not contain its origin. When annular closing is applied to a grey-scale image, isolated dark spots will be removed. The tophat by annular closing, given by $\Psi_{\text{ancl}}(f, B) - f$, yields the removed isolated dark spots. These spots serve as candidate cell nuclei markers [46].

The procedure is illustrated in Figure 3.1 for a single cervical cell and an annular structuring element. The control over the size and relative isolation of the nuclei is achieved by changing the inner radius and outer radius of the annular structuring element.

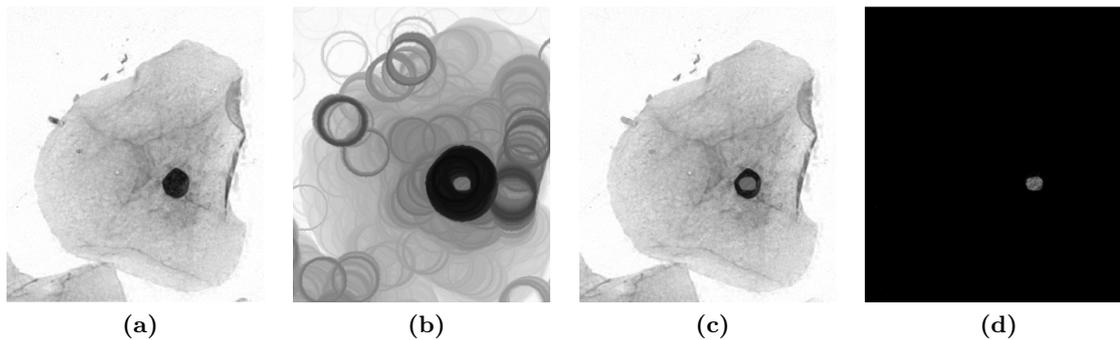


Figure 3.1: Extracting an inner marker for a free-lying cell nucleus. (a) Original image. (b) Grey-scale erosion with an annular structuring element. (c) Pixel-wise maximum of (a) and (b). (d) Arithmetic difference between (c) and (a).

In order to detect a nucleus, an annular structuring element with an inner radius larger than that of the nucleus is needed. The nuclei of normal intermediate and parabasal cells measure approximately $8\mu\text{m}$ in diameter and may enlarge up to $15\mu\text{m}$ in the case of malignant or rare benign disorders changes [47]. Hence, to detect all nuclei within this range, a set of independent annular closings with structuring elements with a range of inner diameters is needed. This is then the basis for the more sophisticated inner marker extraction algorithm presented in Algorithm 1.

3.3 Marker-controlled watershed segmentation of the detected nucleus-like objects (steps 2-3)

Marker-controlled watershed segmentation is used to delineate the boundaries of the cell nuclei detected by the inner marker extraction algorithm. Rather than flooding from the regional minima, as is the case for the traditional watershed transform, flooding is initiated from the markers. The procedure for delineating the detected nuclei-like objects is as follows. First, a watershed segmentation of the original image (f) with respect to the inner markers is performed to obtain the outer marker. Next a watershed segmentation of the gradient magnitude image with respect to the union of the inner

markers and the outer marker is performed. The resulting watershed lines then delineate

Algorithm 1 Proposed nuclei inner marker extraction algorithm

Input: Grey-scale image (f), and parameters λ_0, α, r_1 and r_2

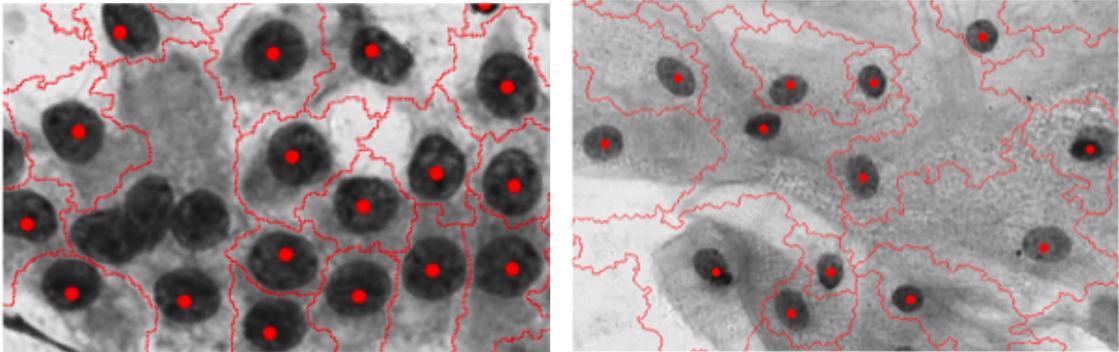
Output: Inner markers binary mask (X_m) for free-lying nuclei-like objects.

- 1: Let B_0 be a disk structuring element of radius λ_0
 - 2: **for** $\lambda = r_1 \rightarrow r_2$ **do**
 - 3: Let B_{an} be an annular structuring element with inner and outer radii of λ and $\lambda + \alpha$ respectively.
 - 4: $g = \Psi_{ancl}(f, B) - f$
 - 5: $X_1 = g > 0$
 - 6: $X_2 = (X_1 \ominus B_0) \oplus B_0$
 - 7: $X = X \cup X_2$
 - 8: **end for**
 - 9: $X_m =$ set of centroids of the connected components in X .
-

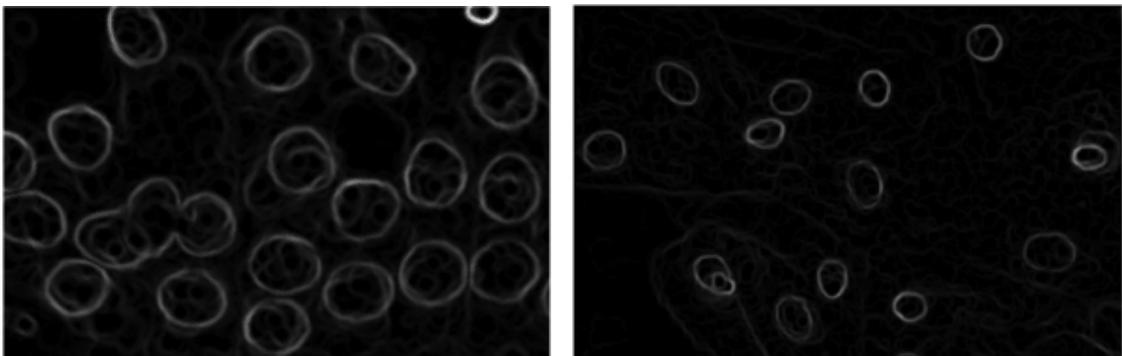
the nuclei-like objects. This idea is illustrated in Figure 3.2. The reason behind using the gradient image is that the nuclei boundaries are located on high gradient points. Therefore the watershed transform results in regions with boundaries corresponding to those of the nuclei.

The gradient magnitude image can be quite noisy. For this reason, the original image is median filtered to remove impulse noise (size of the kernel is selected to be 3×3), its magnitude of gradient is computed, and the result is Gaussian filtered.

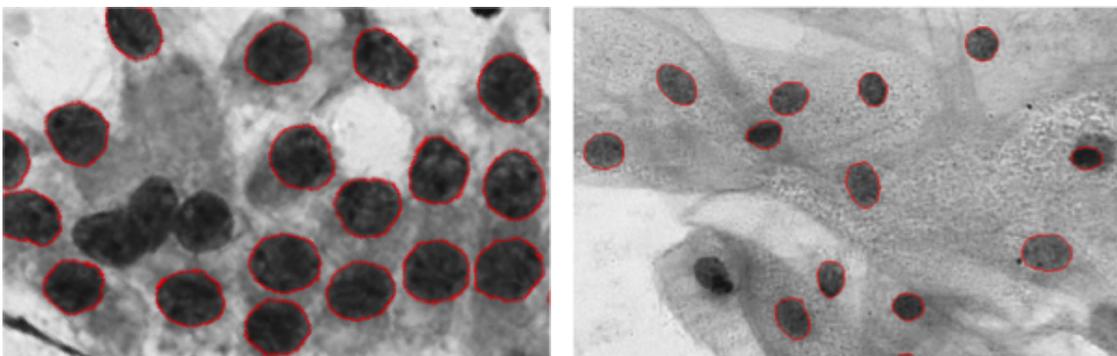
There exist many different approximations for the computation of the magnitude of the gradient of an image. Several approaches with similar results have been tested. Among all these methods, the very simple and straight forward Gaussian derivative operator is used. The standard deviation of Gaussian filter is selected to be 0.9. The gradient magnitude image is approximated by sum of the absolute values of the convolution of the image with the vertical and horizontal Gaussian derivative operators.



(a) Watershed segmentation of the original image with respect to the inner markers (shown as disks) yielding the outer marker.



(b) Gaussian filtered gradient magnitude of the median-filtered original image in (a).



(c) Watershed segmentation of (b) with respect to the union of the inner markers and the outer marker.

Figure 3.2: Segmentation of the detected nuclei-like objects.

3.4 Artefact Rejection (steps 4-5)

The presence of overlapping cells, overlapping and/or folded cytoplasm, blood and cellular debris presents a formidable challenge for nuclei segmentation [21]. Ensuring that the segmented objects are nuclei of cells relevant for MAC analysis is essential. The artefact rejection strategy is based on the size, shape and granularity of the objects. Quantitative measures of the area, elliptical shape, and the texture/granularity are computed for each segmented nucleus-like object.

3.4.1 Size criterion

The squamous epithelium of the female genital tract is composed of three principal layers [47]: the basal cell layer (immature), the intermediate cell layers, and the superficial cell layers (most mature) (see Figure 3.3). As the cell maturation progresses toward the surface, the amount of cytoplasm per cell increases [47]. The nuclei of superficial cells are pyknotic and considerably smaller than intermediate and parabasal cells with a nuclear diameter of about $4\mu\text{m}$ [47]. In a normal Pap smear usually only the upper few layers of the squamous epithelium are removed and so the immature cells near the base of the epithelium are not sampled [16]. Given that the aim of the algorithm is to detect intermediate cell nuclei (presenting a fine network of chromatin and chromocenters suitable for further MAC analysis), a threshold value can be defined to remove the superficial cell nuclei from the segmentation result.

The nuclei of normal intermediate cells ranges approximately between $8\mu\text{m}$ to $15\mu\text{m}$ in diameter and are oval-shaped [47], accordingly the minimum area can be defined. The minimum area is deemed to be the area of a circle with radius r_{min} . This artefact rejection step removes not only the superficial squamous cell nuclei but also many of the small objects belonging to image artefacts.

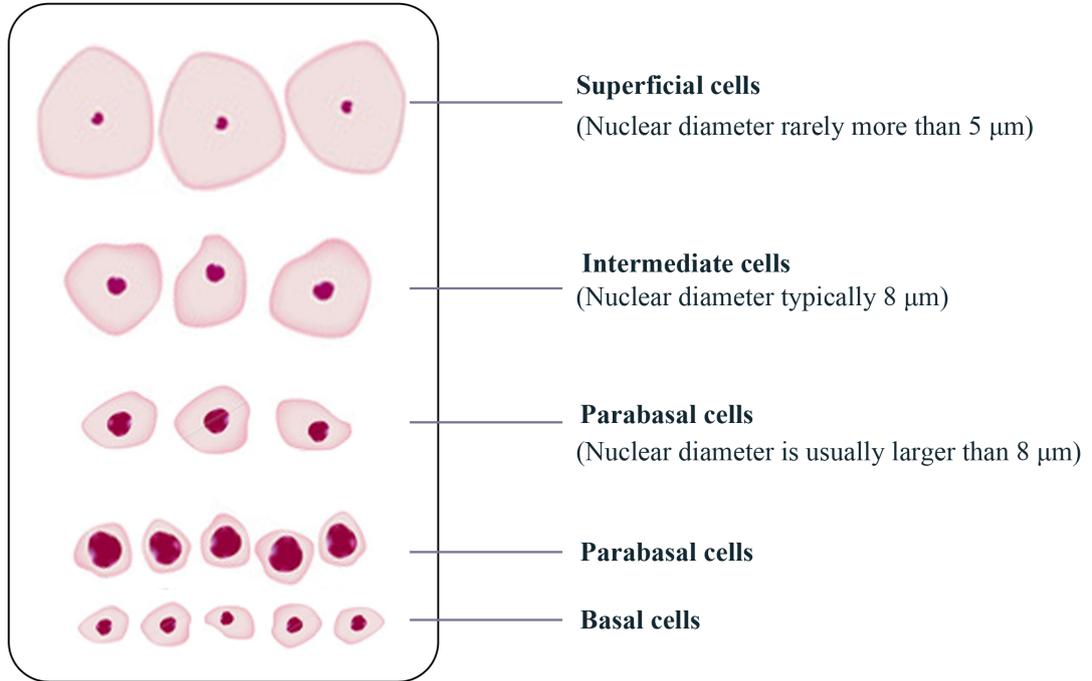


Figure 3.3: Cells of the squamous epithelium (freehand adaptation of Koss (2006, Figure 5-4)).

3.4.2 Shape criterion

Nuclei of the intermediate cervical cells are elliptical in shape [47]. To assess whether a candidate nucleus object is elliptical, the elliptic variance feature is computed. Some other features we looked at to fulfill the shape criterion are Danielsson's G shape factor, ellipticity, elliptic variance and Dice-area (see Appendix E). The elliptic variance descriptor (E_{var}) [48] measures how closely the borders of the fitted ellipse agree with those of the segmented object (see Figure 3.4).

The simplest way to generate a signature for coarseness of the boundary of an object is to compute the radial distances of the object boundary from the object centroid. Suppose $\mathbf{g} = [g_x, g_y]^T$ is the centroid of the object and object boundary has N data points $\mathbf{p}_i = (x_i, y_i)^T$. The covariance matrix of the data points is calculated using:

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^N (\mathbf{p}_i - \mathbf{g})(\mathbf{p}_i - \mathbf{g})^T$$

Then the radial distances (d'_i) of boundary points from the centroid are calculated as:

$$d'_i = \sqrt{(\mathbf{p}_i - \mathbf{g})^T \cdot \mathbf{C} \cdot (\mathbf{p}_i - \mathbf{g})}$$

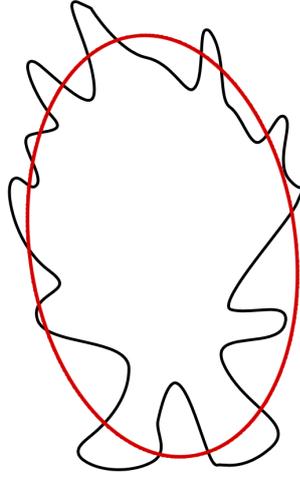


Figure 3.4: Ellipse fitting to an arbitrary shape

The mean (μ'_R) and standard deviation (σ'_R) of the radial distances over all boundary points are:

$$\mu'_R = \frac{1}{N} \sum_{i=1}^N (d'_i) \text{ and } \sigma'_R = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \mu'_R)^2}$$

and finally E_{var} is defined:

$$E_{var} = \frac{\sigma'_R}{\mu'_R} \quad (3.2)$$

It should be noted that for elliptical objects E_{var} is close to 0.

3.4.3 Texture criterion

The feature chosen to measure the degree of the granularity of the texture in a nucleus is the Tamura coarseness feature. Tamura [49] devised a texture model corresponding to visual perception. It is useful and robust in the sense that it does not depend directly on the exact grey-levels in the object and so has robustness to non-uniformity of illumination and staining variations (provided that these do not greatly affect the size and number of texture primitives). The Tamura coarseness aims to pick the biggest size where the texture is present. The primitive elements (textures) are larger in size but smaller in number for a coarse texture, while a fine texture contains a large number of small primitives. The computational procedure can be summarized in the following steps.

1. Take averages at each pixel (x,y) over neighbourhoods whose sizes are powers of

two $2^k \times 2^k$.

$$A(x,y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+x^{k-1}-1} \frac{f(i,j)}{2^{2k}} \quad (3.3)$$

where $f(x,y)$ is the grey-level at (x,y) . The averaging procedure at different levels is depicted in figure 3.5.

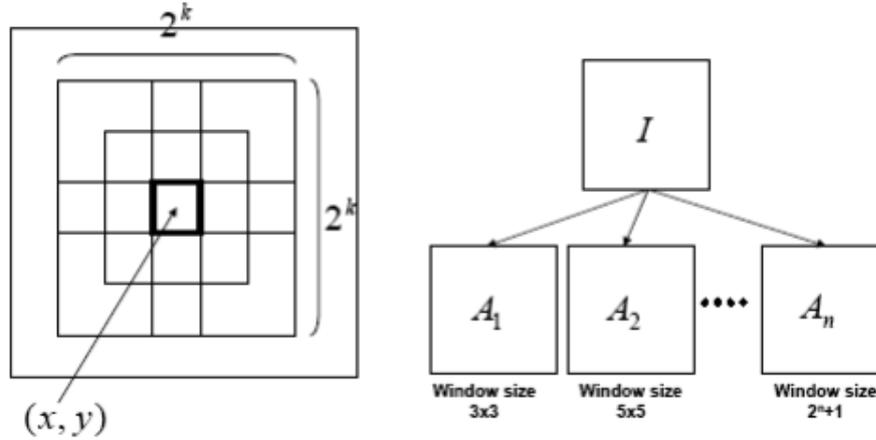


Figure 3.5: Computing averages at different scales

- At each pixel, take the absolute difference between pairs of non-overlapping averages on opposite sides of the point in both horizontal and vertical directions. The difference in the horizontal case is:

$$E_{k,h}(x,y) = \left| A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y) \right| \quad (3.4)$$

- At each pixel, pick the best size of k which maximizes the difference $E_k(x,y)$ in either direction and set the best size to $S_{opt}(x,y) = 2^k$.
- Calculate the coarseness measure by taking the average of S_{opt} over the entire image:

$$F_{crs} = \frac{1}{m \times n} \sum_i^m \sum_j^n S_{opt}(i,j) \quad (3.5)$$

Where, m and n are the width and height of an image, respectively.

Figure 3.6 shows 6 images of cell nuclei with their corresponding Tamura coarseness measures.

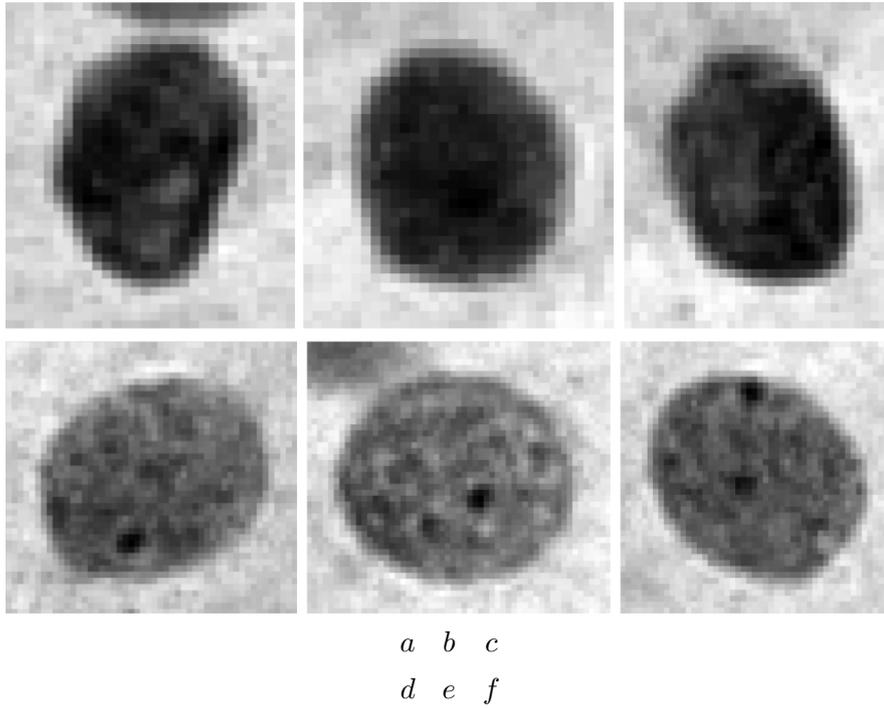


Figure 3.6: Measure of Tamura coarseness for 6 different cell nuclei.

Cell nuclei	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>
Coarseness measures	9.10	9.52	9.56	10.41	10.72	10.9

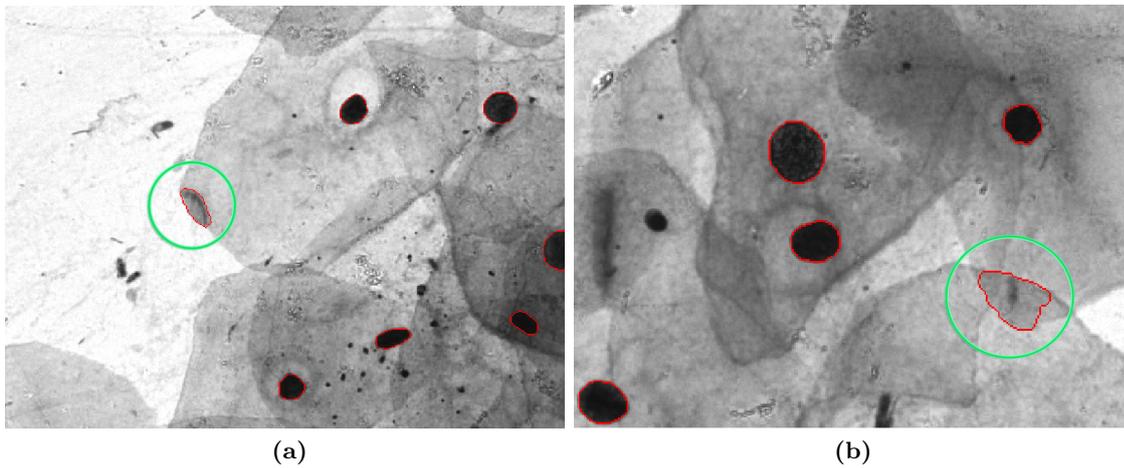


Figure 3.7: Two examples of artefacts rejected by the Tamura coarseness feature.

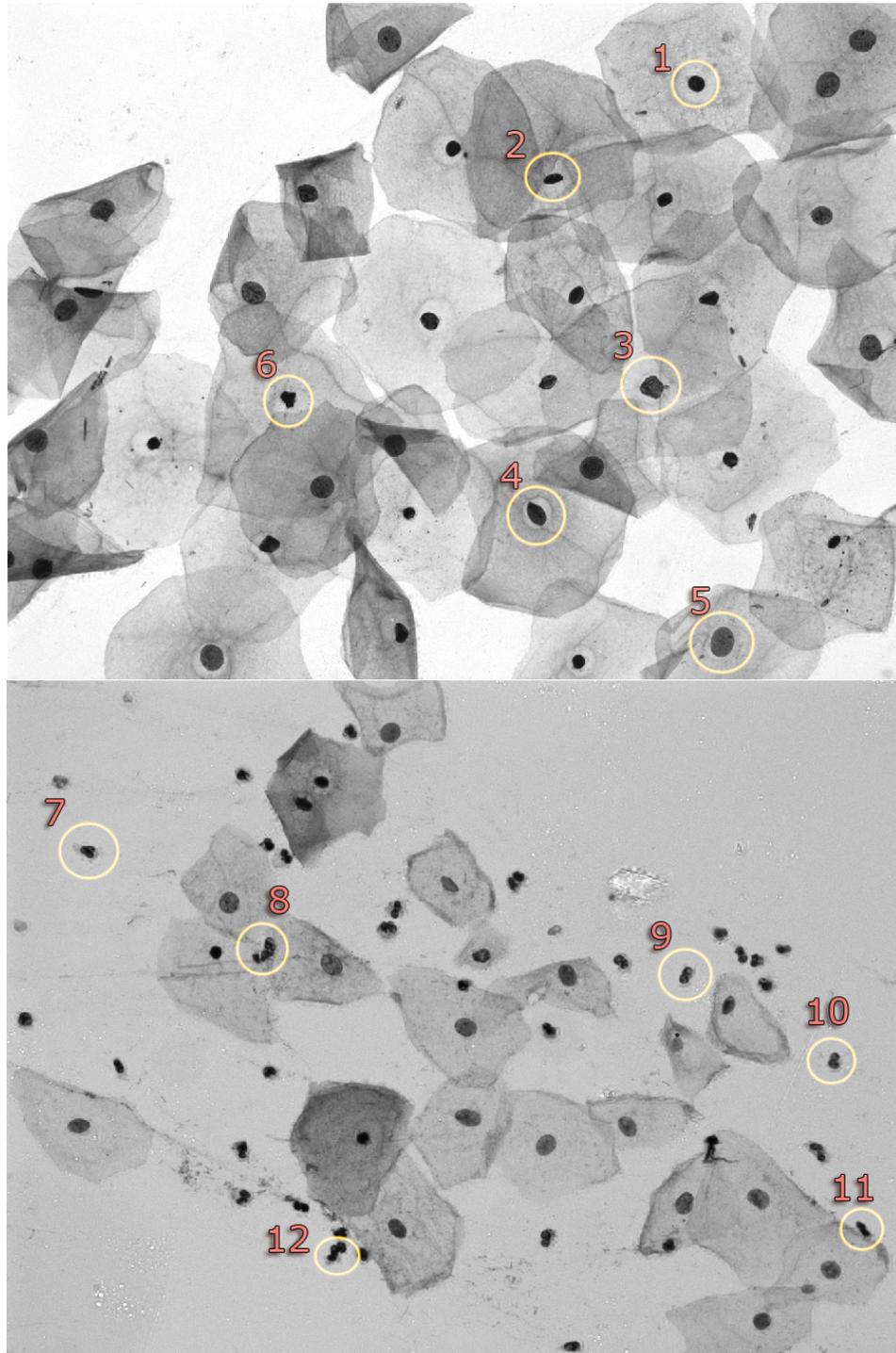


Figure 3.8: Two sample FOVs from a Pap smear slide. Twelve samples of the segmented objects before the artefact rejection step are highlighted. The quantitative measures of shape and granularity for these objects are in Table 3.2

Table 3.2: Shape and granularity features for 6 different cell nuclei.

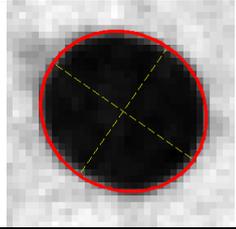
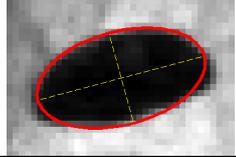
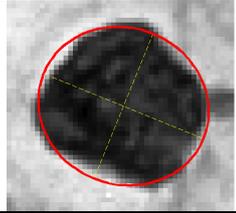
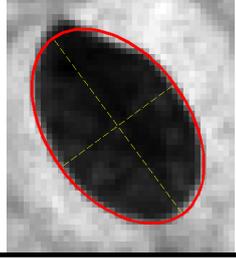
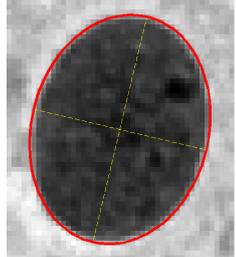
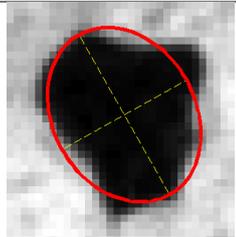
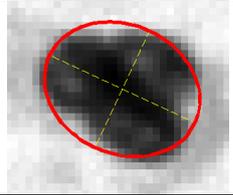
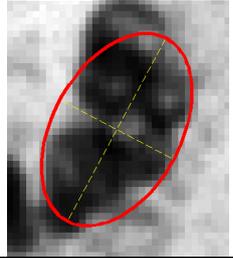
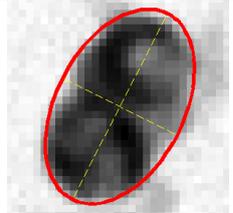
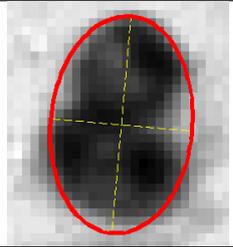
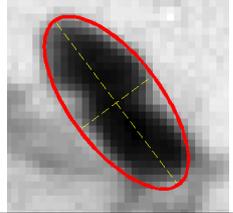
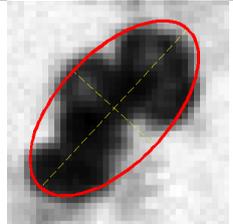
<i>No</i>	<i>Cell Nuclei</i>	<i>Area</i>	E_{var}	f_{course}
1		480	0.031	9.48
2		331	0.69	8.79
3		924	0.07	11.95
4		740	0.043	12.23
5		1325	0.033	10.13
6		390	0.104	10.13

Table 3.3: Shape and granularity features for 6 non-nuclei artefacts.

<i>No</i>	<i>Artefacts</i>	<i>Area</i>	<i>E_{var}</i>	<i>f_{course}</i>
7		311	0.1138	10.78
8		460	0.129	11.10
9		392	0.072	9.80
10		336	0.137	9.31
11		288	0.109	12.80
12		464	0.209	11.81

3.5 Empirical evaluation of the proposed segmentation algorithm

It was noted in the review in section 2.1 that many of the proposed segmentation algorithms have not been formally empirically evaluated. Herein we present an empirical evaluation of our proposed method relative to expert manual segmentation.

