Image segmentation with prior guidance: application to cervical cytology images

by

Ratna Saha

Thesis
Submitted to Flinders University
for the degree of

Doctor of Philosophy

College of Science and Engineering

March 2020

Contents

List of abbreviations **xiv**

Abstract **xv**

Declaration **xvii**

Publications **xviii**

Dedication **xx**

Acknowledgements **xxi**

1 Introduction **1**

1.1 Motivation 1

1.2 Cervical cancer 5

1.3 Screening 8

1.4 Automated screening systems 9

1.5 Nuclei segmentation in automated screening 13

1.6 Prior knowledge in segmentation 16

1.7 Research objectives 17

1.8 Thesis outline 17

2 Literature review **19**

2.1 Computer-aided cervical cytology screening 19

2.1.1 Preprocessing 19

2.1.2 Image segmentation 24

2.1.3 Feature generation 33

2.1.4 Feature selection 35
2.1.5 Classification 37
2.2 Prior guided image segmentation 40
2.2.1 User interaction 40
2.2.2 Shape and appearance prior 41
2.2.3 Statistical shape model 43
2.2.4 Spatial prior 45
2.2.5 Boundary information 48
2.3 Related works 49
2.3.1 Thresholding and morphological analysis based approaches 49
2.3.2 Deformable model, contour and shape based approaches . 51
2.3.3 Watershed based segmentation approaches 53
2.3.4 Region and clustering based approaches 54
2.3.5 Machine learning based segmentation 55

3 Shape prior in fuzzy c-means clustering 57
3.1 Introduction 57
3.2 Circular shape constrained fuzzy clustering (CiscFC) 59
3.2.1 An overview 59
3.2.2 The proposed circular shape function (CSF) 60
3.3 Dataset 62
3.4 Implementation 63
3.4.1 Background subtraction 64
3.4.2 Implementation of CiscFC 67
3.4.3 False positive reduction 68
3.4.4 Nucleus shape regularization 69
3.5 Evaluation metrics 70
3.6 Experimental setup 71
3.7 Background subtraction technique selection 72
3.8 Parameter tuning 73
3.8.1 Thresholds for area and shape features 73
3.8.2 Thresholds for CSF 74

3.8.3 Optimal number of clusters 77
3.9 Results 78
3.10 Discussion 81
3.11 Conclusion 86

4 Shape prior in graph based segmentation 88
4.1 Introduction 88
4.2 Circular shape prior in efficient graph based segmentation (CircEGS) 90
4.2.1 An overview of EGS method 90
4.2.2 The proposed circular shape prior 91
4.3 Dataset 93
4.4 Evaluation metrics 93
4.5 Experimental setup 95
4.6 Parameter tuning 95
4.7 Results 97
4.7.1 Quantitative evaluation 97
4.7.2 Border precision evaluation 100
4.8 Discussion 101
4.9 Conclusion 104

5 Superpixel merging with gradient guidance 107
5.1 Introduction 107
5.2 Dataset 108
5.3 Methodology 109
5.3.1 Nuclei marker detection 109
5.3.2 Superpixel merging criteria 112
5.3.3 Superpixel generation 112
5.3.4 A novel superpixel merging technique with pairwise regional contrast and gradient boundary 113
5.3.5 Shape regularization 114
5.4 Evaluation metrics 115
5.5 Experimental setup 115

5.6 Finding thresholds for parameters 116
5.7 Results 119
5.8 Discussion 121
5.8.1 SRM superpixels 122
5.9 Summary and conclusion 129

6 Abnormality detection in cervical cells 131
6.1 Introduction 131
6.2 Dataset 132
6.3 Methodology 133
6.3.1 Nucleus segmentation 134
6.3.2 Feature extraction and normalization 135
6.3.3 Feature selection 136
6.3.4 Classifiers 137
6.4 Evaluation measures 137
6.5 Experimental results and discussion 139
6.6 Conclusion 147

7 Conclusion 149
7.1 Thesis summary and contribution 149
7.2 Future work 151

List of figures

1.1 Pie charts representing the estimated number of (a) new cases and (b) deaths for the most common cancers in women worldwide, in 2018. Data source: GLOBOCAN 2018. Graph production: Global cancer observatory (`https://gco.iarc.fr/`). 2

1.2 Cervical cancer incidence in Australia from 1991 to 2017. Data source: Australian Institute of Health and Welfare (AIHW), Australian Cancer Database. 3

1.3 Annual rates of new cervical cancer incidence in United States from 1999 to 2016. Data source: US Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute (`www.cdc.gov/cancer/dataviz`). 4

1.4 Bar chart representing the estimated age-standardized incidence and death rates due to cervical cancer in women, based on income levels. Data source: GLOBOCAN 2018. Graph production: Global cancer observatory (`https://gco.iarc.fr/`). 5

1.5 Anatomical illustration of female reproductive system. Adapted from (Malm 2013, Bengtsson & Malm 2014). 6

1.6 Development of squamous cells through different layers. Adapted from (Sellors & Sankaranarayanan 2003). 7

1.7 Illustration of endocervical columnar epithelium. Adapted from (Sellors & Sankaranarayanan 2003). 7

1.8 Examples of Pap smear images. Data source: Overlapping Cervical Cytology Image Segmentation Challenge - ISBI 2014 (see Chapter 3 for details). 9

2.1 An example of star shape object. 42

3.1 Four examples of overlapping synthetic cervical smear images from ISBI 2014 challenge dataset (in light background) (a), (c), (e), and (g), followed by corresponding nuclei ground truths (in dark background) (b), (d), (f), and (h). 63

3.2 The flowchart of the proposed cervical nuclei segmentation framework. 64

3.3 Cervical cell image background subtraction using thresholding based technique, (a) original image, (b) complemented image, (c) plot of the complemented image histogram with dominant peaks (marked with red $*$), and, (d) binarized image using the second highest peak. 65

3.4 Normalized histograms of: (a) an original image from the training dataset, (b) same image after applying morphological opening-by-reconstruction and closing-by-reconstruction. 66

3.5 Background extraction and finding cell clump, (a) original image, (b) morphologically processed image, (c) foreground cell clump, and (d) complement image of the cell clump. 67

3.6 An example of false positive reduction. (a) Complemented foreground image used for clustering (nuclei with high graylevel values), (b) detected nucleus candidates from clustering output, and (c) nucleus cluster after false positive reduction process using area and shape features. 69

3.7 An example of nucleus shape regularization. (a) Complemented foreground image, (b) nucleus candidates after false positive reduction, and (c) nucleus candidates after shape regularization. . . 70

3.8 Examples of two background removal techniques: top row represents results for thresholding based and bottom row presents morphological reconstruction based, (a), (d) complemented foreground images, (b), (e) brightest clusters as nucleus, and (c), (f) nucleus clusters after false positive reduction. Differences are highlighted with yellow rectangles. 73

3.9 Comparison of background subtraction techniques with two example images. Top row presents image foregrounds found with threshold based method and final nucleus clusters, and bottom row shows foregrounds found using morphological reconstruction based technique and corresponding nuclei clusters. 74

3.10 Determination of t_1 and t_2 using Otsu multiple thresholding. (a), (d) Complemented foreground images, (b), (e) candidates in nucleus class after thresholding, and (c), (f) filtered candidates. . . . 75

3.11 MSER technique used to determine t_1 and t_2. (a), (d) Complemented foreground images, (b), (e) candidates after applying MSER algorithm, and (c), (f) area and shape filtered candidates. 76

3.12 Plot of segmentation accuracy measures of the proposed framework for ISBI test dataset: (a) Dice similarity coefficient, (b) pixel based precision and (c) pixel based recall. 82

3.13 Histograms of circularity values of all ground truth nuclei, from ISBI (a) training and (b) test datasets. 83

3.14 Circularity value histograms of segmented nucleus using the proposed framework with: (a) standard FCM and (b) CiscFC techniques, from ISBI test dataset. 84

3.15 Nucleus segmentation examples from the test dataset: first column represents the original images marked with nuclei ground truth, second column presents nuclei segmented using the standard FCM clustering, and the last column shows nuclei segmented using the proposed CiscFC technique. Differences between CiscFC and the standard FCM are marked with green rectangles. 85

3.16 An example of missed nucleus detection, (a) nucleus in question (circled in red), (b) detected target nucleus (circled in red) after applying the proposed method and (c) target nucleus removed after false positive reduction. 86

3.17 Examples of missed nucleus detection and segmentation due to uneven staining and poor contrast. 86

4.1 Schematic illustration of circularity definition. 92

4.2 Plots of evaluation measures for k in the range $[50, 650]$ using ISBI training set using, (a) EGS and (b) CircEGS techniques. 96

4.3 Plot of H_t metric for EGS and CircEGS techniques. 100

4.4 Circularity histograms of: (a) ground truth, and nuclei segmented using (b) standard EGS and (c) CircEGS techniques. 102

4.5 Nucleus segmentation examples. First column presents original images marked with nucleus ground truth, second column shows the segmented boundaries using the EGS method, and third column contains the segmentation boundaries using CircEGS technique. Differences between CircEGS and EGS techniques are marked with green rectangles. 105

4.6 Examples of nucleus segmentation failures: top row shows nucleus ground truth boundaries, and bottom row presents corresponding CircEGS technique segmented nucleus boundaries. Failed segmentation cases and nucleus in ground truth, are indicated with red arrows. 106

5.1 The flowchart of the proposed nucleus segmentation framework. . 109

5.2 Visual representation of (a) a 8 bit Pap smear image, and its 8 bit planes as: (b) bit 1 plane, (c) bit 2 plane, (d) bit 3 plane, (e) bit 4 plane, (f) bit 5 plane, (g) bit 6 plane, (h) bit 7 plane, and (i) bit 8 plane. 110

5.3 An example of nucleus marker detection process: (a) original image, (b) MSB plane with marker candidates, (c) MSB plane after applying morphological operation and filtering, (d) MSER output with nuclei marker candidates, (e) MSER output after applying morphological operation and filtering, (f) centroids of final nuclei candidates - found by combining outputs in (c) and (e), are marked with red stars and used as nuclei markers in the proposed framework. 111

5.4 An example of gradient boundary generation, where blue square in the center indicates nucleus marker, green lines represent intensity profiles, and red stars stand for gradient boundary points. 113

5.5 Test cases for a neighbor superpixel during the proposed merging process. The centroid of a neighbor superpixel (in green) is: (a) inside, (b) on and (c) outside, the gradient boundary (in white). . 114

5.6 Superpixel merging process illustration. (a) Parent superpixel (in blue) containing nucleus marker, (b) - (e) show neighbor superpixels (in green) tested for merging with the parent, and (f) is the final merged superpixel as segmented nucleus. 116

5.7 Plot of Dice similarity coefficient for pairwise regional contrast threshold in the range $[256, 2306]$ with step of 50, using the proposed superpixel merging framework on ISBI training set. 118

5.8 Plot of Dice similarity coefficient for Euclidean distance in the range $[1, 20]$, using the proposed superpixel merging framework on ISBI training set. 119

5.9 Estimation of Q values for SRM segmentation. (a) Finding Q_s for coarse SRM segmentation, and finding Q_b from (b) average and (c) maximum region size analysis for finer SRM segmentation. . . 124

5.10 Plot of (a) Dice similarity coefficient (DSC), (b) pixel-based precision (PR_{pix}) and (c) pixel-based recall (RC_{pix}) - for the proposed framework using SRM and SLIC methods. 127

5.11 Visual comparison of ground truth (in black, left column), and segmented boundaries found by the proposed superpixel merging framework with SLIC (in yellow, middle column) and SRM (in blue, right column) superpixels. Differences in segmented boundaries found with SLIC and SRM superpixel generation techniques are highlighted with green rectangles. 128

5.12 Examples of failure cases with proposed framework with SRM superpixels. Top row presents failures due to missed marker detection, and bottom row shows failed cases for superpixel generation. Ground truth boundaries are in black, segmented boundaries are in blue, and differences are indicated with red arrows. 129

6.1 Example of cervical cell images from Herlev dataset: (a)-(c) normal and (d)-(g) abnormal. 133

6.2 Block diagram of the proposed cervical cell classification framework. 134

6.3 Comparison of ground truth (in black, first column), and segmented nuclei boundaries found using: CircEGS (in white, second column), CiscFC (in cyan, third column), and SPmerg (in yellow, last column) techniques. 140

6.4 Plot of AUC over number of ranked features with LDA, KNN, SVM-linear, Ensemble-Bag, and SVM-RBF classifiers, for Herlev dataset segmented with: (a) CircEGS, (b) CiscFC and (c) SPmerg techniques. 144

6.5 ROC curves illustrating the best and lowest AUC scores found in two folds of classification in a run, with CircEGS segmented nuclei and SVM linear classifier. 146

List of tables

3.1 Parameter values for area and shape features of a candidate nucleus. 74

3.2 Finding the optimal number of clusters. 77

3.3 Quantitative object based evaluation of the proposed CiscFC technique in comparison with the standard FCM and recent state-of-the-art methods. 79

3.4 Quantitative comparison of segmentation accuracy of the proposed CiscFC methodology, with the standard FCM and recent state-of-the-art techniques. 80

3.5 Evaluation of shape regularization impact on performance measures. 80

3.6 Identifying stability of the proposed CiscFC framework with respect to different initial partition matrices. Last row shows the mean and standard deviation over 5 runs. 81

4.1 Empirical values for parameter k and weight parameter λ. 97

4.2 Object level evaluation of the proposed CircEGS technique, in comparison with the standard EGS and recent state-of-the-art methods. 98

4.3 Pixel level evaluation of the proposed CircEGS technique. 99

4.4 Hausdorff distance evaluation. 100

4.5 H_t metric with tolerance threshold t set to 1, 2, ..., 5 (pixels). . . 101

4.6 Influence of circularity on nucleus segmentation measures, in terms of DSC, pixel based precision (PR_{pix}) and recall (RC_{pix}), and Hausdorff distance (H). 103

5.1 Threshold values to detect candidate nuclei markers. 117

5.2 Object based evaluation of the proposed superpixel merging framework and comparison with the state-of-the-art methods. 120

5.3 Evaluation of segmentation accuracy measures for the proposed superpixel merging framework, in comparison to the state-of-the-art techniques. 121

5.4 Comparison of nucleus segmentation and detection evaluation for the proposed superpixel merging framework with SLIC and SRM techniques. 125

6.1 An example of confusion matrix for a two class problem. 138

6.2 Evaluation of nucleus segmentation and detection for Herlev dataset. 141

6.3 Comparison of nucleus segmentation performance of the proposed framework, with the state-of-the-art methods using Herlev dataset. 142

6.4 Fuzzy entropy based hierarchical ordering/ranking of 13 features of 5-folds in a run. 143

6.5 Two-class classification performance of Herlev dataset - segmented with CircEGS, CiscFC and SPmerg methods, with 10 independent runs of 5-fold cross validation and 10 nuclei features. 145

6.6 Confusion matrices for the best classifiers in 2 class classification problem of Herlev dataset with 10 features, accumulated over 5-folds in a run. 145

6.7 Classification results of 5-folds in a run, found for CircEGS segmentation and SVM linear classification with 10 nuclei features. . 145

6.8 Comparison of Herlev dataset classification performance of the proposed approach with state-of-the-art techniques. 147

List of algorithms

1 Efficient graph based segmentation algorithm 92
2 Superpixel merging algorithm 115

List of abbreviations

AUC	Area under ROC curve
CAD	Computer-aided diagnosis
CLAHE	Contrast limited adaptive histogram equalization
EGS	Efficient graph based segmentation
FCM	Fuzzy c-means
HPV	Human papilloma virus
KNN	K-nearest neighbor
LDA	Linear discriminant analysis
MSER	Maximally stable extremal region
MST	Minimum spanning tree
ROC	Receiver operator characteristics
ROI	Region of interest
SLIC	Simple linear iterative clustering
SRM	Statistical region merging
SVM	Support vector machine

Abstract

Cervical cancer is one of the most common type of gynecological cancer. In cervix tissues, pre-cancerous changes in cell morphology and structure develop over a span of 8 to 10 years. Pap smear is the most effective and popular screening worldwide for early detection of invasive cervical cancer. However, manual screening task of cytotechnologist is complex, time-intensive, tedious, and requires high level of expertise. Computer-aided screening of cervical smear images has the potential to assist the cytotechnologists and improve Pap test outcome, in a turn decrease cervical cancer morbidity and mortality rate. Nucleus features are crucial for characterization of cervical cytology images and diagnostic decisions. Hence, accurate segmentation of nuclei is the essential primary step towards computer-assisted cervical cell analysis. This study aims to improve cervical nuclei segmentation performance, and subsequently develop an abnormality detection or cervical cell classification framework exclusively based on nuclei features.

The main contributions of this thesis are the development of three novel nucleus segmentation techniques and a nuclei feature based classification framework. The first two segmentation frameworks are developed on pixel grid level and the last one on superpixel based image representation. Incorporation of prior knowledge in image segmentation techniques is useful in partitioning images with noise and low contrast. The proposed segmentation algorithms are designed to utilize prior guidance, for precise segmentation of cervical nuclei from overlapping Pap smear images.

Particularly, in the first contribution, a circular shape function is proposed at pixel level, utilizing spatial location of pixels. This shape function was integrated into fuzzy clustering technique. In the second contribution, circular shape guidance is included in efficient graph based segmentation algorithm. The adaptive shape guidance is designed at pixel level to include circularity measure of regions, in addition to image intensity information. In the third contribution, a gradient guided superpixel level merging framework is proposed. The novel merging criterion consists of pairwise regional contrast and gradient boundary guidance. Superpixels produced by SLIC and SRM segmentation techniques are employed in this merging framework. Lastly, a nuclei feature based classification framework

is proposed for abnormality detection in cervical cells, where nuclei are segmented using all three proposed segmentation techniques.

The effectiveness of the proposed segmentation approaches is validated on ISBI 2014 and Herlev datasets, using Dice similarity coefficient, pixel and object based precision and recall as performance measures. The experimental results indicate that the proposed segmentation techniques can precisely segment nuclei from overlapping cervical cytology images, while keeping high level of precision and recall. Performance of the proposed classification framework is validated on Herlev dataset, and evaluated using accuracy, sensitivity, specificity, AUC score, and Spearman rank order coefficient. Promising classification results of the proposed framework indicate that, in the presence of precisely segmented nuclei boundaries, it is possible to characterize cervical cells using only nuclei features.

Declaration

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Ratna Saha

Refereed publications from the study

Refereed conference papers

1. Ratna Saha, Mariusz Bajger and Gobert Lee, SRM superpixel merging framework for precise segmentation of cervical nucleus, *in* '2019 International Conference on Digital Image Computing: Techniques and Applications (DICTA)', Perth, WA, 2-4 Dec, 2019, pp. 1-8.

2. Ratna Saha, Mariusz Bajger and Gobert Lee, Prior guided segmentation and nuclei feature based abnormality detection in cervical cells, *in* '2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)', Athens, Greece, 28-30 Oct, 2019, pp. 742-746.

3. Ratna Saha, Mariusz Bajger and Gobert Lee, Circular shape prior in efficient graph based image segmentation to segment nucleus, *in* '2018 International Conference on Digital Image Computing: Techniques and Applications (DICTA)', Canberra, ACT, 10-13 Dec, 2018, pp. 1-8.

4. Ratna Saha, Mariusz Bajger and Gobert Lee, Segmentation of cervical nuclei using SLIC and pairwise regional contrast, *in* '2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)', Honolulu, HI, 18-21 July, 2018, pp. 3422-3425.

5. Ratna Saha, Mariusz Bajger and Gobert Lee, Spatial Shape Constrained Fuzzy C-Means (FCM) Clustering for Nucleus Segmentation in Pap Smear Images, *in* '2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)', Gold Coast, QLD, 30 Nov-2 Dec, 2016, pp. 1-8.

Refereed journal papers

6. Ratna Saha, Mariusz Bajger, Gobert Lee, 'Circular shape constrained fuzzy clustering (CiscFC) for nucleus segmentation in Pap smear images', *Computers in Biology and Medicine*, **85**:13-23, June 2017.

Dedication

To my father Narendra Nath Saha for showering his blessings from heaven and always being my courage. To my mother Shila Rani Saha for her unconditional and endless love, support and being the source of my strength.

Acknowledgements

This thesis would not have been completed without the help, support and encouragement of some wonderful people around me.

First and foremost, I would like to express my sincere gratitude to my supervisor Dr Mariusz Bajger, who guided and supported me throughout my PhD study, and never stopped trusting in me. He taught me to think critically and mentored me to be a researcher. His endless support, patience, encouragement, and constructive discussions, contributed to the accomplishment of this thesis.

I would like to express my sincere appreciation to my co-supervisor Dr Gobert Lee for her guidance, support and invaluable discussions. She always tried to show me a ray of light to overcome the challenges I faced during my PhD journey, and motivated me to complete this thesis.

I would also like to thank Australian Government Research Training Program Scholarship for the fee offset, and College of Science and Engineering, Flinders University for the College of Science and Engineering Research Award (CSERA). I am also thankful to Flinders University staff for giving me the opportunity to work as a casual academic.

I have been fortunate enough to get immense support, inspiration and motivation from my mother, sisters, extended family, my nephews and nieces, and my amazing friends in Australia and Bangladesh. Words are not enough to express my gratefulness.

I would like to mention the most wonderful daughter Ihita Saha and the most amazing husband Lipon Saha. I do not want to thank you for your unconditional love, care and patience during this tough journey and in all days of highs and lows, but I want to tell that you are my world. Without your love, I would not dare to dream this far.

Last but not the least, thanks to the almighty for abundance of blessings; the reason of being me.

Chapter 1

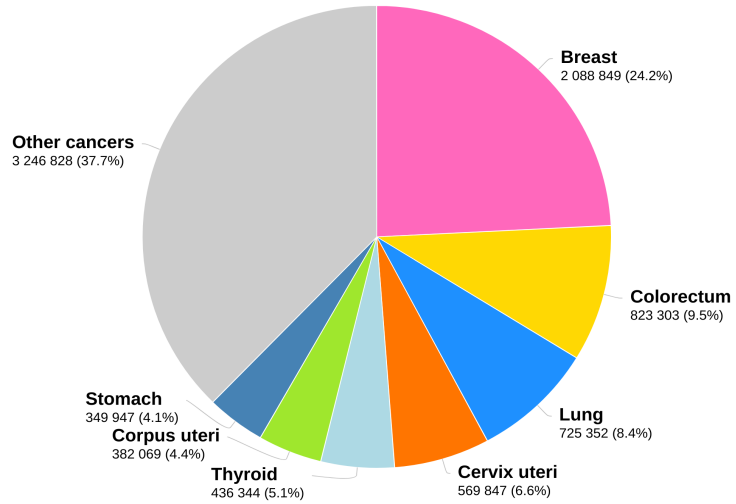
Introduction

This introductory chapter provides the motivation for computer-aided screening system or abnormality detection from cervical cytology images. Section 1.2 and Section 1.3 present brief introduction to cervical cancer and screening systems. Computerized techniques to detect abnormality in cervical cells are presented in Section 1.4. Section 1.5 briefly describes nucleus segmentation techniques in computer-aided screening, and incorporation of prior knowledge in segmentation technique is presented in Section 1.6. Research objectives are described in Section 1.7. Finally, Section 1.8 provides an outline of this thesis.

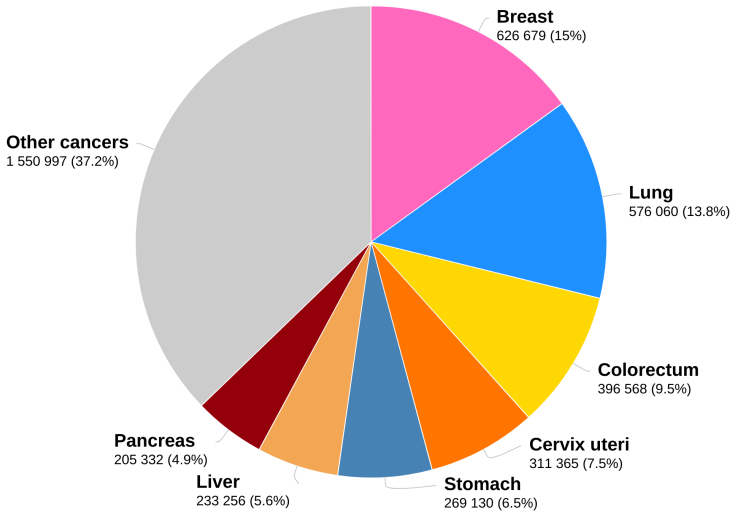
1.1 Motivation

Cervical cancer is the fourth principal cause of female cancer death worldwide (*Cervical cancer statistics* 2019, Bray et al. 2018). World Health Organization (*WHO — Cervical cancer* 2019) and GLOBOCAN 2018 (Bray et al. 2018) estimated $570,000$ (6.6%) new cervical cancer incidences and $311,000$ (7.5%) deaths in 2018. The estimated number of incidence and mortality rates due to common cancers in women in the year 2018 are presented in Figure 1.1, where 'cervix uteri' indicates the instances of cervical cancer (2019 data was not available at the time of writing this thesis).

Cervix tissues go through pre-cancerous or dysplastic changes for long period of time (generally 8 - 10 years) before true development or progression to cancer (Li et al. 2012, Australian Institute of Health and Welfare & Cancer Australia 2012). If abnormality in cervical cells can be detected or diagnosed at early stages, it can be cured and the treatment can be relatively simple, non-invasive



(a)



(b)

Figure 1.1: Pie charts representing the estimated number of (a) new cases and (b) deaths for the most common cancers in women worldwide, in 2018. Data source: GLOBOCAN 2018. Graph production: Global cancer observatory (`https://gco.iarc.fr/`).

and inexpensive. Therefore, early detection of pre-cancerous changes in cervix tissues is important and can help to reduce mortality and morbidity rates due to cervical cancer. However, cervix tissues rarely exhibit any physical symptoms until late stages. Therefore, timely detection of cervical cancer is possible only by performing regular screening.

The Pap smear test (or Papanicolaou test) is the most effective and standard screening test for cervical cancer (Australian Institute of Health and Welfare & Cancer Australia 2012, American Cancer Society 2019). Greek doctor

Georges N. Papanicolaou (Papanicolaou 1942) introduced this test, to detect abnormal changes in cervix. For the test, sample is collected from the cervix, then smeared and fixed on a glass slide, and stained following the standard procedure guideline. The stained microscopic slides are then visually examined by cytotechnologists for evidence of abnormalities in cell morphology and structure.

Over the past few decades, population-based regular screening programs have helped to reduce the rate of cervical cancer incidence and deaths in many countries worldwide (Bray et al. 2018). In Australia, national Pap test screening program from 1991 to 2017 helped to decrease the incidence of new cases from 13.3 to 7.1 per $100,000$ women (see Figure 1.2). Infection with human papilloma virus

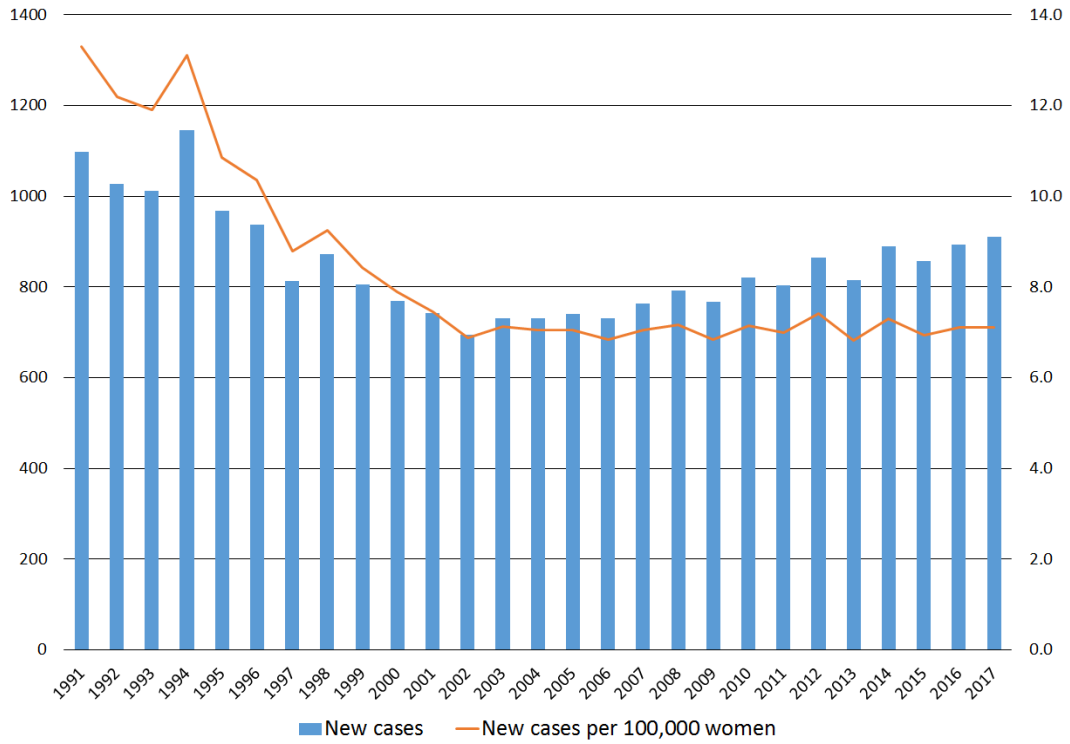


Figure 1.2: Cervical cancer incidence in Australia from 1991 to 2017. Data source: Australian Institute of Health and Welfare (AIHW), Australian Cancer Database.

(HPV) is the primary cause in most of the cervical cancer cases (Australian Institute of Health and Welfare 2019). In light of that and from the success rate of Pap test screening for around 2 decades, Australia's 2-yearly Pap smear test screening has been replaced with a 5-yearly Cervical Screening Test, from 01 December, 2017. The new test is composed of HPV testing and Pap smear test, to get more accurate results with less frequent screening. First through HPV testing, cytologists will search for any presence of HPV in the sample of cells collected from the cervix. By doing this, it will be possible to detect traces of

HPV earlier, before it infects cervical cells and causes cancer. If any trace of HPV is found, then Pap smear test will be performed to detect the cellular changes.

In United States (US), cervical cancer incidence rate has also declined significantly, as a result of regular Pap test (from 9.7 to 7.7 new cases per 100,000 women). Figure 1.3 shows this trend of change in cervical cancer incidence over years in United States. While Australia, New Zealand, North America and many

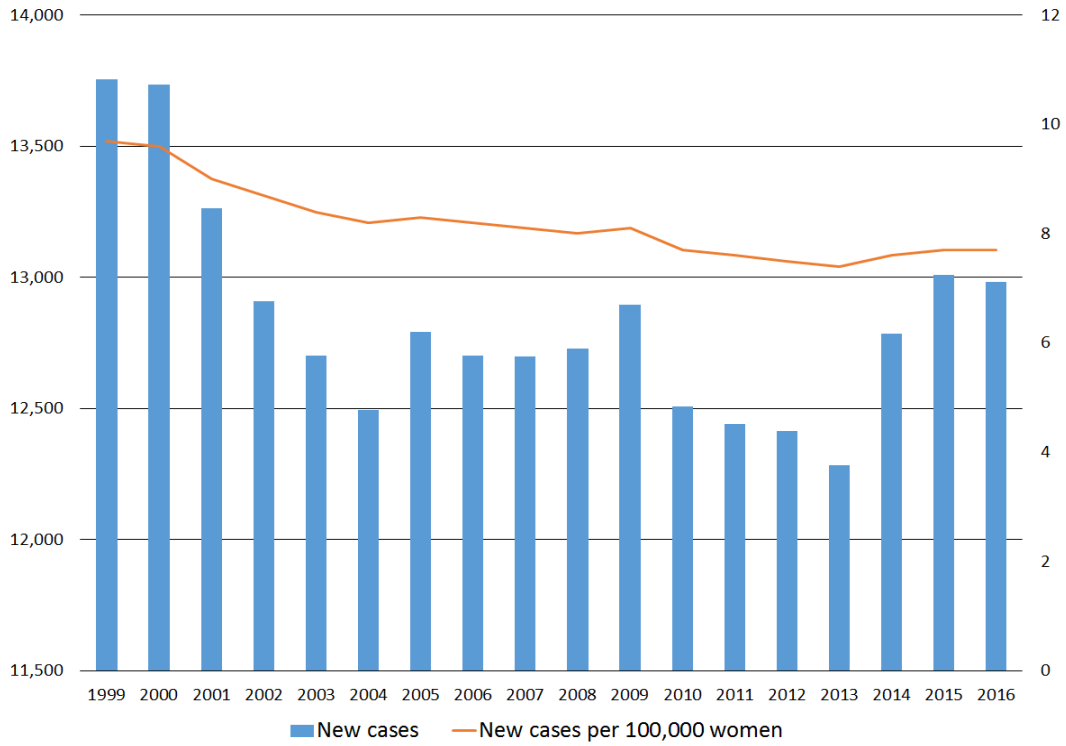


Figure 1.3: Annual rates of new cervical cancer incidence in United States from 1999 to 2016. Data source: US Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute (`www.cdc.gov/cancer/dataviz`).

European countries benefited from population-wide cervical cancer screening programs, cervical cancer still remains one of the most common cancers in women in the countries without nationwide screening (Bray et al. 2018, Australian Institute of Health and Welfare 2019). Specifically, cancer rates are high - in low and middle income countries (Bray et al. 2018) (see Figure 1.4) or where human development index (HDI) is low or middle (American Cancer Society 2018).

The success of screening programs highly depends on resources and expertise (Bray et al. 2018). Visual inspection of cervical cytology microscopic slides is time-intensive, which requires expertise and close attention from a cytotechnologist to localize an abnormal cell among thousands of cells (Malm et al. 2013).

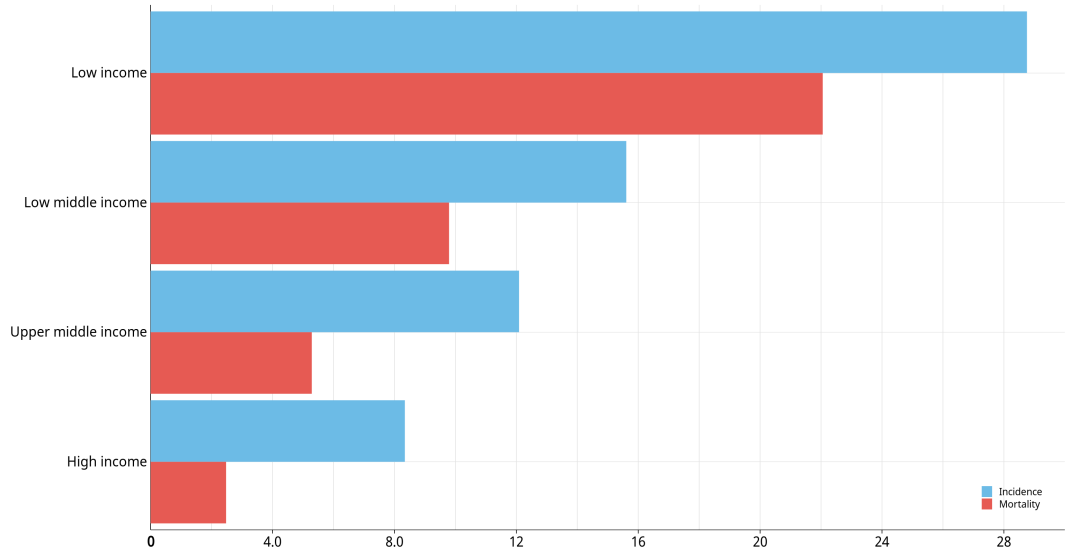


Figure 1.4: Bar chart representing the estimated age-standardized incidence and death rates due to cervical cancer in women, based on income levels. Data source: GLOBOCAN 2018. Graph production: Global cancer observatory (`https://gco.iarc.fr/`).

Correct interpretation of manual cervical screening faces major challenges from smearing (presence of other elements with cervical cells, such as: blood, mucus and other debris), quality of sampling (number of cells in the slide and cell overlapping), poor contrast, and uneven staining (Gençtav et al. 2012). Many low and middle income countries cannot benefit from cervical cancer screening due to: insufficient number of trained and skilled cytotechnologists, lack of organized Pap test teams (Catarino et al. 2015), lack of access to health care services and screening, and insufficient financial resources, which contribute to high mortality rate (above 85%) of cervical cancer (Tareef et al. 2018). Visual analysis of cervical cytology slides is prone to intra observer variability, misdiagnosis and missed-diagnosis. Hence, computer-aided screening and diagnosis system can be beneficial in assisting the cytotechnologists and pathologists. Such systems can help to localize the regions of interest and improve Pap test outcome, and consequently decrease the incidence and death rates due to cervical cancer.

1.2 Cervical cancer

Cervical cancer or cancer of cervix uteri, is the malignant disease that develops in the tissues of cervix. Cervix is an important part of female reproductive system and it is located at the lower part of uterus. Figure 1.5 presents the location

and epithelial distribution in cervix. The cervix is covered with epithelium or

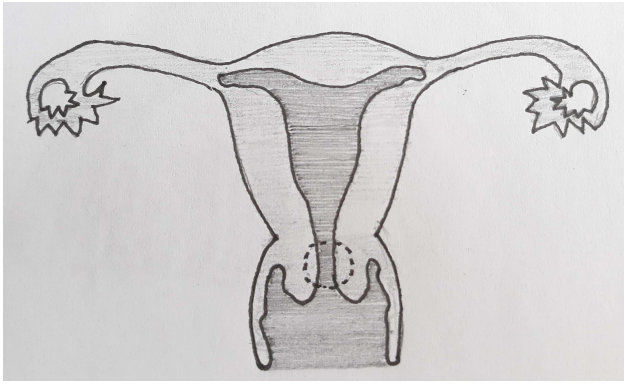


Figure 1.5: Anatomical illustration of female reproductive system. Adapted from (Malm 2013, Bengtsson & Malm 2014).

sheet of cells, which consists of squamous and columnar cells. The outer layer of cervix or ectocervix is covered with flat and thin squamous cells. Columnar or glandular cells are found in endocervix or cervical canal area. The meeting point of squamous and columnar epithelium is called squamo-columnar junction or transformation zone. Around 80% of cervical cancer cases starts in squamous cells of the cervix (*Cervical cancer* 2019) and is called squamous cell carcinoma. Adenocarcinoma is a less common type of cervical cancer that arises in columnar cells.

The squamous epithelium layer of ectocervix consists of four layers: basal, parabasal, intermediate, and superficial. The basal layer is a single layer of round cells and is attached to the basement membrane. Cells in this layer are small and contain large nucleus and small cytoplasm. When basal cells get mature, they divide and form the parabasal cell layer. Continuous division and maturation of these cells move through intermediate layer and finally form the superficial layer. Cells in superficial layer are large and flat with small nuclei and transparent cytoplasm. As the cells move through the layers, size, shape and characteristics of cells change, with reduction in size of nucleus and increase in cytoplasm size. Different layers of squamous cell are illustrated in Figure 1.6.

The endocervical lining contains columnar cells. Columnar cells are composed of a single basal layer. These cells have column-like tall shape, with nuclei located closer to the basement membrane. Figure 1.7 presents the illustration of columnar cells.

Usually, long phase of untreated precancerous condition of cervix is followed by an invasive cervical cancer. The progression of abnormality in squamous cells

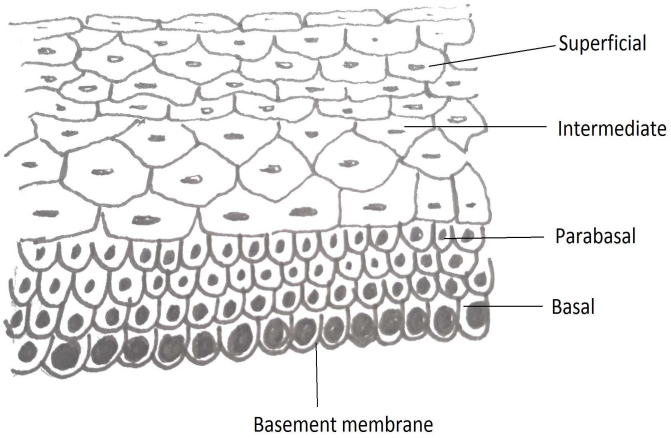


Figure 1.6: Development of squamous cells through different layers. Adapted from (Sellors & Sankaranarayanan 2003).