3.5.1 Image data

The data used in this study is a subset of 889 fields of view (FOVs) captured by a cytopathologist from 68 Pap smear slides. Each FOV was acquired using a CCD camera mounted on a light microscope. The images were captured with a $40\times$ objective lens. Each FOV image is of size 1024×1344 pixels with square pixels of size $0.25\mu\text{m}$. The grey-scale resolution is 8 bits per pixel.

3.5.1.1 Ground truth generation

Eleven slides, each containing a minimum of 100 non-superficial cervical cell nuclei, were randomly selected from among the 68 slides. For each slide three FOVs were randomly selected to yield a total of 33 FOVs.

Two graphical user interfaces (GUI) were developed to permit a user to review each FOV and to place a marker on individual nuclei and also trace the boundary of nuclei. These two GUIs were designed specifically for the two purposes of marking and manual border delineation of cell nuclei (see Figures 3.9 and 3.10). They provide the following functions: cell nuclei marking function (by clicking within the nucleus area), border delineation function (using the mouse to trace the cell nuclei boundaries) and a measurement tool for checking the desired particles within the image. Figures 3.9 and 3.9 show two GUIs designed and used for ground truth generation.

Three untrained subjects were recruited to independently review the FOVs using the GUI and to mark each free-lying nucleus. Prior to performing this task each was shown examples of intermediate cell nuclei in another FOV (not one of the 33 FOVs they had to review). Each subject was specifically instructed to mark elliptical objects, of approximately the right size, with a well-defined boundary, and with a granular texture. The set of all objects selected by at least two of the three subjects were taken to be the ground truth for free-lying intermediate cell nuclei. Two image analysis experts (authors) then used the GUI to independently trace the boundary of each ground truth nucleus. These manual segmentations were taken to be the ground truth for the boundaries of the free-lying intermediate cell nuclei.

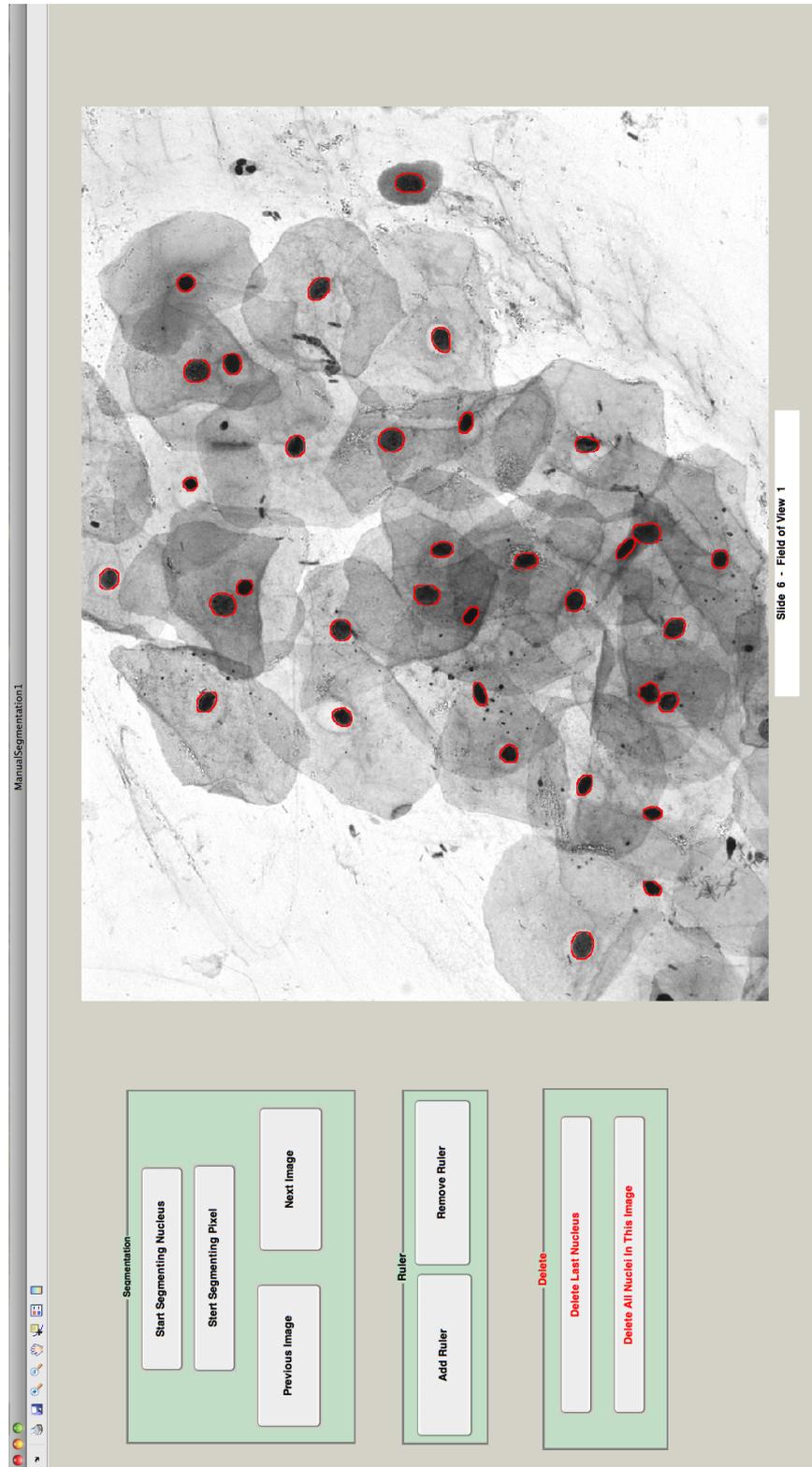


Figure 3.9: GUI for Ground truth generation

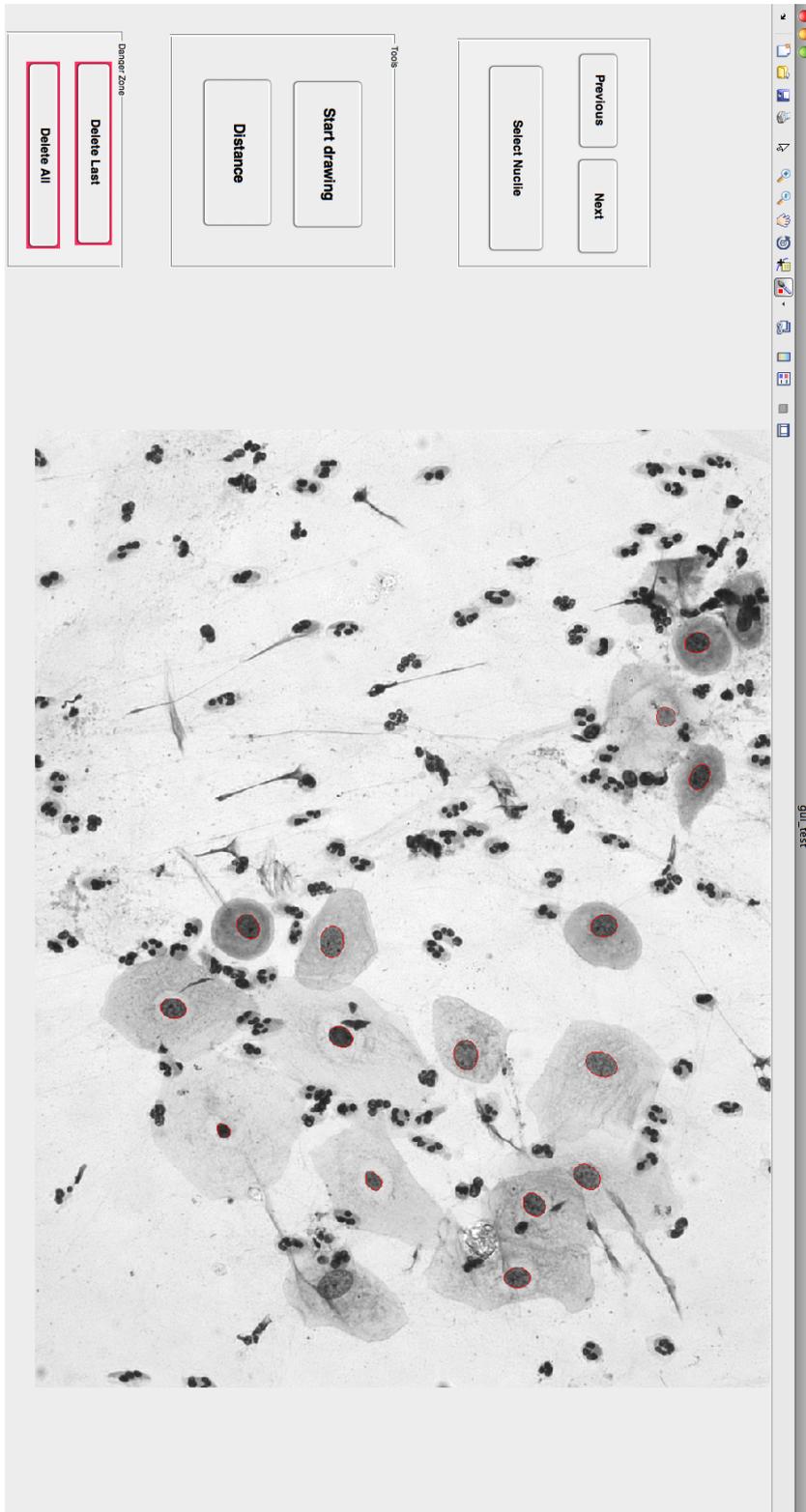


Figure 3.10: GUI for Ground truth generation

3.5.2 Method and Experiment

The proposed segmentation algorithm was applied to the 33 FOV images. The parameters for different steps of the algorithm were selected after several experiments on a small subset of images independent of the 33 selected FOV images. The minimum and maximum values of the inner radius (r_1 and r_2) of the annular structuring elements (see algorithm 1) were set to 22 and 33 pixels. The values were selected based on the nucleus diameter range for the intermediate cell nuclei defined by Koss [47]. The outer radius of each annulus was set to be two pixels more than the inner radius (i.e. $\alpha = 2$) to guarantee the extraction of inner markers of the adjacent free-lying nuclei. Finally the size of the disk-structuring element (λ_0) was set to 3 pixels (for noise removal).

The parameters for the artefact rejection step were also tuned on an independent data set. Based on the defined size range of intermediate cell nuclei in [47], the threshold value of 400 for segmented regions was selected for artefact rejection (The minimum area is deemed to be the area of a circle with radius r_{min}). The threshold value of 0.095 and 13.2 were selected, respectively, for the elliptic variance and Tamura coarseness features.

Objects selected by the algorithm were compared to the ground truth nuclei obtained manually and used to compute the sensitivity and specificity of the algorithm for the detection of free-lying intermediate cell nuclei.

The accuracy of segmentation by the proposed algorithm for each detected mask was assessed through comparison of the boundaries of the segmented object to the two corresponding boundaries (boundaries traced by two experts) in the ground truth datasets. The similarity between pairs of masks and the similarity is computed. More specifically the similarity between pairs of masks was computed in terms of the Dice similarity coefficient (DSC) scores defined by equation E.3.

3.5.3 Results

The sensitivity and specificity of the algorithm for the detection of free-lying intermediate cell nuclei is 94.71% and 85.3% respectively. Box-plots of the DSC scores for the comparison of the proposed automatic segmentation to the two manual segmentations, and for the comparison between the two manual segmentations are shown in Figure 3.11.

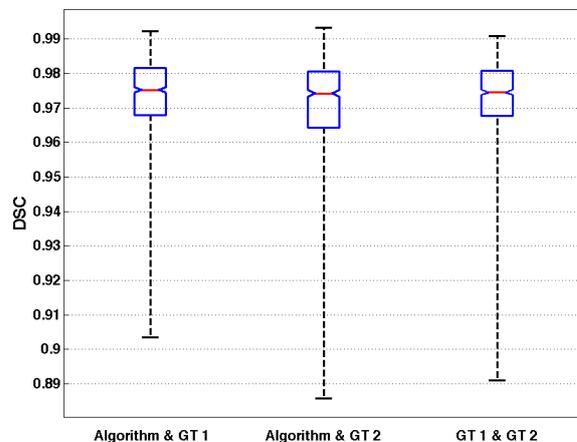


Figure 3.11: DSC scores for the comparison of the proposed automatic segmentation to the two manual segmentations, and for the comparison between the two expert segmentations (GT1 and GT2).

The agreement between the algorithm and the two manual segmentations is $97.30 \pm 1.35\%$ and $96.96 \pm 1.7\%$ respectively (mean DSC \pm standard deviation). The overall agreement between the two expert segmentations is $97.26 \pm 1.2\%$.

3.5.4 Discussion

The sensitivity of the algorithm to the detection of free-lying intermediate cell nuclei is remarkably high. However its specificity, whilst still quite high, could be improved. A review of false positives indicates that some of them are due to segmentation failures as the result of severe background noise and artefacts. However other apparent failures in fact represent genuine free-lying nuclei overlooked by the three recruited subjects. Understanding the reasons for an algorithm's failure modes can be used to improve its robustness. The main failure modes responsible for detection of false positive cases are discussed below.

The DSC scores for boundary delineation evaluation (see Figure 3.11) show that nuclei boundaries obtained using the marker-controlled watershed transform are highly accurate and consistent with the two experts' visual perception of the intermediate cell nuclei boundaries.

3.5.4.1 Failure modes

In the experiment 111 false positive cases were detected. In many of the cases the false positives are visually similar in appearance to a real nucleus. Therefore the marker

detection algorithm extracts them as a nucleus-like object. The aim of the artefact rejection step is to discard a segmented artefact. The ellipticity and size features effectively eliminate any object with an irregular size or shape. After reviewing the false positive we discovered we concluded that the Tamura coarseness feature, whilst quite effective in most cases, did not reject some of the artefacts as expected. Figure 3.12 shows two such cases. The results suggest that other features be considered in future to improve the specificity.

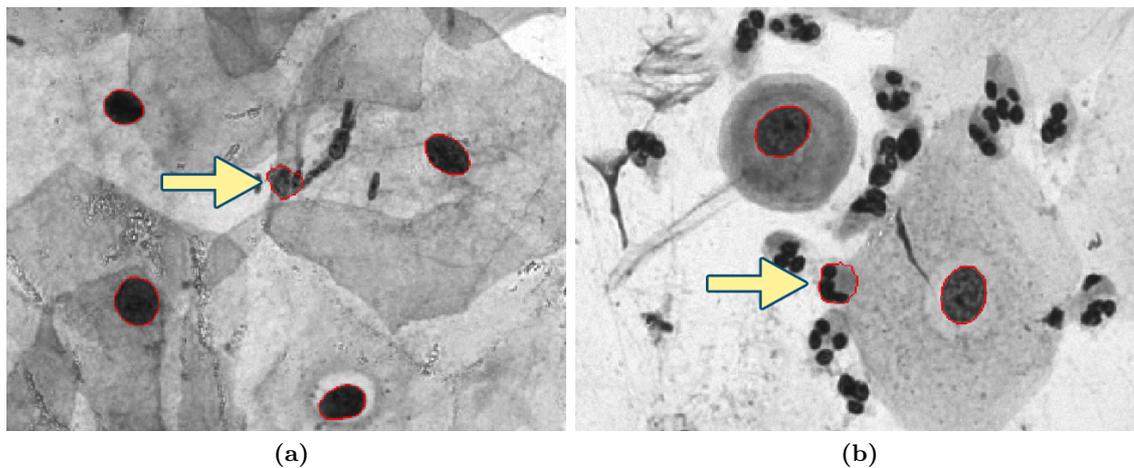


Figure 3.12: (a) and (b) show two examples of false positive detection by the algorithm.

There were very few cases where there was no marker for a nucleus. Among the total 755 nuclei present in our database, the algorithm did not produce any marker for 14 nuclei. This basically happens when the nucleus is surrounded by many other objects, especially when the adjacent objects have considerably lower intensity compared to the nucleus. An example of this type of failure is depicted in Figure 3.13a. All the other nuclei markers were successfully extracted and no multiple seeds for a single object were detected. The other mode of failure was due to inaccurate border delineation by the watershed algorithm. This happens when there is very weak gradient information on the boundary of the nucleus (see Figures 3.13b and 3.13c). A total of 6 nuclei were missed due to watershed failure. Finally the particular thresholds defined for artifact rejection features were responsible for eliminating some of the true segmented free-lying nuclei. First of all the size feature has caused the rejection of the some of the segmented nuclei that have slightly lower size than the predefined threshold. The ellipticity and the Tamura coarseness features were responsible for discarding the rest of the true segmented nuclei. Overall, 20 segmented regions that were truly segmented were rejected by the artefact rejection features. Figure 3.14 shows some examples of these failure modes.

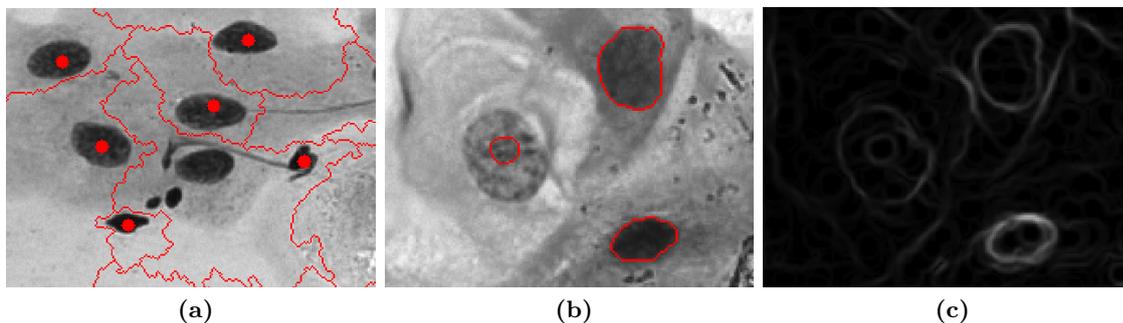


Figure 3.13: Examples of missed nuclei due to presence of background noise surrounding the nucleus and watershed transform failure: (a) The marker extraction algorithm has failed to provide a marker for the nuclei. (b) Example of failure of the morphological watershed technique because of sharp edge information inside the object, as seen in the gradient image (c)

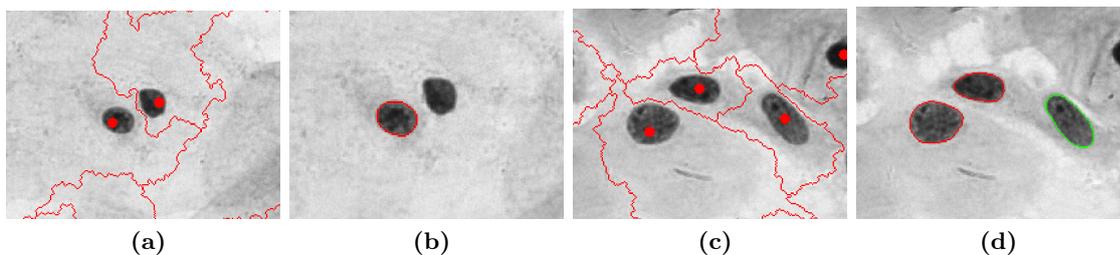


Figure 3.14: Examples of missed nuclei due to unsuitable choice of parameter values: (a) and (b) The two objects have approximately the same size, however, the segmented nucleus has a size slightly above the threshold value while the size of the other nucleus is slightly below, hence discarded as an artefact. (c) and (d) Example of an accurately segmented object which was discarded as an artefact in the artefact rejection step. This is because the value for Tamura coarseness feature of this nucleus is lower than the defined threshold value.

3.6 Summary

This chapter presented a new algorithm for segmentation of free-lying cell nuclei in Pap smears. The novelty of the algorithm stems from a robust marker selection method for selecting candidate free-lying nuclei-like objects for subsequent marker-controlled watershed segmentation to obtain the nuclear boundaries. The algorithm also implements artefact rejection based on size, shape, and nuclear granularity to ensure only the nuclei of intermediate squamous epithelial cells are retained. In addition, this chapter presented an empirical evaluation of the performance of the algorithm and discussed the results. The sensitivity and specificity of the algorithm to the detection of free-lying intermediate cell nuclei together with the accuracy of the segmentation of each nucleus detected by the algorithm were presented.

4

Texture features characterising chromatin distribution

THIS chapter is concerned with the quantitative characterisation of the texture of chromatin. A review of chromatin texture features proposed to date was presented in chapter 2. A novel set of structural chromatin texture features are presented in this chapter. These are derived from a segmentation of the chromatin into blob-like primitives. By considering each of the chromatin particles as individual regions (objects), many features can be extracted that describe the distribution and arrangement of those particles. This is the structural way of describing texture where the texture is characterised by a description of its primitives and placement rules [20].

Using the segmented chromatin particles, two sets of features can be measured: Features that solely characterise the blob-like particles in terms of measurements such as area, perimeter, dynamics, etc. and features that characterise the spatial relationships existing between the chromatin particles.

The remainder of this chapter is organised as follows. Section 4.1 presents the utilised chromatin segmentation algorithm. Section 4.2 presents the proposed structural texture features to characterise chromatin texture/distribution. Finally, Section 4.3 summarises the chapter.

4.1 Chromatin segmentation

Chromatin is the material in the cell nucleus consisting of DNA and associated proteins. It is so-named because of its ability to take on stain. In a Pap smear nuclear chromatin is visualised under light microscope as a mosaic of interchanging light and dark regions within the nucleus (dark and light regions with high and low optical density¹,

¹The optical density also called extinction of a material is the logarithmic ratio of the incident radiation to the transmitted radiation through that material.

respectively). Regions of high optical density are well defined as particles [50].

The method chosen to segment the chromatin is proposed by Mehnert [20]. It can be used to segment the dark and light particles (blobs) in a grey-level image of nucleus. The grey levels in the image can represent either the intensity or optical density. Figure 4.1 shows a representation of a nucleus as a topographic relief. The dark particles or chromatin blobs correspond to valleys and light particles correspond to mountains. In a typical topographic relief, the valleys are associated with minima and the mountains are associated with maxima.

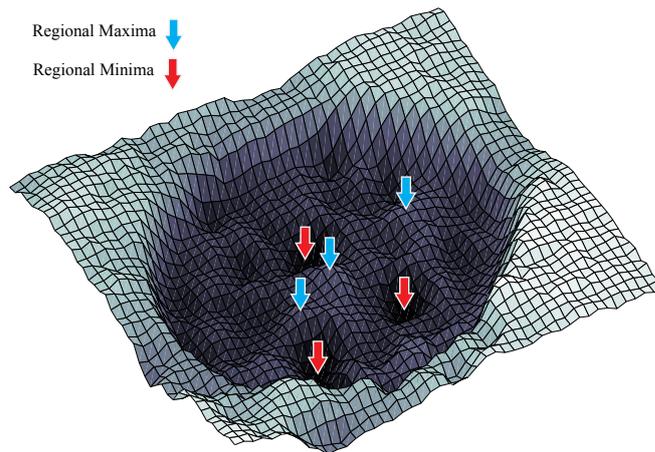


Figure 4.1: Representation of a cervical cell nucleus as a topographic relief (grey-level represents height).

The chromatin segmentation algorithm proposed by Mehnert [20] has the following steps:

1. Locating the regional minima;
 - Applying a 3x3 median filter (to attenuate impulse noise);
 - Up-sampling by a factor of 3 (to facilitate the rendering of watershed lines in a subsequent step);
 - Locating the regional minima (inner markers).
2. Computing the watershed transform of the filtered image with respect to the inner markers (to produce an outer marker that delineates a zone of influence around each regional minimum); and
3. Computing the magnitude of gradient image for the filtered image; and
4. Applying the watershed transform to the gradient image with respect to both the inner markers and the outer marker (to delineate each dark chromatin blob).

Figure 4.2 shows examples of the segmentations produced by this algorithm. Light particles can also be obtained by applying the algorithm to the negative of the image.

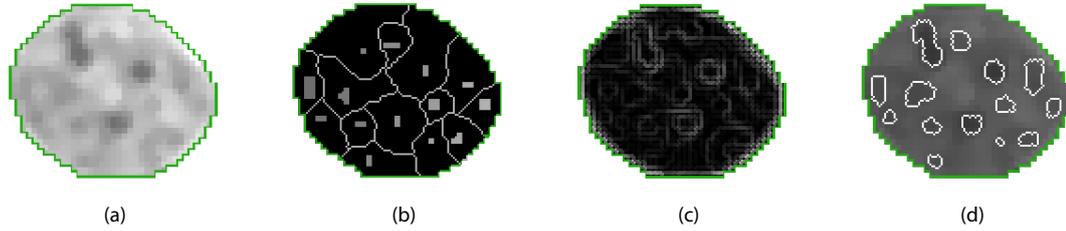


Figure 4.2: Chromatin segmentation. (a) Original image of cell nucleus. (b) Watershed transform of (a) with respect to regional minima (inner markers) yielding the outer marker. (c) The morphological gradient image of (a). (d) Watershed segmentation of (c) with respect to the union of the inner markers and the outer marker.

4.2 Proposed structural texture features

In this section we present a novel set of structural chromatin texture features derived from a segmentation of the chromatin. By considering each of the chromatin particles as individual regions (objects), many features can be extracted that describe the distribution and arrangement of those particles. Using the segmented chromatin particles, two sets of features can be measured: features that characterise the spatial relations existing between the chromatin particles and features that solely characterise the blob-like particles through measurements such as area, perimeter, dynamics, etc.

4.2.1 Graph-based features

The graph-based features are derived from ordinary delaunay graph (the dual of the ordinary Voronoi diagram), several graphs related to the area-Voronoi diagram—generalised Delaunay graph, generalised Gabriel graph, generalised relative neighbourhood graph, and the generalised minimum spanning tree—and from the area-Voronoi diagram itself. These graphs describe different adjacency relationships between the chromatin particles. Formally, a graph is represented by $G = (V, E)$ where V is called the vertex set and E is the edge set of the graph. Vertices correspond to particles and edges denote adjacency (see [51] for details concerning the computation of the graphs).

Our proposed graph-based features can be categorised into 6 groups: Delaunay graph based features, features based on the area-Voronoi diagram, features based on the generalised Delaunay graph derived from the area-Voronoi diagram, generalised Gabriel and relative neighbourhood graph based features, and finally features based on the generalised minimum spanning tree. Definitions and basic geometric properties of these graphs can be found in [51]. Figure 4.3 shows the area-Voronoi diagram constructed by computing the watershed transform of the Euclidian distance transform of the segmented chromatin particle masks and the generalised Delaunay graph derived from the area-Voronoi diagram.

Figure 4.4 shows the graphs driven from the area-Voronoi diagram. The generalised Delaunay graph (DG) is obtained directly from the area-Voronoi diagram. Generalised

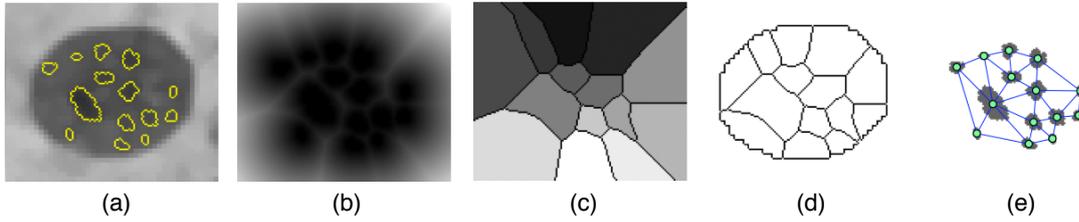


Figure 4.3: Illustration of how the area-Voronoi diagram and generalised Delaunay graph are computed. (a) Segmented chromatin particles superimposed on the original nucleus image. (b) Euclidian distance transform of the set complement of the union of the chromatin particle masks in (a). (c) Area-Voronoi diagram. (d) The area-Voronoi diagram clipped to the nuclear mask. (e) The region adjacency graph defined on the Voronoi regions then defines a generalised Delaunay graph.

Gabriel graph (GG), generalised relative neighbourhood graph (RNG), and generalised minimum spanning tree (MST) are all obtained as sub-graphs of the generalised Delaunay graph. In particular the generalised Delaunay graph is treated as an ordinary Delaunay graph, with vertices now corresponding to particle centroids, and the sub-graphs are obtained by removing select edges. This means that they each have the same vertices and the edges satisfy $MST \subseteq RNG \subseteq GG \subseteq DG$. Several features can be computed from the area-Voronoi diagram and associated graphs to quantify the arrangement of the chromatin particles inside the nucleus. Features extracted from the Delaunay, Gabriel and relative neighbourhood graphs measure the distances between chromatin particles and describe variations in the connectivity of the nodes. Voronoi-based features describe the shape and size of the Voronoi polygons/regions. The minimum spanning tree connects all the vertices in a given graph in a way that the sum of the edge lengths are the minimum possible.