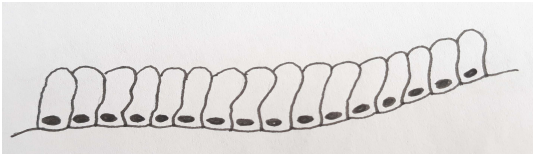


Figure 1.7: Illustration of endocervical columnar epithelium. Adapted from (Sellors & Sankaranarayanan 2003).

is characterized through grades of dysplasia or cervical intra-epithelial neoplasia (CIN). Dysplasia or CIN indicates the intermediate changes that appear between normal cells and invasive carcinoma (Reagan et al. 1953). Dysplasia can be categorized into three types: mild, moderate and severe. According to CIN terminology, CIN1 correspond to mild, CIN2 to moderate, and CIN3 to severe dysplasia and carcinoma-in-situ (CIS) (Richart 1967). Different grades of dysplasia are featured by nucleus enlargement, increased nucleus-cytoplasmic ratio, and change in shape and size of the cell (Sellors & Sankaranarayanan 2003). Mild dysplastic cells contain slightly enlarged nucleus than normal cells, with low nuclear staining intensity. In moderate dysplasia, irregular chromatin distribution or granulation is visible along with dark enlarged nuclei. Nuclei become darker, larger and often deformed for severe dysplastic cells. More granulation and diminishing cytoplasm can be seen in case of severe dysplasia. CIS included in CIN3 grade - is the very early stage of cancer, where abnormal cells have not yet affected the surrounding tissues. From the above characterization of abnormality, it can be summarized that the higher the CIN grade, the higher the risk of developing invasive cervical cancer.

The CIN is the widely used system for reporting Pap test results. How-

ever, The Bethesda System (TBS) (Nayar & Wilbur 2015) has also been used in many laboratories (particularly in the United States). The term squamous intra-epithelial lesion (SIL), is the main feature of TBS reporting. In TBS terminology, mild dysplasia or CIN1 is designated as low-grade squamous intra-epithelial lesion (LSIL), and CIN2 and CIN3 are termed as high-grade squamous intra-epithelial lesion (HSIL).

1.3 Screening

Screening refers to the regular interval specialized disease test of a person, who does not have any physical symptoms. According to GLOBOCAN 2018 (Bray et al. 2018), the lowest cervical cancer incidence rates are in Australia and New Zealand, Western Asia, Northern America, and Western Europe, and the highest rates are in sub-Saharan Africa, Melanesia, Micronesia, and South-Eastern Asia. Population wide availability of cervical cancer screening is the main reason for this variation across geographic regions of the world (American Cancer Society 2018). Regular screening can reduce cancer mortality rate through early detection of any abnormal changes and identifying precancerous lesions at treatable stages (Smith et al. 2019). Over the past five decades, existence of nation-wide screening programs using Papanicolaou or Pap test in many countries - helped to decrease the incidence rate by 50% or more.

Pap test is the primary tool of cervical cancer screening worldwide. In mid-1990's, liquid-based cytology (LBC) was introduced to improve the performance of conventional cervical cytology slide preparation procedure (Smith 2011). In LBC based Pap test, sample is collected by medical practitioner from cervical epithelium using a specialized brush, spatula or swab. The specimen is immediately inserted in preservative liquid, labeled and transported to cytology laboratory. The sample is smeared on a microscopic glass slide and fixed using an alcohol-based fixative solution. This process is done to stop drying out and distortion of cervical cells. Then the slide is stained and visually examined under light microscope to detect signs of malignancy or cancer precursors. In LBC system, SurePath and ThinPrep are the most widely used techniques worldwide nowadays, for cervical cytology slide preparation (Bengtsson & Malm 2014). Some typical examples of Pap smear images are presented in Figure 1.8.

In laboratory, the screening is performed by cytotechnologists. The Pap test screening result is designated as 'normal', in absence of any cervical epithelial

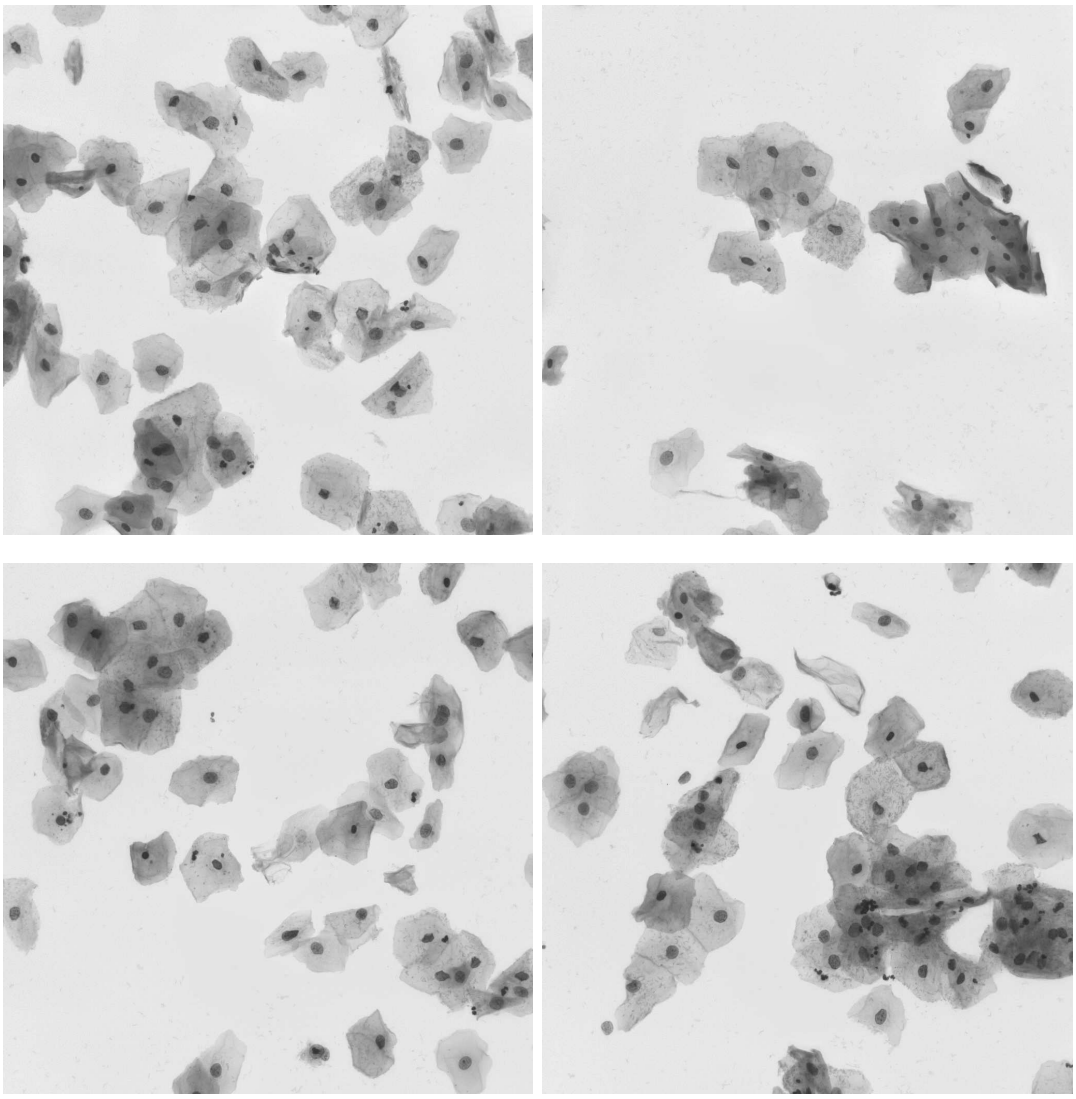


Figure 1.8: Examples of Pap smear images. Data source: Overlapping Cervical Cytology Image Segmentation Challenge - ISBI 2014 (see Chapter 3 for details).

abnormality. If cytotechnologist finds any 'abnormal' or suspicious cells, it is reported. In many laboratories, suspicious findings are further reviewed and confirmed by cytopathologists, along with the stages of the lesions (Bengtsson & Malm 2014).

1.4 Automated screening systems

Generally, visual analysis or manual screening is done at a low optical resolution using 10× lens first. When something abnormal is detected, typically 40× lens is used to take more closer look of the suspicious cell. On average, it takes

about $5 - 10$ minutes for a cytotechnologist, to analyze one sample (Bengtsson & Malm 2014). Full concentration of cytotechnologists is always required to find local precancerous changes in a cell from few hundred thousand cells on a specimen (Malm et al. 2013). According to the recommendation from (Elsheikh et al. 2013), cytotechnologists' workload for Pap test screening should not be more than 70 slides analysis (or 7 hour) in a 24-hour period, to avoid the adverse effect of mental fatigue and discomfort on Pap test outcome.

A large number of attempts have been taken place over the past few decades - to automate the cervical cancer screening system. The first attempt of automated screening for Pap smear was the Cytoanalyzer, in 1950's (Tolles & Bostrom 1956). Nucleus size and optical density were the basis for designing this system. Nucleus density was plotted against nucleus diameter for each cell in a sample, and the graph was analyzed to detect abnormality in the specimen. However, fixed logic circuit was not robust enough to detect abnormality and produced a large number of false alarms, and this was the serious limitation of Cytoanalyzer (Spencer & Bostrom 1962). In addition, inadequacies in cytological preparation and presence of a large proportion of leukocytes in specimens, were also responsible for the failure of this system.

Another early attempt for automated prescreening was CYBEST (Cyto Biological Electronic Screener by Toshiba) (Watanabe 1974) during 1970's. This system was based on the morphological features such as: nucleus area, nucleus density and nucleus/cytoplasmic ratio. The scanning was through two different modes: coarse scan to detect suspicious cells and fine scan of suspicious cells to confirm malignancy. In coarse scanning, each 1mm $\times$ 1mm area of a slide was examined at 4μm resolution. Then 3 features were extracted and malignancy was found using linear discriminant analysis. Fine scanning was carried at 1μm resolution for a suspicious cell using more precise measurements. However, this project was not successful, since automatic focusing used in CYBEST system could not produce reliable good images with sufficiently good focus for all cell nuclei (Bengtsson & Malm 2014).

During 1980's, quite a few automated screening systems were developed like: BioPEPR (Zahniser et al. 1979), FAZYTAN (Erhardt et al. 1980), Cerviscan (Tucker 1976, Tucker & Husain 1981), LEYTAS (van Driel-Kulker & Ploem-Zaaijer 1989). In BioPEPR system (Zahniser et al. 1979), cervical cell analysis was performed using basic morphological features: nucleus area, nucleus optical density, nucleus texture, and nucleus/cytoplasmic ratio. A hierarchical tree strategy was used to prioritize the features for abnormality detection. This sys-

tem could analyze and produce output for 120 to 180 smears per hour, with 10% false positive and 1% false negative rates. However, fixed cut-off percentages for different features to detect abnormality - was the main drawback of this system, and the efficiency of this system was highly dependent on sample's quality. For Cerviscan (Tucker 1976, Tucker & Husain 1981), specimens needed to be prepared with a different technique, namely suspension-polylysine-hematoxylin method. The system produced 4% false negatives and 18% false positives, however performance of Cerviscan for hematoxylin stained slides was poor (Tucker & Husain 1981). FAZYTAN (Erhardt et al. 1980) and LEYTAS (van Driel-Kulker & Ploem-Zaaijer 1989) were both TV based imaging instruments. FAZYTAN system was developed in 1980, with the aim of fast pre-screening. It was based on a host minicomputer with array processors and digital image storage, and an optimized TV microscope. The system was efficient, however it lacked cost effectiveness. When LEYden Television Analysis System (LEYTAS) was developed in 1989, both hardware and software fields improved. LEYTAS was a fully automated screening system that resulted in 0.3% false negative rate, 13% false positive rate and 2.7% rejection rate. However, specimen preparation was time-consuming and required a considerable amount of human intervention. Most of these 1980's systems were unsuccessful because of the complexity of the task, limited computing capacity available and lack of cost effectiveness. Hence, those projects ended up as being only operational prototype, rather than getting into the commercial market (Bengtsson 2003).

During 1990's, computing power has improved a lot compared to 80's and an interest for cytology automation was also increased (Bengtsson & Malm 2014). A semi-automatic system for slide preparation and screening has been reported in (Carothers et al. 1994). CYTOPRESS was used to prepare mono-layered slides and CERVIFIP was used to scan and locate suspicious objects in the specimen. Around 85% of all the samples were found positive or containing suspicious objects, and were further reviewed by a human operator. This system was built to reduce human workload and decrease screening time, however it still required a considerable amount of human intervention.

Despite of long research history over the decades to automate cytological screening, the US Food and Drug Administration (FDA) did not approve any commercial screening system until mid 90's. In 1995, FDA approved two commercial re-screening systems, namely PAPNET testing system (Neuromedical Systems, Inc., Suffern, NY) and the Neopath Auto Pap 300 QC system (NeoPath, Inc., Redmont, WA). These systems were approved to re-screen the smears that

were found as negative or normal through manual screening by cytotechnologists (Russell et al. 2005).

PAPNET was a neural network based automated screening system to analyze conventional Pap smear images (Koss et al. 1994). The system scanned slides and selected suspicious cells to be reviewed by trained cytotechnologists. The cell selection process was composed of an algorithmic classifier and neural network classifier. Output of the classifiers was an abnormality based ranking, and 64 most abnormal samples (or sample regions) were stored on a magnetic tape for re-screening (Bengtsson 2003). The Auto Pap 300 QC system (Patten Jr. et al. 1996) was developed for automatic analysis and classification of Pap smear specimens. It was developed as a quality control re-screener to review the specimens classified as normal during manual screening. This system consisted of: field-of-view computing, high speed video microscopy, and image interpretation software. The samples were classified considering morphological features associated with epithelial abnormalities.

Tripath imaging was the first company whose cervical screening product received FDA approval in 1998. This company was the merger of NeoPath, Neuromedical and AutoCyte. Tripath was acquired by BD (Becton, Dickinson and Company, Franklin Lakes, NJ) in 2006. Their product SlideProfiler is now named as BD FocalPoint (*BD FocalPoint GS imaging system* 2019). The FocalPoint system automatically classifies and ranks specimens for manual review.

Another aspect of automation of cytology was to develop new techniques for sample preparation. ThinPrep (Hutchinson et al. 1991, Zahniser & Hurley 1996) was a liquid-based technique for thin layer cervical slide preparation, which received FDA approval in 1996. While using ThinPrep technique, samples were collected in a preservative solution (PreservCyt), which can preserve cell morphology during transportation to a laboratory. At the laboratory, the cells were transferred to a glass slide and processed using ThinPrep processor. ThinPrep Imaging System (TIS) (*ThinPrep Imaging System* 2019) was developed to screen ThinPrep cervical cytology sample slides and highlight suspicious areas for manual review by cytotechnologist. CytoProcessor (DATEXIM, Caen, France) (*CytoProcessor* 2019) is a recent fully automated system for virtual diagnosis, which works on ThinPrep slides. While compared to TIS, CytoProcessor has shown better performance to detect abnormality in cervical cell images (Crowell et al. 2019).

In 1999 another technique, named AutoCyte-Prep (Bishop J et al. 1998, How-

ell et al. 1998) system also got FDA approval. In AutoCyte-Prep technique, the sample was collected in a container with AytoCyte transportation fluid. The samples were centrifuged in a laboratory and slides were prepared automatically using PrepStain machine. AutoCyte-Prep system is currently known as BD SurePath (*BD SurePath* 2019) and PrepStain machine for slide preparation as BD PrepStain Slide Processor (*BD PrepStain Slide Processor* 2019).

The most important factor of automated cervical cancer screening is recognizing the malignant cells correctly, so that it can be treated before turning into invasive cancer. In case of visual screening, cytotechnologists classify around 96% samples as normal, thus a small percentage of samples is further reviewed by cytopathologists (Bengtsson & Malm 2014). In automated screening, an expert needs to review the slide, even if a few malignant cells are present. Therefore, these systems should be designed to to keep the false positive rate as low as possible, in addition to maintaining the lowest possible false negative rate. For example, false positive rate of 0.1% would contribute to 20 or 100 false alarms for LBC based or conventional Pap smear slides, respectively (a liquid-based cytological slides hold around 20,000 and conventional Pap smear slide contains around 100,000 cervical cells). This phenomenon may yield the automated screening system relatively impractical and this was the case with some early systems like Cytoanalyzer, BioPEPR, Cerviscan, CYTOPRESS, and LEYTAS.

Computational performance is a great challenge when developing a practical automated screening system. If a framework requires on average 3 seconds per cell for screening, then it will need around 17 hours for a LBC based and 4 days for a conventionally prepared cervical smear slide. However, the output of a functional screening system is expected to be one slide per few minutes. Thus, such a framework would require improvement of the processing performance by 100 to 1000 times to be used for practical purpose. Therefore, when devising a new computer-aided cervical screening system, proper consideration should be given to the areas of speed optimization and computational complexity.

1.5 Nuclei segmentation in automated screening

Morphological and textural features of nucleus are crucial while determining abnormality in cervical cells (Holmquist, Bengtsson, Olsen, Stenkvist & Noguchi

1976, Plissiti et al. 2011*b*). Automated screening systems (discussed in Section 1.4) also used nuclei based features to detect suspicious cells. Therefore, precise segmentation of nucleus is the most important step in automated screening and computer-aided diagnosis (CAD) systems. Many methods have been proposed over the years to segment nucleus from single and overlapping cytology images. Nucleus segmentation methods used in different computer assisted systems can be roughly divided into five categories: thresholding and morphological analysis, deformable model and shape-based approaches, watershed segmentation, region and clustering based segmentation approaches, and machine learning based segmentation techniques. A brief description of these categories is presented in this section.

1. **Thresholding and morphological analysis based approaches:** Thresholding and morphological analysis based approaches (Cahn et al. 1977, Borst et al. 1979, Bengtsson et al. 1979, Yang-Mao et al. 2008, Plissiti et al. 2011*a*,*b*, Pai et al. 2012, Ushizima et al. 2015, Plissiti et al. 2015, Guan et al. 2015, Tareef et al. 2015, Riana et al. 2015, Lee & Kim 2016, Phoulady, Goldgof, Hall & Mouton 2016, Phoulady et al. 2017) are widely used traditional approaches to segment nuclei from cytology images. These techniques use intensity histogram characteristics, intensity based or adaptive thresholding, and different morphological operations to segment nuclei. Generally, these techniques are faster compared to other approaches, however segmentation accuracy may be compromised. These type of segmentation approaches may fail for images, in the presence of cell overlap - where overlapping parts of cytoplasm appear with similar average intensity as nuclei.

2. **Deformable model, contour and shape based approaches:** In medical image segmentation, parametric deformable models: snake, active contour and active shape models, geometric deformable models: level set technique, contour, and shape - based segmentation methods (Bengtsson et al. 1981, Bamford & Lovell 1998, Plissiti et al. 2006, Tsai et al. 2008, Harandi et al. 2010, Plissiti & Nikou 2012*b*, Bergmeir et al. 2012, Li et al. 2012, Lu et al. 2013, Nosrati & Hamarneh 2015*a*, Husham et al. 2016, Phoulady, Zhou, Goldgof, Hall & Mouton 2016, Zhang, Kong, Liu, Wang, Chen & Sonka 2017, Dong et al. 2019, Kostrykin et al. 2019) have been widely used. Deformable models are based on curves or contours that can move/deform towards an object boundary, because of internal forces and external/image

constrains. Usually, deformable models are designed as solutions to energy minimization problems. Level set based geometric deformable models are particularly popular, since they do not require explicit parameters or depend on topological constraints. Level set and active contour based techniques can produce precise segmentation boundaries, if provided with initial boundaries close enough to the actual/true object boundaries. Therefore, these techniques are often used as a final step to regularize segmented boundaries. In shape-based techniques, an object shape derived from the image is used as a constraint for generating segmentation output.

3. **Watershed based segmentation approaches:** Watershed is a long-standing, yet powerful classic technique in image segmentation paradigm. Marker controlled watershed segmentation of nucleus has been proven successful in several studies (Malpica et al. 1997, Cheng & Rajapakse 2009, Bai et al. 2009, Cloppet & Boucher 2010, Jung & Kim 2010, Béliz-Osorio et al. 2011, Gençtav et al. 2012, Moshavegh et al. 2012, Tareef et al. 2018). In most of these studies, touching or overlapping nuclei from fluorescence microscopy images were split using watershed segmentation, however were not well-explored for cervical cell nuclei segmentation. Accurate marker extraction is the main challenge for marker controlled watershed segmentation techniques.

4. **Region and clustering based approaches:** Region and clustering based segmentation approaches (Wu et al. 1998, Isa 2005, Kale & Aksoy 2010, Lu et al. 2013, Chankong et al. 2014, Zhang, Kong, Chin, Liu, Chen, Wang & Chen 2014, Nosrati & Hamarneh 2015*b*, Lu et al. 2015, Oprisescu et al. 2015, Gautam et al. 2017, Bora et al. 2017, Roy et al. 2020), are also popular for segmenting nuclei. These techniques are designed to segment an image either by merging or splitting image pixels/regions based on some predefined criteria. Most of the region-growing based techniques require some seed or starting point, hence these techniques are dependent on some nucleus detection methods. These techniques also suffer from high rate of false positives and false negatives.

5. **Machine learning based segmentation:** In the paradigm of nucleus segmentation, machine learning based methods become popular in recent years (Jung et al. 2010, Song et al. 2015, Zhang, Sonka, Lu, Summers & Yao 2017, Tareef, Song, Huang, Wang, Feng, Chen & Cai 2017, Liu et al. 2018, Jith et al. 2018, Lin et al. 2019, Araújo et al. 2019). In these techniques,

image pixels are classified into different classes based on some features. Traditional classification techniques, Support Vector Machine (SVM) classifiers, convolutional neural networks, and deep learning frameworks, are used in these machine learning based segmentation approaches. Usually, these techniques produce rough segmentation boundaries, and require another shape regularization step to achieve finer segmentation results.

1.6 Prior knowledge in segmentation

Performance of CAD systems largely depends on the quality of segmentation of the target object. Presence of noise, clutter, poor contrast, occlusion, and illumination conditions in an image, make the cell and nuclei segmentation task more challenging. In particular, segmentation techniques exclusively depending on image information, are affected by image characteristics. Inclusion of prior knowledge available about the target object into segmentation techniques, helps to obtain more precise and plausible segmentation outcome (Leung et al. 2004, Slabaugh & Unal 2005, Veksler 2008, Hamarneh & Li 2009, Lu et al. 2013, Grosgeorge et al. 2016).

Various features that can distinguish between the target object and the background, can be used as prior knowledge. Therefore, the nature of prior information varies from one application to other. Usually, user interaction, boundary/contour information, shape of target object (geometric, statistical, or physical), appearance, and spatial distance/location, are used as prior information. Prior information have been incorporated into: graph based, watershed, clustering based, deformable models, and edge-based segmentation techniques. Mathematical modeling of prior information and incorporation of that prior guidance in a segmentation framework, are active fields of research, and popular among variety of applications (Cootes et al. 1995, Noordam et al. 2000, Boykov & Jolly 2000, Cootes et al. 2001, Boykov & Jolly 2001, Stegmann & Gomez 2002, Ahmed et al. 2002, Paragios 2003, Tsai et al. 2003, Leung et al. 2004, Rother et al. 2004, Freedman & Zhang 2005, Slabaugh & Unal 2005, Chuang et al. 2006, Cremers et al. 2007, Cai et al. 2007, Zhu-Jacquot & Zabih 2007, Wang et al. 2008, Veksler 2008, Das et al. 2009, Bai et al. 2009, Lempitsky et al. 2009, Hamarneh & Li 2009, Li et al. 2011, Lu et al. 2013, Wang, Song, Soh & Sim 2013, Wang, Zhang & Ray 2013, Grosgeorge et al. 2013, Bai et al. 2014, Adhikari et al. 2015, Nosrati & Hamarneh 2015*a*, Grosgeorge et al. 2016).

1.7 Research objectives

Precise nuclei segmentation is the most important step in computer assisted abnormality detection in cervical cells. Though many CAD systems have been developed over the decades for identifying abnormal cells from Pap smear images, and researchers developed a range of segmentation techniques to detect/segment cervical nuclei, there is still possibility to explore alternative approaches and improve the segmentation accuracy. It is evident from the literature that segmentation accuracy can be improved with the aid of prior influence. Therefore, the main objectives of this thesis are:

1. to model prior information and incorporate the prior guidance in segmentation techniques,
2. to apply prior guided segmentation techniques for cervical nuclei segmentation to improve segmentation accuracy and plausibility,
3. to develop a cervical cell abnormality detection framework based on the developed segmentation techniques with prior information, and primarily based on nuclei features only.

1.8 Thesis outline

This thesis consists of seven chapters, where the first two chapters (including this introductory Chapter 1), describe the general context and methodological background of this thesis. The next four chapters present the contribution of this thesis, while the last chapter contains concluding remarks. More specifically, this thesis is organized as follows:

Chapter 2 provides a review of key stages of automated cervical screening system, along with popular techniques involved in each stage. It also presents an overview of prior guided segmentation techniques for different applications and related work on cervical nuclei segmentation.

Chapter 3 presents a novel technique where a circular shape function is defined and incorporated into Fuzzy c-means (FCM) clustering technique. This technique is used to segment nuclei from overlapping Pap smear images.

In Chapter 4, a novel circular shape prior guided graph-based segmentation is presented. The shape prior was modeled using circularity definition of a region

and incorporated in the merging predicate of efficient graph based segmentation technique. The technique was evaluated by segmenting cervical cell nuclei.

Chapter 5 contains a superpixel based merging framework to segment nuclei from cervical smear images. A novel pairwise regional contrast and gradient boundary based merging criteria, is presented in this chapter.

In Chapter 6, a cervical cell classification framework is presented based on prior guided segmentation techniques developed in Chapters 3, 4 and 5. Intensity, shape and texture features extracted exclusively from segmented nuclei were used for this abnormality detection framework.

Chapter 7 provides the conclusion, and outlines future work and contributions of this thesis.

Chapter 2

Literature review

In this chapter, the main steps involved in computer-aided screening systems are reviewed. Prior guided image segmentation techniques and an overview of currently available techniques for cervical nuclei segmentation, are also presented. In Section 2.1, an overview of the key stages in automated cervical cytological screening and the major techniques used in this thesis, are presented. Section 2.2 describes different priors and how those priors were incorporated in image segmentation techniques. A brief review of cervical nuclei segmentation techniques in literature, is outlined in Section 2.3.

2.1 Computer-aided cervical cytology screening

In automated or computer assisted cervical cytological image analysis, the computational system should be intelligent enough to at least match the performance of an expert cytotechnologist. The key stages involved in a typical computer-aided cervical screening system are: preprocessing, segmentation, feature generation, feature selection, and classification. The popular methods used in literature and the major techniques used in this thesis for these stages, are reviewed in this section.

2.1.1 Preprocessing

A staining procedure is performed on cervical cell specimens, for better visualization of different tissue features under the microscope. This procedure stains nuclei and cytoplasm in different shades of colors (dark-blue, pink to orange,

blue-green, light-green). However, automated analysis of these specimens is not straightforward, due to presence of variability in the specimens for concentration of stain or uneven staining and the image acquisition process (Hayakawa et al. 2019). Cell overlapping in Pap smear images also affects the image contrast. Therefore, preprocessing is often used as the first step in computer-aided cytological analysis to reduce the noise and variability in specimen image. Different preprocessing techniques and their combinations are used in literature. A brief review of these techniques is presented here.

2.1.1.1 Noise removal

Denoising process is generally carried using different filter techniques and morphological operations on the given image. Depending on the application, it is possible to enhance ROI by defining different neighborhood size for the corresponding operation.

Bilateral (Tomasi & Manduchi 1998) and non-local means (NL-means) (Buades et al. 2005) filters, are effective tools for edge preserving smoothing of images. Bilateral filter (BF) considers both intensity/photometric and spatial/geometric similarity of a pixel in it's spatial neighborhood. In NL-means filtering process, the new/estimated value for a pixel is calculated as a weighted mean of all image pixels. The weight for each pixel is determined systematically by using the similarity between - the pixels and the square neighborhoods of those pixels. BF and NL-means filter were adapted in (Ushizima et al. 2015) and (Li et al. 2012, Guan et al. 2015, Zhao et al. 2016), respectively, for denoising cervical smear images.

Wiener filter (Pratt 1972) is an image restoration technique to recover the image from noise. It removes the additive noise and deblurs the image, based on a statistical approach. This filter considers local image variance while smoothing the image; less smoothing for large variance and more smoothing for small variance. This adaptive Wiener filter is used in Chapter 5 of this thesis to denoise the image, while finding image gradient boundaries.

Median filter (Sun & Neuvo 1994) is very effective in reducing noise and preserving object edges in an image. This is a non-linear spatial filter that replaces the value of a pixel by the median of the graylevels in the neighborhood of that pixel. The result of preprocessing by using median filter can be varied by changing the size of neighborhood/filter kernel. This filter is particularly effective for *impulsive* or *salt-and-pepper* noise. Median filter was used to denoise images in (Tsai et al. 2008, Chankong et al. 2014, Zhang, Kong, Chin, Liu, Chen, Wang

& Chen 2014, Husham et al. 2016, Gautam et al. 2017, Bora et al. 2017). In Chapter 6 of this thesis, median filter is used to remove noise from cervical smear images.

2.1.1.2 Contrast enhancement

Contrast enhancement or illumination correction, plays an important role in digital image analysis. It is used to improve the perception of the target object in an image. Some of the common contrast enhancement techniques are: contrast stretching, histogram equalization (HE), morphological operations, and contrast limited adaptive histogram equalization (CLAHE).

Contrast stretching and HE are the simplest techniques for contrast enhancement. In low-contrast images, contrast stretching increases the dynamic range of image intensity levels to span the full range of intensity values (Gonzalez & Woods 2008). In this technique, a linear mapping function is used to map input image pixels to output. Contrast stretching technique was used in (Zhang, Kong, Chin, Liu, Chen, Wang & Chen 2014, Win et al. 2018) to improve contrast in cervical cytology images. HE is a non-linear mapping - used to distribute the image intensity level over the whole intensity levels (Shih 2010). This technique aims to transform the density distribution in image histogram to a uniform one. This process enhances the contrast for intensity values near to histogram maxima, and decreases the contrast near histogram minima. However, HE may not enhance the local details in a ROI, and there is also a possibility to enhance the noise, since this is a global approach.

Some morphological operations are also popular to enhance the image contrast. Black top hat/top hat closing operation is computed as the difference between the closing of the image and the image itself (Gonzalez & Woods 2008). In (Gençtav et al. 2012), this operation was performed with a disk structuring element for illumination correction. In (Kaur & Sahambi 2016), black top hat was combined with white top hat transform to enhance the contrast between white and black regions in cell images.

CLAHE (Pizer et al. 1987) is a modified version of adaptive histogram equalization technique. This technique subdivides the image into small non-overlapping tiles/regions, and histogram equalization process is performed on each tile. Contrast enhancement on noise is controlled by clipping the histogram with a user specified clip limit. This technique is popular for image contrast enhancement in literature, for example it was used in (Plissiti et al. 2011*a*,*b*, 2006, Ushizima

et al. 2015, Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017, Win et al. 2018) while segmenting cervical cell images. CLAHE technique is used in Chapter 6 of this thesis, to enhance contrast in cervical cytology images after removing noise with median filter.

2.1.1.3 Color processing

In computer-aided image analysis systems, color plays an important role. By changing the color space of a given image, it is possible to ease and improve the task of ROI detection and image segmentation in certain applications. In literature, RGB images have been converted into different color systems, and often a single channel from the image color system is used for subsequent stages of analysis, to avoid the interference of colors in image processing.

Extracting a single channel from RGB images may be useful in some applications. In case of Pap smear images, the single color channel should be selected in such a way that there is reliable color difference between nucleus/cytoplasm and cytoplasm/background. An investigation on selecting the optimal color channel was reported in (Holmquist, Imasoto, Bengtsson, Olsen & Stenkvist 1976), for Papanicolaou stained cervical smear images. It was found in (Holmquist, Imasoto, Bengtsson, Olsen & Stenkvist 1976) that green channel (specifically the yellowish green around 570 nm wavelength) from RGB color space provides the optimal contrast between nucleus and cytoplasm. Thus, cell nuclei is more identifiable in this channel. In (Win et al. 2018, Silva et al. 2019), the green channel from RGB color images was extracted for further processing of cervical smear images. In (Bergmeir et al. 2012, Riana et al. 2015), weighted sum of each channel of an RGB image was used to generate the grayscale image, with the same luminance as the original RGB images remain in transformed images.

HSV transform is also used in image analysis with a target of image segmentation. This transform consists of three components: hue (H), saturation (S) and value (V). The H channel represents the color using an angle from 0° to 360°. The S channel indicates the amount of gray (0 to 100%) in a particular color, and channel V describes the brightness of the color (ranges from 0 to 100%). For cervical smear images, this transformation is used while identifying abnormal cells from specimens - prepared using liquid based cytology technique. The contrast between abnormal cells and non-cellular artifacts present on the smear is an important aspect in case of abnormality detection. In V channel of HSV color space, the contrast is enhanced between nucleus/leukocyte and cytoplasm. RGB

image was converted to HSV color space in (Zhang et al. 2011, Zhang, Kong, Liu, Wang, Chen & Sonka 2017), then V channel was used for further stages of segmentation frameworks.

CIELAB or CIE $L*a*b*$ color space - defined by the International Commission on Illumination (CIE) in 1976, is the most exact color representation to approximate human vision. Color is expressed using three values: $L*$ for lightness (values range from black (0) to white (100)), $a*$ from green $(-)$ to red $(+)$, and $b*$ from blue $(-)$ to yellow $(+)$. This color processing is widely used in literature for cervical cell image processing, where smears are prepared using conventional Papanicolaou or H&E staining technique. In these staining processes, cell regions are colored with shades of red and blue, whereas the background region remains colorless (Gençtav et al. 2012, Zhang, Kong, Chin, Liu, Chen, Wang & Chen 2014). Thus, lightness can be used to differentiate the cell regions in this color space. In (Gençtav et al. 2012, Li et al. 2012, Guan et al. 2015, Zhang, Kong, Liu, Wang, Chen & Sonka 2017), RGB cervical smear images were converted to CIELAB color space, and $L*$ dimension was extracted for subsequent processing. This technique is used in Chapter 6 of this thesis, for the conversion of RGB cervical cell images to grayscale. In (Zhang, Kong, Chin, Liu, Chen, Wang & Chen 2014, Zhang, Kong, Ting Chin, Liu, Fan, Wang & Chen 2014), $a*$ channel from CIELAB color space was used, followed by a contrast enhancement procedure.

2.1.1.4 Background removal

The background removal process aims to divide a Pap smear image into foreground (containing cervical cells or cell clumps) and background (remaining area) regions. In subsequent stages of computer-aided image analysis systems, this approach can help to focus on the foreground images (ignoring the background). Thresholding, different morphological filtering and clustering based techniques are mostly used for background removal process in literature.

Thresholding based background removal was used in (Gençtav et al. 2012, Plissiti et al. 2011*b*, 2006). Otsu thresholding (Otsu 1979) is the most widely used thresholding technique. This thresholding was used in (Bergmeir et al. 2012), and Otsu thresholding with morphological dilation was used in (Plissiti et al. 2011*a*), to detect the ROI from a given image. In this thesis, an Otsu multi-thresholding is utilized for background subtraction, in Chapter 3.

In many applications, background is removed using a combination of superpixel level representation of image (instead of pixel grid level) and clustering.

Simple linear iterative clustering (SLIC) (Achanta et al. 2010, 2012) superpixel generation algorithm, is used widely to segment image into local smooth regions/superpixels. In (Oliveira et al. 2017), SLIC technique was used to generate superpixels, and density-based spatial clustering of applications with noise (DBSCAN) technique was used to group similar superpixels, with the aim to separate background and foreground. Thresholding techniques on SLIC superpixels were used in (Plissiti et al. 2015, Lee & Kim 2016), to remove background and detect the cell mass. In (Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017, Tareef, Song, Huang, Wang, Feng, Chen & Cai 2017, Tareef et al. 2018), SLIC superpixel representation of images was used for cell cluster segmentation and ROI localization. Quick shift (Vedaldi & Soatto 2008) is another superpixel generation technique. In (Lu et al. 2013, Lu et al. 2015), cell clumps were found using quick shift algorithm with subsequent edge detection technique. In (Silva et al. 2019), mean shift clustering superpixel generation, followed by k-means clustering and morphological opening was used. Morphological filtering based k-means clustering was used in (Guan et al. 2015) to find the foreground image. In (Phoulady, Goldgof, Hall & Mouton 2016, Phoulady et al. 2017), Gaussian mixture model with Expectation Maximization (EM) algorithm was used for finding cell clump/foreground.

Grayscale morphological reconstruction can be used to smooth the extreme value regions, without affecting the shape or size of an object in the image. Morphological opening by reconstruction and closing by reconstruction can be used subsequently, to clean up the image by minimizing intra class and maximizing inter class variance in image histogram (Gonzalez & Woods 2008). In (Tareef et al. 2015), morphological reconstruction (opening and closing) and regional maxima based background removal process was used. This technique is used in Chapter 3 to remove background from cervical smear images.

2.1.2 Image segmentation

Image segmentation is the technique of partitioning or pixel classification, which aims to segment objects or regions from the background. Image segmentation involves assigning a class or label to pixels in the image in such a way that pixels in the same class share some common characteristics, for example: similar mean intensity, texture, or color (Gonzalez & Woods 2008). In computer-aided systems for image analysis, image segmentation can help to delineate the ROIs. Numerous image segmentation techniques have been proposed and implemented in past

decades, to tackle a wide variety of problems. Traditional and commonly used techniques include: thresholding, region-based segmentation, clustering-based segmentation, and graph-based segmentation. Beside these, many combinatorial image segmentation techniques also exist to overcome the limitations associated with the basic ones.

2.1.2.1 Thresholding

Thresholding is the most simple and convenient technique for image segmentation and is still widely used. The basic idea of this technique is the assumption that objects in an image are identifiable using image intensity values (Sonka et al. 1993). With simple thresholding, an input image f can be transformed to an output (segmented) binary image g as:

$$g(i,j) = \begin{cases} 0, & f(i,j) < \tau, \\ 1, & f(i,j) \geq \tau, \end{cases} \tag{2.1}$$

where, τ is a threshold for segmentation. In the transformed image, pixels labeled with 1 correspond to objects/foreground and pixels labeled with 0 indicate the background.

Based on the threshold selection technique, thresholding can be categorized as: global and local/adaptive thresholding. Global thresholding can use single or multiple thresholds - calculated from the image histogram. Local threshold is location specific and its value depends on local image characteristics, that may vary over the entire image. Choosing a correct/optimal threshold value is a critical task for a successful threshold based image segmentation. Since, the global threshold is determined from the whole image, the likelihood of success of this method is usually low (Sonka et al. 1993). Most common reasons of failure for global threshold based segmentation are: intensity variation in objects and background due to noise, non-uniform illumination, or a number of other artifacts. Local/adaptive threshold can complement global thresholding in these situations.

An adaptive iterative threshold selection method was proposed in (Ridler & Calvard 1978), and this technique works as follows.

1. Assume 4 corner points of the image f as background pixels, and set μ_0 to the mean graylevel value of these 4 pixels. Assume others as foreground pixels. Set μ_1 to their mean graylevel value

2. Set the threshold, $\tau = \frac{\mu_0+\mu_1}{2}$, and segment the image f using Equation (2.1)

3. Recompute μ_0 and μ_1

4. Go to step 2 and iterate until τ converges or no longer changes significantly

Otsu thresholding (Otsu 1979) is the most popular method of thresholding based segmentation. In this technique, the optimal threshold is determined by maximizing a criterion function, where inter-class variance is considered to define the function. Consider f as an image of size N and graylevels $i \in [1, 2, \cdots, L]$. For two class image segmentation, C_0 and C_1 presents the background and foreground regions separated with a threshold t $(1 \leq t \leq L)$. The probability $p(i)$ of occurrence of graylevel i is expressed as: $p(i) = \frac{n_i}{N}$, where n_i is the number of occurrences of graylevel i in image f. Then, the class probabilities are defined as:

$$\omega_0 = \omega(t) = \sum_{i=1}^{t} p(i) \quad \text{and} \quad \omega_1 = 1 - \omega(t) = \sum_{i=t+1}^{L} p(i),$$

and means of the classes are:

$$\mu_0 = \sum_{i=1}^{t} \frac{i \cdot p(i)}{\omega_0} \quad \text{and} \quad \mu_1 = \sum_{i=t+1}^{L} \frac{i \cdot p(i)}{\omega_1}. \tag{2.2}$$

The total mean of the graylevel image can be written as:

$$\mu_T = \sum_{i=1}^{L} i \cdot p(i). \tag{2.3}$$

Using Equation (2.2), Equation (2.3) can be presented as:

$$\mu_T = \omega_0 \, \mu_0 + \omega_1 \, \mu_1.$$

The inter-class variance of the classes C_0 and C_1 is as:

$$\sigma_b^2 = \omega_0 \, (\mu_0 - \mu_T)^2 + \omega_1 \, (\mu_1 - \mu_T)^2.$$

The optimal threshold can be found by searching the full range of t values $[1, L]$, and selecting the value that maximizes σ_b^2. In Chapter 3 of this thesis, Otsu thresholding is used to separate cervical cell/clump from background.