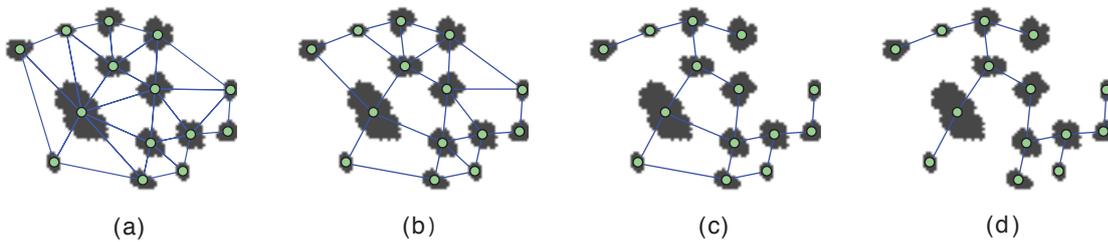


Figure 4.4: Graphs derived from the area-Voronoi diagram. (a) the generalised Delaunay graph (b) the generalised Gabriel graph (c) the generalised relative neighbourhood graph (d) the generalised minimum spanning tree. See the text for details.

4.2.2 Margination features

Our margination features characterise the distances of the segmented chromatin blobs to the nucleus boundary. Six of these features are computed from the cookie-cutting-distance [20] illustrated in Figure 4.5. The slopes of the cumulative frequency curve of the cookie-cutting-distance at the 25th and 37th percentiles are calculated ($M1$ and $M2$) as two measures of margination for each nucleus¹.

Two other features for characterising margination are derived by computing the mean and standard deviation of the distances in the cookie-cutting-distance image ($M3$ and $M4$). These two features also include information about the size of the nucleus, which is an indicator of the state of cell in certain disease processes [52]. In addition, features $M3$ and $M4$ are both normalized to the maximum value of the nucleus boundary distance transform to yield pure measures of margination ($M5$ and $M6$). One additional feature is the mean of the minimum distances between the geometrical centres of chromatin particles and the nucleus boundary ($M7$).

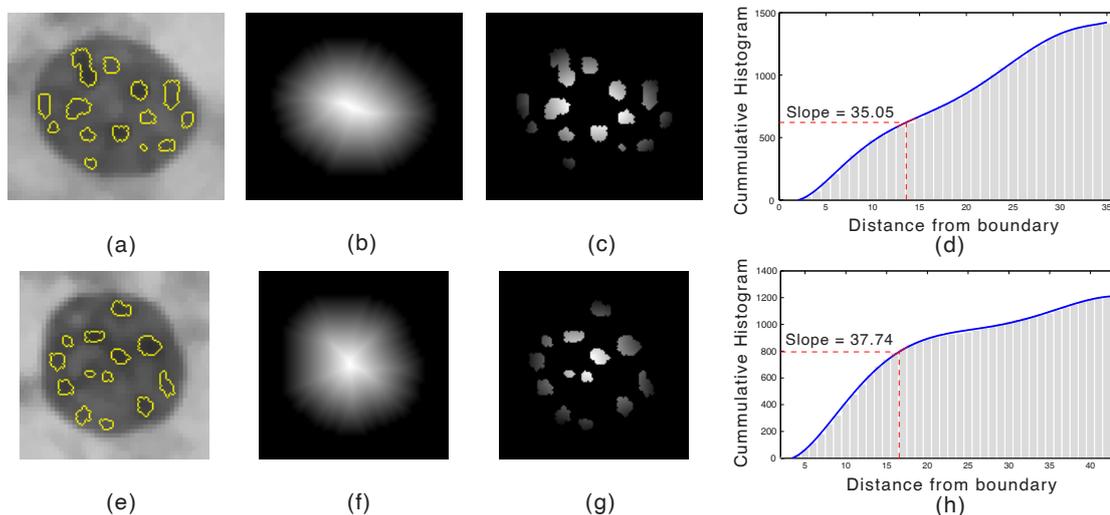


Figure 4.5: Illustration of how the margination features $M1$ and $M2$ are computed. (a),(e) Segmented chromatin particles. (b),(f) Euclidean distance transform of the nucleus boundary. (c),(g) Portions of (b),(f) cut out by the segmented chromatin particle masks. This is called cookie-cutting-distance. (d),(h) Slope of the cumulative frequency curve of the cookie-cutting-distance at the 37th percentile.

4.2.3 Clustering features

Four quantitative measures for chromatin particle clustering are proposed. For the first measure, the distance transform of the area-Voronoi diagram circumscribed by nucleus

¹These percentiles were chosen empirically using the cell images that were not used in the empirical evaluation described in the next chapter.

boundary is calculated. Portions of the resultant image are cut out by the segmented blob masks. Afterwards, the cumulative frequency histogram of this image is constructed and then the slopes of the cumulative frequency curve at the 25th and 37th percentile² are calculated to give the first two measures $C1$ and $C2$. Features $C3$ and $C4$ are computed from the area-Voronoi diagram and the distance transform image used in its construction (Figure 4.7a and 4.7b) as follows. Firstly for every pair of Voronoi regions that share a common edge, the convex hull is computed for the union of the corresponding pairs of particles. Secondly the union of these convex hulls is filled to form a binary mask (Figure 4.7d). Thirdly this mask is intersected with the divide lines of the area-Voronoi diagram (Figure 4.7e). Finally features $C3$ and $C4$ are the mean and standard deviation, respectively, of the distance transform values corresponding to the residual divide line pixels in this intersection (Figure 4.7f).

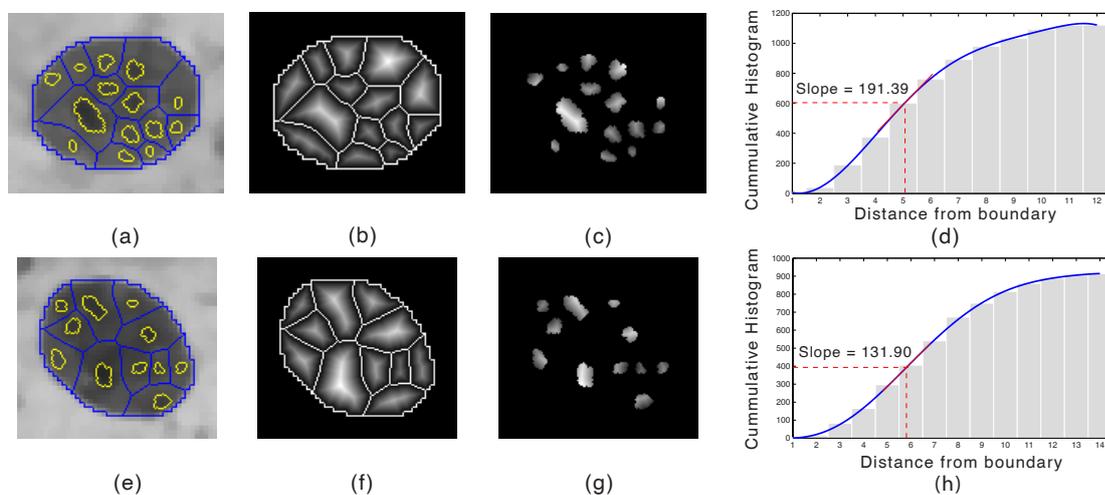


Figure 4.6: Illustration of how the features $C1$ and $C2$ are computed. (a),(e) Area-Voronoi diagram and segmented chromatin particles. (b),(f) Distance transform of the Voronoi regions. (c),(g) Portions of (b),(f) cut out by the segmented chromatin particle masks. (d),(h) Slope of the cumulative frequency curve of the distances in (c),(g) at the 37th percentile.

²See footnote 1.

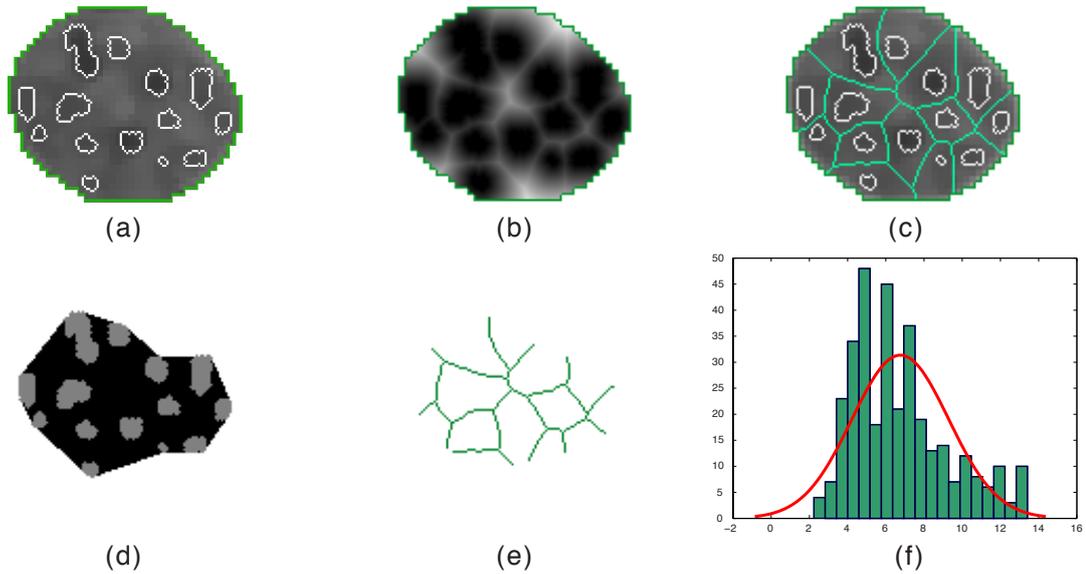


Figure 4.7: Illustration of how the features $C3$ and $C4$ are computed. (a) Segmented chromatin particles. (b) Distance transform of the set complement of the union of the chromatin particle masks. (c) Area-Voronoi diagram. (d) Binary mask (with particles superimposed) formed from the union of the convex hulls (see the text for details). (e) Voronoi region divide lines clipped according to the mask in (d). (f) Frequency histogram of the distance transform values corresponding to the residual divide line pixels in (e).

4.2.4 Contextual chromatin particle features

Contextual features provide information based on object-object or object-scene interactions. The information related to the segmented chromatin blobs can be integrated with the higher level contextual information. For example, the area ratio of each chromatin particle to its watershed ZOI captures information about object area as well as background contextual information. Table 4.1 lists the contextual chromatin particle features measured in this study.

4.2.5 Statistics of chromatin particle morphometric features

Morphometric features describe the geometry (shape, size, position and boundary) of the chromatin blob and are computed from its binary mask. By computing the mean and standard deviation of the extracted features for all the blobs inside a nucleus, a structural texture measure for the nucleus is obtained.

Table 4.1: Summary of the extracted nucleus features

Morphometric Features	
Area, perimeter, mean-radius, elliptic variance [53]	
Densitometric Features	
Photometric features [54]: Integrated optical density (<i>IOD</i>), mean optical density (<i>MOD</i>), variance of optical density, skewness of optical density, kurtosis of optical density	
Texture Features	
Statistical	Fractal texture features [42]: Fractal area 1, fractal area 2, fractal dimension
	Run length features [55]: SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE
	Histogram based features [42]: Mean, standard deviation, and skewness of the grey-level histogram
	GLCM features [54]: Contrast, correlation, energy, entropy, local homogeneity, maximum probability, cluster shade, cluster prominence and information measure of correlation (<i>H-IMC1</i> , <i>H-IMC2</i>)
	Complex Daubechies wavelet features of nuclei [56]: Mean, standard deviation, and skewness of the grey-level histogram. GLCM features – contrast, correlation, energy, entropy, local homogeneity, maximum probability, cluster shade, cluster prominence and information measure of correlation (<i>W-IMC1</i> , <i>W-IMC2</i>)
Statistics of the chromatin particle morphometric features	
Mean of chromatin particle areas and perimeters, mean of chromatin particle areas normalized to the nucleus area, mean of chromatin particle perimeters normalized to the nucleus perimeter, chromatin particle compactness (P2A)	
Contextual chromatin particle features	
Structural	Margination: $M_1, M_2, M_3, M_4, M_5, M_6, M_7$ (See section 4.2.2) Clustering: C_1, C_2, C_3, C_4 (See section 4.2.3)
	Blob features: Heterogeneity (the ratio between the area of the segmented dark and light regions, and the nucleus area) [52], the ratio between the area of the segmented dark regions and the nucleus area, number of segmented dark particles, number of segmented light particles, area ratio of each chromatin particle to its watershed ZOI, average distance between the geometrical center of the nucleus and pixels of all chromatin particles, distance between the geometrical center of the nucleus and center of mass of the chromatin particles
	Discrete texture features [42, 57]: Medium and high DNA amount, medium and high DNA area, medium and high DNA compactness, medium-high DNA compactness, center of gravity (the distance from the geometrical center of the blob to the center of gravity of the optical density function)
	Area-Voronoi diagram: Mean and standard deviation (SD) of the areas of the Voronoi regions, mean and SD of the areas of the Voronoi regions normalized to the nucleus area, area disorder [55], Voronoi regions roundness [55], mean of area ratio of each chromatin particle to its Voronoi region [55]
	Delaunay graph and generalised Delaunay graph: Mean of the Delaunay triangle areas, mean and SD of the Delaunay edge lengths, average of the mean and SD of the edge lengths connected to each vertex (chromatin particle), maximum Delaunay edge length, mean of the number of connections per chromatin particle, number of vertices (chromatin particles) on the graph boundary
	Gabriel graph and relative neighbourhood graph: Mean and SD of the graph edge lengths, maximum edge length, average of the mean and SD of the graph edge lengths connected to each vertex (chromatin blob), mean of the number of connections per vertex (chromatin blob)
	Minimum spanning tree: Mean and SD of the edge lengths, total edge length, edge disorder [55], minimum to maximum edge ratio, mean of the number of connections per vertex, percentage of vertices connected to one vertex (<i>MST1</i>), percentage of vertices connected to two vertices (<i>MST2</i>), percentage of vertices connected to more than two vertices (<i>MST3</i>)

4.3 Summary

The chapter presented a set of novel structural texture features for quantifying nuclear chromatin patterns in cells on a conventional Pap smear. The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. Several new features were introduced to quantify the qualitative description of chromatin used by cytoprofessionals, such as margination and clustering, using an structural approach to texture analysis requiring an initial segmentation of chromatin.

5

Evaluation of the proposed texture features for screening Pap smear slides

THIS chapter presents an empirical evaluation of the performance of the structural chromatin texture features proposed in the previous chapter. In particular two experiments, using clinical Pap smears sourced from the department of pathology, Regional Cancer Centre (RCC), Thiruvananthapuram in India, are detailed. The first is a comprehensive feature selection experiment that sought to determine the most discriminatory subset of features, from among the proposed features and features drawn from the literature, for discriminating between normal and abnormal smears. The second is a classification experiment that sought to evaluate the performance of a variety of classifiers, built using the feature sets obtained in the first experiment, for discriminating between normal and abnormal slides on the basis of only normal-appearing cells.

The remainder of this chapter is organized as follows. Section 5.1 presents the image data used in the experiments. Section 5.2 details the feature selection methods considered in this study. Section 5.3 details the first experiment carried out to obtain the most discriminatory subset of features. Section 5.4 then details the second experiment.

5.1 Pap smear images and ground truth

The image data used in this study originate from a set of 68 conventional Pap smear slides sourced from the Regional Cancer Centre (RCC), Thiruvananthapuram in India. Each slide was reviewed by a cytopathologist and assigned a cytological diagnosis according to the Bethesda system [58]. The cytopathologist subsequently acquired representative FOVs from each slide. In the case of abnormal smears this included FOVs with and without diagnostic cells. The cytopathologist also labelled individual cells in each FOV

accordingly. The FOV images were acquired using a monochrome CCD camera mounted on a light microscope with a $40\times$ objective lens with a numerical aperture of 0.95. Each image has a grey-scale resolution of 8 bits per pixel and is of size 1024×1344 pixels with square pixels of size $0.25\mu\text{m}$.

5.2 Feature selection methods considered in this study

5.2.1 The curse of dimensionality and the need for feature selection

The first step in designing a classifier is to have a dataset for training and evaluating the performance of the classifier. It is unequivocal that this data must contain representative observations from each class, and that the sample size determines the number of features required to build a classifier to discriminate between different classes [59]. As illustrated earlier in the chapter 4, a wide range of features can be computed and those can be used for building the classifier. However, increasing the number of features to build a classifier tends to increase the misclassification error. In addition, the prediction variability increases and the classifier becomes very sensitive to the outliers. Hence, there would be no guarantee for the designed classifier to have roughly the same classification performance on a new data set [59].

It is important to point out that building a classifier with too many features on a too small training dataset can lead to a “perfect” classification performance on the training dataset, but very poor performance on unseen test data.

Feature selection is a dimensionality reduction process, which aims to select an optimum subset of features from the original potentially discriminating set of features. There are two main reasons for using feature selection. First of all, irrelevant and redundant features are detrimental for machine learning algorithms. For a given fixed number of training samples, the predictive power of the classifier decreases as the dimensionality increases. Feature selection techniques are intended to avoid the curse of dimensionality by removing irrelevant and redundant features. Secondly, feature selection techniques greatly reduce the measurement and computational cost of classifying high-dimensional data [59].

The appropriate selection of relevant feature subset makes it possible to achieve excellent performance on classification of malignancy. In this study, the following three feature selection methods were investigated: the state-of-the-art multiple support vector machines with recursive feature elimination (MSVM-RFE) [60], $L1$ -regularization path for generalised linear models [61], and the guided regularized random forest (GRRF) [62].

5.2.2 MSVM-RFE

The feature extraction technique utilising Support Vector Machine based on recursive feature elimination (RFE) was first proposed by Guyon [63]. It returns a ranking of the features of a classification problem by training a SVM with a linear kernel and removing the feature with smallest ranking score.

The criterion used by the SVM-RFE is the weight magnitude and it is related to each feature's support to the discrimination function. The recursive feature elimination algorithm, recursively identifies the features with the smallest weights in magnitude and removes them.

Multiple SVM-RFE [60] is another feature selection method that uses backward feature elimination procedure similar to SVM-RFE. However, instead of using a single linear SVM at each recursive step, multiple linear SVMs are trained on sub-samples of the training data obtained from multiple runs of k -fold cross validation. The feature ranking score is then computed from statistical analysis of the weight vectors of the multiple linear SVMs. Similar to the SVM-RFE method, the feature with the smallest ranking score is omitted at each step.

5.2.3 Guided Regularized Random Forest (GRRF)

Decision trees are increasingly used for the purpose of the feature selection. Bagged random trees known as Random forest (RF) have also been widely utilised for measuring the feature importance. Those feature importance scores can be used in the feature selection process to shortlist the most discriminating features and taking into account only the features with high importance scores. The feature shortlisting procedure consists of multiple iterations. The feature with the lowest importance score will be eliminated from feature set per iteration. The number of features to be eliminated per iteration can be increased to make the algorithm more time and cost efficient (the authors implemented the algorithm to eliminate 20% of the features per iteration) [64].

The main drawback of the ordinary random forest is that at every node, a feature that optimizes an information-theoretic criterion will be selected to split the data. However, the redundancy of that feature to the features selected in the previous splits is not considered in this procedure.

A regularization framework can address the issue and can avoid selecting a new feature for data splitting at each node when that feature produces a similar gain (importance scores) to the feature set that has been selected in the previous splits. The RF gives an importance score for each feature, though it does not end up in a feature subset. Hence, the regularization framework can be easily added to the RF to ease the feature subset selection. This implementation of the regularization framework on random forest known as regularized random forest (RRF) [62] has been recently proposed for the feature selection purposes. Nevertheless, evaluations of the features are usually done by a portion of the training data at each tree node, which might cause in-stability in feature selection procedure. To overcome this problem, an enhanced version of the RRF, referred to as guided RRF has been proposed by Deng [62]. In this method, first the importance scores for each variable from an ordinary random forest is computed. Then those scores are used to guide the feature selection (in this case RRF). Given that the importance scores in ordinary RF are computed on all of the trees and also on all of the training data, the GRRF is likely to perform more efficiently in comparison to RRF in terms of feature selection performance.

5.2.4 L1-regularization procedure with generalised linear model

Generalised linear models (GLM) generalise the classical linear models based on the normal distribution. In an ordinary linear model the relation between the response variable Y_i and the independent predictor variables X_{ij} is given by:

$$\eta_i = Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ij} \quad (5.1)$$

where η is the linear predictor and β s are the coefficients of the linear combination whose values are unknown and have to be estimated. The coefficient β_j is the measure of the impact of the predictor X_{ij} on the target Y_i . In an ordinary linear model, for each set of values for the predictors, the response variable is continuous and normally distributed [65].

In a generalised linear model, by contrast, response variables belong to the exponential family of distributions, such as the Gaussian, Binomial, Poisson, Gamma, or inverse Gaussian. In addition, the relation between the response and predictor variables does not necessarily have the simple linear form as in equation 5.1. In fact, the independent predictive variables can take non-linear transformations of the other predictor variables. However, the linear predictor is still linear in the coefficients (parameters) that need to be estimated.

GLMs model the random variable Y by providing a relationship between its expected value and the linear predictor η .

$$\eta_i = g(\mu) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ij} \quad (5.2)$$

where g is called the link function. The link function relates the linear predictor η to the expected value of the response variable. The link function for a classical linear model is the identity function $g(\mu) = \mu = E(Y)$. Examples of link functions for GLMs include the identity, log, reciprocal, logit and probit.

The next step after selecting a model is to estimate its parameters and to assess the precision of the estimates. The $L1$ -regularization procedure uses maximum likelihood method for estimating the parameters of GLM (parameters β in the linear predictor η), with a penalization on the size of the $L1$ -norm of the coefficients. The $L1$ regularization procedure selects variables according to the amount of penalization on the $L1$ -norm of the coefficients. To achieve this, the algorithm makes use of the predictor-corrector method to compute solutions along the entire path of the coefficient estimates [61].

5.3 Experiment I: Determination of the most discriminatory subset of features for the classification of Pap smears

The aim of this experiment was to determine the most discriminatory subset of features, from amongst the proposed structural chromatin texture features and a wide range of

features drawn from the literature, that have the most discriminating power for discriminating between normal and abnormal Pap smears. A normal smear is deemed to be one that has been labeled as NILM (Negative for Intraepithelial Lesion for Malignancy). An abnormal smear is one that has been labeled as either LSIL (Low grade squamous intraepithelial lesion) or HSIL (High grade squamous intraepithelial lesion).

5.3.1 Cytology slides for experiment I

As a compromise between having a reasonable number of cells per slide for MAC analysis and having a reasonable number of slides for the feature extraction and feature selection procedure, it was decided to impose the following two criteria. Firstly each slide should have a minimum of 80 non-superficial cervical cells with a Negative for Intraepithelial Lesion or Malignancy (NILM) label. Secondly, 10% to 40% of the cells in the abnormal slides should be diagnostic cells. This left a total of 44 slides: 25 NILM, 10 LSIL, and 9 HSIL.

5.3.2 Feature extraction

The free-lying nuclei in each field of view were automatically segmented using the segmentation algorithm proposed in chapter 3. Then, 159 features¹ were computed for each nucleus (see Table 4.1). Figure 5.1 shows the breakdown of the types of features investigated (see Appendix C). These features include the proposed structural features by the authors and other competing cytology features available in the literature for the purpose of detecting MACs (see Appendix C).

Statistical texture features used in this study include fractal texture features [42], run-length features [55], grey-level co-occurrence matrix (GLCM) features [54], and complex Daubechies wavelets features [56]. The wavelet statistical and co-occurrence features of nuclei were computed from decomposed images of the segmented nuclei by Daubechies wavelets transform at first level of decomposition. However, instead of using a window varying in size for each level of decomposition, a down sampled binary mask of the original image was used to identify and retain only the nuclei regions in the decomposed images for the purpose of feature extraction. Details of the extracted features can be found in Table 4.1.

The mean and standard deviation of each nucleus feature was computed for each slide yielding a total of 318 slide-based features. Finally, an additional random feature was included to gauge the efficacy of the feature selection (the feature does not have any discriminatory power).

¹Contextual features were excluded for nuclei, and textural and densitometric features were excluded for chromatin blobs.



Figure 5.1: Pie chart showing the contribution of the four main object feature categories in this study.

5.3.3 Feature selection and classification

In this research work, the MSVM-RFE, GRRF and L1-regularization procedure with generalised linear model feature selection methodologies have been investigated on the data. Also, a 5-fold cross validation procedure has been implemented to optimize the parameters and to decrease the bias in feature selection as much as possible. The CV loops have been iterated 100 times and the feature selection performance evaluation has been averaged over all these iterations.

In order to optimize the classification performance it is necessary to couple each classification approach with the most relevant feature selection method. In this experiment we paired the SVM-RFE with a linear kernel SVM classifier, the GRRF with an ordinary RRF classifier, and the L1-regularization procedure with a generalised linear model. Given the limited number of slides available in this study, all the slides were used for feature selection. However an internal cross validation (in-loop feature selection) technique[66] was employed to avoid overfitting and over-optimistic estimation of the performance result. Feature selection by the MSVM-RFE was achieved using 100 iterations of double 5-fold cross validation, each iteration with random combinations of samples in the training and the test set. The inner cross validation loop was used for feature selection and tuning the optimal parameters for the SVM classifier with a linear kernel. The outer cross validation loop was used to evaluate the classifier model performance in terms of the AUC (area under the receiver operating characteristic curve). For GRRF feature selection and the L1-regularization path for generalised linear models, 100 repeated 5-fold cross validation was employed. Figure 5.2 illustrates the use of 100 iterations of 5-fold cross validation for in-loop feature selection and classification.

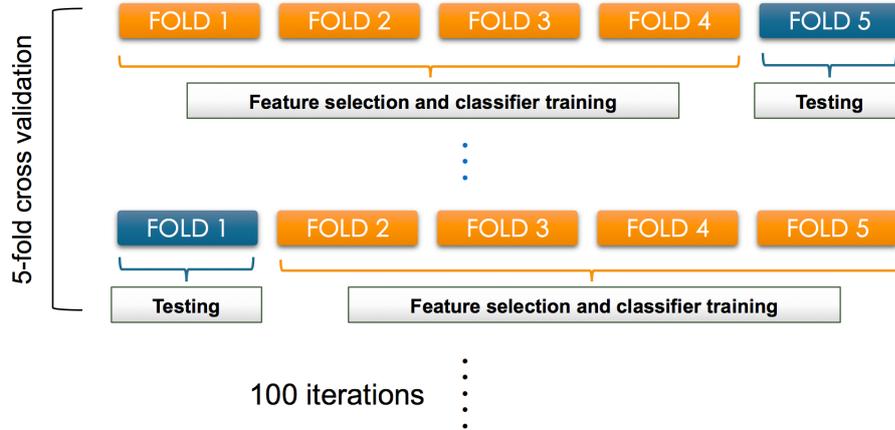


Figure 5.2: Illustration of 100 times iterated 5-fold cross validation used for feature selection.

5.3.4 Results of experiment I

This section provides the classification performance results together with the four top-ranked features for each of the three feature selection approaches mentioned in section 5.2 and used in the experiment I.

Table 5.1 shows the frequency with which each feature appeared in the best feature subset of size three in the 100 iterations of 5-fold CV by SVM-RFE and GRRF feature selection methods. The number of times a variable has been selected in 100 iterations of 5-fold CV in a fitted generalised linear model is also shown in this Table.

Table 5.1: Frequency table for top 5 ranking features in experiment I.

SVM-RFE	Feature	$F1$	$F2$	$F4$	$F3$	$F7$
	Frequency	477	190	184	120	66
GRRF	Feature	$F1$	$F2$	$F5$	$F6$	$F8$
	Frequency	500	144	87	38	66
L1-Reg	Feature	$F1$	$F6$	$F2$	$F5$	$F9$
	Frequency	500	418	398	338	334
<p>$F1$: SD of IOD, $F2$: Mean of $M2$, $F3$: SD of $C1$, $F4$: SD of nucleus area, $F5$: SD of $MST3$, $F6$: Mean of $W-IMC1$, $F7$: SD of $C2$, $F8$: Mean of $M1$, $F9$: SD of Heterogeneity</p>						

The classification results are presented in Table 5.2. Features “Standard deviation of IOD ” and “Mean of $M2$ ” are the among the top ranked features in all three feature

selection (FS) methods. The top four features appearing most frequently among the feature selection methods are: SD of IOD , Mean of $M2$, SD of $MST3$ and Mean of $W-IMC1$. Figure 5.3 shows boxplots for each of these features for both the normal and abnormal slides. The best AUC (0.954 ± 0.019) was achieved using the $L1$ -regularization procedure with a generalised linear model.

Table 5.2: Classification results for experiment I.