2.1.2.2 Region-based Segmentation

Image regions are defined as connected homogenous subsets of an image, with respect to some similarity criterion (Sonka et al. 1993, Gonzalez & Woods 2008). Region-based segmentation of an image f is the partition of f into homogeneous regions R_i, where $i \in [1, 2, \cdots, m]$. The homogeneity criteria can be based on graylevel, color, shape, or texture in a region. Region merging and region splitting are two basic approaches of region-based segmentation.

In region merging based segmentation process, the initial segmentation starts with a single pixel p in an image. Consider a pixel q adjacent to pixel p. A similarity measure $S(p, q)$ is defined such that it produces a high value if pixels p and q are similar, and low value, otherwise. Pixel q can be added to the region of pixel p, if $S(p, q) > T$, where T is a threshold. This merging technique proceeds to the other neighbors of p and the process iterates, till no regions are left to be merged. A similarity measure can be defined by considering, for example: edge strength of the boundary, texture of the regions, or intensity difference within a region. Major steps involved in region merging based segmentation can be summarized as follows:

- take an initial (over) segmented image, where each pixel constitutes a region,
- merge the neighboring regions that are similar based on some criteria, and
- repeat this process until no regions remain to be merged.

Region splitting based image segmentation process starts from a single region (usually the entire image). The criteria for splitting can be found considering, for example: the variance of its graylevel values or texture, the occurrence of strong internal edges, standard deviation within a region, or distance between mean values of the regions. This type of segmentation process works as follows:

- take an (under) segmented image,
- split each region that is not homogeneous according to certain criterion,
- repeat the process until all regions are homogeneous.

Region based methods are robust, because: regions consist of many pixels and it is easy to avail more information to characterize a region, texture features can be used to detect a region, and region based techniques perform better in

noisy images. However, region based segmentation techniques have some drawbacks too. Decisions about region membership or defining similarity measure for segmentation, and object segmentation from multiple disconnected regions, are often difficult. Finding seed points for region growing is crucial, and the order in which regions are treated also influence the segmentation outcome and must be carefully determined.

Statistical Region Merging (SRM) proposed in (Nock & Nielsen 2004), is a probability theory based segmentation technique that follows region merging approach. Image segmentation is considered as an inference problem by this technique, and the key idea of SRM is to reconstruct the regions of an observed image, based on the theoretical (true) statistical regions. Two essential components of this algorithm are the merging predicate (a statistical test) and the ordering of regions followed to merge. The merging predicate for a couple of regions (R_1, R_2) is defined as:

$$P(R_1, R_2) = \begin{cases} true, & \text{if } \left|\overline{R_2} - \overline{R_1}\right| \leq \sqrt{b^2(R_1) + b^2(R_2)}, \\ false, & \text{otherwise,} \end{cases}$$

where

$$b(R) = g\sqrt{\frac{1}{2Q|R|} \ln \frac{2}{\delta}},$$

$\overline{R_1}, \overline{R_2}$ are the mean intensities across regions R_1 and R_2, respectively, g is the number of image intensity levels. Q is the parameter that controls statistical complexity of the segmented image and the coarseness or scale of segmentation, $|.|$ stands for cardinality, and δ denotes the probability error. The merging predicate solely depends on the merging threshold, and the merging threshold inversely depends on the value of Q (higher value of Q produces lower threshold). Thus, SRM produces smaller regions with higher value of Q, and vice versa. The order of merging follows a simple invariant: if two parts of the true regions are inspected, that means all tests inside those regions have previously being done. The simplest choice for the merging order $O(p, q)$, is to use pixel color channel (or intensity) values as:

$$O(p, q) = |I_p - I_q|,$$

where, I_p and I_q are the intensities for image pixels p and q, respectively. SRM

approach can cope with a significant noise corruption, handles occlusions and performs scale sensitive segmentation, but it suffers from over-merging error. Regions produced from SRM are accurate enough to be used as seeds for region refinement, merging methods or object detection. In Chapter 5 of this thesis, SRM segmentation is used to generate superpixels and used in a superpixel merging framework to segment cervical nuclei.

2.1.2.3 Clustering-based segmentation

In clustering-based segmentation techniques, an image is represented in terms of clusters of pixels or disjoint groups, based on some specific similarity criterion. The objective function for partitioning is designed to encourage intra-cluster similarity and inter-cluster dissimilarity. Intensity values, texture measures defined on a local neighborhood, or distance between two data points/pixels, are some examples of measures that can be used to define the similarity criterion. K-means clustering, Simple Linear Iterative Clustering (SLIC), and Fuzzy c-means (FCM) clustering, are some commonly used clustering-based segmentation techniques.

K-means clustering (MacQueen 1967) is one of the simplest and popular clustering techniques used for image segmentation. This technique employs squared error criterion as the objective function for partitioning an image into k number of clusters. This criterion is defined as:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \left\| x_i^{(j)} - c_j \right\|^2,$$

where n_j is the number of pixels in j^{th} cluster, $x_i^{(j)}$ is the i^{th} pixel belonging to j^{th} cluster, c_j is the centroid of j^{th} cluster, and $\|.\|$ is the similarity measure. Usually, Euclidean distance is used to calculate the similarity between a data point and a cluster center. K-means clustering is an iterative approach that starts with random initial partitions as clusters. Then, the data points are reassigned to clusters by minimizing the sum of distances from each data point to cluster centers, until convergence. The algorithm can be summarized as follows:

1. Initialize number of cluster k and randomly define k cluster centers
2. Assign each data point to the closest cluster center
3. Recalculate the cluster centers using current cluster memberships

4. Repeat the process until convergence or squared error criterion satisfies a minimum threshold value

The quality of final clustering highly depends on the random selection of initial cluster centers, is the major limitation of k-means clustering.

SLIC (Achanta et al. 2012) technique is an adaptation of popular k-means clustering, to generate image superpixels (group of pixels with similar characteristics). Desired number of superpixel k, is the only parameter required for this method and this technique produces approximately equal sized superpixels. If an image contains N pixels, each superpixel will be of approximately N/k pixels. In initialization step, k initial centers are sampled on a regular grid interval of $S = \sqrt{N/k}$. Since, the approximate size of superpixel is $S \times S$, search for similar pixels is done in a $2S \times 2S$ region around the cluster center (Achanta et al. 2012). Each image pixel is associated with the closest cluster center whose search area overlaps that pixel. Color similarity and spatial proximity are considered while calculating the distance measure, to determine the closest cluster for each pixel. The distance measure for clustering is defined as:

$$D = \sqrt{{d_c}^2 + \left(\frac{d_s}{S}\right)^2 m},$$

where $d_c = \sqrt{(I_j - I_i)^2}$ is the intensity similarity, and $d_s = \sqrt{(x_j - x_i)^2 - (y_j - y_i)^2}$ is the spatial proximity of cluster center $j \in [1, 2, \cdots, k]$ and an image pixel i. The positive constant m controls the relative importance of color and spatial proximity. Once all pixels are assigned to the closest clusters, cluster centers are recomputed from the mean of all pixels in the corresponding cluster. This process is repeated iteratively until convergence. Finally, if any disjoint pixels exist at the end of the clustering process, those are assigned to the closest neighboring cluster. In Chapter 5 of this thesis, image superpixels are generated using SLIC technique and cervical nuclei are found by merging SLIC superpixels based on some merging criterion.

FCM clustering algorithm was proposed in (Dunn 1973) and further improved in (Bezdek 1981). In fuzzy clustering, each of the data points belongs to all the clusters with varying degrees of membership in range $[0, 1]$. This method is based on an iterative optimization algorithm and the minimization of a quadratic objective function. Let $x = [x_1, x_2, ..., x_N]$ denote an image with N pixels. FCM clustering algorithm divides the image into C clusters, by calculating the cluster

centers, the fuzzy partition matrix, and by minimizing the objective function J:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} U_{ij}^{m} \left\| x_i - c_j \right\|^2,$$

where C is the number of clusters, N is the number of pixels, m is the fuzziness weighting factor ($1 \leq m < \infty$) which controls the degree of fuzzy overlap, U_{ij} is the fuzzy partition matrix, and it defines the membership degree of pixel x_i in the cluster c_j, and $\|.\|$ is any criteria to measure the similarity (typically the Euclidean distance is used). When the pixels close to the cluster centers are assigned high membership values, and low membership values are assigned to the pixels far from the cluster centers, the objective function is minimized. An initial guess of the fuzzy partition matrix U is made to start the clustering process, and U converges by iterating and updating the partition matrix and cluster centers as follows:

$$U_{ij} = \left(\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{(m-1)}} \right)^{-1},$$

and

$$c_j = \frac{\sum_{i=1}^{N} U_{ij} x_i}{\sum_{i=1}^{N} U_{ij}}.$$

Convergence is checked by comparing the cluster centers or the partition matrix, at successive iterations. Image segmentation is obtained by assigning a pixel to the class with the highest membership degree. In Chapter 3 of this thesis, FCM clustering is modified by incorporating a circular shape function and used for cervical nuclei segmentation.

2.1.2.4 Graph-based segmentation

A graph G is composed of a pair (V, E), where V is the set of vertices/nodes of G. The set E contains the edges of G, where edges are represented as the pairs of elements in V. The graph G is called a directed graph, if the edges in set E have directions, otherwise the graph is called undirected graph. Graph based image segmentation methods are based on graphs, where image pixels stands for the vertices and edges are defined by an adjacency relationship of the image pixels.

2.1.2.5 Graph cut

In graph cut segmentation, the graph $G = (V, E)$ is considered as a directed weighted graph. In this algorithm, nodes and edges are divided in two types: neighborhood and terminal (in other words, source s and sink t). Node p is connected with node q, if q is in the set of neighborhood nodes of p and neighborhood nodes are connected by neighborhood edges. Edges connecting node p with terminal nodes s or t are called terminal edges. A subset of edges $C \in E$ is called a cut, if the terminal nodes are completely separated by inducing $G = (V, E - C)$. If all the edges in the cut are removed, there will be no path from terminal s to t. Thus, the cut partitions all the nodes into two disjoint subsets S and T, where $s \in S$, $t \in T$ and the graph cut cost $|C|$ is the sum of weights w of all the edges in the cut, defined as:

$$|C| = \sum_{i \in S,\, j \in T} w(i, j).$$

Graph cut segmentation (Boykov et al. 1999, 2001) is defined as an energy minimization problem. An image I is considered as a graph with a set of pixels P, $p \in P$ and a set of labels L. A labeling function $f : I \ni p \mapsto f_p \in L$ needs to be set to minimize some energy function $E(f)$. The function $E(f)$ can be expressed as:

$$E(f) = \sum_{p \in P} R_p(f_p) + \sum_{p \in P,\, q \in N_p} B_{pq}(f_p, f_q),$$

where N_p is the set of neighboring pixels of p, the cost of labeling $f_p \in L$ to pixel p is $R_p(f_p)$, and the cost of assigning different labels $f_p, f_q \in L$ on the neighborhood pixel pair p and q is $B_{pq}(f_p, f_q)$. Thus, for graph cut cost energy function, $R_p(f_p)$ and $B_{pq}(f_p, f_q)$ can also be termed as regional and boundary terms, respectively. For the disjoint subsets S and T of the graph G, binary label is assigned to pixels: p is assigned $f_p = 1$ (object) if $p \in S$, otherwise $f_p = 0$ (background) if $p \in T$. The goal of segmentation methods using graph cut is to assign graph edge weights in such a way that the minimum cut cost $|C|$ and minimum energy $E(f)$ are equal.

In (Felzenszwalb & Huttenlocher 2004), an efficient graph based segmentation algorithm is proposed - based on pairwise region comparison. Image is considered as a graph and edge weight is the measure of dissimilarity between neighboring nodes. The segmentation process starts by considering each pixel being a single

component/region and the algorithm produces minimum spanning tree (MST) by comparing the inter component difference to the intra/within component differences for merging as:

$$D(C_1, C_2) = \begin{cases} true, & \text{if } Dif(C_1, C_2) > MInt(C_1, C_2), \\ false, & \text{otherwise,} \end{cases}$$

where

$$MInt(C_1, C_2) = min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)),$$

and

$$\tau(C) = \frac{k}{|C|}.$$

Here, $C_1, C_2 \subseteq V$ are two components, $Dif(C_1, C_2)$ is the difference between two components and $MInt(C_1, C_2)$ is the minimum internal difference. Parameter k controls the scale of segmentation outcome; large value of k yields large segment. The segmentation algorithm is closely related to Kruskal's algorithm for constructing MST. In Chapter 4 of this thesis, a weighted circular shape guidance is incorporated in efficient graph based segmentation technique, and the modified algorithm is used to segment cervical cell nucleus.

2.1.3 Feature generation

In the area of digital image analysis, a feature is a characteristic or measurable quantity of an object/ROI in an image, that can help to distinguish that object from other objects or background in the image. Choosing features with informative and enough discriminatory properties, is the crucial step for effective recognition and classification performance of an application. Prediction models use feature vector or set of features for prediction/object classification. Intensity, morphological/shape, and texture features are commonly used for classification tasks (Theodoridis & Koutroumbas 2008).

2.1.3.1 Intensity features

Intensity features - calculated from image graylevel/intensity values in ROI, are the simplest and widely used features for pattern recognition. In cervical cell

classification, intensity features are widely used, since cell nucleus and cytoplasm have different contrasts than surrounding background (Jantzen & Dounias 2006). For example, mean/average intensity, standard deviation, variance, skewness, contrast, local minima and maxima, kurtosis of ROI, are some intensity based features used in literature for cervical cell analysis (Jantzen & Dounias 2006, Plissiti et al. 2011*b*, Gençtav et al. 2012, Win et al. 2018). Algorithms exclusively based on intensity features may fail in the presence of intensity inhomogeneity and poor contrast. However, intensity inhomogeneity resulting from chromatin distribution can be useful for abnormality detection using texture features.

2.1.3.2 Morphological features

Morphological/shape features are calculated from detected/segmented boundaries of objects. In cervical cell analysis, change in cell morphology often reflects tissue specific condition or state of abnormality. Area, major and minor axis length, circularity/roundness, compactness, perimeter, elongation, solidity, eccentricity, equivalent circular diameter, and actual diameter, are the most common morphological features used for cervical cell classification task (Jantzen & Dounias 2006, Plissiti et al. 2011*b*, Gençtav et al. 2012, Chankong et al. 2014, Mariarputham & Stephen 2015, Win et al. 2018). Though morphological/shape features can assist in detecting the presence of abnormality, these features are highly dependent on segmentation accuracy, which is often not precise enough. To improve the accuracy of segmentation, three novel segmentation techniques with prior guidance are presented in Chapters 3, 4 and 5 of this thesis. Morphological features extracted from precisely segmented cervical nuclei using the proposed techniques, are used to detect abnormality in cervical cells in Chapter 6.

2.1.3.3 Textural features

Texture analysis of an image region is based on graylevel/intensity distribution of pixels in ROI (Theodoridis & Koutroumbas 2008). Textural features are useful to identify different classes/stages, in pattern recognition or image classification based applications. In cervical cell image analysis, texture features have been widely used to identify/characterize abnormality in cervical cells and different stages of abnormality.

First order statistic based textural features depend on the statistical properties of the intensity histogram of ROI and are calculated from central moments.

Variance, skewness, and kurtosis are the most frequently used first-order measures (Theodoridis & Koutroumbas 2008). Some first order measures were used for cervical cell analysis in (Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017, Bora et al. 2017).

Second order textural features are extracted from second order histograms - that contain spatial information about the ROI. Gray level co-occurrence matrix (GLCM) or gray-tone spatial-dependence matrix (Haralick et al. 1973) is commonly used to extract second order textural features from an image. Angular second moment/energy/uniformity, homogeneity, contrast, entropy, inverse different moment, correlation, are some commonly used measures calculated from GLCM (Haralick et al. 1973, Theodoridis & Koutroumbas 2008). Some of these textural features were used in (Plissiti et al. 2011*b*, Chankong et al. 2014, Mariarputham & Stephen 2015, Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017, Bora et al. 2017, Win et al. 2018) for cervical cell classification.

Graylevel run length matrix (GLRM) (Galloway 1975) is popular for extracting radial features. Short run emphasis, long run emphasis, graylevel non-uniformity, run length non-uniformity, and run percentage, are some textural features extracted from GLRM (Theodoridis & Koutroumbas 2008, Silva et al. 2019). These features can be calculated from different directions (0°, 45°, 90°, 135°). Histogram of oriented gradients (HOG) features (Dalal & Triggs 2005) and local binary pattern (LBP) features (Ojala et al. 2002), are two other popular local texture descriptors, and also used in cervical cell classification applications (Plissiti et al. 2011*b*, Mariarputham & Stephen 2015, Phoulady, Zhou, Goldgof, Hall & Mouton 2016).

2.1.4 Feature selection

Feature selection or reduction - is the process of selecting a subset of features from a given number of features, in order to reduce feature space dimension and mutual correlation, while retaining most discriminatory information in a context (Theodoridis & Koutroumbas 2008). The aim for feature selection is to maximize inter-class difference and minimize intra-class difference in feature vector space. A classifier's error rate first decreases and then increases with the increment of feature dimension or number of features - for a finite sample size, and this is called "peaking" phenomenon (Raudys & Jain 1991). Optimal number of features is the point, where probability of misclassification or classification error

is minimal and then starts to increase. Finding the optimal number of features is most critical for small sample size (Hua et al. 2004). There is a possibility that the designed classifier could be overfitted to the sample data, if the used feature set is too big. Therefore, feature subset selection or determining the optimal number of feature, is crucial for the performance and generalization properties of classifiers.

In scalar feature selection, any class separability criterion such as: divergence, receiver operator characteristics (ROC) curve, or Fishers discriminant ratio (FDR), is computed individually for each of the generated features. Features are then ranked/ sorted with respect to this criterion. Finally, the best scoring features are selected to form the feature vector.

Feature vector selection is the process to find the best combination of features - obtained through scalar feature selection process. Search-based feature set selection techniques involve a search strategy, to select feature subsets, and an objective function to determine class separability power of that feature subset. Exhaustive search, sequential backward or backward selection, floating search selection, are the common feature vector selection techniques.

In exhaustive search strategy, all combinations of n features out of m given features ($n \leq m$) are considered. The total number of possible feature vectors is $2^n - 1$, and this can be very large even for small n (Theodoridis & Koutroumbas 2008). Therefore, the use of exhaustive search in practical applications is computationally expensive and often not feasible.

To overcome the issue with exhaustive search, sub-optimal search techniques (for example sequential backward selection, sequential forward selection, or floating search methods) can be used. These techniques do not examine all possible combinations of feature sets. Sequential backward selection technique starts with the whole features set, eliminate one feature at a time from possible combinations. The combination with the best value is selected as the sub-optimal feature set, based on some criterion function. Sequential forward selection starts from the single best feature, adds one feature at a time, and select the best combination. However, both of these techniques suffer from "nesting effect" (Theodoridis & Koutroumbas 2008); once a feature is discarded or added, there is no way to reconsider it. Floating search methods (backward and forward) provide flexibility to reconsider a feature that was previously considered. These methods improve the performance compared to the sequential feature selection techniques, however complexity is increased substantially.

A fuzzy entropy and similarity classifier based feature selection technique was proposed in (Luukka 2011). This technique is based on the measure of degree of fuzziness defined in (Luca & Termini 1972), and the definition of this corresponds to the Shannon's probabilistic entropy measure (Shannon 1948). Fuzzy entropy measure for a fuzzy set A can be defined as in (Luukka 2011):

$$H(A) = -\sum_{i=1}^{n} (P_A(x_i) \ln P_A(x_i) + (1 - P_A(x_i)) \ln (1 - P_A(x_i))),$$

where n is the number of elements in A. Similarity values $P_A(x_i) \in [0, 1]$ are calculated by comparing ideal and sample vectors. In other words, $P_A(x_i)$ denotes the probability/membership degree of x_i in A. Features with high similarity value will have low fuzzy entropy value. That means, when the uncertainty is high, high entropy value is expected. Using this process, features can be ranked based on fuzzy entropy values, with the assumption that the feature with the lowest entropy is the most informative. Since, this technique ranks features in a feature set, computation cost is considerably low compared to the aforementioned feature selection techniques. In Chapter 6 of this thesis, fuzzy entropy and similarity classifier based feature selection technique is used to rank features (intensity, shape and texture) - extracted from segmented cervical nuclei, and generate a feature subset with top ranking features. Subsequently, this feature subset is used for abnormality detection in cervical cells.