FS Method	Top 4 features	Classifier	AUC (Mean \pm SE)
MSVM-RFE	$F1, F2, F3, F4$	Linear SVM	0.934 ± 0.038
$L1$ -Regularization Path	$F1, F2, F5, F6$	GLM	0.954 ± 0.019
GRRF	$F1, F2, F5, F6$	RRF	0.935 ± 0.029

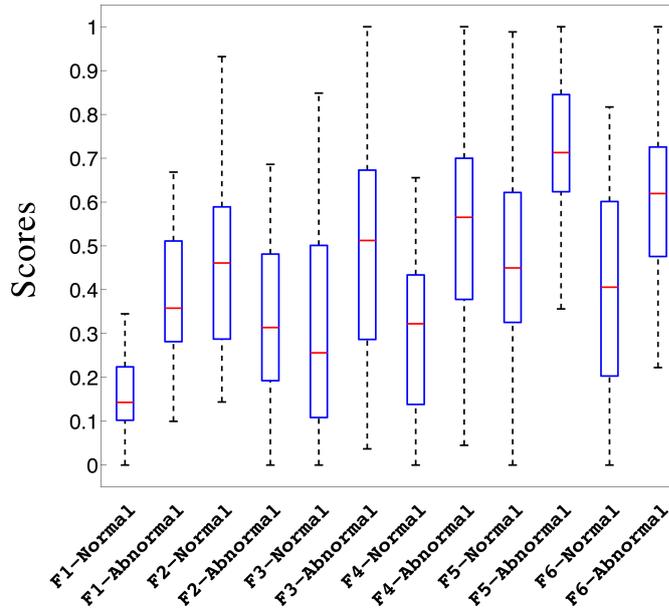


Figure 5.3: Boxplots of the features $F1, F2, F3, F4, F5, F6$ for experiment 1.

The feature subset selected in this experiment will be utilised to model and generate a MAC-based classifier in experiment II.

5.4 Experiment II: MAC-based classifiers for Pap smear screening

The aim of this experiment was to evaluate the performance of a variety of classifiers, built using the feature sets obtained in the first experiment, for discriminating between

normal and abnormal slides on the basis of only normal-appearing cells. Although, the diagnostic (LSIL and HSIL) cells were not excluded from the slides used for the feature selection in the first experiment, they have been excluded from the data in the second experiment. This makes sense, because in building a classifier specifically designed to detect MACs, it is crucial to generate a database of cell nuclei including all nuclei from normal slides and only normal-looking cell nuclei from abnormal slides.

The method for generating the database needed for the second experiment together with a cell nuclei count constraint applied to the slides for acquiring the proper slides for the MACs analysis is detailed in the following section.

5.4.1 Cytology slides for experiment II

In this experiment only the cells marked by the cytologist as NILM on normal and abnormal slides were segmented and slides with a minimum of 80 such cells retained. This left a total of 44 slides: 26 NILM, 9 LSIL, and 9 HSIL. Figure 5.4 shows an example of a digitalized Pap smear FOV. All the NILM cell nuclei marked by cytopatologist are segmented, though the 4 diagnostic cell nuclei are excluded from the FOV. After selecting the suitable slides for the MAC-based classification, a variety of classifiers were built based on the best feature subset obtained in experiment I and the classification performance of each in detecting MACs is empirically evaluated.

5.4.2 Classification performance evaluation of the best subset of features derived in experiment I

Among the six top-ranked features, ($F1$ to $F6$), having the most discriminating power, identified in the experiment I, the features having the minimum of two votes between three feature selection methods are considered for the experiment II.

Prior to the classifier design, the number of features for building the classifier must be carefully determined to avoid the curse of dimensionality. For this purpose, only feature subsets of size one or two were considered for the experiment two. Performance of two classifiers were evaluated using 100 iterations of 5-fold cross validation to better evaluate the discriminatory power of the best subset of features to detect the MACs. The built and studied classifiers in the experiment II are as follows:

- SVM with a linear kernel
- Logistic Regression (LR)

See Appendix F for details of the two classifiers. These classifiers were built using the best subset of features obtained in experiment I. Performance of two classifiers were evaluated using 100 iterations of 5-fold cross validation. The results of experiment II is depicted in the following section.

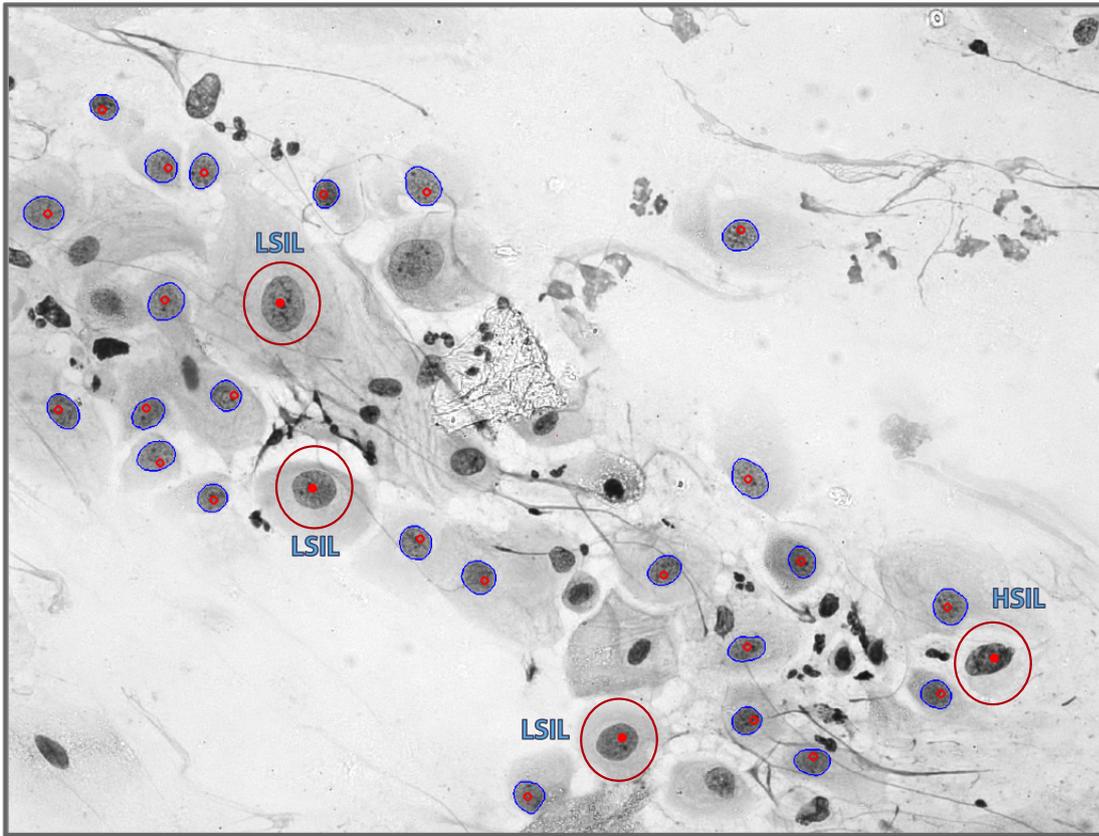


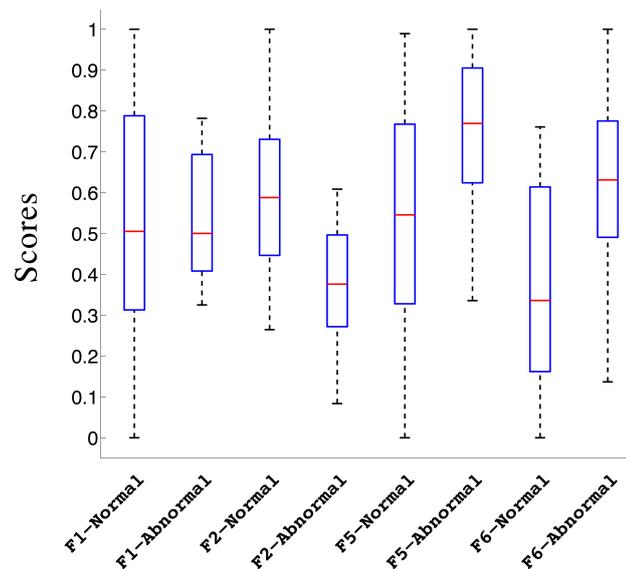
Figure 5.4: An example of a digitalized Pap smear FOV, in which the diagnostic cells were excluded from the segmentation process and only the NILM labelled cell nuclei remained and used for the second experiment.

5.4.3 Results of experiment II

This section presents the classification performance results for the feature subsets of size 1 and 2 obtained from experiment II. These results for each classifier are depicted in Table 5.3. Boxplots for each feature for both the normal and abnormal slides are shown in Figure 5.5. The feature “Mean of $M2$ ”, which is a measure of chromatin margination inside the nucleus, achieved the best classification performance in both classifiers for the detection of MACs. The feature “SD of IOD ”, whilst quite discriminatory in the first experiment, does not show any discriminatory power in the second experiment. This is not surprising given that the Pap stain is not stoichiometric and that none of the cells in experiment 2 were diagnostic. The non-stoichiometric property means that the amount of the stain uptake does not necessarily correspond to the DNA amount present in the cell nuclei. This is also supported visually in Figure 5.5, in which the boxplot scores for feature “SD of IOD ” show a complete overlap of the scores between normal and abnormal slides.

Table 5.3: Classification results for experiment II.

Classifier	Features	AUC (Mean±SE)
SVM	$F2$	0.815 ± 0.015
	$F5$	0.706 ± 0.053
	$F6$	0.792 ± 0.026
	$F2, F6$	0.822 ± 0.024
Logistic Regression	$F2$	0.803 ± 0.026
	$F5$	0.591 ± 0.054
	$F6$	0.784 ± 0.027
	$F2, F6$	0.752 ± 0.038

**Figure 5.5:** Boxplots of the features $F1, F2, F5, F6$ for experiment 2. The relatively high overlap of feature $F1$ scores between normal and abnormal slides depicts the low discrimination power of $F1$ for dataset in experiment II.

5.5 Summary

The chapter presented an empirical evaluation of the performance of the proposed structural chromatin texture features in the Chapter 4. Overall, 159 features were computed for each nucleus including the proposed features and a wide range of features drawn from the literature. Two experiments were carried out to evaluate the discrimination

power of those features for discriminating between normal and abnormal smears. The results provided empirical evidence that it is possible to detect malignancy-associated changes in Pap smears. The high performance of the classifiers for the detection of MACs compares favourably with those reported in the literature for other stains.

6

Summary and Conclusion

THIS chapter reviews the work presented in this thesis, summarizes the major contributions and findings, outlines the limitations of the research undertaken, and finally proposes an avenue of future work.

6.1 Thesis summary

Chapter 1: This chapter provided a background to the project and illustrated the problem with the conventional method of screening, the Papanicolaou test. It was noted that existing research suggests that the MACs phenomenon may be the solution to the problem of sampling error in conventional Pap screening and that automated screening can address interpretation errors associated with manual screening. The chapter states the major aim of the thesis:

.. to fully explore the structural approach to chromatin pattern description and to evaluate the efficacy of the features derived from it for discriminating between normal and abnormal Pap slides.

and its objectives:

1. To develop a robust algorithm for detecting and segmenting cell nuclei in digitized Pap smear images obtained using bright-field microscopy;
2. To develop structural texture features that quantitatively characterise the pattern (arrangement, size, shape, etc.) of the nuclear chromatin;
3. To determine the most discriminatory subset of features for discriminating between normal and abnormal slides using real clinical data; and

4. To evaluate the performance of a classifier(s), based on the selected features, using real clinical data.

Chapter 2: This chapter presented two literature reviews pertinent to chapter 3 and chapter 4. The first was a review of the cell and cell nucleus segmentation methods in Pap smear images published in the literature. The second presented a detailed review of the standard texture features available in the literature devised to quantify chromatin texture/distribution.

Chapter 3: This chapter presented the new algorithm for robustly detecting and segmenting free-lying intermediate cell nuclei in bright-field microscope images of Pap smears. The novelty of the algorithm stems from a robust marker selection method for selecting candidate free-lying nuclei-like objects for subsequent marker-controlled watershed segmentation to obtain the nuclear boundaries. The algorithm also implements artefact rejection based on size, shape, and nuclear granularity to ensure only the nuclei of intermediate squamous epithelial cells are retained. In addition, this chapter presented an empirical evaluation of the performance of the algorithm and discussed the results. The sensitivity and specificity of the algorithm to the detection of free-lying intermediate cell nuclei together with the accuracy of the segmentation of each nucleus detected by the algorithm were presented. Finally, the algorithm's failure modes were detailed and the summary and conclusion of the chapter were presented.

Chapter 4: This chapter dealt with the problem of quantitative characterisation of chromatin texture and presents a set of novel structural texture features to describe nuclear chromatin patterns in cells on a conventional Pap smear. These features derived from a segmentation of the chromatin into blob-like primitives.

Chapter 5: This chapter presents an the evaluation of the performance of the proposed structural chromatin texture features in the previous chapter. Overall, 159 features were computed for each nucleus including the proposed features and a wide range of features drawn from the literature. Two experiments were carried out for evaluating the features. The aim of the first experiment was to determine the most discriminatory subset of features for discriminating between normal and abnormal Pap smear slides. The aim of the second study was to evaluate the performance of a variety of classifiers built using the feature subset obtained in the first experiment to discriminate between the normal and abnormal slides. The results provided empirical evidence that it is possible to detect malignancy associated changes in Pap smears. The high performance of the classifiers for the detection of MACs compares favourably with those reported in the literature for other stains.

6.2 Key contributions and findings

- A new automated algorithm for robustly detecting and segmenting free-lying cell nuclei in bright-field microscope images of Pap smears. The proposed segmentation

algorithm makes use of gray-scale annular closings to identify free-lying nuclei-like objects together with marker-based watershed segmentation to accurately delineate the nuclear boundaries. The algorithm also employs artefact rejection based on size, shape, and granularity to ensure only the nuclei of intermediate squamous epithelial cells are retained. The empirical results show that the algorithm has both a high sensitivity and specificity for the detection of free-lying nuclei, and that the algorithm is able to delineate the boundaries of these with high accuracy. The achieved results are remarkable and encouraging for the development of an automated screener for Pap smears based on the MAC phenomenon.

- A set of novel structural texture features for quantifying chromatin texture. The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. They fall into two broad categories. The first consists of statistics of morphometric features computed for individual blob. The second consists of contextual features which can be further subdivided into the following classes of features: margination, clustering, blob features, discrete texture features, and features derived from the Voronoi diagram and its associated graphs. The most prominent of the proposed features were margination and clustering.
- Empirical evidence that it is possible to detect malignancy-associated changes (MACs) in Papanicolaou stain which in turn suggests the possibility of developing a fully automated Pap smear screener based on MACs. The results of a classification experiment, using only normal-appearing cells from both normal and abnormal slides, demonstrate that a single structural texture feature measuring chromatin margination yields a classification performance of 0.815 ± 0.019 . The result compares favourably with the experimental results published in the literature for other stains developed for automated cytology.

6.3 Conclusion

An automated algorithm for robustly detecting and segmenting free-lying cell nuclei in bright-field microscope images of Pap smears was presented. This is an essential initial step in the development of an automated screening system for cervical cancer based on malignancy associated change (MAC) analysis. The empirical results provided in chapter 3 demonstrate that the algorithm has a high sensitivity and specificity for the detection of free-lying nuclei, and that the algorithm is able to delineate the boundaries of these with high accuracy.

A set of novel structural texture features for quantifying nuclear chromatin patterns in cells on a conventional Pap smear was presented. The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. The experimental results of chapter 5 demonstrate the efficacy of the structural approach to chromatin texture analysis and that a combination of these structural texture features and conventional features can be used to discriminate between normal (NILM) and abnormal (LSIL and HSIL) slides with high accuracy. The designed MAC classifier outperforms

the classifiers reported in the literature. However, the majority of the studies for the detection of MACs in cervical smears used stoichiometric stains. Stoichiometric stains, whilst quite useful for cyto-diagnostic purposes, are not the stain of choice for cervical cancer screening programs. Papanicolaou stain is still the most commonly used staining technique and therefore developing an automated conventional Pap smear screener will provide significant advantages. The results of the present study demonstrate that it is possible to detect malignancy-associated changes in Papanicolaou stain. This concurs with the study by Mehnert [20]. The most discriminating single feature, Mean of $M2$, for detecting malignancy-associated changes is a structural feature measuring chromatin margination. This in turn suggests the possibility of developing a fully automated Pap smear screener based on MACs.

6.4 Limitations

Brightfield microscopy suffers from limited depth of field. This shortcoming produces images where portions of the field of view image that lie within the depth of focus are in-focus whereas the remaining portions are out of focus [67]. Many methods have been proposed for obtaining an extended depth of field (EDF) image from the “z-stack” of images that collectively contain all in-focus information of the specimens in a field of view.

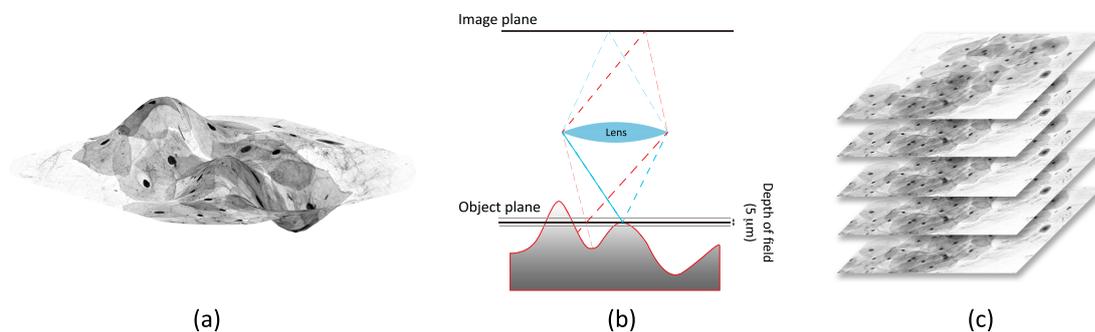


Figure 6.1: Focus stacking for extended depth of field in optical microscopy. (a) Topography of the specimens on the sample slide has wide variation in z -position. (b) Due to limitations in the extended depth of focus of the microscopic imaging system, only some portion of the image is in focus at each depth within the sample. (c) Multilayer scanning at different planes of the specimen on the sample slide permits one to navigate between various planes and capture z -stack images.

In this study the complex wavelet based image fusion technique [68] was used to obtain EDF images. However, several other alternatives exist in the literature that outperform the implemented method in our study. Recent approaches to EDF produce results with significant improvements in terms of quality and sharpness and artefact suppression [67, 69, 70].

Although all the cells in the second experiment are not necessarily included in the first

experiment, there is a large overlap between the cells used in these two studies. Ideally, the efficacy of the proposed structural texture features should have been assessed on an independent data set. Therefore, a much larger number of Pap smear slides are required to make conclusive claims about the feasibility of developing a fully automated Pap smear screener that works based on MACs.

The total number of slides used in this study was limited to 68 because of practical limitations of data collection. The literature suggests that at least 500 nuclei are necessary to achieve a relatively constant MAC determination [21]. Unfortunately some of the collected slides have hardly more than 100 nuclei (minimum of 15 and maximum of 700 nuclei).

The evaluation of the nucleus segmentation algorithm proposed in this thesis was made relative to two collected ground-truth data sets which were limited to 33 FOVs randomly selected from 11 slides. The use such limited number of FOVs was due to limitations in recruiting image analysis experts. Ideally, several independent cytopathologists should have selected and delineated each cell nuclei.

6.5 Opportunities for further research

Numerous possibilities remain for developing improved chromatin segmentation algorithms and developing features that may offer improved classification performance for the detection of MACs.

- **New segmentation approaches**

Developing a fully automated mean-shift filtering algorithm for segmentation of the chromatin blobs in cervical cell nuclei.

To date, the mean-shift algorithm has been widely used for image segmentation in different biomedical applications. In the case of cervical cell segmentation, Edwards et al. [71] used mean-shift filtering to smoothen the cervical images prior to segmentation by an adaptive threshold algorithm. Mean shift has also been used by Bell et al. [72] for the segmentation of sub-cellular structures, like the nucleolar organizer regions (AgNORs) in silver stained cytopathology images. However, the mean shift segmentation approach has not yet been investigated for the segmentation of the nuclear chromatin into its blob-like primitives in Pap smears. This could be a subject the future work.

A more thorough evaluation of the proposed nucleus segmentation algorithm is required and this can be achieved by having a bigger cervical smear image data set. This is the immediate future for the work presented. The sensitivity of the algorithm to the detection of free-lying intermediate cell nuclei is high. However its specificity, whilst still quite high, could be improved. The development of a more sophisticated artefact rejection strategy would enhance the specificity of the nucleus detection.

- **Ideas for new features**

Of particular interest is the extraction of new features that can express several desired properties of a cell nucleus in a single feature. As an example, the proposed features $M3$ and $M4$, as two measures of margination, implicitly include information about the size of the nucleus. In order to incorporate the information from several features (describing the property of a nucleus or a single chromatin particle) a 3D feature space can be constructed (see figure 6.2). The differences in the spatial distribution of the discriminatory information within this 3D feature space suggest the possibility of extracting several optimized features. One way to characterize the distribution of the data is to use co-occurrence based methods of texture analysis. For this purpose, feature values should be quantized to 8 or 16 discrete levels. The 3D co-occurrence matrix is a $N_i \times N_j \times N_k$ matrix where each element $p[i,j,k]$ of the matrix represents the summation of the nuclei having the feature vector values of (i,j,k) . Several features can then be extracted from this 3D co-occurrence matrix.

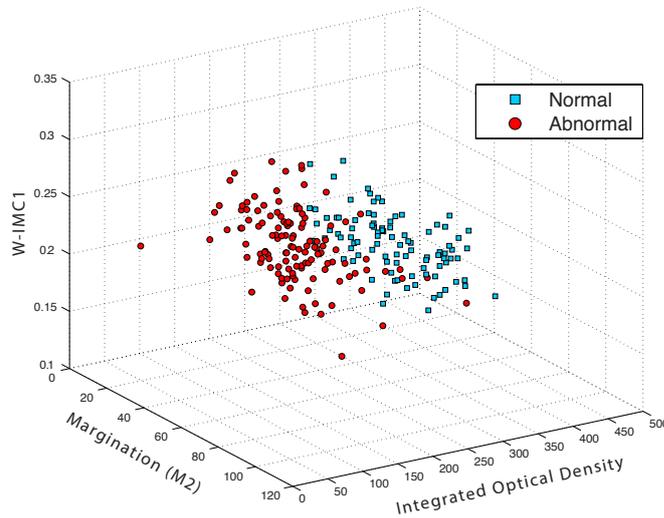


Figure 6.2: 3D feature space

Appendices

A

Fundamental operations from Mathematical Morphology

The definitions below are drawn from “Morphological image operators” by Henk Heijmans [73].

A.1 Mathematical morphology for binary images

Let \mathbb{E}^n be either \mathbb{Z}^n or \mathbb{R}^n representing the space of pixel coordinates. The dilation of a set (binary image) $X \subset \mathbb{E}^n$ by the structuring element $B \subset \mathbb{E}^n$ is defined

$$X \oplus B = \bigcup_{b \in B} X_b$$

where $X_b = \{x + b \mid x \in X\}$. The erosion of a set (binary image) X by the structuring element B is defined

$$X \ominus B = \bigcap_{b \in B} X_{-b}$$

The opening of X by B is defined

$$X \circ B = (X \ominus B) \oplus B.$$

The closing of X by B is defined

$$X \bullet B = (X \oplus B) \ominus B.$$

A.2 Mathematical morphology for grey-scale images

Let $\text{Fun}(\mathbb{E}^n, \mathcal{T})$ denote the set of all functions of the form $f : \mathbb{E}^n \rightarrow \mathcal{T}$ where \mathbb{E}^n represents the domain space of pixel coordinates (usually \mathbb{Z}^n or \mathbb{R}^n) and \mathcal{T} is either \mathbb{Z} or \mathbb{R} representing grey-levels.

Let $f, g \in \text{Fun}(\mathbb{E}^n, \mathcal{T})$. The dilation of the grey-scale image f by the structuring function g is defined point-wise

$$(f \oplus g)(x) = \max_{h \in \mathbb{E}^n} \{f(x - h) + g(h)\}.$$

The erosion of the grey-scale image f by the structuring function g is defined point-wise

$$(f \ominus g)(x) = \min_{h \in \mathbb{E}^n} \{f(x + h) - g(h)\}.$$

The opening and closing operations, respectively, of the grey-scale image f by the structuring function g are defined

$$f \circ g = (f \ominus g) \oplus g,$$

$$f \bullet g = (f \oplus g) \ominus g.$$

When the structuring function only takes on the value 0 on its domain then it suffices to represent the function by a set B (its domain). The expressions for dilation and erosion then simplify to the following:

$$f \oplus B = \max_{h \in B} \{f(x - h)\},$$

$$f \ominus B = \min_{h \in B} \{f(x + h)\}.$$

The set B is called a structuring element or flat structuring function.

B

Segmentation methods underlying methods in Chapters 2 and 3

B.1 Thresholding

One of the most simple and frequently used methods for image segmentation is thresholding. Thresholding utilizes the histogram of pixel intensities of the image. It is useful for the segmentation of images where the grey-values of the background pixels lie below or above the grey-values of the object. In other words, the object of interest is brighter or darker than the background. Therefore a threshold value, T , can be applied to the image to convert the grey-scale image to a binary image by replacing all the grey-values greater than T with the value 1 and setting all other grey-levels to 0. Thus by thresholding the original image $f(x,y)$, the segmentation label image $g(x,y)$ is obtained:

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \leq T \end{cases}$$

Figure B.1a depicts thresholding as a plane cutting the 3D image surface. As seen, the final location of region boundaries are affected by the choice of T . In this Figure, a larger T value decreases the object area and a smaller T will increase it. In the simplest case where the histogram of the image has two dominant peaks, a suitable threshold value lies somewhere between the two peaks (corresponding to foreground and background pixels)(see Figure B.1b).

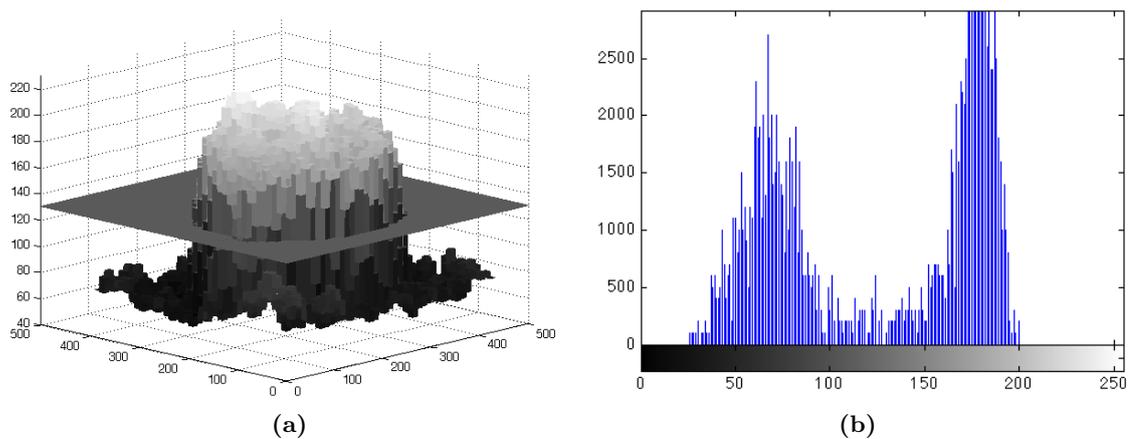


Figure B.1: Global Thresholding: (a) The grey-level image depicted as a topographic surface is globally thresholded (Topographic image is generated by interpreting the grey-values of each pixel as heights). (b) The bi-modal histogram of the image. The threshold value ($T = 130$) lies somewhere in between the two peaks

Numerous parametric and non-parametric algorithms have been proposed to determine the optimal threshold value by locating the valley in the grey-scale histogram. The methods that determine this single threshold value for the entire image are known as global thresholding [24]. These algorithms do not make use of any spatial information and lack robustness to noise and uneven illumination. To overcome these problems several local thresholding algorithms have been proposed. Local thresholding methods [74, 75, 76] compute separate thresholds for each pixel using additional information derived from the surrounding neighbourhood; e.g. Niblack [76] determined a local threshold value on the basis of the local mean and standard deviation of grey-values in the image. The grey-value of each pixel is compared with the average grey-values in some neighbourhood. If the grey-value of the pixel is significantly larger than the average it is classified as foreground, otherwise it is classified as background. The major problem with the local thresholding techniques is that they are mostly dependent on many parameters.

B.2 Active contours and Deformable models

Active contours and their associated techniques can be categorized as deformable models. These models have been extensively studied and widely used in medical image segmentation with promising results. An active contour is a planar curve (unbroken border initialized somewhere in the image) with an associated energy function. The position of the initial contour must be localized such that it roughly surrounds the object of interest. The energy function is defined such that it attains its minimum when the contour lies upon the desired object.

The snake is an active contour model introduced by Kass [77], and is able to deform

elastically. A contour is defined parametrically in the continuous domain [77] by $v(s) = (x(s), y(s))$ where $x(s)$ and $y(s)$ are the image coordinates and s is the normalized length of the contour where $s \in [0, 1]$ (see Figure B.2). The energy of the contour usually includes two separate energy terms: internal energy (E_{int}) and external energy (E_{ext}). These two energy functionals are combined to give the total energy of the contour as follows

$$E_{\text{contour}}(v(s)) = \int_{s=0}^1 [E_{\text{int}}(v(s)) + E_{\text{ext}}(v(s))] ds. \quad (\text{B.1})$$

The external energy can be divided into two energy terms as

$$E_{\text{ext}} = E_{\text{image}}(v(s)) + E_{\text{con}}(v(s)) \quad (\text{B.2})$$

where E_{image} denotes the external image energy and E_{con} denotes the external constraint energy. The energy of the contour is given by

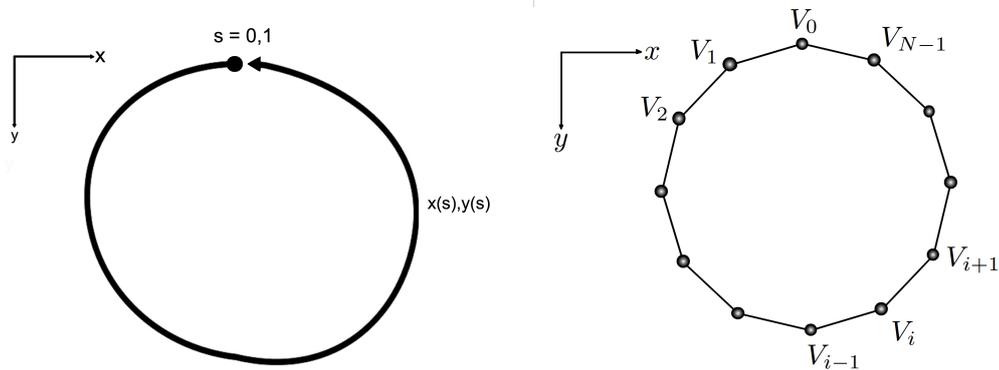
$$E_{\text{contour}}(v(s)) = \int_{s=0}^1 [E_{\text{int}}(v(s)) + E_{\text{image}}(v(s)) + E_{\text{con}}(v(s))] ds. \quad (\text{B.3})$$

The energy function in equation B.3 is a continuous description and has been implemented using B-splines [78] and finite elements [79]. Given that these methods are computationally expensive [26], methods based on describing a contour in discrete form (a set of points, that are then joined to form a closed polygon) have been used and yield similar results (see Figure B.2). The energies of the contour are then only calculated at the discrete sample points as follows

$$E_{\text{contour}}(v) = \sum_{i=0}^{N-1} [E_{\text{int}}(v_i) + E_{\text{image}}(v_i) + E_{\text{con}}(v_i)]. \quad (\text{B.4})$$

During the deformation/evolution process of the contour, the internal energy imposes smoothness and continuity constraints (it is usually defined in terms of the first and second-order derivatives of the contour). The image energy term attracts the snake to the object's boundaries (e.g. this attraction could be defined in terms of the magnitude of the image gradient in which the contour is attracted to the prominent object edges). The external constraint energy term permits an external constraint to be placed on the contour (e.g. adding spring and repulsive forces to selected points on the contour), and can be used to shift the snake from one local energy minimum to another. Figure B.3 illustrates the active contour segmentation performance on a single cell nucleus.

In order to solve the two major problems of proper initialization of contour and poor convergence to boundary concavities, an external force was introduced by Xu [80]. This external force is called gradient vector flow (GVF), and is computed based on the diffusion or gradual change of the gradient vectors of a grey-level or binary edge map derived from the image [80]. Figure B.4 shows an example of nuclei segmentation by using GVF snake. After introducing the GVF, Xu [81] also proposed a generalized form of GVF which is called generalized gradient vector flow (GGVF) to improve active



(a) Continuous contour: The contour length is normalized to $[0,1]$. The contour is described by $v(s) = (x(s),y(s))$ where x and y are the coordinates.

(b) Discrete contour: The contour is divided into N discrete points. The contour is described by $v(i) = (x(i),y(i))$ where $i = 0, \dots, N + 1$, x and y are the coordinates.

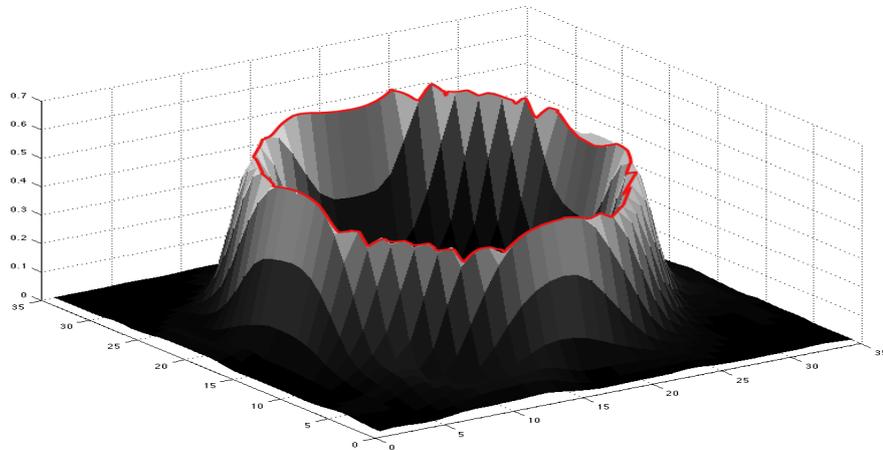
Figure B.2: Parametric representation of a contour

contour convergence to long, thin boundary indentations, but at the same time keep the desired properties of GVF (e.g. extended capture range).

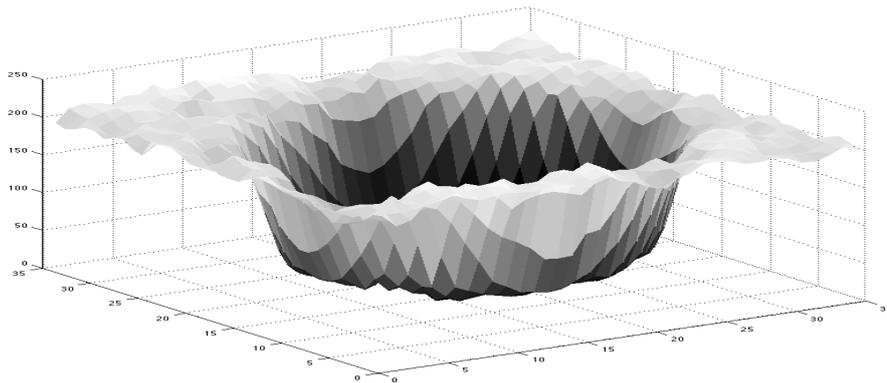
Recently, many new active contour models have been introduced in the literature. These include:

- The high contrast segmentation framework (HCS), based on variational snakes and is efficient for nuclei segmentation, in which a modified internal energy function is introduced [82].
- Distance mapping active contour, in which distance mapping is used to create a gradient vector flow [83].
- Multi-direction gradient vector flow using a new anisotropic diffusion filter for removing the noise before applying the multi-direction GVF snake [84].
- Active contours using special processing named Selective Binary and Gaussian Filtering Regularized Level Set (SBGFRLS) method [85].

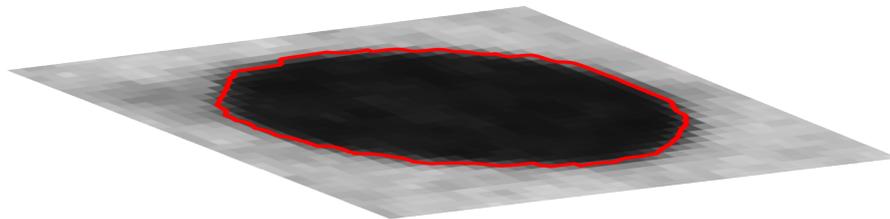
To summarise, algorithms based on active contours or deformable models are highly dependent on strong prior knowledge about the shape and location of objects in an image to guide the segmentation process. Moreover they require the initial contour to be reasonably close to the true object boundaries. As a consequence they can fail in images containing clustered and overlapping cells [27, 28, 29, 30]. These methods can also become trapped in local minima yielding the incorrect segmentation.



(a) The energy function is minimised and the contour lies upon the boundary in the magnitude of the gradient.

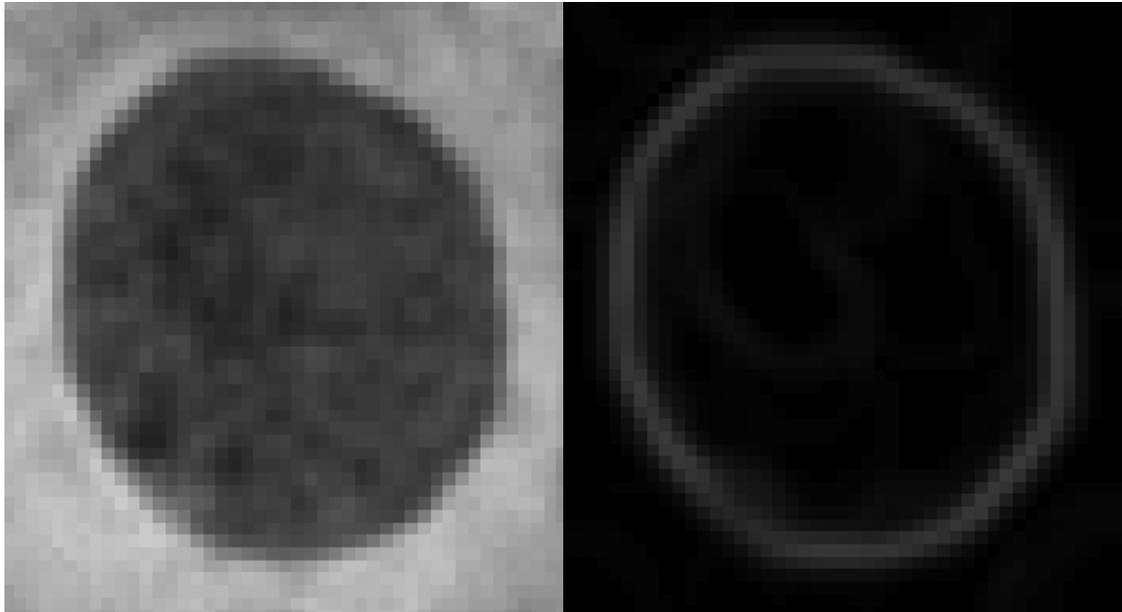


(b) Topographic map of the cell nucleus.



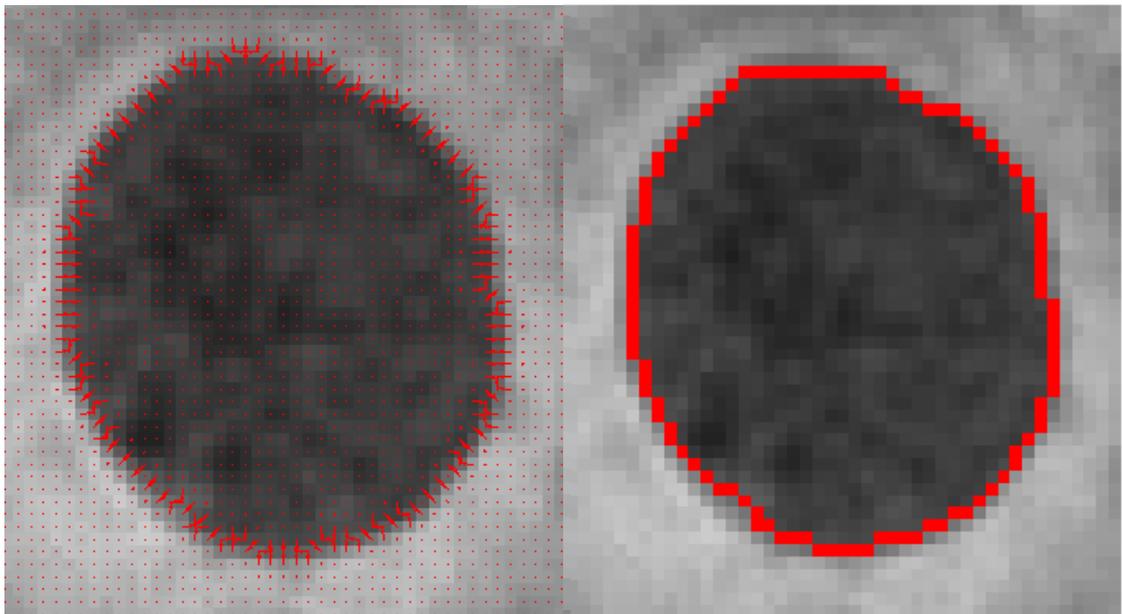
(c) Detected boundary of the nucleus obtained by the active contour (overlaid on the original image).

Figure B.3: Active contour segmentation



(a) An example of a cell nucleus.

(b) Magnitude of gradient.



(c) The GVF overlaid on the image.

(d) Segmented nucleus by active contour after 30 iterations.

Figure B.4: Cell nuclei segmentation by GVF snake.

B.3 Seeded Region Growing

The seeded region growing algorithm [86] was introduced by Adams and Bischof. It starts from a set of seed regions representing the desired image regions and uses a predefined similarity criterion to append neighboring pixels. This is continued until the entire image has been partitioned. Unfortunately, construction of a seeding method is not straightforward; in fact it is the most difficult part of the segmentation. The seeding algorithm should be able to provide exactly one seed per object. To overcome the problem Adams and Bischof [86] propose that a user (based on personal judgment) manually mark the seeds. Therefore the final segmentation results in the same number of regions as the seeds. However, it is inherently dependent on the order of pixel scanning. Mehnert and Jackway [87] introduced an improved seeded region growing algorithm that retains the advantages of the Adams and Bischof algorithm whilst being pixel order independent. The feature in common with all seeded region growing algorithms is that the final segmentation result is highly dependent on the chosen similarity criteria and an appropriate seed extraction method.

B.4 Watershed Transform

The watershed transform has proved to be a powerful and efficient segmentation tool in mathematical morphology [88]. The watershed transform is a special case of seeded region growing.

The intuitive idea behind this method originates from geography and describes catchment basins formed by rain falling on a landscape. The water falling onto the surface will flood each local minima until total immersion and if we prevent the merging of waters coming from different sources by dams, we can partition the landscape into catchment basins separated by watershed lines (dams) as illustrated in Figure B.5. As shown in Figure B.6(a) a grey-level image can be regarded as a topographic surface by considering grey levels as altitude information. For example, the gradient of this image has two catchment basins and the boundaries determined by the watershed transform correspond to the transition regions between the two basins. A reliable result is achieved if the catchment basins highly correspond to the regions of interest in the original or gradient image and the watershed lines represent the desired region boundaries. For image segmentation, the watershed transform is typically applied to the gradient magnitude image. However, gradient images are noisy and contain many minima due to local irregularities and intensity variations in the image. Therefore the watershed transform leads to over-segmentation. The marker-controlled watershed devised by Beucher and Meyer [89] offers an efficient solution to the over-segmentation problem. To avoid over-segmentation due to numerous sources of flooding, flooding of the topographic surface should only be allowed from a priori defined set of markers (instead of flooding from every minimum in the image).

The most crucial and difficult part of marker controlled watershed-based segmentation is the extraction of object markers. If the marker extraction algorithm fails to mark

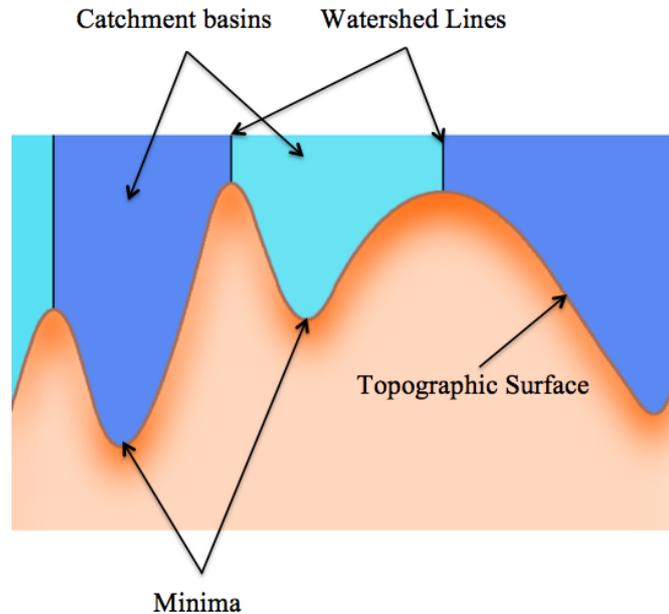
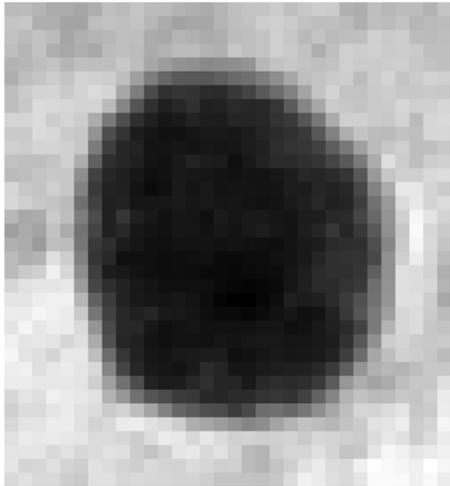


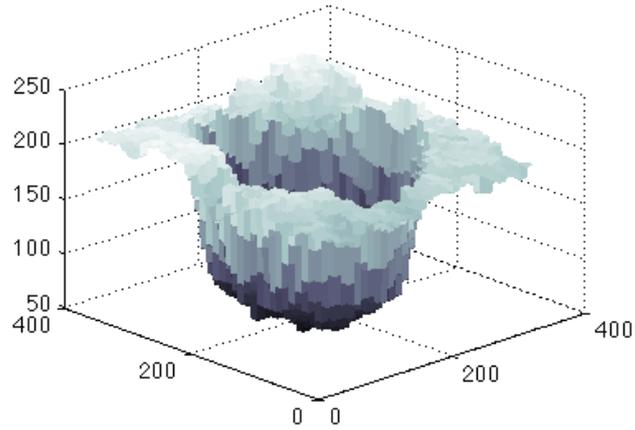
Figure B.5: Simulating the flooding of a topographic surface from its regional minima.

an object, it will be missed in the final segmentation. In practice, a lot of effort has been made to develop fully automated algorithms for the extraction of appropriate markers. The thresholding method [90] is sensitive to noise and uneven illumination. Utilizing the distance transform [91] for extracting the markers will lead to multiple markers for a single object particularly when the image objects are irregular in shape. The use of Bayes classifier to identify pixel groups as internal markers has good performance in certain applications, but it is complex. The method proposed by Lezoray and Cardot [92] employs pixel classification techniques to extract the object markers. The result in this method is dependent on the number of the classes the pixels belong to. Many of the recent marker extraction techniques are based on mathematical morphology such as h-minima [32], top hat transforms and the skeleton of the gradient image. The approaches based on grey-scale morphological reconstruction have achieved remarkable results in the application of cell nuclei segmentation in Pap smear images. However, there is still much scope for improvement.

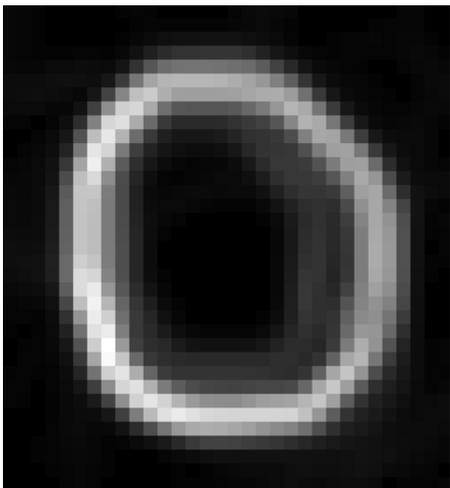
The watershed transform can accurately delineate the object boundaries and is robust to slight optical changes. However, due to the lack of a boundary smoothness constraint, the watershed transform can produce a jagged boundary in some cases. A marker-controlled watershed algorithm with a new marking function has been proposed by Cheng et al [45] to avoid jagged boundaries of segmented regions.



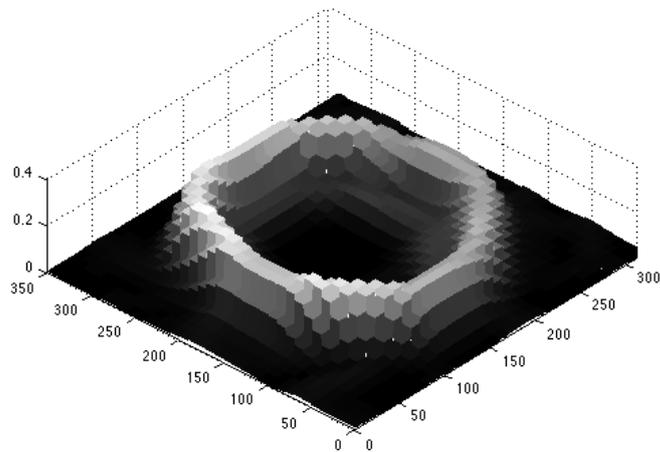
(a) Original image.



(b) The image as a topographic surface.



(c) Gradient image.



(d) The gradient image as a topographic surface.

Figure B.6: Computing the magnitude of gradient image: Watershed transform considers the gradient image as a topographic surface. (a) Original image. (b) The original image as a topographic surface. (c) Gradient image. (d) The gradient image as a topographic surface. Watershed transform is usually applied to gradient magnitude image. The topographic surface of the gradient image has two catchment basins and the boundary of the object corresponds to the transition regions between the two basins.

C

Cytology features

Rodenacker and Bengtsson [54] proposed a taxonomy of features for cytology images based on four main categories: morphometric, densitometric, textural and contextual.

Morphometric features describe the geometry (shape, size, position and boundary) of the object and are computed from its binary mask. Densitometric features describe the intensity or optical density of the object and are computed from the histogram of grey-values within the object. Contextual features are derived from the spatial relationship between objects; e.g. the number of neighbors an object has, and statistics of the distances to neighboring objects. Texture features describe the spatial variation of grey-levels within an object. Such features can be broadly classified into statistical features and structural/syntactic features. Statistical texture features used in this study include fractal texture features, run-length features, histogram based features, grey-level co-occurrence matrix (GLCM) features, and complex Daubechies wavelets features. The wavelet features used in this study to characterize nuclear texture were computed from the first level of the wavelet decomposition[56]. A down-sampled binary mask of the original image is used to identify and retain only the nuclei regions in the decomposed images for the purpose of feature extraction.

The structural texture features used in this study to characterize nuclear texture are derived from a segmentation of the nuclear chromatin into blob-like texture primitives. In particular they are derived from morphometric and contextual features computed for these blobs.

Conceptually these features can be computed for both nuclei and chromatin blobs/particles.

D

Candidate cell nuclei marker extraction methods

D.1 Morphological reconstruction

There are two ways to perform morphological reconstructions either based on conditional morphological operations or geodesic morphological operations. This section discusses the morphological reconstruction as a set of image operators referred to as geodesic. Morphological reconstruction is a powerful operator for filtering, segmentation, and feature extraction from binary and grey-scale images. In binary case this method can extract the connected components of an image which have been marked by another image. However, this operator can be extended for grey-scale images as well to extract regions of interest (i.e. cell nuclei candidate positions within cervical image).

D.1.1 Grey-scale reconstruction:

The definition of elementary geodesic transformations is needed for understanding the concept of image reconstruction. Two important geodesic transformations are geodesic dilation and geodesic erosion.

The grey-scale geodesic dilation of size 1 is defined [93]

$$\delta_g^{(1)}(f) = \delta^{(1)}(f) \wedge g \tag{D.1}$$

where f is the marker image, g is the mask image, $f \leq g$ and have the same domain, the operator \wedge stands for point-wise minimum, $\delta^{(1)}(f)$ is the elementary dilation of f . The grey-scale geodesic dilation of size $i \geq 0$ can be computed by iterating i geodesic dilations of size 1 as follows:

$$\delta_g^{(i)}(f) = \delta_g^1(\delta_g^1(\delta_g^1 \dots (\delta_g^1(f))))). \quad (\text{D.2})$$

The morphological grey-scale reconstruction by dilation of a mask image g from a marker image f where $f \leq g$ is defined as the geodesic dilation of f with respect to g iterated until stability and is denoted by $R_g^\delta(f)$ [93] :

$$R_g^\delta(f) = \delta_g^{(i)}(f), \quad (\text{D.3})$$

where i is such that $\delta_g^{(i)}(f) = \delta_g^{(i+1)}(f)$.

The grey-scale geodesic erosion of size $i \geq 0$ can be computed by iterating i geodesic erosions of size 1 as follows:

$$\varepsilon_g^{(1)}(f) = \varepsilon^{(1)}(f) \vee f, \quad (\text{D.4})$$

$$\varepsilon_g^{(i)}(f) = \varepsilon_g^1(\varepsilon_g^1(\varepsilon_g^1 \dots (\varepsilon_g^1(f))))), \quad (\text{D.5})$$

where the operator \vee stands for point-wise maximum. The morphological grey-scale reconstruction by erosion of a mask image g from a marker image f where $f \geq g$ is defined as the geodesic erosion of f with respect to g iterated until stability and is denoted by $R_g^\varepsilon(f)$ [93] :

$$R_g^\varepsilon(f) = \varepsilon_g^{(i)}(f), \quad (\text{D.6})$$

where i is such that $\varepsilon_g^{(i)}(f) = \varepsilon_g^{(i+1)}(f)$.

D.1.2 Regional minima:

The regional extrema of an image are significant morphological features, because they often represent relevant objects within the image; minima typically correspond to the dark objects and maxima to the bright objects [93]. In morphology, a regional minimum (M) of an image (f) at elevation of (t), is a connected component of pixels each with the value of t , surrounded by pixels with values greater than t . In mathematical terms the regional minimum can be defined as follows:

M is a regional minimum at level $t \Leftrightarrow M$ is connected and

$$\begin{cases} \forall p \in M & \text{if } f(p) = t \\ \forall q \in \delta^{(1)}M \setminus M & \text{if } f(q) > t \end{cases} \quad (\text{D.7})$$

Figure D.1 depicts an example of a regional minimum in an image matrix of 5×7 .

5	3	4	5	5	5	4
6	6	2	2	2	2	3
4	4	2	7	2	5	5
5	4	2	2	2	5	5
6	4	7	7	6	4	4

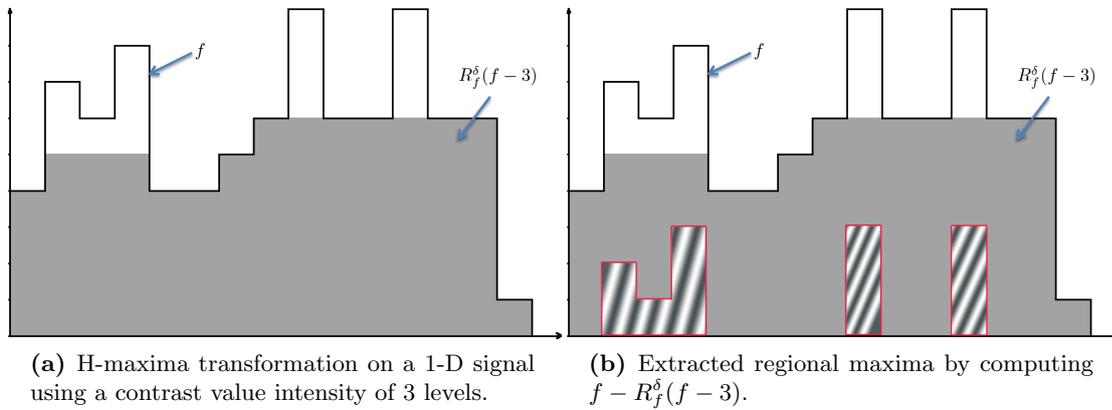
Figure D.1: Regional minimum at level 2 in an image matrix (red connected components).

D.1.3 H-extrema:

The regional extrema of an image represent both relevant and irrelevant features within the image. H-extrema transformations extract the image extrema by using a contrast criterion. To be more precise, the h-maxima transformation will suppress all maxima values whose depth is equal or lower than a specific threshold level of (h) [93]. The h-maxima can be computed by performing the reconstruction by dilation of f from $f - h$ as follows:

$$HMAX_h(f) = R_f^\delta(f - h). \tag{D.8}$$

Figure D.2 shows a simple example of h-maxima transformation on a 1D signal.



(a) H-maxima transformation on a 1-D signal using a contrast value intensity of 3 levels. **(b)** Extracted regional maxima by computing $f - R_f^\delta(f - 3)$.

Figure D.2: H-maxima transformation on a 1-D signal and extracting the regional maxima by computing $f - R_f^\delta(f - h)$.

Similarly the h-minima transformation in mathematical terms can be defined as follows:

$$HMIN_h(f) = R_f^\epsilon(f + h). \tag{D.9}$$

and Figure D.3 depicts a h-minima transformation with contrast level of 3 on a 1-D signal.