2.1.5 Classification

Once the feature selection process is completed, the feature set can be used with a classifier to detect or grade an abnormality in cervical cells. Based on the learning process used, classification techniques can be divided into supervised and unsupervised. In supervised machine learning, the classifier learns about the context from a training set. Unsupervised machine learning is based on some underlying structure or distribution in the data, and it does not require any training data. Bayesian classifier, Fisher linear discriminant analysis (LDA), k-nearest neighbor (KNN), support vector machine (SVM), artificial neural network (ANN), decision tree, random forest, and Ensemble are some popular examples of supervised machine learning techniques. Clustering (k-means, spectral, hierarchical) based techniques are popular examples of unsupervised techniques for classification. In this thesis, LDA, KNN, SVM, and Ensemble classifiers are used in Chapter 6, to classify cervical cytology cells.

Fisher linear discriminant analysis: LDA (Fisher 1936) is a type of linear classifier that considers the classification problem as linearly separable. The linear discriminant function, or decision hyperplane, for a d dimensional feature set x can be defined as:

$$y = w^T x,$$

where w is the weight vector. LDA technique aims to find the value of w maximizing the distance between the mean of the two classes and minimizing the variation within each class, which guarantees maximum class separability. Training data is used to find the value of w, then test data or new observation is classified based on the training model.

K-nearest neighbor classifier: KNN (Fukunaga & Narendra 1975) is one of the simplest supervised machine learning algorithm. This technique classifies an object using majority vote of its neighbors, where neighbors are from a set of known/correct classification or training set. The new object is allocated to the most common class among its k nearest neighbors (if $k = 1$, then the new case is allocated to the class of its single neighbor). Various distance metrics (including Euclidean or Mahalanobis distance), can be used to find k nearest neighbors of the test instance in the training set (Theodoridis & Koutroumbas 2008). With a large number of training samples, this classifier exhibits good performance. In addition, overall effect of noise can be reduced by choosing a large k value, and a suitable value for k can be found using cross validation, grid search or other techniques.

Ensemble classifier: The aim of ensemble methodology is to construct a predictive model from a set of classifiers, and classify a new instance using voting of classifiers predictions (Rokach 2005). Both weighted and unweighted voting can be used. Performance of an ensemble classifier depends on the accuracy and diversity of the classifiers in the set. Ensembles classifiers can be constructed using various methods as: Bayesian voting, bagging, boosting, feature selection ensembles, error-correcting output coding, random forest, and randomness injection (Dietterich 2000).

Bootstrap **agg**regat**ing** (bagging) (Breiman 1996) is the most well known method for ensemble methodology. In this technique, multiple versions of a predictor are generated and aggregated, at each split of decision tree. This technique

can reduce the variance of a decision tree and improve accuracy of classification task using different classifiers.

Support vector machine: SVM is a popular classification technique than can be applied to both linearly and non-linearly separable classes. This algorithm was originally proposed in (Vapnik & Lerner 1963). The goal of SVM algorithm is to design a hyperplane (or set of hyperplanes) that correctly classifies the training data. The hyperplane can be described analytically by equation:

$$f(x) = w^T x + w_0 = 0,$$

where $x = [x_1, x_2, \cdots, x_N]$ is the feature vector of the training set and w_0 is a threshold. Consider, two linearly separable classes ω_1 and ω_2. All hyperplanes are characterized by its direction (determined from w) and exact position in space (determined from w_0). The hyperplane that separates the classes with maximal possible margin from both classes, is the optimal hyperplane. The distance of a point from a hyperplane can be found as:

$$z = \frac{|f(x)|}{\|w\|}.$$

To maximize the distance z, w needs to be scaled in a way that:

$$w^T x + w_0 \geq 1, \quad \forall x \in \omega_1,$$

$$w^T x + w_0 \leq -1, \quad \forall x \in \omega_2.$$

This is a non-linear optimization task. The Karush-Kuhn-Tucker (KKT) condition can be used to solve this problem with Lagrange multipliers λ_i, as:

$$w = \sum_{i=1}^{N} \lambda_i \, y_i \, x_i, \qquad \sum_{i=1}^{N} \lambda_i \, y_i = 0.$$

In order to design non-linear SVM classifiers, different kernel functions (polynomial, hyperbolic tangent, Gaussian or radial basis function) can be used to map the input data into a higher dimensional space.

2.2 Prior guided image segmentation

Accurate image segmentation is still a challenging task in many medical image analysis applications (that deal with for example: cytology images, CT and MRI images of human organs, glands, blood vessels, or lymph nodes, mammograms, or histopathology images), despite of great advances in segmentation techniques over past decades. Traditional image segmentation methods based on the notion of homogeneity or only image information, have great chance to fail in cases of images with noise, intensity inhomogeneity, poor contrast, or presence of complex objects. Prior information about the target object can be helpful in the aforementioned cases. If any particular object or ROI needs to be segmented from an image, prior information of that ROI can guide the segmentation process and make it more robust and accurate (Tsai et al. 2003, Vu & Manjunath 2008, Ibragimov et al. 2014, Nosrati & Hamarneh 2016). Many attempts have already been taken in the direction of incorporating prior guidance into segmentation tasks. Some commonly used priors are: user interaction, shape prior, boundary information, appearance prior, statistical and physical model, spatial prior, and topological prior.

2.2.1 User interaction

The simplest way to characterize the target object for the segmentation process, is through user interaction. Input from the user can be provided by: specifying seed points, object boundaries, or bounding boxes for the target object. Seed points can be specified by labeling some image pixels as being part of the object of interest and some outside of the targeted object. In case of bounding box prior, the user can specify a sub-region by drawing a box around the target object.

User input was incorporated into graph cut energy function in (Boykov & Jolly 2000, 2001). In this interactive graph cut segmentation, user selects some pixels as part of object or background - named hard constraints. Soft constraints are determined from the likelihood of a pixel being a part of the object or background. User defined segmentation constraints were incorporated in level set technique in (Paragios 2003, Cremers et al. 2007, Ben-Zadok et al. 2009).

In case of boundary specification, user can provide input for the initial seeds along the target object boundary. In (Freedman & Zhang 2005), user needs to draw circles around the foreground and squares in the background. These inputs were treated as landmark data, and Procrustes method was used to match

template curve to the data. Live-wire (LW) (Barrett & Mortensen 1997), is an interactive boundary extraction tool that requires minimal user provided seeds on the target object boundary. Segmentation using LW involves user interaction by positioning the cursor on the boundary of the object to select a point. For the next cursor position, LW shows real time optimal path from the first point to the current cursor position. In Live-lane (LL) approach (Falcão et al. 1998), user needs to select only one initial point. Then, steer the cursor within a certain width lane around the boundary, where the lane width varies adaptively with the speed of cursor movement. During user's action, successive points are selected automatically in an intermittent way. LW segments are calculated and displayed in real time for each pair of successive points, thus making LW less dynamic than LL.

GrabCut (Rother et al. 2004) is one of the most popular techniques in the direction of bounding box prior. In this method, users can drag a rectangle around the target object to indicate the background region. Similar rectangular bounding box was used in (Lempitsky et al. 2009) to impose a strong topological prior. This prior prevents the segmentation from shrinking and splitting, thus ensures that the segmented object is sufficiently close to the edges of the bounding box.

2.2.2 Shape and appearance prior

Shape is the geometric information that remains after location, scale and rotational effects are filtered out from an object (Stegmann & Gomez 2002). Hence, shape is a powerful prior to segment ROI from an image. Desirable properties for shape representations are: invariance for geometric transformations (as scaling, rotation, or translation), robustness in terms of initialization, ability to generate prior model from training data even with small training set (Litvin & Karl 2005). Appearance is an important visual prior while distinguishing a target object from the background. Distribution of features derived from intensity, color, or texture can be incorporated as appearance cues into segmentation methods.

In (Slabaugh & Unal 2005), an elliptical shape prior was incorporated in graph cut based segmentation, through an iterative refinement process of an initial ellipse - found through user interaction. The initial ellipse was used to form a binary shape mask M. Both data and boundary terms of the proposed technique had two instances; one from the image data and another from the shape prior. Image data term was calculated from the average of pixels' intensities inside and

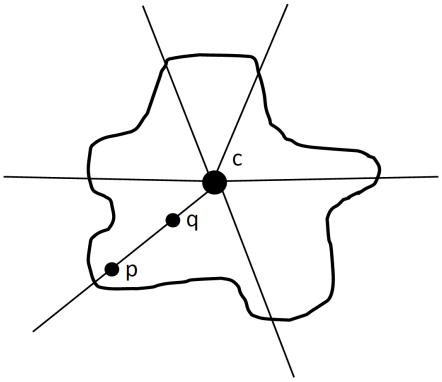


Figure 2.1: An example of star shape object.

outside the mask M. Shape prior based data term was calculated using the shape mask M as:

$$S_{\text{ellipse}} = \sum_{i \in P} |M_i - f_i| \,,$$

where P is the set of pixels and f_i is the assigned label to pixel i. The value of M_i is 0, if the pixel i is inside the mask and 1 if the pixel i is outside. A graph was computed over the set of pixels in the narrow band around the shape mask. After the minimum cut was found from the graph, the current shape mask was updated by finding the best fitting ellipse to the graph cut points. Then, a new band was formed around the updated mask, and this process iterated until convergence. In (Leung et al. 2004), an elliptic shape function was incorporated in fuzzy c-means clustering. The shape function was comprised of both color information and spatial distance. The dissimilarity measure of fuzzy c-means clustering was redefined to include the shape function.

A generic star shape prior was proposed in (Veksler 2008) and incorporated into graph cut based segmentation. This prior can be applied to diverse classes of convex objects with an assumption of the object's center position. An object has a star shape, if for any point p inside the object, all the points on the straight line connecting center c and p are also inside the object (see Figure 2.1). The shape constraint was incorporated in the energy function as:

$$E(f) = \sum_{p \in P} R_p(f_p) + \lambda \sum_{(p,q) \in N} B_{pq}\left(f_p, f_q\right) + \sum_{(p,q) \in N} S_{pq}^{\text{star}}(f_p, f_q),$$

where R_p is the region term, B_{pq} is the boundary term, and S_{pq}^{star} is the shape prior. P is the set of all image pixels, (p, q) is the ordered pixel pairs, N is the

neighborhood (4 or 8 connected), f_p and f_q are the labels (0 or 1) assigned to pixels p and q, and parameter λ is the weight to balance the relative importance. Shape prior term S_{pq}^{star} was defined as:

$$S_{pq}^{\text{star}}(f_p, f_q) = \begin{cases} 0, & \text{if } f_p = f_q, \\ \infty, & \text{if } f_p = 1 \text{ and } f_q = 0, \\ \beta, & \text{if } f_p = 0 \text{ and } f_q = 1. \end{cases}$$

Star shape prior tries to remove the shrinking bias of graph cut segmentation towards shorter boundary segments, using the value of parameter β, in the absence of strong data term.

Prior shape and appearance knowledge was considered for watershed based segmentation in (Hamarneh & Li 2009). Shape histogram and statistical analysis of image intensities, were used to design the prior knowledge about the shape and appearance. Aligned binary shape images were used to generate shape histogram. Appearance knowledge was captured as mean and variance from the intensity patches of the target object.

Usually, shape based methods impose a constant weight on shape prior, though all the pixels in an image do not require support from the shape prior. An adaptive way to determine the importance of shape prior at pixel level was presented in (Wang, Zhang & Ray 2013). The shape weight term was computed from image intensity based probability map as:

$$S_{pq} = e^{-(\alpha_p - \alpha_q)^2},$$

where α denotes the likelihood of a pixel (in the range [0, 1]) to be in foreground. Probability map α was obtained from the smoothed images by applying Gaussian filter to the original images. This adaptive shape prior was incorporated with two existing graph based segmentation energy functions (Freedman & Zhang 2005, Veksler 2008), and demonstrated superior performance compared to the baseline method.

2.2.3 Statistical shape model

Statistical analysis of shapes and building models for image segmentation, are important and active fields of research. In statistical shape analysis, geometrical features measured from set of shapes are analyzed using statistical methods. From

this analysis, statistical shape models can be built. Usually, intra-class shape variation is the basis for developing a shape model. Shapes can be represented either explicitly or implicitly depending on the segmentation method used, and this is the foremost decision for designing statistical shape models.

2.2.3.1 Shape representations

The simplest and generic method for shape representation is the use of a set of landmarks to represent shapes explicitly (Heimann & Meinzer 2009). Landmark points correspond to the position of an object feature. Location/coordinate of all landmark points are concatenated as a vector to describe the shape (Stegmann & Gomez 2002). For n landmark points, the shape can be presented as: $S = [x_1, y_1, x_2, y_2, \cdots, x_n, y_n]^T$.

Medial or skeleton is another form of explicit shape representation using center lines and corresponding radii as medial primitives. This medial representation was modeled using a coarse to fine representation of figural shapes in (Pizer et al. 1999). This approach was extended to 3D in (Pizer et al. 2003) and named *m-rep*, where a single figural shape was used to segment objects. In this model, an object was represented by a hierarchy of single figure *m-rep* models.

Zero-level set is an implicit shape representation technique, which originated from level set approach introduced in (Osher & Sethian 1988). In this approach, pixels are represented by their Euclidean distance to the boundary and zero-level set is considered as the contour or shape. Commonly, signed distance function (SDF) is used as the level set function, while describing shapes (Leventon et al. 2000, Paragios et al. 2002). Assume, a level set function is $\phi : \Omega \rightarrow R^+$ refers to the shape S, that defines a region R in the image plane Ω, and $\phi = 0$ is the zero-level set. To represent shape S, function ϕ can be defined as follows:

$$\phi_S(x,y) = \begin{cases} 0, & (x,y) \in S, \\ D((x,y),S), & (x,y) \in R_S, \\ \text{-}D((x,y),S), & (x,y) \in [\Omega - R_S]. \end{cases}$$

where, $D((x,y),S)$ is the minimum Euclidean distance between point (x,y) and shape S.

2.2.3.2 Shape model construction

Statistical shape model can be constructed by finding the mean shape and modes of variation from a training set (Heimann & Meinzer 2009). Point Distribution Model (PDM) was introduced in (Cootes et al. 1992), where landmark points from training images were used to build statistical shape model. PDM was further explored in (Cootes et al. 1995) and termed Active Shape Model (ASM). This model can deform iteratively to find the best match of an example in a new image. Objects are represented by a set of landmark points on the boundary. The main modes of variation are found by applying Principal Component Analysis (PCA) on the training shapes. A training shape can be approximated as:

$$S = \overline{S} + \sum_{i=1}^{k} P_i \, w_i,$$

where $\overline{S}$ is the mean shape. P_i are the principal components and w_i are their corresponding weights. Given an initial estimate of the object's position, the model iterates to deform towards the shape and try to optimize an energy function.

In (Tsai et al. 2003, Zhu-Jacquot & Zabih 2007, Grosgeorge et al. 2013), shapes were represented using level sets. In (Tsai et al. 2003), training shapes were aligned using similarity transformation, and a parametric shape model was generated from the training shapes using SDF. The model was used in a region-based curve evolution framework for medical image segmentation. In (Zhu-Jacquot & Zabih 2007, Grosgeorge et al. 2013), shape model of the target object was constructed using mean shape and Eigen shapes (found by PCA analysis of the aligned training shapes). In (Zhu-Jacquot & Zabih 2007), statistical shape model was added in graph cut based energy function. Shape fitting and segmentation problem was solved using expectation maximization (EM) approach. In (Grosgeorge et al. 2013), shape model was registered to the image via minimal user interaction with identifiable landmarks by experts (only 2 landmarks are required in this case), and incorporated in the boundary term of graph cut energy function.

2.2.4 Spatial prior

Spatial prior or spatial information of target object, is popular for prior knowledge representation. Many attempts have already been taken to incorporate spatial prior in image segmentation, specially in clustering based techniques.

In (Noordam et al. 2000), geometrical information was included in FCM clustering, with an aim to develop a semi-supervised segmentation technique. Geometrical guidance was determined from the local neighborhood of each pixel and incorporated in the objective function for clustering. In (Ahmed et al. 2002), a new bias term was introduced to control/regulate the clustering process to compensate the effect of intensity inhomogeneity, and included in the objective function of FCM clustering. The new term allows the clustering of a pixel being influenced by the clusters of its immediate neighborhood ($k \times k$ window around the target pixel). This neighborhood effect works as a regularizer and enforces the clustering to be homogenous. The new bias term for pixel $\{x_k\}_{k=1}^{N}$ was defined in (Ahmed et al. 2002) as:

$$B = \frac{\alpha}{N_R} \sum_{i=1}^{C} \sum_{k=1}^{N} u_{ik}^{p} \left(\sum_{x_r \in N_k} \|x_r - v_i\|^2 \right),$$

where α is the parameter to control neighborhood effect, N_k is the set of neighbors in a window around x_k, N_R is the cardinality of N_k, C is the number of clusters, N is the total number of pixels, u_{ik}^{p} is the fuzzy partition matrix with weighting exponent p, and v_i is the i^{th} cluster center. However, lack of robustness to noise and outliers, optimal selection of parameter α to balance the effectiveness of segmentation and robustness to noise, and computation cost for finding neighborhood of a pixel, are some limitation of this technique.

Two variants of the method presented in (Ahmed et al. 2002) were introduced in (Chen & Zhang 2004), to reduce the computation cost by simplifying the neighborhood term. Neighborhood term was replaced by mean and median filtered image. This filtration process can be done in advance, thus execution time of the clustering process is reduced.

Another approach to alleviate the limitations in (Ahmed et al. 2002), was proposed in (Cai et al. 2007). A novel similarity measure S_{ij} containing both local spatial and graylevel information from neighborhood was defined to replace α. Consider, i^{th} pixel is the centroid/central pixel in a local window and j^{th} pixel is in the set of neighbors centering at i^{th} pixel. The local spatial relationship was given by

$$S_{s_{ij}} = \exp\left(\frac{-\max(|x_j - x_i|, |y_j - y_i|)}{\lambda_s} \right),$$

where (x_i, y_i) is the spatial location of i^{th} pixel and λ_s is the scaling factor. Local

graylevel similarity measure was found as:

$$S_{g_{ij}} = \exp\left(\frac{-\|l_i - l_j\|^2}{\lambda_g \times \sigma_{g_i}^2}\right),$$

where l_i and l_j are the graylevels of i^{th} and j^{th} pixels, respectively. λ_g is the scale factor, and $\sigma_{g_i}^2$ is the local density function of i^{th} pixel. The density function $\sigma_{g_i}^2$ was calculated as:

$$\sigma_{g_i}^2 = \sqrt{\frac{\sum_{j \in N_i} \|l_j - l_i\|^2}{N_R}},$$

where N_R is the cardinality of neighborhood. The similarity measure S_{ij} to replace α, was defined as:

$$S_{ij} = \begin{cases} S_{s_{ij}}\, S_{g_{ij}}, & \text{if} \quad j \neq i. \\ 0, & \text{otherwise.} \end{cases}$$

In (Chuang et al. 2006), a spatial function was incorporated into the fuzzy partition matrix of FCM clustering. The summation of membership function values in a square window - centered at the target pixel, was used to compute this function. In (Wang et al. 2008), the usual distance metric was replaced with a novel dissimilarity measure. Both local and non-local information were taken into account while devising this measure. The weighted summation of Euclidean distance in a local neighborhood was used as the local information. The weighted average of all pixels' Euclidean distances in the image was treated as non-local information. A weighting factor was used to balance the contribution of local and non-local information.