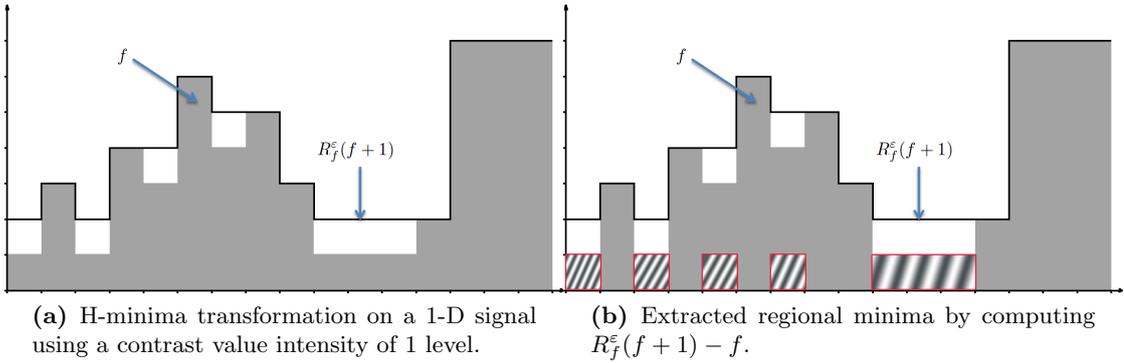


Figure D.3: H-minima transformation on a 1-D signal and extracting the regional minima by computing $R_f^e(f+h) - f$.

As it is discussed in section D, in principle a set of regional minima as markers for locating the cell nuclei in Pap-stained images can be found. These regional minima can be computed using the h-minima transformation with required contrast level h (see Figure D.3).

D.1.4 Morphological Top-hat Transform

The Top-hat transform originally proposed by Meyer [94] is a mathematical morphology operator that uses morphological opening or closing for extracting bright (respectively dark) objects from an uneven background. Top-hat transformation can be formulated in two ways: White Top-hat (*WTH*) and Black Top-hat (*BTH*). White Top-hat can identify small bright regions inside a grey-scale image with dark background. The *WTH* transformation of a grey-scale image f is defined as:

$$T_{\text{white}}(f) = f - (f \circ B) \quad (\text{D.10})$$

In order to identify the bright regions, the Top-hat transform should be applied using a flat structuring element that is somewhat larger than the size of these regions. The \circ is the grey-scale opening operator and is basically a min operation that removes the bright regions smaller than the size of structuring element. Subtracting this image from the original image produces an image that contains the desired bright regions, together with some other small bright noise objects.

Similarly, the *BTH* transformation is defined as:

$$T_{\text{black}}(f) = (f \bullet B) - f \quad (\text{D.11})$$

The *BTH* is the dual of the *WTH*, and subtracts the original image from the closing of the image with structuring element B . *BTH* extracts small dark regions from a bright background. The Top-hat transform is usually followed by a thresholding operation to provide the binary mask which serves as markers or seeds for the segmentation. Figure

D.4 shows the effect of WTH and BTH transforms on a 1-D signal. More generally it is possible to define tophats using other types of morphological openings and closings; e.g. annular closing and opening.

D.1.5 Improved Morphological Top-hat Transform

Jackway [44] introduced an improved morphological top hat transform for the detection of bright and dark objects in the presence of certain types of noise.

Let B_{d_o} and B_{d_i} represent the outer and inner structuring elements, respectively with diameters d_o and d_i , where $d_o > d_i$. The brim which is the annular structuring element is obtained by applying the set difference ($/$ operator) between B_{d_o} and B_{d_i} . The improved white top-hat transformation denoted by $IWTH$ is constructed by first dilating the original image with the annular structuring element, and then eroding the result by the inner structuring element. The result is then subtracted from the original image to produce the $IWTH$, defined as [44]:

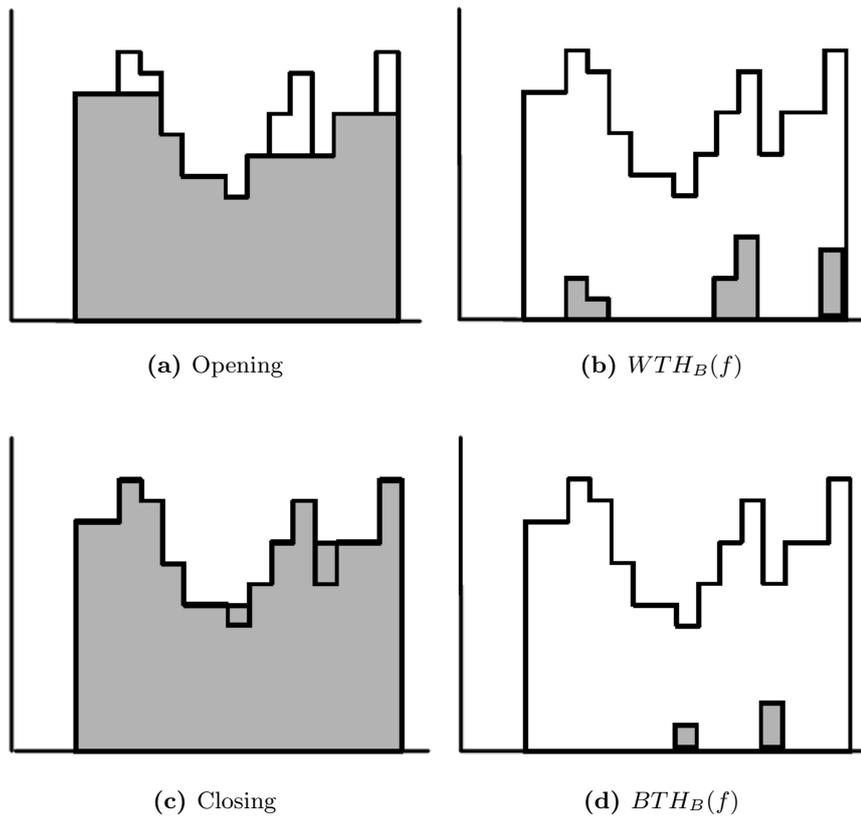


Figure D.4: (a) Morphological Opening operation on image f (b) 1D Profiles of highest peaks extracted by White Top Hat Transform (c) Morphological Closing operation on image f (d) 1D profiles of valleys extracted by the black Top Hat transform.

$$H_W = f - ((f \oplus (B_{d_o} \setminus B_{d_i})) \ominus B_{d_i}) \quad (\text{D.12})$$

IWTH can be used for bright blob detection. Similarly, the improved black top hat (*IBTH*) can be introduced as follows:

$$H_B = ((f \ominus (B_{d_o} \setminus B_{d_i})) \oplus B_{d_i}) - f \quad (\text{D.13})$$

IBTH can be used for black object extraction. Improved top-hat transform can well extract the desired regions because of the superiority of effective background suppression. However, in order to perform marker extraction with the use of morphological top hat transform, an appropriately sized structuring element and a threshold value have to be tuned to the application.

E

Shape Criteria

E.1 Danielsson's G shape factor

Danielsson (1978) devised the G shape factor [95] for a binary image $X \subset \mathbb{Z}^2$ (see Figure E.1):

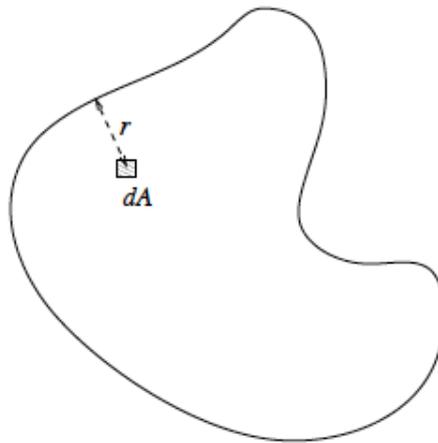


Figure E.1: G Shape-factor calculation

$$G = \frac{A}{9\pi(\bar{d})^2} \quad (\text{E.1})$$

where

$$\bar{d} = \left(\int_A \int r dA \right) / A, \quad (\text{E.2})$$

and A is the area of X . The G factor is dimensionless and is equal to one for an ideal circle. The shape factor is more than one for all other shapes. For a digital image the quantity \bar{d} is estimated by taking the mean of the distance transform of X .

E.2 Dice similarity coefficient

Dice similarity coefficient (DSC) [96] is a spatial overlap index, which measures the similarity between two point sets A and B . The coefficient value ranges from 0 to 1, in which 1 represents a perfect overlap and 0 represents no overlap at all. The DSC is given by

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (\text{E.3})$$

where the operator $| \cdot |$ indicates the number of pixels in the enclosed set, and A and B are two regions.

E.3 Eccentricity

Eccentricity is a scalar that specifies the eccentricity of the ellipse that has the same second-moments as the region. The eccentricity is the ratio of the distance between the foci of the ellipse (l_3) and its major axis length (l_1) [97]. Figure E.2 shows an ellipse with minor axis and major axis length of (l_2) and (l_3) respectively. F_1 and F_2 show the foci of the ellipse.

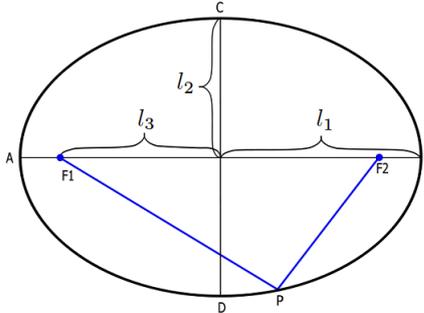


Figure E.2: Parametric description of an ellipse

The eccentricity is given by

$$E = \frac{l_3}{l_1} \quad (\text{E.4})$$

and it can be verified that $(l_2)^2 + (l_3)^2 = (l_1)^2$ [97]. Therefore the eccentricity in terms of major axis length and minor axis length is

$$E = \sqrt{1 - \frac{(l_2)^2}{(l_1)^2}} \quad (\text{E.5})$$

The value is between 0 and 1. The two values of 0 and 1 are degenerate cases; an ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment.

F

Classifiers used in the experiment II

F.1 SVM classifier with a linear kernel

In real life, sometimes we need to assign an object to a specific category from a variety of different categories based on some special characteristics of that object. As a case in point, in this study we want to categorize a particular Pap slide as either normal or abnormal. In computer science, these kind of situations are described as classification problems. Classification of 2-class data is called binary classification and can be described as follows.

Assume that we have a set of labelled objects denoted by the ordered pairs (x_i, y_i) , $i = 1, \dots, n$, where $x_i \in \mathbb{R}^d$ are known as feature vectors and $y_i \in \{-1, +1\}$ are class labels. The classification task is to generate a rule that assigns any new object (point) x to one of the classes [98].

The Support Vector Machine has become one of the most popular classification methods in medical applications.

F.1.1 Theory of linearly separable binary classification

Assume that the data is linearly separable, so a line can be drawn on a graph of x_1 vs x_2 in case of having just two features. Likewise, a hyperplane can be drawn on graphs of x_1, x_2, \dots, x_D when $D > 2$. That hyperplane is given by:

$$\mathbf{w} \cdot \mathbf{x} + b = 0$$

where \mathbf{w} is the normal vector of the hyperplane, and the term $\frac{b}{\|\mathbf{w}\|}$ is the perpendicular distance between the hyperplane and the origin.

The Support Vectors in essence are those objects closest to the separating hyperplane. The main objective of the support vector machine (SVM) is to orient the separating hyperplane to a direction that maximises the distance between the support vectors of each class and the hyperplane. Figure F.1 shows an illustration of two linearly separable classes of data by a hyperplane [98].

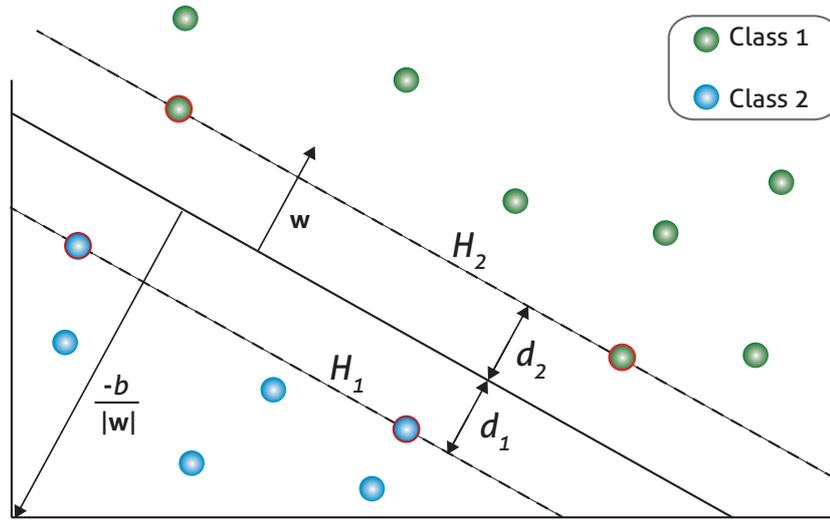


Figure F.1: Illustration of how a two-class data can be separated with a hyperplane.

Figure F.1 depicts that implementing a support vector machine can be cut down to selecting the appropriate values for the variables \mathbf{w} and b . Hence the training data for two classes can be described as follows:

$$\mathbf{w} \cdot \mathbf{x}_i + b \geq +1 \quad \text{for } y_i = +1 \quad (\text{F.1})$$

$$\mathbf{w} \cdot \mathbf{x}_i + b \leq -1 \quad \text{for } y_i = -1 \quad (\text{F.2})$$

and the generalization of these formulas can be given as follows:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \quad \forall i \quad (\text{F.3})$$

The support vectors for two classes are also shown with red circles in Figure F.1. The two planes containing these support vectors for classes 1 and 2 are named H_1 and H_2 respectively. Also, these two planes can be computed as follows:

$$\mathbf{w} \cdot \mathbf{x}_i + b = +1 \quad \text{for } H_1 \quad (\text{F.4})$$

$$\mathbf{w} \cdot \mathbf{x}_i + b = -1 \quad \text{for } H_2 \quad (\text{F.5})$$

The variables d_1 and d_2 are the perpendicular distances of separating hyperplane to H_1 and H_2 planes respectively and are known as SVM margins. The SVM aims to orient

the hyperplane in a way that maximises the margin. This margin is equal to $\frac{1}{\|\mathbf{w}\|}$. To maximise this margin it is necessary to minimise $\|\mathbf{w}\|$. Also minimising the term $\|\mathbf{w}\|$ is equal to minimising the term $\frac{1}{2}\|\mathbf{w}\|^2$. Finally, we can use quadratic programming will help to find:

$$\frac{1}{2}\|\mathbf{w}\|^2 \quad \text{such that} \quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \quad \forall i \quad (\text{F.6})$$

By imposing and solving a Lagrange multiplier and it's dual the optimum values of the \mathbf{w} and b corresponding to the maximum margin can be computed. Those values are as follows:

$$\mathbf{w} = \sum_{i=1}^L \beta_i y_i x_i \quad \text{and} \quad b = \frac{1}{N_s} \sum_{s \in S} (y_s - \sum_{m \in S} \beta_m y_m x_m \cdot x_s) \quad (\text{F.7})$$

N_s is the number of support vectors. m and s belong to the support vector set points ($m \in S$ and $s \in S$). Finally, β values are Lagrange coefficients.

F.2 Logistic regression

Logistic regression is an easy to implement and widely used method for medical data classification. This method has been poorly named as a regression approach, however, because the aim of the logistic regression is to discriminate between different classes of the data. Logistic regression in essence aims to identify the relationship between a binary response and one or more predictor attributes. In logistic regression the response variable is binary, and that distinguishes logistic regression from ordinary linear regression [65].

Unlike the linear model that tries to predict the mean response, logistic regression aims to predict the logit (*log-odds*) of the response having one particular value versus the other value. If response value only takes values 0 and 1 then the logit of the response with value 1 would be $\frac{P(Y=1)}{[1-P(Y=1)]}$. Taking the logarithm of this ratio results in a response that varies between $(-\infty, +\infty)$. This in turn suggests the following model:

$$\text{logit}[P(Y = 1|X)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (\text{F.8})$$

where Y is a random variable denoting the response, vector $X^T = (X_1, X_2, \dots, X_k)$ denote a collection of k independent predictor variables, and where the β_i are the parameters to be estimated.

The logit transform can be defined as:

$$\text{logit}(x) = \ln\left(\frac{x}{1-x}\right) \quad (\text{F.9})$$

By substituting the equation F.9 in equation F.8 the probability $P(Y = 1|X)$ will be:

$$P(Y = 1|X) = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} \quad (\text{F.10})$$

The equation F.10 is called the logistic regression model, and a maximum likelihood estimation method can be used to identify the parameters in the model. Assume a simplified version of this equation as follows:

$$f(z) = \left(\frac{e^z}{1 + e^z} \right) \quad (\text{F.11})$$

This logistic function is depicted in the Figure F.2. By assuming that the value 0 denotes the normal slides and value 1 denotes abnormal slides, the function $f(z)$ represents the estimate for a given value of z .

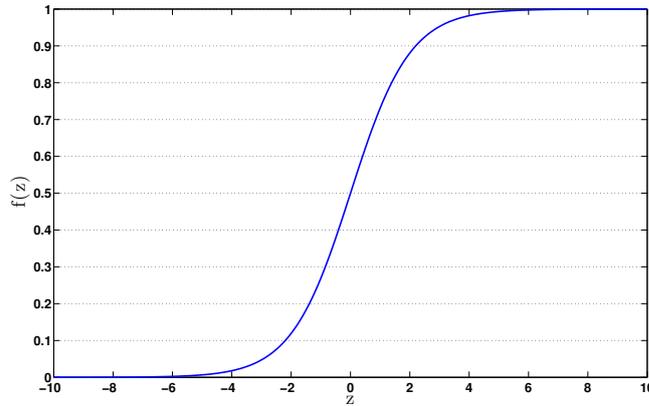


Figure F.2: The logistic function $f(z) = \left(\frac{e^z}{1+e^z} \right)$ restricted to the domain $[-10, +10]$.

This idea can be used to utilize a fitted logistic regression model as a classifier. In fact, for any given observation of x_i , in the equation F.10, the fitted logistic regression will end up with an estimated probability (see equation F.11) that shows the likelihood of being 1. This estimated probability can be used for binary classification purposes (classification of Pap smears). For this purpose, it is necessary to define a *cut – point* value and compare the estimated probabilities with this value. Observations having the estimated values of higher than the *cut – point*, belong to the abnormal class and vice versa.

The optimal decisions in the logistic regression are based on the posterior class probabilities $p(y|x)$, and for a binary classification problem such as Pap smear classification these decisions can be written as:

$$\begin{cases} y = 1 & \text{if } \ln \frac{P(y=1|x)}{P(y=0|x)} > 0 \\ y = 0 & \text{Otherwise} \end{cases} \quad (\text{F.12})$$

Generally the probabilities $P(y|x)$ are not known but the possible decisions can be parametrized according to:

$$\ln \frac{P(y = 1|x)}{P(y = 0|x)} = f(x, w) = w_0 + \mathbf{x}^T \mathbf{w}_1 = 0 \quad (\text{F.13})$$

The parameters can be calculated using a maximum likelihood estimation method. Figure F.3 shows a possible linear decision boundary for classification of a two-class data.

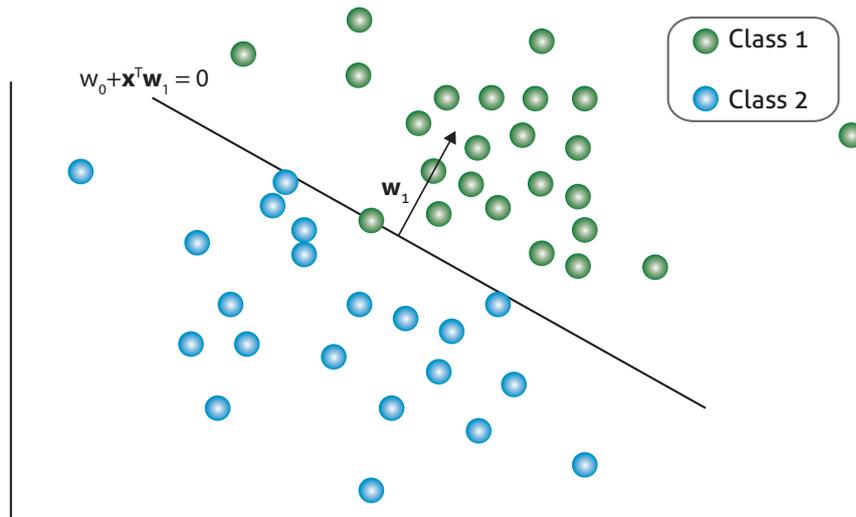


Figure F.3: Logistic regression linear decision boundary for classification of a two-class data.

G

Publications

Paper I

Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis

Ramin Moshavegh, Babak Ehteshami Bejnordi, Andrew Mehnert,
K. Sujathan, Patrik Maim and Ewert Bengtsson

Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis

Ramin Moshavegh, Babak Ehteshami Bejnordi, Andrew Mehnert, K. Sujathan, Patrik Malm and Ewert Bengtsson

Abstract—This paper presents an automated algorithm for robustly detecting and segmenting free-lying cell nuclei in bright-field microscope images of Pap smears. This is an essential initial step in the development of an automated screening system for cervical cancer based on malignancy associated change (MAC) analysis. The proposed segmentation algorithm makes use of gray-scale annular closings to identify free-lying nuclei-like objects together with marker-based watershed segmentation to accurately delineate the nuclear boundaries. The algorithm also employs artifact rejection based on size, shape, and granularity to ensure only the nuclei of intermediate squamous epithelial cells are retained. An evaluation of the performance of the algorithm relative to expert manual segmentation of 33 fields-of-view from 11 Pap smear slides is also presented. The results show that the sensitivity and specificity of nucleus detection is 94.71% and 85.30% respectively, and that the accuracy of segmentation, measured using the Dice coefficient, of the detected nuclei is $97.30 \pm 1.3\%$.

I. INTRODUCTION

THE Papanicolaou test is the primary screening test for cervical cancer. It involves the microscopic examination of cells sampled from in and around the cervix for signs of precancerous and cancerous changes; e.g. large nucleus relative to cytoplasm. The test is labor-intensive and complex requiring the exhaustive review of tens of thousands of cells. One in every 10 to 20 positive cases is missed in routine screening [1]. There are two reasons for this [1]. The first is inappropriate interpretation due to factors such as fatigue, habituation, and inexperience. The second is due to sampling error wherein diagnostic cells do not make it onto the glass slide in the first place. Whilst automation using a computer and robotic microscope can address the issue of inappropriate interpretation it cannot address sampling error. Research suggests that the malignancy-associated change (MAC) phenomenon may be the solution. MACs are subtle sub-visual changes in the appearance of normal-looking cells from an abnormal Pap smear. The aim of MAC analysis is not to perform an exhaustive review of all of the cellular material to identify

diagnostic cells but rather to look for MACs in a sub-population of cells sampled from the slide.

An essential first step in the development of an automated screener based on MACs is robust automatic segmentation of free-lying cell nuclei in digitized Pap smear images. Numerous algorithms have been published for this purpose. They can be categorized according to the primary underlying segmentation methodology used: global and adaptive thresholding [2], watershed transform [3], boundary detection algorithms and deformable models [4-7], and edge enhancement based techniques [8, 9]. The algorithms based on thresholding lack robustness to noise, uneven illumination, and variations in staining intensity. The algorithms based on the watershed transform can yield oversegmentation (when too many potential cell nuclei candidates are selected) and inaccurate boundary segmentation (because of the lack of sufficient gradient information). The algorithms based on boundary detection and deformable models are highly dependent on strong prior knowledge about the shape and location of objects in an image to guide the segmentation process. Moreover they require the initial contour to be reasonably close to the true object boundaries. As a consequence they can fail in images containing clustered and overlapping cells. The algorithms based on edge enhancement are designed to segment the cytoplasm and nucleus of free-lying cells and likewise do not perform well in images containing clustered and overlapping cells. Another criticism of many of these proposed algorithms is that their segmentation accuracies have not been objectively evaluated.

In this paper we present a new algorithm for robustly detecting and segmenting free-lying intermediate cell nuclei in bright-field microscope images of Pap smears. The novelty of the algorithm stems from a robust marker selection method for selecting candidate free-lying nuclei-like objects for subsequent marker-controlled watershed segmentation to obtain the nuclear boundaries. The algorithm also implements artifact rejection based on size, shape, and nuclear granularity to ensure only the nuclei of intermediate squamous epithelial cells are retained. The remainder of this paper is organized as follows. Section II introduces the proposed algorithm. Section III presents an empirical evaluation of its performance. Finally Section IV is the summary and conclusion.

R. Moshavegh, B. Ehteshami and A. Mehnert are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. A. Mehnert is also with MedTech West, Sahlgrenska University Hospital, Gothenburg, Sweden (e-mail: andrew.mehnert@chalmers.se).

P. Malm and E. Bengtsson are with the Centre for Image Analysis, Uppsala University, Uppsala, Sweden.

K. Sujathan is with the Department of Pathology, Regional Cancer Centre (RCC), Thiruvananthapuram, India.

II. PROPOSED SEGMENTATION ALGORITHM

The proposed algorithm is designed to detect and segment free-lying intermediate squamous epithelial cell nuclei because these are the most abundant on a Pap smear and are the most suitable for MAC analysis [10]. The “nuclei of the intermediate cells measure about $8\mu\text{m}$ in average diameter, are round or oval, with a clearly defined nuclear membrane surrounding well-preserved homogeneous nucleoplasm” [11]. The algorithm is summarized in Table I and the individual steps discussed below.

TABLE I
PROPOSED SEGMENTATION ALGORITHM FOR FREE-LYING
INTERMEDIATE CELL NUCLEI.

Input:	Gray-scale image containing a field-of-view (FOV) from a Pap smear slide.
Output:	Binary image containing connected components, each corresponding to a free-lying intermediate cell nucleus.
Steps:	<ol style="list-style-type: none"> 1. Extract inner markers for free-lying nuclei-like objects (these locate the interiors of candidate objects). 2. Apply the marker-controlled watershed transform on the FOV image with respect to the inner markers (this yields an outer marker that lies between the candidate objects). 3. Apply the marker-controlled watershed transform on the gradient image with respect to the inner and outer markers (this yields the object boundaries/masks). 4. Compute the area and quantitative measures of shape and texture granularity for each segmented nuclei-like object. 5. Reject objects that are too small or large to be intermediate cells, that do not have an elliptical shape, and that do not have a granular texture.

Step 1: Extraction of inner markers

The detection of free-lying nuclei is achieved using the gray-scale annular closing operator. The operator is defined

$$\Psi_{\text{anclo}}(f, B) = (f \ominus B) \vee f \quad (1)$$

where f is a gray-scale image, \ominus denotes erosion, \vee denotes pointwise maximum, and B is a symmetric structuring element which does not contain the origin [12]. The effect of this operator is to remove isolated dark spots in the gray-scale image. The operator is an algebraic closing and, like a conventional morphological closing, is increasing, extensive, and idempotent. Thus the arithmetic difference $\Psi_{\text{anclo}}(f, B) - f$ is non-negative and yields the removed isolated dark spots. These serve as candidate cell nuclei markers. This idea is illustrated in Fig. 1 for a single cervical cell and an annular structuring element. The inner radius and outer radius dimensions of the annular structuring element provide control over the size and relative isolation of the nuclei that can be detected.

The squamous epithelium is made up of three principal layers [11]: the basal cell layer (immature), the intermediate cell layers, and the superficial cell layers (most mature). The nuclei of superficial cells are pyknotic and considerably smaller than intermediate and parabasal cells with a nuclear

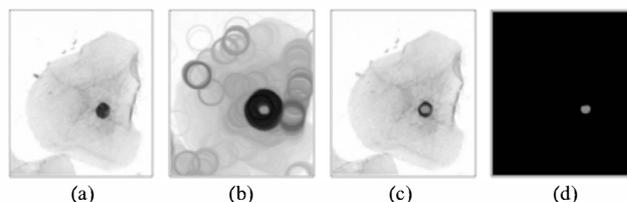


Figure 1. Extracting an inner marker for a free-lying cell nucleus. (a) Original image. (b) Gray-scale erosion with an annular structuring element. (c) Pixel-wise maximum of (a) and (b). (d) Arithmetic difference between (c) and (a).

diameter of about $4\mu\text{m}$ [11]. In a normal Pap smear usually only the upper few layers of the squamous epithelium are removed and so the immature cells near the base of the epithelium are not sampled [13].

In order to detect a nucleus, an annular structuring element with an inner radius larger than that of the nucleus is needed. The nuclei of normal intermediate and parabasal cells measure approximately $8\mu\text{m}$ in diameter and may enlarge up to $15\mu\text{m}$ in the case of malignant or rare benign changes [11]. Hence, to detect all nuclei within this range, a set of independent annular closings with structuring elements with a range of inner diameters is needed. This is then the basis for the more sophisticated inner marker extraction algorithm presented in Table II.

Steps 2-3: Marker-controlled watershed segmentation of the detected nucleus-like objects

Marker-controlled watershed segmentation [14] is used to segment the boundaries of the nuclei-like objects located by the inner markers. First a watershed segmentation of the original image f with respect to the inner markers is performed to obtain the outer marker. Next the original

TABLE II
PROPOSED NUCLEI INNER MARKER EXTRACTION ALGORITHM

Inputs:	Gray-scale image (f), and parameters λ_0, α, r_1 and r_2
Output:	Binary mask (X_m) containing an inner marker for each free-lying nucleus-like object detected.
1.	Let B_0 be a disk structuring element of radius λ_0 .
2.	for $\lambda = r_1$ to r_2 do
3.	Let B_{an} be an annular structuring element with inner and outer radii of λ and $\lambda + \alpha$ respectively
4.	$g = \Psi_{\text{anclo}}(f, B_{\text{an}}) - f$
5.	$X_1 = g > 0$
6.	$X_2 = (X_1 \ominus B_0) \oplus B_0$
7.	$X = X \cup X_2$
8.	end for
9.	$X_m =$ set of centroids of the connected components in X

Note: \oplus and \ominus denote dilation and erosion respectively.

image f is median filtered, its magnitude of gradient is computed, and the result is Gaussian filtered. A watershed segmentation of this filtered gradient with respect to the union of the inner markers and the outer marker yields the desired segmentation of the nuclei-like objects. This idea is illustrated in Fig. 2.

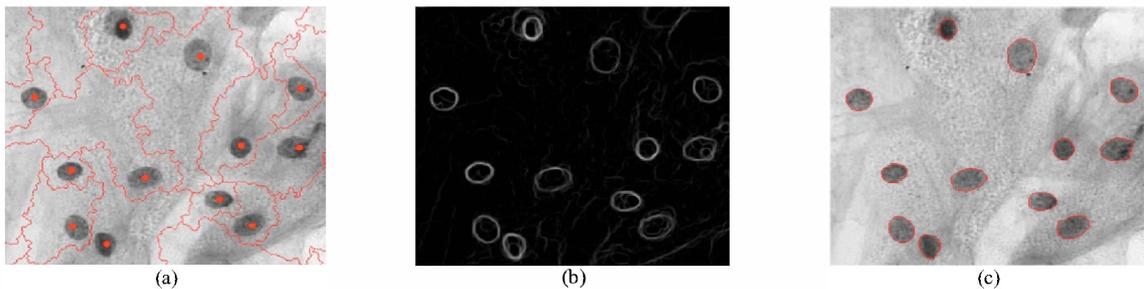


Figure 2. Segmentation of the detected nuclei-like objects. (a) Watershed segmentation of the original image with respect to the inner markers (shown as disks) yielding the outer marker. (b) Gaussian filtered gradient magnitude of the median-filtered original image in (a). (c) Watershed segmentation of (b) with respect to the union of the inner markers and the outer marker.

Steps 4-5: Artifact rejection

Quantitative measures of the area, elliptical shape, and the texture/granularity are computed for each segmented nucleus-like object. These values are then used to reject objects that are too small or large to be intermediate cells, do not have an elliptical shape, or do not have a granular texture.

The minimum area is deemed to be the area of a circle with radius r_{\min} . To measure how elliptical a segmented nucleus-like object is, an ellipse is fitted to the object pixels such that it has the same normalized second central moments as the segmented object. The elliptic variance descriptor (E_{var}) [15] is then used to measure how closely the borders of the fitted ellipse agree with those of the segmented nucleus-like object. For elliptical objects E_{var} is close to 0 (see Fig. 3).

The granularity or coarseness of texture of each segmented object is computed using the coarseness feature devised by Tamura [16] (derived from a texture model corresponding to visual perception). This feature reflects the size and number of texture primitives. It is useful and robust in the sense that it does not depend directly on the exact gray-levels in the object and so has robustness to non-uniformity of illumination and staining variations (provided that these do not greatly affect the size and number of texture primitives). Fig. 3 depicts two candidate nuclei with different E_{var} and Tamura coarseness values.

III. EMPIRICAL EVALUATION

The performance of the algorithm was evaluated relative to expert manual segmentation.

A. Image data

The data used in this study is a subset of 889 fields-of-view (FOVs) captured by a cytopathologist from 68 Pap smear slides. Each FOV was acquired using a CCD camera mounted on a light microscope. The images were captured with a 40 \times objective lens. Each FOV image is of size 1024 \times 1344 pixels with square pixels of size 0.25 μm . The gray-scale resolution is 8 bits per pixel.

Eleven slides, each containing a minimum of 100 non-superficial cervical cell nuclei, were randomly selected from among the 68 slides. For each slide three FOVs were randomly selected to yield a total of 33 FOVs.

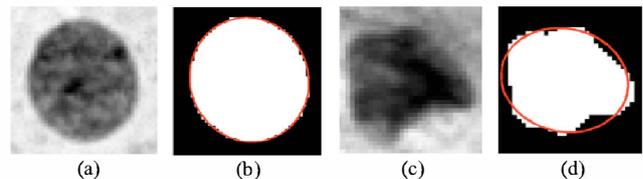


Figure 3. Assessing the shape and texture of a candidate nucleus. (a) and (b) show the result for a candidate nucleus with an elliptical shape ($E_{var} = 0.026$ and $Tamura\ coarseness = 11.17$). (c) and (d) show the result for a candidate nucleus with a non-elliptical shape ($E_{var} = 0.096$ and $Tamura\ coarseness = 10.75$).

B. Ground truth

A graphical user interface (GUI) was developed to permit a user to review each FOV and to place a marker on individual nuclei. Three untrained subjects were recruited to independently review the FOVs using the GUI and to mark each free-lying nucleus. Prior to performing this task each was shown examples of intermediate cell nuclei in another FOV (not one of the 33 FOVs they had to review). Each subject was specifically instructed to mark elliptical objects, of approximately the right size, with a well-defined boundary, and with a granular texture. The set of all objects selected by at least two of the three subjects were taken to be the ground truth for free-lying intermediate cell nuclei.

Two image analysis experts then used the GUI to independently trace the boundary of each ground truth nucleus. These manual segmentations were taken to be the ground truth for the boundaries of the free-lying intermediate cell nuclei.

C. Method

The proposed segmentation algorithm was applied to the 33 FOV images. The parameters for different steps of the algorithm were selected after several experiments on a small subset of images independent of the 33 selected FOV images. The minimum and maximum values of the inner radius (r_1 and r_2) of the annular structuring elements (see Table I) were set to 22 and 33 pixels. The values were selected based on the nucleus diameter range for the intermediate cell nuclei defined by Koss [11]. The outer radius of each annulus was set to be two pixels more than the inner radius (i.e. $\alpha = 2$) to guarantee the extraction of inner markers of the adjacent free-lying nuclei. Finally the size of the disk-structuring element (λ_*) was set to 3 pixels (for noise mitigation).

Objects selected by the algorithm were compared to the ground truth nuclei obtained manually and used to compute the sensitivity and specificity of the algorithm for the detection of free-lying intermediate cell nuclei.

To evaluate the accuracy of the segmentation of each nucleus detected by the algorithm, the resulting boundary was compared to the two corresponding boundaries (one from each expert) in the ground truth. More specifically the similarity between pairs of masks was computed in terms of the Dice similarity coefficient (DSC) scores [17]. The coefficient ranges between 0 and 1. A value of 1 indicates perfect agreement and a value of 0 indicates no agreement.

D. Results

The sensitivity and specificity of the algorithm for the detection of free-lying intermediate cell nuclei is 94.71% and 85.30% respectively. Boxplots of the DSC scores for the comparison of the proposed automatic segmentation to the two manual segmentations, and for the comparison between the two manual segmentations are shown in Fig. 4.

The agreement between the algorithm and the two manual segmentations is $97.30 \pm 1.3\%$ and $96.96 \pm 1.7\%$ respectively (mean DSC \pm standard deviation). The overall agreement between the two expert segmentations is $97.26 \pm 1.2\%$.

E. Discussion

The sensitivity of the algorithm to the detection of free-lying intermediate cell nuclei is high. However its specificity, whilst still quite high, could be improved. A review of false positives indicates that some of them are due to segmentation failures as the result of severe background noise and artifacts. However other apparent failures in fact represent genuine free-lying nuclei overlooked by the three recruited subjects.

The DSC scores for boundary delineation evaluation (see Fig. 4) show that nuclei boundaries obtained using the marker-controlled watershed transform are highly accurate and consistent with the two experts' visual perception of the intermediate cell nuclei boundaries.

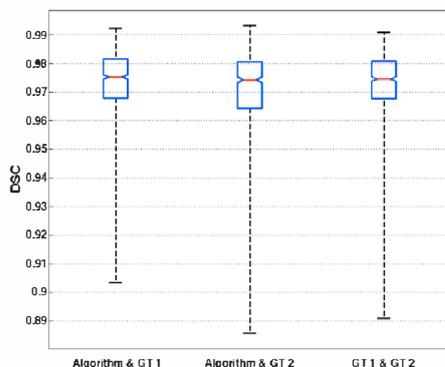


Figure 4. DSC scores for the comparison of the proposed automatic segmentation to the two manual segmentations, and for the comparison between the two expert segmentations (GT1 and GT2).

IV. SUMMARY AND CONCLUSION

In this paper we presented a new automated algorithm for detecting and segmenting free-lying cell nuclei in bright-field microscope images of Pap smears. The empirical results show that the algorithm has a high sensitivity and specificity for the detection of free-lying nuclei, and that the algorithm is able to delineate the boundaries of these with high accuracy. This work represents the first step in the development of an automated screener for Pap smears based on the malignancy-associated change phenomenon.

REFERENCES

- [1] R. M. DeMay, "Common problems in papanicolaou smear interpretation," *Archives of Pathology & Laboratory Medicine*, vol. 121, pp. 229-238, 1997.
- [2] H. S. Wu, J. Gil, and J. Barba, "Optimal segmentation of cell images," *Vision, Image and Signal Processing, IEE Proceedings*, vol. 145, pp. 50-56, 1998.
- [3] O. Lezoray and H. Cardot, "Cooperation of color pixel classification schemes and color watershed: a study for microscopic images," *Image Processing, IEEE Transactions on*, vol. 11, pp. 783-789, 2002.
- [4] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours," *Signal Process*, vol. 71, pp. 203-213, 1998.
- [5] A. Garrido and N. P. de la Blanca, "Applying deformable templates for cell image segmentation," *Pattern Recognition*, vol. 33, pp. 821-832, 2000.
- [6] M. E. Plissiti, C. Nikou, and A. Charchanti, "Accurate localization of cell nuclei in Pap smear images using gradient vector flow deformable models," *Proceedings of 3rd International Conference on Bio-inspired Signals and Systems (BIOSIGNALS 2010)*, pp. 284-289, 2010.
- [7] H. S. Wu, J. Barba, and J. Gil, "A parametric fitting algorithm for segmentation of cell images," *Biomedical Engineering, IEEE Transactions on*, vol. 45, pp. 400-407, 1998.
- [8] C. H. Lin, Y. K. Chan, and C. C. Chen, "Detection and segmentation of cervical cell cytoplasm and nucleus," *Int. J. Imaging Syst. Technol*, vol. 19, pp. 260-270, 2009.
- [9] S. F. Yang-Mao, Y. K. Chan, and Y. P. Chu, "Edge Enhancement Nucleus and Cytoplasm Contour Detector of Cervical Smear Images," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 38, pp. 353-366, 2008.
- [10] L. M. Isenstein, D. J. Zahniser, and M. L. Hutchinson, "Combined malignancy associated change and contextual analysis for computerized classification of cervical cell monolayers," *Anal Cell Pathol*, vol. 9, pp. 83-93, 1995.
- [11] L. G. Koss and M. R. Melamed, "Diagnostic cytology of organs," in *Koss' Diagnostic Cytology and Its Histopathologic Bases*, vol. 1, 5 ed: Lippincott Williams & Wilkins, 2006, p. 186.
- [12] C. Ronse and H. J. A. M. Heijmans, "A Lattice-Theoretical Framework for Annular Filters in Morphological Image Processing," *Applicable Algebra in Engineering, Communication and Computing*, vol. 9, pp. 45-89, 1998.
- [13] E. S. Cibas and B. S. Ducatman, *Cytology: Diagnostic principles and clinical correlates*, 3 ed.: Saunders, 2009.
- [14] F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, pp. 21-46, 1990.
- [15] M. Peura and J. Iivarinen, "Efficiency of Simple Shape Descriptors," *Aspects of Visual Form, Proceedings of the Third International Workshop on Visual Form*, pp. 443-451, 1997.
- [16] H. Tamura, S. Mori, and T. Yamawaki, "Textural Features Corresponding to Visual Perception," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, pp. 460-473, 1978.
- [17] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecological Society of America*, vol. 26, pp. 297-302, 1945.

Paper II

Novel Chromatin Texture Features for the Classification of Pap Smears

Babak Ehteshami Bejnordi, Ramin Moshavegh, K. Sujathan,
Patrik Maim, Ewert Bengtsson and Andrew Mehnert

Novel Chromatin Texture Features for the Classification of Pap Smears

Babak Ehteshami Bejnordi^{ac}, Ramin Moshavegh^{ac}, K. Sujathan^b,
Patrik Malm^d, Ewert Bengtsson^d, and Andrew Mehnert^{ac}

^aDepartment of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden;

^b Department of Pathology, Regional Cancer Centre (RCC), Thiruvananthapuram, India;

^c MedTech West, Sahlgrenska University Hospital, Gothenburg, Sweden;

^d Centre for Image Analysis, Uppsala University, Uppsala, Sweden

ABSTRACT

This paper presents a set of novel structural texture features for quantifying nuclear chromatin patterns in cells on a conventional Pap smear. The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. The results of a comprehensive feature selection experiment, including the set of proposed structural texture features and a range of different cytology features drawn from the literature, show that two of the four top ranking features are structural texture features. They also show that a combination of structural and conventional features yields a classification performance of 0.954 ± 0.019 (AUC \pm SE) for the discrimination of normal (NILM) and abnormal (LSIL and HSIL) slides. The results of a second classification experiment, using only normal-appearing cells from both normal and abnormal slides, demonstrates that a single structural texture feature measuring chromatin margination yields a classification performance of 0.815 ± 0.019 . Overall the results demonstrate the efficacy of the proposed structural approach and that it is possible to detect malignancy associated changes (MACs) in Papanicolaou stain.

Keywords: Cervical cancer, feature extraction, feature selection, classification, chromatin texture, malignancy associated changes, structural texture features, Pap smear classification

1. INTRODUCTION

The Papanicolaou (Pap) test is the primary screening test for cervical cancer.¹ It involves the microscopic examination of cells sampled from the cervix. It is laborious and time-consuming involving the review of possibly hundreds of thousands of cells for visual signs of cancerous or precancerous changes.² One in every 10 to 20 positive cases is missed in routine screening.¹ Two major factors affect the accuracy of the Pap test. The first is sampling error wherein no diagnostic cells make it on to the slide. The other is interpretation error for reasons including fatigue, inexperience, and habituation. Computer-assisted interpretation can potentially address the issue of interpretation error. The malignancy-associated change (MAC) phenomenon may potentially address sampling error. Several studies have demonstrated the phenomenon.³⁻⁵ MACs are subtle sub-visual, i.e. visually imperceptible, changes in the appearance of otherwise normal-looking cells from an abnormal Pap slide. MACs cannot be detected on a cell-by-cell basis but rather in a population of cells. The MAC approach obviates the need to perform an exhaustive review of all of the cellular material to identify diagnostic cells.

The MAC features that appear to have the most discriminatory power are those characterizing nuclear texture or more specifically the chromatin pattern in the nucleus. Traditionally MAC features have been based on a stochastic approach to defining texture. However such features do not correspond well to the terms used by cytopathologists to describe chromatin texture such as heterogeneity, granularity, margination, condensation, compaction, clumping, diffuse, blobs and particles.² This motivates interest in a structural approach to chromatin texture analysis. This paper presents a set of novel structural texture features for quantifying chromatin texture.

B. Ehteshami Bejnordi: (✉) ehteshami@live.se

R. Moshavegh: (✉) raminmoshavegh@gmail.com

A. Mehnert: (✉) andrew.mehnert@chalmers.se

Medical Imaging 2013: Digital Pathology, edited by Metin N. Gurcan, Anant Madabhushi,
Proc. of SPIE Vol. 8676, 867608 · © 2013 SPIE · CCC code: 1605-7422/13/\$18
doi: 10.1117/12.2007185

Proc. of SPIE Vol. 8676 867608-1

The features are derived from an initial segmentation of the chromatin into blob-like texture primitives. The paper also presents an investigation of the most discriminatory subset of features, from among the proposed features and a wide range of features drawn from the literature, for discriminating between normal and abnormal Pap smears using the MAC approach.

2. METHODS

The types of features that can be extracted from cytology images can be broadly classified into four classes: morphometric, densitometric, textural, and contextual.⁶ Whilst each of these can conceptually be computed for both nuclei and chromatin blobs/particles, in this study contextual features were excluded for nuclei, and textural and densitometric features were excluded for chromatin blobs. Morphometric features describe the geometry (shape, size, position and boundary) of the object and are computed from its binary mask. Densitometric features describe the intensity or optical density of the object and are computed from the histogram of gray-values within the object. Contextual features are derived from the spatial relationship between objects; e.g. the number of neighbors an object has, and statistics of the distances to neighboring objects. Texture features describe the spatial variation of gray-levels within an object. Such features can be broadly classified into statistical features and structural/syntactic features. Statistical texture features used in this study include fractal texture features, run-length features, histogram based features, gray-level co-occurrence matrix (GLCM) features, and complex Daubechies wavelets features. The wavelet features used in this study to characterize nuclear texture were computed from the first level of the wavelet decomposition.⁷ A down-sampled binary mask of the original image is used to identify and retain only the nuclei regions in the decomposed images for the purpose of feature extraction.

The structural texture features used in this study to characterize nuclear texture are derived from a segmentation of the nuclear chromatin into blob-like texture primitives. In particular they are derived from morphometric and contextual features computed for these blobs. The approach is described in the next section.

2.1 Novel structural chromatin texture features

Our proposed structural chromatin texture features are based on an initial segmentation of the chromatin into blob-like texture primitives.² For a given nucleus this involves: (i) applying a 3×3 median filter (to attenuate impulse noise); (ii) up-sampling by a factor of 3 (to facilitate the rendering of watershed lines in a subsequent step); (iii) locating the regional minima (inner markers); (iv) computing the watershed transform of the filtered image with respect to the inner markers (to produce an outer marker that delineates a zone of influence (ZOI) around each regional minimum); (v) computing the magnitude of gradient image for the filtered image; and (vi) applying the watershed transform to the gradient image with respect to both the inner markers and the outer marker (to delineate each dark chromatin blob). Our structural features (see Table 1) in essence describe the attributes and arrangement of these resulting chromatin blobs. They fall into two broad categories. The first consists of statistics of morphometric features computed for individual particles in a nucleus. The second consists of contextual features computed from the particles in a nucleus. These contextual features can be further subdivided into the following classes of features: margination, clustering, blob features, discrete texture features, and features derived from the Voronoi diagram and its associated graphs.

The margination features characterize the distances of segmented chromatin blobs to the nucleus boundary. They are computed from the cookie-cutting distances² illustrated in Figure 1. In this study a 6th order polynomial was fitted to the counts in the cumulative frequency histogram of these cookie-cutting distances, and the slopes at the 25th and 37th percentiles were used as margination features $M1$ and $M2$ respectively for each nucleus. The mean and standard deviation of these cookie-cutting distances were used as features $M3$ and $M4$ respectively. The four features implicitly include information about the size of the nucleus. For this reason it was decided in this study to include features $M5$ and $M6$ defined to be $M3$ and $M4$, respectively, each normalized to the maximum value of the distance transform of the nucleus. The last margination feature, $M7$, is defined to be the minimum distance between the centroids (geometrical centers) of the chromatin particles and the nucleus boundary.

Table 1: Summary of the extracted nucleus features

Morphometric Features	
	Area, perimeter, mean-radius, elliptic variance ⁸
Densitometric Features	
	Photometric features: ⁶ Integrated optical density (<i>IOD</i>), mean optical density (<i>MOD</i>), variance of optical density, skewness of optical density, kurtosis of optical density
Texture Features	
Statistical	Fractal texture features: ⁹ Fractal area 1, fractal area 2, fractal dimension
	Run length features: ³ SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE
	Histogram based features: ⁹ Mean, standard deviation, and skewness of the gray level histogram
	GLCM features: ⁶ Contrast, correlation, energy, entropy, local homogeneity, maximum probability, cluster shade, cluster prominence and information measure of correlation (<i>H-IMC1</i> , <i>H-IMC2</i>)
	Complex Daubechies wavelet features of nuclei: ⁷ Mean, standard deviation, and skewness of the gray level histogram. GLCM features – contrast, correlation, energy, entropy, local homogeneity, maximum probability, cluster shade, cluster prominence and information measure of correlation (<i>W-IMC1</i> , <i>W-IMC2</i>)
Statistics of the chromatin particle morphometric features	
	Mean of chromatin particle areas and perimeters, mean of chromatin particle areas normalized to the nucleus area, mean of chromatin particle perimeters normalized to the nucleus perimeter, chromatin particle compactness (P2A)
Contextual chromatin particle features	
Structural	Margination: <i>M1</i> , <i>M2</i> , <i>M3</i> , <i>M4</i> , <i>M5</i> , <i>M6</i> , <i>M7</i> (See section 2.1)
	Clustering: <i>C1</i> , <i>C2</i> , <i>C3</i> , <i>C4</i> (See section 2.1)
	Blob features: Heterogeneity (the ratio between the area of the segmented dark and light regions, and the nucleus area), ¹⁰ the ratio between the area of the segmented dark regions and the nucleus area, number of segmented dark particles, number of segmented light particles, area ratio of each chromatin particle to its watershed ZOI, average distance between the geometrical center of the nucleus and pixels of all chromatin particles, distance between the geometrical center of the nucleus and center of mass of the chromatin particles
	Discrete texture features: ^{9,11} Medium and high DNA amount, medium and high DNA area, medium and high DNA compactness, medium-high DNA compactness, center of gravity (the distance from the geometrical center of the blob to the center of gravity of the optical density function)
	Area-Voronoi diagram: Mean and standard deviation (SD) of the areas of the Voronoi regions, mean and SD of the areas of the Voronoi regions normalized to the nucleus area, area disorder, ³ Voronoi regions roundness, ³ mean of area ratio of each chromatin particle to its Voronoi region ³
	Delaunay graph and generalized Delaunay graph: Mean of the Delaunay triangle areas, mean and SD of the Delaunay edge lengths, average of the mean and SD of the edge lengths connected to each vertex (chromatin particle), maximum Delaunay edge length, mean of the number of connections per chromatin particle, number of vertices (chromatin particles) on the graph boundary
	Gabriel graph and relative neighborhood graph: Mean and SD of the graph edge lengths, maximum edge length, average of the mean and SD of the graph edge lengths connected to each vertex (chromatin blob), mean of the number of connections per vertex (chromatin blob)
Minimum spanning tree: Mean and SD of the edge lengths, total edge length, edge disorder, ³ minimum to maximum edge ratio, mean of the number of connections per vertex, percentage of vertices connected to one vertex (<i>MST1</i>), percentage of vertices connected to two vertices (<i>MST2</i>), percentage of vertices connected to more than two vertices (<i>MST3</i>)	

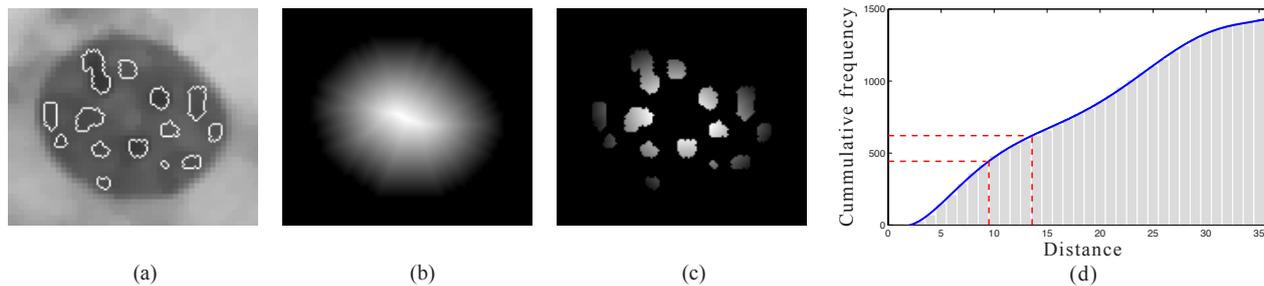


Figure 1: Illustration of how the features $M1$ and $M2$ are computed. (a) Segmented chromatin particles. (b) Euclidian distance transform of the nucleus. (c) Portions of (b) cut out by the segmented chromatin particle masks. (d) Cumulative frequency histogram of these cut-out distances with a polynomial fit through the counts superimposed. Features $M1$ and $M2$ are the slopes at the 25th and 37th percentiles respectively.

A subset of the contextual features are those derived from the Voronoi diagram and a generalization called the area Voronoi diagram. In this study the Delaunay graph (dual of the Voronoi diagram) is computed for the centroids of the particles in the nucleus and several features are then directly computed from it. The area Voronoi diagram is computed from the watershed transform of the distance transform of the set complement of the union of the chromatin particle masks. The region adjacency graph defined on the resulting watershed regions yields a generalization of the Delaunay graph as shown in Figure 2. Several sub-graphs—Gabriel graph, relative neighborhood graph, and the minimum spanning tree¹²⁻¹⁴—are also computed. Each type of graph defines a different adjacency relationship between the particles. Details concerning the explicit features computed from the Voronoi diagram and associated graphs are given in Table 1.

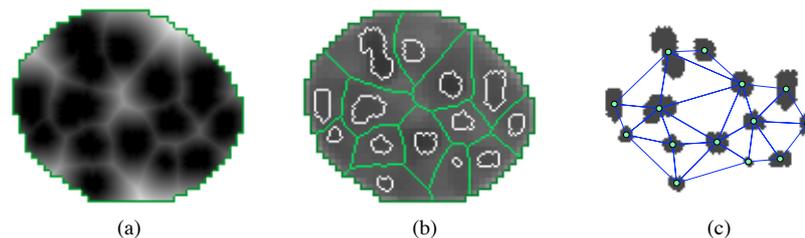


Figure 2: Illustration of how the area Voronoi diagram and generalized Delaunay graph are computed. (a) Distance transform of the set complement of the union of the chromatin particle masks in Figure 1a. (b) Area Voronoi diagram and segmented chromatin particles. (c) The region adjacency graph defined on the Voronoi regions then defines a generalized Delaunay graph.

The clustering features characterize the way the chromatin particles are clustered in the nucleus. They are computed from the area Voronoi diagram. Features $C1$ and $C2$ are computed similarly to features $M1$ and $M2$ except that the cookie-cutting distances are obtained from the distance transform of the Voronoi regions as shown in Figure 3. Features $C3$ and $C4$ are computed from the area Voronoi diagram and the distance transform image used in its construction (see Figure 2a and 2b) as follows. Firstly for every pair of Voronoi regions that share a common edge, the convex hull is computed for the union of the corresponding pairs of particles. Secondly the union of these convex hulls is filled to form a binary mask (Figure 3b). Thirdly this mask is intersected with the divide lines of the area Voronoi diagram. Finally features $C3$ and $C4$ are the mean and standard deviation, respectively, of the distance transform values corresponding to the residual divide line pixels in this intersection (Figure 3c).

Details concerning the blob features and the discrete texture features are given in Table 1.

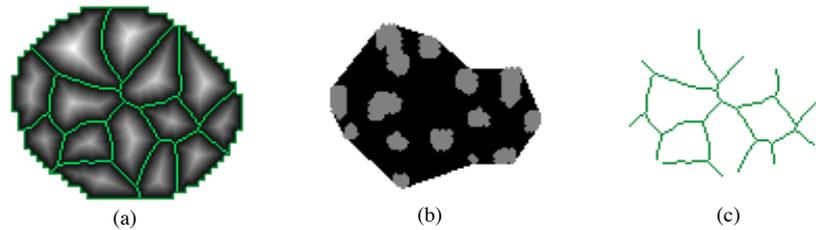


Figure 3: Illustration of how the features $C1$, $C2$, $C3$, and $C4$ are computed. (a) Distance transform of the Voronoi regions in Figure 2b. (b) Binary mask (with particles superimposed) formed from the union of the convex hulls (see the text for details). (c) Voronoi region divide lines clipped according to the mask in (b).

2.2 Pap smear images and ground truth

The image data used in this study originate from a set of 68 conventional Pap smear slides. Each slide was reviewed by a cytopathologist and assigned a cytological diagnosis according to the Bethesda system.¹⁵ The cytopathologist subsequently acquired representative FOVs from each slide. In the case of abnormal smears this included FOVs with and without diagnostic cells. The cytopathologist also labeled individual cells in each FOV accordingly. The FOV images were acquired using a monochrome CCD camera mounted on a light microscope with a $40\times$ objective lens. Each image has a gray-scale resolution of 8 bits per pixel and is of size 1024×1344 pixels with square pixels of size $0.25\mu\text{m}$.