In (Wang, Song, Soh & Sim 2013), the objective function of FCM clustering was modified. Noise and outliers were identified from a measure of input data distribution. Local neighborhood information of a pixel was considered, while developing the improved adaptive similarity measure. Both of these were included in the modified objective function. In (Adhikari et al. 2015), local spatial information from adjacent neighborhood was incorporated into fuzzy membership function. Conditional spatial membership variable was introduced to define the affinity of a pixel x_k in i^{th} cluster v_i as:

$$f_{ik} = \frac{\sum_{j \in N(x_k)} U_{ij}}{M},$$

where $N(x_k)$ is the square neighborhood with x_k as the center, U_{ij} is the global membership value, and M is the number of pixels in the neighborhood. This conditional variable f_{ik} was used to define a local spatial membership function. A weighted membership of both global and local functions, was considered in the modified clustering technique.

2.2.5 Boundary information

Knowledge about the boundary/edges of a target object is a powerful cue, while delineating object of interest from the background. Usually boundaries are the locations, where an abrupt change in intensity or large contrast is observed. Change of labels (foreground/background) are expected in these areas. Traditional variational methods of segmentation (Malladi et al. 1995, Caselles et al. 1995) are based on this idea. Directional change in intensity in an image is expressed with image gradient, where gradient magnitude expresses the variation in local contrast.

In (Malladi et al. 1995, Caselles et al. 1995), boundary information was used with level set and active contour based techniques. In both methods, energy functionals were minimized by aligning objects boundaries with high intensity gradients. A graph-based image segmentation technique - "random walks", was proposed in (Grady 2006). Gradient magnitude was used in this technique, while formulating the segmentation problem.

In (Singaraju et al. 2008), a modification of "random walks" (Grady 2006) was proposed. Gradient magnitude and the direction of transition were considered for detecting object boundary using the graph based framework. In this technique, edge weights were calculated considering the direction of edges. For a pair of neighboring nodes v_i and v_j with intensity values x_i and x_j (where $x_i > x_j$), weight of the directed edges e_{ij} and e_{ji} were given by:

$$w_{ij} = e^{-\beta_1(x_i - x_j)^2} \quad \text{and} \quad w_{ji} = e^{-\beta_2(x_i - x_j)^2},$$

for some $\beta_1 \geq \beta_2 \geq 0$ as asymmetric penalty for gradient directionality. In this approach, edge weight for transition from dark to bright is more than the edge weight for bright to dark transition.

In some segmentation techniques, boundary or edge information was not directly encoded in the segmentation algorithm. In (Wählby et al. 2004), marker controlled watershed segmentation was applied on gradient magnitude image -

calculated from first derivative based Sobel operator. After initial segmentation, results were further refined considering gradient magnitude edge strength. In (Schindler & Suter 2008), an edge map with closed edge chain or contour was generated, and used as a cue for object detection. The edge map was obtained from superpixel image found using SRM segmentation technique. In (Li et al. 2011), a new fuzzy level set algorithm was presented, where results from spatial fuzzy clustering were used as the initial contour for the level set technique. Final segmentation boundaries were found by locally regularized evolution.

2.3 Related works

Researchers are working for several decades to automate the screening process of cervical smear cells. In (Sarwar et al. 2019), cervical cell segmentation techniques, from the year 1977 to present (more than four decades) are reviewed. Precise or shape preserving nucleus segmentation, is the most important factor in cellular morphology calculation and computer-aided approach of cytological image analysis. Comprehensive reviews of nucleus detection and segmentation techniques for digital pathology images can be found in (Irshad et al. 2014, Xing & Yang 2016, Hayakawa et al. 2019)

In Chapter 1 of this thesis, nucleus segmentation techniques for automated systems, were roughly divided in five categories. Overview of some segmentation techniques from those categories, is presented in this section.

2.3.1 Thresholding and morphological analysis based approaches

In (Cahn et al. 1977), a segmentation threshold was determined considering the stability of the cell perimeter. For this, an initial threshold was found from a sparse histogram. This threshold value was modified using a cell perimeter information based iterative testing, and used to detect the target object in an image, ignoring the background.

In (Borst et al. 1979), three histograms from: extinction value distribution, gradient and circumference, were found from filtered Pap smear images. A single combination histogram was generated from three histograms. Then, two thresholds were computed from the combination histogram, to separate nucleus from cytoplasm and cytoplasm from background.

In (Bengtsson et al. 1979), two different techniques were used to find thresholds for nucleus and cytoplasm. An adaptive graylevel thresholding from image histogram was used to find the threshold for nucleus segmentation. Cytoplasm was segmented using 2-dimensional thresholding in the bivariate histogram. In (Yang-Mao et al. 2008), an adaptive optimal threshold was defined to detect nucleus and cytoplasm edge pixels. Gradient histogram of image was used to determine this threshold.

Morphological reconstruction in combination with the detection of regional minima was used in (Plissiti et al. 2011*a*,*b*), to find initial nuclei centroids. In (Plissiti et al. 2011*a*), candidate nuclei boundaries were found using 8-radial profiles in equal arc length intervals, and this process started from initial candidate nuclei centroids. In (Plissiti et al. 2011*b*), nuclei centroids were used as markers for watershed transform. This marker controlled watershed transform was used to find nuclei boundaries from morphological gradient image. In (Pai et al. 2012), a framework for nucleus and cytoplasm contour detection was presented. An adaptive thresholding - based on graylevel histogram, was used for initial contour detection. Final nuclei contours were detected using a maximal graylevel gradient difference based technique.

Cell clumps were extracted by applying triangle global search on image superpixels in (Ushizima et al. 2015). Image superpixels were generated using SRM segmentation technique. Nuclei were segmented from image superpixels - using a local thresholding technique presented in (Phansalkar et al. 2011). In (Plissiti et al. 2015), nuclei boundaries were obtained by thresholding the difference between nucleus centroid's intensity and the average intensity of superpixels in the circumference of that centroid. SLIC segmentation was used to generate image superpixels.

In (Guan et al. 2015), nucleus was segmented using a morphological filter based k-means clustering technique. Morphological filter was composed of morphological dilation and erosion with a disk shaped structuring element. This filter was used to remove small dark contamination in the image. Then, the cell image was segmented into nucleus, cytoplasm and background regions, using spatial k-means clustering. In (Tareef et al. 2015), Otsu thresholding was used to detect nuclei from cervical cell images. Before thresholding, images were enhanced using anisotropic diffusion filtering and CLAHE, and filtered by H-maxima transformation. In (Riana et al. 2015), a combination of graylevel thresholding and distance rule definition based segmentation technique was proposed. Graylevel thresholding was used for initial segmentation. Final nuclei boundaries were obtained

using shortest distance analysis.

In (Lee & Kim 2016), first cell mass was segmented using triangle thresholding method from SLIC technique generated superpixels. Then, nuclei candidates were extracted from the cell mass using a local thresholding technique - considering mean and standard deviation of intensities in local window. In (Phoulady, Goldgof, Hall & Mouton 2016, Phoulady et al. 2017), an iterative thresholding based binarization technique was proposed to detect nuclei candidates. Binarization process was performed considering area, intensity and shape features of the segmented regions (Phoulady, Goldgof, Hall & Mouton 2016). In (Phoulady et al. 2017), mean intensity of the segmented region was also considered in addition to the aforementioned features.

2.3.2 Deformable model, contour and shape based approaches

In (Bengtsson et al. 1981), an algorithm was presented to detect overlapping cervical cell nuclei, using information from boundary and density profile of nucleus. Nucleus boundary/contour information was analyzed with smoothed difference chain code to find any significant concavity. The presence of concavity was used detect the nucleus as overlapped nuclei. Otherwise, density profile was analyzed to find any overlap situation. In (Bamford & Lovell 1998), a Viterbi search based dual active contour model was used to segment cervical nuclei. In cell images, image domain/search space was defined by considering the darkest point as centroid. Then, forward dynamic programming was used to find the least cost paths of the search spaces.

In (Plissiti et al. 2006), radial profiles in equal arc length intervals were used to estimate initial nuclei boundaries. The final boundaries were determined by applying active contour based deformable model on the initial estimation. A cytoplasm and nucleus contour detector was proposed in (Tsai et al. 2008). This technique was based on the idea of edge detection and consisted of three approaches. The aim of the first approach/bigroup enhancer was to suppress the noise and emphasize edge pixels. Nucleus contour was detected using maximum color difference (MCD) technique and cytoplasm contour was found using k-means clustering. A geometric active contour based segmentation technique was presented in (Harandi et al. 2010). A rectangle around the image was used as the initial curve for active contour model, and produced individual cells/cell

clumps from the image. After localization, nucleus contour was detected using a thresholding operation on locally histogram equalization process.

A spatially adaptive active shape model was presented in (Plissiti & Nikou 2012*b*), to segment two overlapping nuclei from cervical cell images. Physical shape model and its modal distributions, were used for representing the prior knowledge of the expected shapes. In (Bergmeir et al. 2012), relevant edges in an image was found by utilizing Canny edge detector. Then, cell nuclei were localized using a prior knowledge and voting strategy. Finally, randomized Hough transform for ellipses was used to find candidate nuclei and post-processed using level set technique. A radiating gradient vector flow (RGVF) snake based segmentation technique was proposed in (Li et al. 2012). RGVF snake was modified by introducing a new edge map computation method and stack-based refinement technique. Initial contours were found using spatial k-means clustering, and then modified RGVF snake was used for segmentation.

In (Nosrati & Hamarneh 2015*a*), shape prior information was incorporated in segmentation technique for precise segmentation of overlapping cervical cytology images. A novel continuous variational segmentation framework was proposed, where directional derivatives were used to enforce star shape prior. In (Husham et al. 2016), cervical nuclei were segmented using level set technique. To find the seed points for nuclei, morphological opening and closing operations were applied on binarized image - found using Otsu thresholding. Then, centroid transform was applied to remove unwanted candidates. Nuclei were segmented using level set evolution on the seed points. In (Phoulady, Zhou, Goldgof, Hall & Mouton 2016), a framework was presented to classify cervical tissue as normal or cancerous, using solely nucleus features. Nucleus was quantified using an adaptive nucleus segmentation algorithm - composed of adaptive multi level thresholding and ellipse fitting based shape approximation.

In (Zhang, Kong, Liu, Wang, Chen & Sonka 2017), an image unfolding based nucleus segmentation approach was presented. Image unfolding was done by transforming Cartesian image coordinates to polar coordinates with the center of ROI. Edge direction analysis on the unfolded image and region information from the image, were used to find nuclei contours. An adaptive gradient vector flow (AGVF) snake based framework was proposed in (Dong et al. 2019), for cervical cell image segmentation. Cell was localized in the image using an improved Canny edge detection algorithm. Then, accurate cell edges were generated using adaptive initial contour model and adaptive gradient vector fields. Final cell boundaries (nucleus and cytoplasm) were found using particle swarm

optimization (PSO) based SVM classifier. An implicitly parameterized elliptical shape model and global energy minimization based nuclei segmentation technique was presented in (Kostrykin et al. 2019). The technique utilized both shape and intensity information for segmenting cell nuclei.

2.3.3 Watershed based segmentation approaches

A morphological watershed based clustered nuclei segmentation algorithm was presented in (Malpica et al. 1997). Image background was removed by applying ISODATA thresholding on preprocessed image. Nuclei markers were found using morphological gradient based transformation on background subtracted image. These nuclei markers were used with watershed transform, to segment nuclei.

In (Cheng & Rajapakse 2009), shape marker aided watershed transform was used to segment clustered nuclei. Initial segmentation - to separate nuclei from background, was performed using geometric active contour model. Then, shape markers were found using adaptive H-minima transform. The marking function for the proposed watershed transform, was generated based on Euclidean outer distance transform on initial segmentation. In (Bai et al. 2009), marker based watershed transform was used to find contours of touching cells. Cell markers were found using morphological and thresholding operations. Touching cells were split, using ellipse processing of these contours.

In (Cloppet & Boucher 2010), overlapped and clustered nuclei were segmented using a marker based watershed segmentation. Prior information from the geometric properties and graylevels of nuclei, were employed to generate the markers. In (Jung & Kim 2010), watershed transform with H-minima based marker extraction was used for cell nuclei segmentation. While finding nuclei markers, prior information about shape, size, texture, and concavity were analyzed, along with graylevel criteria. In (Béliz-Osorio et al. 2011), a locally constrained watershed transform was used to segment cervical cells.

In (Gençtav et al. 2012), multi-scale watershed segmentation was applied for segmenting cervical nuclei. Cell clumps were found using automatic thresholding. Then, hierarchical tree (based on scale) was found using multi-scale hierarchical watershed segmentation, which partitions the cell clump into smaller regions. Each node in the tree is a candidate for final segmentation. Tree nodes were then separated as nucleus and cytoplasm from the segmented regions, using a binary classifier considering homogeneity and circularity. In (Moshavegh et al. 2012),

marker-controlled watershed segmentation was used for precise nuclei boundary delineation. Nuclei markers were found with grayscale annular closing operation. In (Tareef et al. 2018), a multi-pass fast watershed based segmentation technique was introduced. Nucleus was segmented with barrier-based watershed transform, using intensity gradient information from pre-processed image.

2.3.4 Region and clustering based approaches

An iterative optimal parametric segmentation algorithm was presented in (Wu et al. 1998). A parametric image was first constructed iteratively by minimizing a cost function/approximation error. Then the segmentation of cell region was found by thresholding the parametric image. In (Isa 2005), a combination of moving k-means clustering and a modified seed based region growing (MSBRG) technique was used, to detect edges in Pap smear images. Moving k-means clustering was used to find a threshold value for detecting the seeds, then MSBRG technique was used to find edges of ROI. In (Kale & Aksoy 2010), initial segmentation was performed with a non-parametric hierarchical region extraction technique, using spectral, shape and gradient information. Then, nucleus and cytoplasm regions were classified using SVM with RBF kernel, to get the final segmentation outcome. In (Lu et al. 2013), connected components were found using maximally stable extremal region (MSER) algorithm. Candidate nuclei were detected from these connected components using some morphological features.

A patch based fuzzy clustering technique was used in (Chankong et al. 2014), to segment cervical cells. FCM clustering was applied on preprocessed image. Then, clustering outputs were labeled as nucleus, cytoplasm, and background considering thresholds for the graylevels of the patches. In (Zhang, Kong, Chin, Liu, Chen, Wang & Chen 2014), adaptive local graph cut method was used for cervical nucleus segmentation. Combination of texture, intensity, boundary, and region information, were used with graph cut technique.

In (Nosrati & Hamarneh 2015*b*), cervical nuclei were detected using a combined approach of MSER algorithm and random decision forest classifier. MSER algorithm was also applied in (Lu et al. 2015) or baseline method in (Lu et al. 2017) on the foreground image to detect nuclei. Region level features (eccentricity, area, mean intensity, and area ratio of detected output and cell clump), were used to find nucleus from MSER outputs. In (Oprisescu et al. 2015), region growing segmentation was applied on the edges to segment nuclei. Edges were

detected using Sobel operator and non-nuclei objects were eliminated based on region size and eccentricity.

An adaptive clustering technique was proposed in (Gautam et al. 2017). First, image superpixel was estimated using mean-shift clustering and SLIC technique. Then, an intensity weighted adaptive thresholding was applied on the superpixels, to segment nuclei from cytoplasm and background. In (Bora et al. 2017), MSER technique was integrated with discrete wavelet transform (DWT) and morphological operations, to segment nuclei. Input image was preprocessed with DWT and median filters. Then, MSER was applied on the preprocessed image to find the connected components/regions. Finally, morphological operation was applied, and area, perimeter and circularity based filtering was performed to eliminate unwanted regions. In (Roy et al. 2020), MSER algorithm was used to extract nuclei candidates from cervical cell mass found with a Chan-Vese model based multiphase level set algorithm. A circularity check based post processing was applied to remove non-nuclei candidates.

2.3.5 Machine learning based segmentation

In (Jung et al. 2010), unsupervised Bayesian classifier based segmentation technique was proposed to separate overlapped nuclei. Distance transform was used to generate a topographic surface - viewed as a mixture of Gaussian. Parametric expectation-maximization (EM) algorithm was employed to learn Gaussian mixture model (GMM). Unsupervised Bayesian classifier was used to investigate the clusters, and the number of overlapped nuclei was found by cluster validation. Geometric property and nuclei shape based prior information was used to improve the segmentation performance. Multi-scale convolutional network was used in (Song et al. 2015), for coarse segmentation of cervical cell images. Then, a superpixel-wise graph partitioning was used on top of coarse segmentation, to achieve finer segmentation results.

In (Zhang, Sonka, Lu, Summers & Yao 2017), fully convolutional network (FCN) was used generate nucleus label mask and probabilistic map. Then graph based approach was exploited using the mask, for finer segmentation of nucleus. Dynamic programming was used to find the optimal path in the graph. Finally, nucleus contour was found by mapping the optimal path in Cartesian coordinate system. In (Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017), shape, texture and boundary features were extracted from SLIC superpixels to form the feature vector. This feature vector was used with SVM classifier to

classify image superpixels as nuclei, cytoplasm and background. In (Tareef, Song, Huang, Wang, Feng, Chen & Cai 2017), convolutional neural network (CNN) based feature learning and classification was used on superpixel image, for segmenting nucleus.

In (Liu et al. 2018), a cervical nuclei segmentation framework was proposed based on mask regional convolutional neural network (mask-RCNN). Coarse segmentation of nuclei was obtained with forward propagation of mask-RCNN. Then, a local fully connected conditional random field (LFCCRF) was used to get finer segmentation results. A deep learning based algorithm was presented in (Jith et al. 2018), to classify cervical cell images. The network consisted of: initial 3 layers of AlexNet, batch normalization layer, and a fully connected layer. This framework directly worked with RGB image patches from single cell images.

In (Zhang et al. 2019), a binary tree like deep CNN was presented to work for small datasets. The proposed network aggregated the features captured by the basic network, and two path fusion attention was used to weigh and indicate the importance of the features. In (Lin et al. 2019), a CNN based method was presented, which was guided by cell image appearance and cell morphology. Appearance and morphological information were extracted from the image patches/cell masks, centered at nuclei centroid. In (Araújo et al. 2019), a deep learning based abnormal cell segmentation framework was proposed. In the framework, a intensity variation based preprocessing step was employed to discard poorly sampled Pap smear images. Then, a patch based CNN or CNN model with sliding window, was used for segmentation.

Chapter 3

Shape prior in fuzzy c-means clustering

The main focus of this thesis is to incorporate prior or guidance in segmentation methods, in order to achieve more precise segmentation outcome. In this chapter, a novel circular shape function is proposed and incorporated in fuzzy c-means (FCM) clustering, with the aim to segment nuclei more precisely from overlapping cervical smear images. Section 3.1 provides a brief overview of FCM clustering and its spatial variant techniques in literature, used as a tool for image segmentation. The standard FCM clustering, the novel circular shape function and how this function is incorporated in fuzzy clustering, are described in Section 3.2. Section 3.3 provides a description of the dataset used for conducting experiments in this chapter. Section 3.4 presents the step-wise implementation details of the proposed framework. Evaluation metrics, experimental set-up, selection of background subtraction method, and tuning of required parameters, are presented in Section 3.5, Section 3.6, Section 3.7, and Section 3.8, respectively. Section 3.9 provides results for the proposed framework and direct comparison with recent state-of-the-art cervical nucleus segmentation techniques. A formal discussion about the framework and conclusion of this chapter is provided in Section 3.10 and Section 3.11, respectively.

3.1 Introduction

Among the fuzzy clustering techniques, Fuzzy c-means (FCM) is the commonly used clustering for image segmentation, medical image analysis, and pattern recognition applications (Noordam et al. 2000, Ahmed et al. 2002, Chuang et al.

2006, Cai et al. 2007, Wang et al. 2008, Wang, Song, Soh & Sim 2013, Adhikari et al. 2015, Verma et al. 2016, Haddad et al. 2018). However, intensity information is the only criteria to consider in standard FCM clustering, thus it lacks the spatial information of the data. Segmentation outcome from standard FCM clustering is influenced by intensity inhomogeneity, noise, and other imaging artifacts.