2.3 Experiments

To evaluate the performance of the proposed structural chromatin texture features two experiments were performed. The aim of the first experiment was to determine the most discriminatory subset of features, from among the proposed features and a wide range of features drawn from the literature, for discriminating between normal and abnormal Pap smear slides. A normal smear is deemed to be one that has been labeled as NILM (Negative for Intraepithelial Lesion for Malignancy). An abnormal smear is one that has been labeled as either LSIL (Low grade squamous intraepithelial lesion) or HSIL (High grade squamous intraepithelial lesion). The aim of the second experiment was to evaluate the performance of a variety of classifiers built using the feature subset obtained in the first experiment to discriminate between the normal and abnormal slides.

In the first experiment, the free-lying nuclei in each digitized FOV were automatically segmented using the algorithm we have previously described.¹⁶ For each nucleus we computed the 159 features listed in Table 1. The mean and standard deviation of each feature over all segmented nuclei for a slide yielded a total of 318 slide-based features. An additional random feature was included to gauge the efficacy of feature selection. As a compromise between having a reasonable number of cells per slide for MAC analysis and having a reasonable number of slides for our study we decided to impose the following two criteria. Firstly each slide should have a minimum of 80 non-superficial cervical cells with a NILM label. Secondly, 10% to 40% of the cells in the abnormal slides should be diagnostic cells. This left a total of 44 slides: 25 NILM, 10 LSIL, and 9 HSIL. The following three feature selection methods were investigated: state-of-the-art multiple support vector machines with recursive feature elimination (MSVM-RFE),¹⁷ L1-regularization path for generalized linear models,¹⁸ and the recently introduced guided regularized random forest (GRRF).¹⁹ To optimize classification performance it is necessary to couple each classification approach with the most relevant feature selection method. In this experiment we paired the MSVM-RFE with a linear kernel SVM classifier, the GRRF with an ordinary RRF²⁰ classifier, and the L1-regularization procedure with a generalized linear model. Given the limited number of slides available in this study, all the slides were used for feature selection. However an internal cross validation (in-loop feature selection) technique²¹ was employed to avoid overfitting and overoptimistic estimation of the performance. Feature selection by the SVM-RFE was achieved using 100 iterations of double 5-fold cross validation, each iteration with random combinations of samples in the training and the test set. The inner cross validation loop was used for feature selection and tuning the optimal parameters for the SVM classifier with a linear kernel. The outer cross validation loop was used to evaluate the classifier model performance in terms of the AUC (area under the receiver operating

characteristic curve). For GRRF feature selection and the L1-regularization path for generalized linear models, 100 repeated 5-fold cross validation was employed.

In the second experiment only the cells marked by the cytologist as NILM on normal and abnormal slides were segmented and slides with a minimum of 80 such cells retained. This left a total of 44 slides: 26 NILM, 9 LSIL, and 9 HSIL. To avoid the curse of dimensionality, only feature subsets with one or two features identified in experiment 1 were considered. Two classifiers were evaluated using 100 iterations of 5-fold cross validation: SVM with a linear kernel and a logistic regression (LR) classifier.

3. RESULTS

Experiment 1: The classification performance results together with the four top-ranked features for the three feature selection approaches are presented in Table 2. “SD of *IOD*” and “Mean of *M2*” are the top two features in all three feature selection (FS) methods. The top four features appearing most frequently among the feature selection methods are: SD of *IOD*, Mean of *M2*, SD of *MST3* and Mean of *W-IMC1*. Figure 4a shows boxplots for each of these features for both the normal and abnormal slides. The best AUC (0.954 ± 0.019) was achieved using the L1-regularization procedure with a generalized linear model.

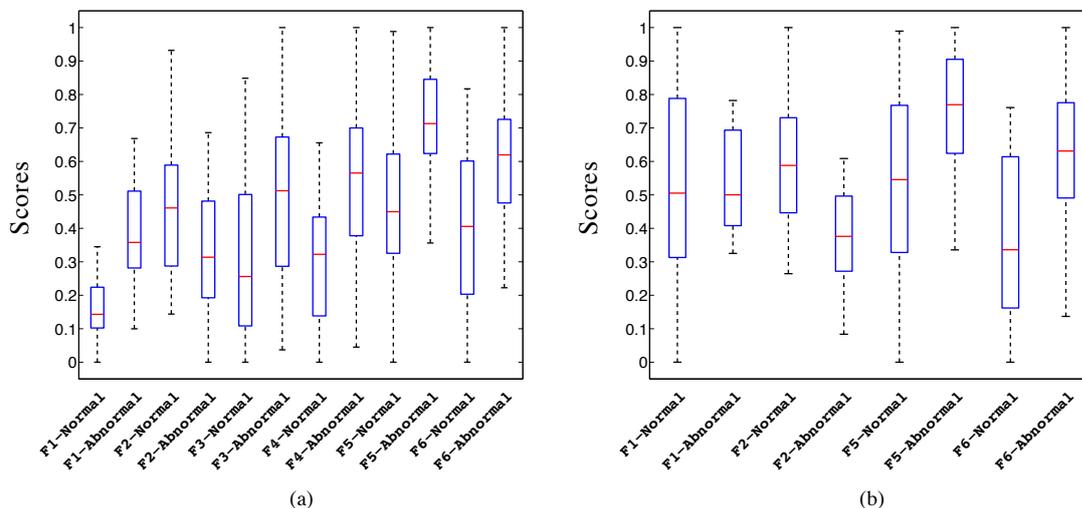


Figure 4: (a) Boxplots of the features *F1*, *F2*, *F3*, *F4*, *F5*, *F6* for experiment 1. (b) Boxplots of the features *F1*, *F2*, *F5*, *F6* for experiment 2.

Experiment 2: The classification performance results are presented in Table 3. Boxplots for each feature for both the normal and abnormal slides are shown in Figure 4b. The feature “Mean of *M2*”, which is a measure of chromatin margination inside the nucleus, achieved the best classification performance in both classifiers for the detection of MACs. This is also supported visually in Figure 4b. The feature “SD of *IOD*”, whilst quite discriminatory in the first experiment, does not show any discriminatory power in the second experiment. This is not surprising given that the Pap stain is not stoichiometric and that none of the cells in experiment 2 were diagnostic.

Table 2: Classification results for experiment 1.

FS Method	Top 4 features	AUC (Mean±SE)
MSVM-RFE	$F1, F2, F3, F4$	0.934±0.038
L1-Reg	$F1, F2, F5, F6$	0.954±0.019
GRRF	$F1, F2, F5, F6$	0.935±0.029
<i>F1: SD of IOD, F2: Mean of M2, F3: SD of C1, F4: SD of nucleus area, F5: SD of MST3, F6: Mean of W-IMC1</i>		

Table 3: Classification results for experiment 2.

Classifier	Features	AUC (Mean±SE)
SVM	$F2$	0.815±0.015
	$F5$	0.706±0.053
	$F6$	0.792±0.026
	$F2, F6$	0.822±0.024
Logistic Regression	$F2$	0.803±0.026
	$F5$	0.591±0.054
	$F6$	0.784±0.027
	$F2, F6$	0.752±0.038

4. CONCLUSIONS AND DISCUSSION

A set of novel structural texture features for quantifying nuclear chromatin patterns in cells on a conventional Pap smear was presented. The experimental results demonstrate the efficacy of this structural approach and that a combination of these structural texture features and conventional features can be used to discriminate between normal (NILM) and abnormal (LSIL and HSIL) slides with high accuracy. They also demonstrate that it is possible to detect malignancy associated changes (MACs) in Papanicolaou stain. The most discriminating single feature, Mean of $M2$, for detecting malignancy associated changes (MACs) is a structural feature measuring chromatin margination. This in turn suggests the possibility of developing a fully automated Pap smear screener based on MACs.

Most of the proposed contextual features for describing chromatin arrangement can be applied at the nucleus level to quantitatively characterize the spatial arrangement of cells on a slide. This will be the subject of future work.

REFERENCES

- [1] DeMay, R. M., "Common problems in Papanicolaou smear interpretation," *Archives of Pathology Laboratory Medicine* **121**(3), 229–38 (1997).
- [2] Mehnert, A., "Image analysis for the study of chromatin distribution in cell nuclei with application to cervical cancer screening," PhD Thesis, School of Information Technology and Electrical Engineering, The University of Queensland (2004).
- [3] Guillaud, M., Cox, D., Adler-Storthz, K., Malpica, A., Staerkel, G., Maticic, J., van Niekerk, D., Poulin, N., Follen, M., and MacAulay, C., "Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: Objectivity, reproducibility, malignancy-associated changes, and human papillomavirus," *Cytometry Part A* **60A**(1), 81–89 (2004).
- [4] Mommers, E., Poulin, N., Meijer, C. M., Baak, J. A., and van Diest, P. J., "Malignancy-associated changes in breast tissue detected by image cytometry," *Analytical Cellular Pathology* **20**, 187–195 (Jan 2000).
- [5] Sacile, R., Montaldo, E., Ruggiero, C., Nieburgs, H., and Nicolo, G., "A decision support system to detect morphologic changes of chromatin arrangement in normal-appearing cells," *IEEE Transactions on NanoBio-science* **2**, 118–123 (June 2003).
- [6] Rodenacker, K. and Bengtsson, E., "A feature set for cytometry on digitized microscopic images," *Analytical Cellular Pathology* **25**(1), 1–36 (2003).
- [7] Niwas, S., Palanisamy, P., Sujathan, K., and Bengtsson, E., "Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using complex Daubechies wavelets," *Signal Processing* (2012).
- [8] Peura, M. and Iivarinen, J., [*Efficiency of Simple Shape Descriptors*], World Scientific (1997).
- [9] Doudkine, A., MacAulay, C., Poulia, N., and Palcic, B., "Nuclear texture measurement in image cytometry," *Pathologica* **87**, 286–299 (1995).

- [10] Young, I. T., Verbeek, P. W., and Mayall, B. H., "Characterization of chromatin distribution in cell nuclei," *Cytometry* **7**(5), 467–74 (1986).
- [11] Sameti, M., Ward, R., Morgan, J., and Palcic, B., "Image feature extraction in the last screening mammograms prior to detection of breast cancer," *IEEE J Selected Topics Signal Processing* **3**(1), 46–52 (2009).
- [12] Okabe, A., Boots, B., Sugihara, K., Chiu, S. N., and Kendall, D. G., [*Definitions and Basic Properties of Voronoi Diagrams*], John Wiley and Sons, Inc. (2008).
- [13] Toussaint, G. T., "The relative neighbourhood graph of a finite planar set," *Pattern Recognition* **12**(4), 261–268 (1980).
- [14] Jaromczyk, J. and Toussaint, G., "Relative neighborhood graphs and their relatives," *Proceedings of the IEEE* **80**, 1502–1517 (Sep 1992).
- [15] Diane, S., Ritu, N., Kurman, R., Davey, D., and Wilbur, D., [*The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*], 2nd edition NewYork, Springer (2004).
- [16] Moshavegh, R., Ehteshami Bejnordi, B., Mehnert, A., Sujathan, K., Malm, P., and Bengtsson, E., "Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis," *Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2012)* , 5372–5375 (Sep 2012).
- [17] Kai, B. D., Rajapakse, J., Haiying, W., and Azuaje, F., "Multiple SVM-RFE for gene selection in cancer classification with expression data," *IEEE Transactions on NanoBioscience* **4**, 228–234 (Sep 2005).
- [18] Park, M. Y. and Hastie, T., "L1-regularization path algorithm for generalized linear models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**(4), 659–677 (2007).
- [19] Deng, H. and Runger, G., "Gene selection with guided regularized random forest," *CoRR* , arxiv, 28 Sep 2012, <http://arxiv.org/abs/1209.6425> (30 Sep 2012).
- [20] Deng, H. and Runger, G., "Feature selection via regularized trees," *CoRR* , arxiv, 21 March 2012, <http://arxiv.org/abs/1201.1587> (12 April 2012).
- [21] Markowetz, F. and Spang, R., "Molecular diagnosis. classification, model selection and performance evaluation," *Methods of Information in Medicine* **44**(3), 438–443 (2005).

Bibliography

- [1] J. Ferlay, H. Shin, F. Bray, D. Forman, C. Mathers, Globocan 2008, cancer incidence and mortality worldwide (2010).
- [2] F. Bray, B. Carstensen, H. Moller, M. Zappa, M. P. Zakelj, G. Lawrence, M. Hakama, E. Weiderpass, Incidence trends of adenocarcinoma of the cervix in 13 european countries., *Cancer Epidemiol Biomarkers Prev* 14 (9) (2005) 2191–2199.
- [3] P. Boyle, B. Levin, World cancer report 2008, *Cancer Control* 199 (Book, Edited) (2008) 512.
- [4] Official statistics of Sweden, Cancer incidence in sweden 2010, The 53st annual report on the incidence of cancer in Sweden.
- [5] R. DeMay, Common problems in papanicolaou smear interpretation, *Archives of Pathology & Laboratory Medicine* 121 (1997) 229–238.
- [6] M. F. Janicek, H. E. Averette, Cervical cancer: prevention, diagnosis, and therapeutics, *CA: a cancer journal for clinicians* 51 (2001) 92–114.
- [7] L. G. Koss, The papanicolaou test for cervical cancer detection. a triumph and a tragedy, *Jama The Journal Of The American Medical Association* 261 (1989) 737–743.
- [8] G. Papanicolaou, H. Traut, Diagnosis of uterine cancer by the vaginal smear, *New York* (1943) 46.
- [9] G. Papanicolaou, H. Traut, The diagnostic value of vaginal smears in carcinoma of the uterus. 1941., *Archives of pathology & laboratory medicine* 121 (3) (1997) 211.
- [10] D. Penney, J. Powers, M. Frank, C. Willis, C. Churukian, Analysis and testing of biological stains—the biological stain commission procedures, *Biotechnic & histochemistry* 77 (5-6) (2002) 237–275.

-
- [11] N. Thekkek, R. Richards-Kortum, Optical imaging for cervical cancer detection: solutions for a continuing global problem, *Nature Reviews Cancer* 8 (9) (2008) 725–731.
- [12] A. Khumsangmeun, S. Tangjitgamol, S. Pratumkaew, Liquid-based cytology for cervical cancer screening, *Vajira Medical Journal* 53 (3) (2010) 311–316.
- [13] J. Sherris, S. Wittet, A. Kleine, J. Sellors, S. Luciani, R. Sankaranarayanan, M. Barone, Evidence-based, alternative cervical cancer screening approaches in low-resource settings, *International perspectives on sexual and reproductive health* 35 (3) (2009) 147–152.
- [14] M. Arbyn, A. Anttila, J. Jordan, G. Ronco, U. Schenck, N. Segnan, H. Wiener, A. Herbert, L. von Karsa, European guidelines for quality assurance in cervical cancer screening. second edition—summary document., *Ann Oncol* 21 (3) (2010) 448–458.
- [15] M. Desai, Role of automation in cervical cytology, *Diagnostic Histopathology* 15 (7) (2009) 323–329.
- [16] L. Isenstein, D. Zahniser, M. Hutchinson, Combined malignancy associated change and contextual analysis for computerized classification of cervical cell monolayers., *Analytical cellular pathology: the journal of the European Society for Analytical Cellular Pathology* 9 (2) (1995) 83.
- [17] H. Nieburgs, R. Zak, D. Allen, H. Reisman, T. Clardy, Systemic cellular changes in material from human and animal tissues in presence of tumors, in: *Transactions of the Seventh Annual Meeting of the International Society of Cytology Council*, Vol. 137, 1959.
- [18] H. E. Nieburgs, A. F. Goldberg, Changes in polymorphonuclear leukocytes as a manifestation of malignant neoplasia, *Cancer* 22 (1) (1968) 35–42.
- [19] E. Bengtsson, B. Nordin, Densitometry, morphometry and textural analysis as tools in quantitative cytometry and automated cancer screening, *Automated Cervical Cancer Screening* (1994) 21–43.
- [20] A. Mehnert, Image analysis for the study of chromatin distribution in cell nuclei with application to cervical cancer screening, PhD Thesis, School of Information Technology and Electrical Engineering, The University of Queensland, 2004.
- [21] B. Palcic, 1994. nuclear texture: Can it be used as a nuclear texture: can it be used as a surrogate endpoint biomarker?, *Cell Biochem Supplement* 19 (1994) 40–46.
- [22] R. Moshavegh, B. Ehteshami Bejnordi, A. Mehnert, K. Sujathan, P. Malm, E. Bengtsson, Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis, 2012, pp. 5372–5375.

- [23] B. Ehteshami Bejnordi, R. Moshavegh, K. Sujathan, P. Malm, E. Bengtsson, A. Mehnert, Novel chromatin texture features for the classification of pap smears, 2013, to be published.
- [24] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man and Cybernetics* 9 (1) (1979) 62–66.
- [25] O. Lezoray, H. Cardot, Cooperation of color pixel classification schemes and color watershed: a study for microscopic images, *Image Processing, IEEE Transactions on* 11 (7) (2002) 783–789.
- [26] P. Bamford, B. Lovell, Unsupervised cell nucleus segmentation with active contours, *Signal Processing* 71 (2) (1998) 203–213.
- [27] A. Garrido, N. Perez de la Blanca, Applying deformable templates for cell image segmentation, *Pattern Recognition* 33 (5) (2000) 821–832.
- [28] H. S. Wu, J. Barba, J. Gil, A parametric fitting algorithm for segmentation of cell images, *Biomedical Engineering, IEEE Transactions on* 45 (3) (1998) 400–407.
- [29] S. F. Yang-Mao, Y. K. Chan, Y. P. Chu, Edge enhancement nucleus and cytoplasm contour detector of cervical smear images, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 38 (2) (2008) 353–366.
- [30] C. Lin, Y. Chan, C. Chen, Detection and segmentation of cervical cell cytoplasm and nucleus, *International Journal of Imaging Systems and Technology* 19 (3) (2009) 260–270.
- [31] P. Bamford, B. Lovell, A water immersion algorithm for cytological image segmentation, in: *University of Technology Sydney, Sydney, 1996*, pp. 75–79.
- [32] E. Bengtsson, C. Wahlby, J. Lindblad, Robust cell image segmentation methods, *Pattern Recognition and Image Analysis C/c of Raspoznavaniye Obrazov / Analiz Izobrazhenii.* 14 (2) (2004) 157–167.
- [33] N. Lassouaoui, L. Hamami, Genetic algorithms and multifractal segmentation of cervical cell images, in: *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on, Vol. 2, 2003*, pp. 1–4.
- [34] M. Plissiti, C. Nikou, A. Charchanti, Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering, *Information Technology in Biomedicine, IEEE Transactions on* 15 (2) (2011) 233–241.
- [35] L. Cohen, L. Vinet, P. Sander, A. Gagalowicz, Hierarchical region based stereo matching, in: *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR '89., IEEE Computer Society Conference on, 1989*, pp. 416–421.
- [36] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, *Journal of Electronic Imaging* 13 (1) (2004) 146–168.

-
- [37] R. D. Paget, Nonparametric markov random field models for natural texture images, Ph.D. thesis (1999).
- [38] M. Beil, T. Irinopoulou, J. Vassy, G. Wolf, A dual approach to structural texture analysis in microscopic cell images, *Computer methods and programs in biomedicine* 48 (3) (1995) 211–219.
- [39] M. Beil, Description of chromatin structures in cell nuclei, *Acta Stereol.* (1992) 129–134.
- [40] F. Albrechtsen, H. Schulerud, L. Yang, Texture classification of mouse liver cell nuclei using invariant moments of consistent regions, in: *Texture classification of mouse liver cell nuclei using invariant moments of consistent regions*, Springer, 1995, pp. 496–502.
- [41] R. F. Walker, P. T. Jackway, Statistical geometric features-extensions for cytological texture analysis, in: *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, Vol. 2, IEEE, 1996, pp. 790–794.
- [42] A. Doudkine, C. MacAulay, N. Poulia, B. Palcic, Nuclear texture measurement in image cytometry, *Pathologica* 87 (1995) 286–299.
- [43] F. Meyer, Quantitative analysis of the chromatin of lymphocytes: an assay on comparative structuralism., *Blood cells* 6 (2) (1980) 159.
- [44] P. Jackway, Improved morphological top-hat, *Electronics Letters* 36 (14) (2000) 1194–1195.
- [45] J. Cheng, J. Rajapakse, Segmentation of clustered nuclei with shape markers and marking function, *Biomedical Engineering, IEEE Transactions on* 56 (3) (2009) 741–748.
- [46] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, Inc., Orlando, FL, USA, 1983.
- [47] L. Koss, M. Melamed, *Koss' diagnostic cytology and its histopathologic bases*, Vol. 1, Lippincott Williams & Wilkins, 2006.
- [48] M. Peura, J. Iivarinen, Efficiency of simple shape descriptors, in: *Efficiency of simple shape descriptors*, Citeseer, 1997, pp. 443–451.
- [49] H. Tamura, S. Mori, T. Yamawaki, Textural features corresponding to visual perception, *Systems, Man and Cybernetics, IEEE Transactions on* 8 (6) (1978) 460–473.
- [50] D. Komitowski, G. Zinser, Quantitative description of chromatin structure during neoplasia by the method of image processing, *Analytical and Quantitative Cytology and Histology* 7 (3) (1985) 178–182.

- [51] A. Okabe, B. Boots, K. Sugihara, S. N. Chiu, D. G. Kendall, *Definitions and Basic Properties of Voronoi Diagrams*, John Wiley and Sons, Inc., 2008.
- [52] I. T. Young, P. W. Verbeek, B. H. Mayall, Characterization of chromatin distribution in cell nuclei, *Cytometry* 7 (5) (1986) 467–474.
- [53] M. Peura, J. Iivarinen, *Efficiency of Simple Shape Descriptors*, World Scientific, 1997, pp. 443–451.
- [54] K. Rodenacker, E. Bengtsson, A feature set for cytometry on digitized microscopic images, *Analytical Cellular Pathology* 25 (1) (2003) 1–36.
- [55] M. Guillaud, D. Cox, K. Adler-Storthz, A. Malpica, G. Staerker, J. Maticic, D. van Niekerk, N. Poulin, M. Follen, C. MacAulay, Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: Objectivity, reproducibility, malignancy-associated changes, and human papillomavirus, *Cytometry Part A* 60A (1) (2004) 81–89.
- [56] S. Niwas, P. Palanisamy, K. Sujathan, E. Bengtsson, Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using complex Daubechies wavelets, *Signal Processing* (2012).
- [57] M. Sameti, R. Ward, J. Morgan, B. Palcic, Image feature extraction in the last screening mammograms prior to detection of breast cancer, *IEEE J Selected Topics Signal Processing* 3 (1) (2009) 46–52.
- [58] S. Diane, N. Ritu, R. Kurman, D. Davey, D. Wilbur, *The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*, 2nd edition NewYork, Springer, 2004.
- [59] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, Springer series in statistics, Springer-Verlag New York, 2009.
- [60] K. B. Duan, J. C. Rajapakse, H. Wang, F. Azuaje, Multiple svm-rfe for gene selection in cancer classification with expression data, *NanoBioscience, IEEE Transactions on* 4 (3) (2005) 228–234.
- [61] M. Y. Park, T. Hastie, L1-regularization path algorithm for generalized linear models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69 (4) (2007) 659–677.
- [62] H. Deng, G. Runger, Gene selection with guided regularized random forest, *CoRR* (30 Sep 2012) arxiv, 28 Sep 2012, <http://arxiv.org/abs/1209.6425>.
- [63] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene Selection for Cancer Classification using Support Vector Machines, *Machine Learning* 46 (1-3) (2002) 389–422.
- [64] H. Deng., G. Runger, Feature selection via regularized trees, *CoRR* (12 April 2012) arxiv, 21 March 2012, <http://arxiv.org/abs/1201.1587>.

-
- [65] P. McCullagh, J. A. Nelder, *Generalized Linear Models*, Second Edition (Chapman & Hall/CRC Monographs on Statistics & Applied Probability), 2nd Edition, Chapman and Hall/CRC, 1989.
- [66] F. Markowetz, R. Spang, Molecular diagnosis. classification, model selection and performance evaluation, *Methods of Information in Medicine* 44 (3) (2005) 438–443.
- [67] F. Piccinini, A. Tesei, W. Zoli, A. Bevilacqua, Extended depth of focus in optical microscopy: Assessment of existing methods and a new proposal, *Microscopy Research and Technique* 75 (11) (2012) 1582–1592.
- [68] B. Forster, D. Van De Ville, J. Berent, D. Sage, M. Unser, Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images, *Microscopy Research and technique* 65 (1-2) (2004) 33–42.
- [69] F. Aguet, D. Van De Ville, M. Unser, Model-based 2.5-d deconvolution for extended depth of field in brightfield microscopy, *Image Processing, IEEE Transactions on* 17 (7) (2008) 1144–1153.
- [70] S. Liu, H. Hua, Extended depth-of-field microscopic imaging with a variable focus microscope objective, *Optics Express* 19 (1) (2011) 353–362.
- [71] A. D. N. Edwards, G. Hines, A. Hunt, Segmentation of biological cell images for sonification, in: *Proceedings of the 2008 Congress on Image and Signal Processing*, Vol. 2 - Volume 02, CISP '08, IEEE Computer Society, Washington, DC, USA, 2008, pp. 128–132.
- [72] A. A. Bell, J. N. Kaftan, T. E. Schneider, D. Meyer-Ebrecht, A. Böcking, T. Aach, Imaging and image processing for early cancer diagnosis on cytopathological microscopy images: Towards fully automatic agnor detection.
- [73] H. J. A. M. Heijmans, *Morphological Image Operators*, Advances in Electronics and Electron Physics. Supplement 25, Academic Press, Boston, 1994.
- [74] J. Bernsen, Dynamic thresholding of grey-level images, in: *Dynamic thresholding of grey-level images*, Vol. 1, 1986, pp. 1252–1255.
- [75] J. Sauvola, M. Pietikainen, Adaptive document image binarization, *Pattern Recognition* 33 (2000) 225–236.
- [76] W. Niblack, *An introduction to image processing*, Englewood Cliffs, New York (1986) 115–116.
- [77] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, *International journal of computer vision* 1 (4) (1988) 321–331.

- [78] S. Menet, P. Saint-Marc, G. Medioni, Active contour models: overview, implementation and applications, in: Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on, 1990, pp. 194–199.
- [79] L. Cohen, I. Cohen, Finite-element methods for active contour models and balloons for 2-d and 3-d images, Pattern Analysis and Machine Intelligence, IEEE Transactions on 15 (11) (1993) 1131–1147.
- [80] C. Xu, J. Prince, Snakes, shapes, and gradient vector flow, Image Processing, IEEE Transactions on 7 (3) (1998) 359–369.
- [81] C. Xu, J. L. Prince, Generalized gradient vector flow external forces for active contours, Signal Processing 71 (2) (1998) 131–139.
- [82] A. Ogier, T. Dorval, B. Kim, A. Genovesio, Spatially adaptive relaxation for active contour cell segmentation, in: Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on, 2010, pp. 1033–1036.
- [83] A. Sanpanich, W. Iampa, C. Pintavirooj, P. Tosranon, White blood cell segmentation by distance mapping active contour, in: White blood cell segmentation by distance mapping active contour, IEEE, 2008, pp. 251–255.
- [84] J. Tang, A multi-direction gvf snake for the segmentation of skin cancer images, Pattern Recognition 42 (6) (2009) 1172–1179.
- [85] K. Zhang, L. Zhang, H. Song, W. Zhou, Active contours with selective local or global segmentation: A new formulation and level set method, Image and Vision Computing 28 (4) (2010) 668–676.
- [86] R. Adams, L. Bischof, Seeded region growing, Pattern Analysis and Machine Intelligence, IEEE Transactions on 16 (6) (1994) 641–647.
- [87] A. Mehnert, P. Jackway, An improved seeded region growing algorithm, Pattern Recognition Letters 18 (10) (1997) 1065–1071.
- [88] S. Beucher, C. Lantuejoul, Use of watersheds in contour detection, in: Use of watersheds in contour detection, 1979, pp. 2.1–2.12.
- [89] F. Meyer, S. Beucher, Morphological segmentation, Journal of visual communication and image representation 1 (1) (1990) 21–46.
- [90] B. Kamgar-Parsi, Improved image thresholding for object extraction in ir images, in: Improved image thresholding for object extraction in IR images, Vol. 1, IEEE, 2001, pp. 758–761.
- [91] G. Borgefors, Distance transformations in digital images, Computer vision, graphics, and image processing 34 (3) (1986) 344–371.

- [92] O. Lezoray, H. Cardot, Bayesian marker extraction for color watershed in segmenting microscopic images, in: Bayesian marker extraction for color watershed in segmenting microscopic images, Vol. 1, IEEE, 2002, pp. 739–742.
- [93] P. Soille, Morphological image analysis: principles and applications, Springer-Verlag New York, Inc., 2003.
- [94] F. Meyer, Contrast features extraction, special issues of practical metallography, Vol. 8, 1977.
- [95] P. Danielson, A new shape factor, Computer Graphics and Image Processing 7 (2) (1978) 292–299.
- [96] L. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.
- [97] I. N. Bronshtein, K. A. Semendyayev, G. Musiol, H. Muehling, Handbook of Mathematics (4th Edition), Springer - Verlag, 2004.
- [98] T. Fletcher, Support vector machines explained, Tech. rep., University College London (UCL) (2009).