For image segmentation, many researchers attempted to consider spatial information in FCM clustering. Geometrical information was included with FCM clustering in (Noordam et al. 2000), where geometrical guidance was determined from the local neighborhood of each pixel. FCM objective function was modified in (Ahmed et al. 2002), so that the clustering of a pixel is influenced by its immediate neighborhood pixels' clusters. In (Leung et al. 2004), an elliptical shape function was included in FCM clustering to precisely segment lip from human face images. The dissimilarity criteria, which measures graylevel similarity and spatial distance, was modified to incorporate the elliptic shape function. In (Chuang et al. 2006), a spatial function was incorporated into the fuzzy partition/membership matrix, in order to segment MRI images of brain. In (Cai et al. 2007), a new similarity criteria was proposed, which considered both graylevel and spatial knowledge from local neighborhood. A linearly-weighted summed image was computed from this measure, and used for the clustering process. In (Wang et al. 2008), a new dissimilarity measure was introduced that considered both local and non-local information. A weight was used to balance the contribution of information from local and non-local parts. FCM objective function was modified in (Wang, Song, Soh & Sim 2013), to include an adaptive spatial weighting factor. In (Adhikari et al. 2015), pixel membership values were weighted and joint cluster was proposed. Global membership value in standard FCM and local square neighborhood spatial membership values, were combined to compute the weighted membership. An improved intuitionistic fuzzy c-means clustering was proposed in (Verma et al. 2016), where local graylevel and spatial information were incorporated in the objective function through a novel intuitionistic fuzzy factor.

Spatial information from neighborhood was considered in most of the published research, and included in the definition of similarity measure, membership degree, or objective function. Neighboring pixels always has the tendency to influence the centroid, regardless of its location (in a homogenous region or in an edge). Therefore, region's boundary or edges may smooth out, when neighboring pixels spatial information is considered. This type of spatial influence also may

result in poor segmented boundaries (Wang et al. 2008). Some other limitations of these methods are choosing a constant value for the size of neighborhood and parameters to control the neighborhood effect.

FCM clustering and it's spatial variants from literature (Ahmed et al. 2002, Chuang et al. 2006, Cai et al. 2007, Wang et al. 2008, Wang, Song, Soh & Sim 2013, Adhikari et al. 2015, Verma et al. 2016), mostly considered the segmentation task of brain MRI images. Microscopic cell image (nucleus/cytoplasm) segmentation has not been well-explored in literature. For a reliable cell image segmentation, spatial information of each pixel also requires attention along with graylevel intensity. Cell overlapping in cytology images results in poor contrast and intensity inhomogeneity. Hence, standard FCM clustering or its existing spatial variants considering local neighborhood information, may smooth out nuclei boundaries or fail to detect nucleus correctly, in the presence of similar mean intensities in nuclei and cytoplasm parts.

Shape information helped to separate overlapping or touching objects, while segmenting images in (Cheng & Rajapakse 2009, Bai et al. 2009, Cloppet & Boucher 2010, Molnar et al. 2015). Spatial neighborhood information has been used in existing literature (Chuang et al. 2006, Cai et al. 2007, Li et al. 2011, Adhikari et al. 2015) for precise segmentation using FCM clustering, however the partitioning process was never been explored to be guided by spatial shape information. In this chapter, a novel spatial shape constrained fuzzy clustering based segmentation framework is proposed and applied to segment nuclei from overlapping Pap smear images.

3.2 Circular shape constrained fuzzy clustering (CiscFC)

3.2.1 An overview

Standard FCM clustering (Dunn 1973, Bezdek 1981) associates a data point with multiple clusters using varying degree of membership in the range [0,1]. A set of data points is partitioned through iterations of a clustering process, by minimizing an objective function J, with respect to the values in the membership or fuzzy

partition matrix U, and the set of cluster centers $c = \{c_j\}$ as:

$$J = \sum_{i=1}^{D} \sum_{j=1}^{C} U_{ij}^{m} \left\| x_i - c_j \right\|^2 . \tag{3.1}$$

Here, C is the number of clusters, x_i is the i^{th} data point in a D dimensional data-space, $\|.\|$ is the Euclidean distance between a data point and a cluster center, and m is the fuzziness weighting exponent. An overview of the standard FCM clustering algorithm is also available in Section 2.1.2.3.

The standard FCM algorithm is optimized by assigning high membership values to the data points close to their cluster center and low membership values to the distant or dissimilar data points. Convergence is achieved by comparing the objective function at successive iterations. FCM algorithm can be used for image segmentation, by assigning image pixels to the clusters with the highest membership values.

In this chapter, spatial information is incorporated into FCM clustering by defining a circular shape function (CSF). CSF allows the influence from a pixel's spatial location during the clustering of that pixel. This information helps the clustering process for precise segmentation of nuclei boundaries from overlapping cervical smear images. The shape function is designed to produce small values for the pixels associated with the target cluster and large value is produced for the pixels that should be member of other clusters. The cluster with the highest average intensity is identified as the target cluster from initial fuzzy partition matrix (see Section 3.4.2). Image pixels and target cluster seed points are considered to define the shape function. Target cluster graph is constructed using intensity differences and spatial distances, and explored to find the connected components. The maximum intensity nodes from each connected component are selected as seed points.

3.2.2 The proposed circular shape function (CSF)

Let $X = [x_1, x_2, \ldots, x_D]$ be an image with size $D = M \times N$ and m_i, n_i be the spatial coordinates of i^{th} pixel x_i with intensity value f_i. Considering i^{th} pixel, the circular shape function $CSF : D \rightarrow \mathbb{R}^+$ is defined as:

$$CSF(x_i) = \min_{p \in P} \sqrt{(m_i - m_p)^2 + (n_i - n_p)^2}, \tag{3.2}$$

where (m_p, n_p) are the spatial locations of seed point p and P is the total number of seed points. For simplicity, $CSF(x_i)$ will be denoted as CSF_i throughout this chapter.

Once the CSF's value is determined, an intensity modifying function δ is used to enforce the shape constrain, and defined as

$$\delta_i = \begin{cases} 1, & \text{if } CSF_i < t_1, \\ 0.8, & \text{if } t_1 \leq CSF_i \leq t_2, \\ 0.7, & \text{if } CSF_i > t_2, \end{cases} \tag{3.3}$$

for some thresholds t_1 and t_2. Intensity value f_i for i^{th} pixel is then modified to f_i' using δ_i as:

$$f_i' = f_i * \delta_i. \tag{3.4}$$

The thresholds t_1 and t_2 control the range of intensity modification. This intensity modification process transforms the image X to X'. New cluster centers, fuzzy partition matrix and objective function, are calculated in each iteration of clustering, using the image X' as:

$$c_j' = \frac{\sum_{i=1}^{D} U_{ij}' x_i'}{\sum_{i=1}^{D} U_{ij}'}, \tag{3.5}$$

$$U_{ij}' = \left(\sum_{k=1}^{C} \left(\frac{d_{ij}}{d_{jk}} \right)^{\frac{2}{(m-1)}} \right)^{-1}, \tag{3.6}$$

$$J' = \sum_{i=1}^{D} \sum_{j=1}^{C} U_{ij}' d_{ij}^2, \tag{3.7}$$

where, d_{ij} is the Euclidean distance between i^{th} pixel and j^{th} cluster center.

The value of CSF and the intensity modifier δ together enforce circularity of the spatial shape of the target cluster. Since, intensity values in the target cluster are modified, it affects further iterations of the clustering process and eventually influence the target cluster formation. The proposed technique is benefited from the inclusion of spatial information in CSF, and able to identify the pixels with similar intensities using spatial distance. Thus, the proposed algorithm tries to

alleviate the drawback of standard FCM algorithm by introducing the circular shape constraint, and helps to find nuclei more accurately from Pap smear images.

3.3 Dataset

Training and evaluation of the proposed framework is carried on the dataset from Overlapping Cervical Cytology Image Segmentation Challenge - ISBI 2014 (Lu et al. 2015, Lu et al. 2017). This dataset is located at `https://cs.adelaide.edu.au/~carneiro/isbi14_challenge/dataset.html` and available for download. The challenge dataset contains 945 synthetic Pap smear images with varying number of cells from 2 to 10 per image and various degrees of overlap. The challenge dataset was released over two phases. In the first phase, 45 training and 90 test synthetic cytology images were released on January 11, 2014. These 45 training images from the first phase of the challenge are used as the training dataset to determine parameters such as: the number of clusters and threshold values for area and shape features. The second phase of the challenge released 810 test images on February 5, 2014. These 900 test images (90 from the first stage and 810 from the second stage) are used as the test dataset to evaluate the proposed algorithm. All images in the dataset are grayscale images of 512 $\times$ 512 pixels. The training set has 270 nuclei and the test set contains 5400 nuclei. All images in this dataset are accompanied with the ground truths - delineated by expert cytotechnologists. A brief overview of the data preparation process is described below for completeness, however detailed description about the ISBI 2014 challenge dataset can be found in (Lu et al. 2015, Lu et al. 2017).

All the cells to generate the synthetic images were extracted from 16 non-overlapping fields of view (FOV) images. These FOV images were obtained from 4 cervical cytology specimens. Each FOV contains 20 to 60 Pap stained cervical cytology cells, with varying degrees of overlap. AutoCyte PREP technology (Bishop J et al. 1998, Howell et al. 1998) was used to prepare the specimens, hence each specimen was around 20 μm thick in the focal dimension. An extended depth of field (EDF) image was generated from at least 20 focal plane images (acquired with a focal depth separation of 1 μm) for each FOV image. These 16 EDF cervical cytology smear images were used to generate all synthetic images in this dataset. Training dataset was built using 4 EDF images, and the rest 12 EDF images were used to produce the test dataset. Randomly chosen annotated isolated cells and background from EDF images were used to

generate the synthetic images. First, background was formed using annotated background pixels randomly chosen from the EDF images. Then, random rigid transformation (rotation, translation and scaling), and random linear brightness transform were applied on the selected isolated cells. These cells were placed on the background image using an alpha channel. In case of additional isolated cells, same transformation was applied and the cells were placed on the image with an overlap with the existing cells, with the strict criteria that the overlap coefficient must lie in any of the following ranges as: [0, 0.1], [0.1, 0.2], [0.2, 0.3], [0.3, 0.4], [0.4, 0.5]. The process was repeated until the desired number of cells (2 to 10 cells) per image was reached. If A and B are two cells, the overlap coefficient was calculated as: $max\left(\frac{|A\cap B|}{|A|}, \frac{|A\cap B|}{|B|}\right)$, where $|.|$ represents the area of the cell. Figure 3.1 shows four examples of Pap smear images from ISBI 2014 dataset, with corresponding nuclei ground truths.

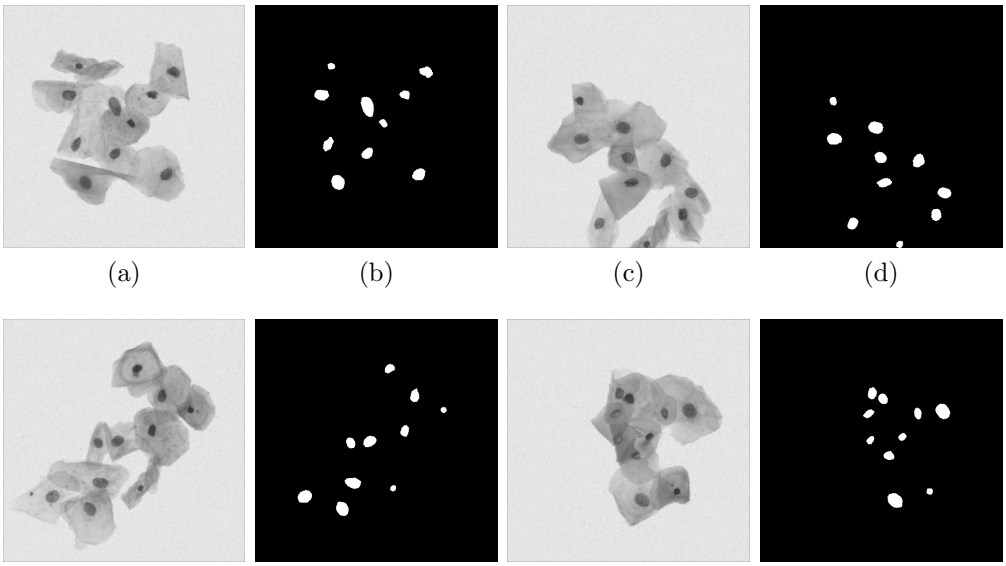


Figure 3.1: Four examples of overlapping synthetic cervical smear images from ISBI 2014 challenge dataset (in light background) (a), (c), (e), and (g), followed by corresponding nuclei ground truths (in dark background) (b), (d), (f), and (h).

3.4 Implementation

The proposed nucleus segmentation framework is composed of three main steps, namely background subtraction, proposed clustering and false positive reduction.

Figure 3.2 illustrates the flowchart of the proposed segmentation framework to segment cervical nuclei. In the background subtraction step, the foreground image

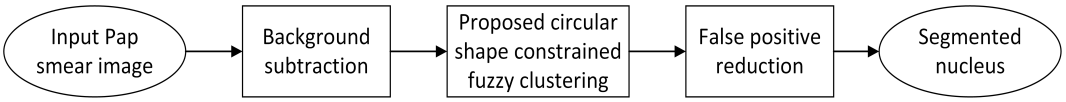


Figure 3.2: The flowchart of the proposed cervical nuclei segmentation framework.

is obtained by removing the background from Pap image. The proposed clustering is then performed using the complemented foreground image and initial fuzzy partition matrix. The highest average intensity cluster is subsequently identified as the nuclei cluster from the clustering output. Finally, in the false positive reduction step, outliers are removed and final nuclei candidates are obtained. The framework steps are presented in detail in this section.

3.4.1 Background subtraction

The goal of this step is to divide the input image into background and foreground, where the cell clump or foreground corresponds to cervical cells, and the background represents the area except the cells. Since the goal is to segment nucleus, the background can be eliminated before the image segmentation process starts. This background elimination process helps in reducing the search space for processing, and the estimated cell clump can be used in further stages. Two techniques based on thresholding and morphological reconstruction, are presented in this section for background subtraction.

3.4.1.1 Thresholding based background subtraction

A threshold based image binarization process is proposed to remove background from input images. This background subtraction process was applied on the complemented image. However, defining a global threshold to binarize an image into foreground and background is challenging. Hence, dominant peaks in complemented image histogram were analyzed to determine a threshold adaptively for binarization. Dominant peaks were found using the technique proposed in (Silva et al. 2010) and ranked in ascending order based on dominant peaks intensity value. The value of second highest peak was chosen as the threshold to binarize the input image into foreground and background. Fig. 3.3 illustrates the process.

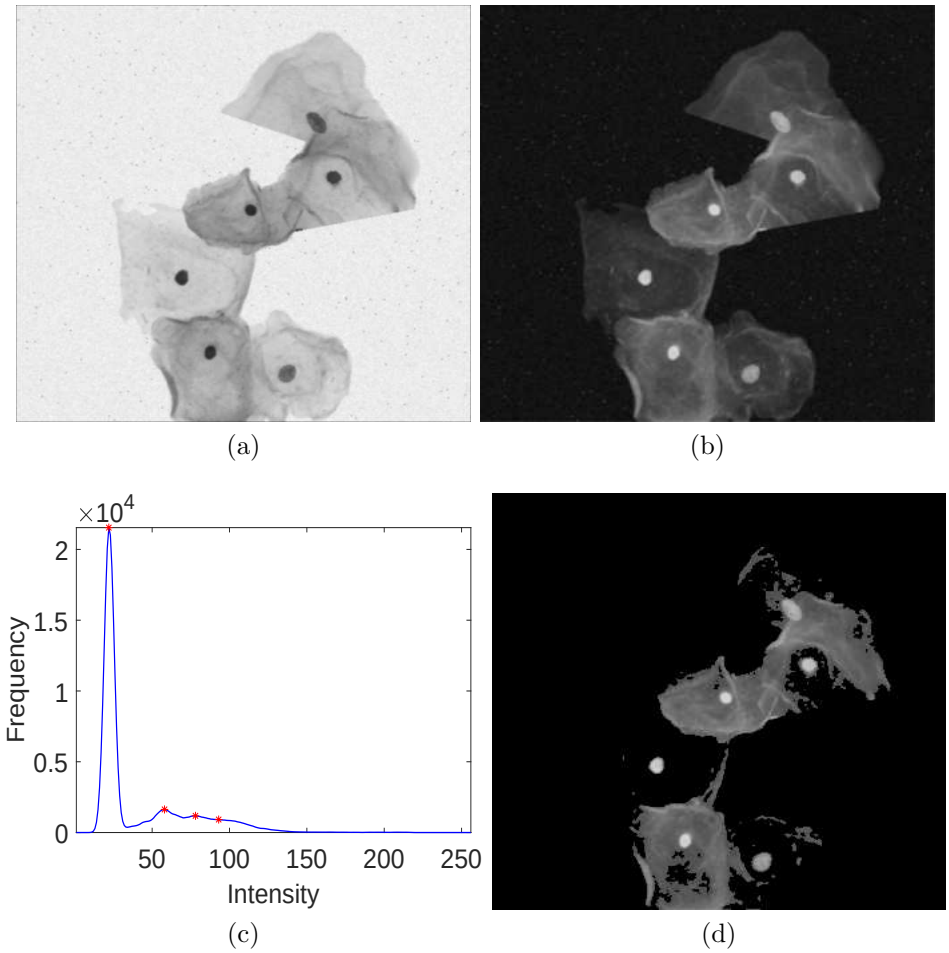
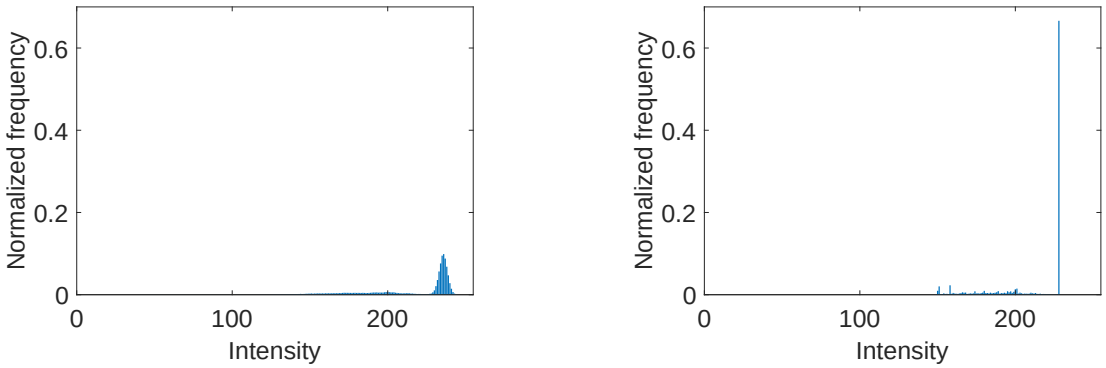


Figure 3.3: Cervical cell image background subtraction using thresholding based technique, (a) original image, (b) complemented image, (c) plot of the complemented image histogram with dominant peaks (marked with red $*$), and, (d) binarized image using the second highest peak.

3.4.1.2 Morphological reconstruction based background subtraction

In original cervical smear images, background visually appears brighter than nuclei and cytoplasm parts. The histogram shown in Figure 3.4a, where background is represented as the high peaks located in the right-most side of the plot, also supports this. If the regional maxima can be used to represent the background, then it can be easily extracted from the image. Morphological reconstruction is effective for extracting regional minima or maxima from grayscale images (Vincent 1993), without changing objects size or shape. Morphological reconstruction operation was applied to increase inter-class dissimilarity and intra-class similarity. Then foreground image was found by removing regional maxima of the morphologically processed image.



(a) (b)

Figure 3.4: Normalized histograms of: (a) an original image from the training dataset, (b) same image after applying morphological opening-by-reconstruction and closing-by-reconstruction.

This morphological reconstruction operation was applied on original images. First, morphological opening-by-reconstruction was utilized to remove bright objects smaller than a structuring element. For an input image X_{in}, opened by reconstruction image X_{obr} was generated by eroding the image with a structuring element S and using the eroded image as the marker image for image reconstruction R_x as follows:

$$X_{obr} = R_x(X_{in} \ominus S). \tag{3.8}$$

Closing by reconstruction was then used to remove dark objects smaller than a structuring element. This was performed by complementing the image, computing its opening by reconstruction and again complementing the result. This can be presented as:

$$\begin{aligned} X_{obrc} &= (X_{obr})^c, \\ X_{cbr} &= R_x(X_{obrc} \ominus S), \\ X_{out} &= (X_{cbr})^c, \end{aligned} \tag{3.9}$$

where $(.)^c$ is the complement operation. Inter-class variance is increased and intra-class variance is reduced, using these morphological operations. This can be observed clearly from the histogram of X_{out} in Figure 3.4b. Regional maxima of the processed image represents the background, and foreground or cell clump can be found by removing it from the original image (shown in Figure 3.5). After removal of the background, foreground image was finally complemented (Figure 3.5d), to use in further stages of the proposed framework.

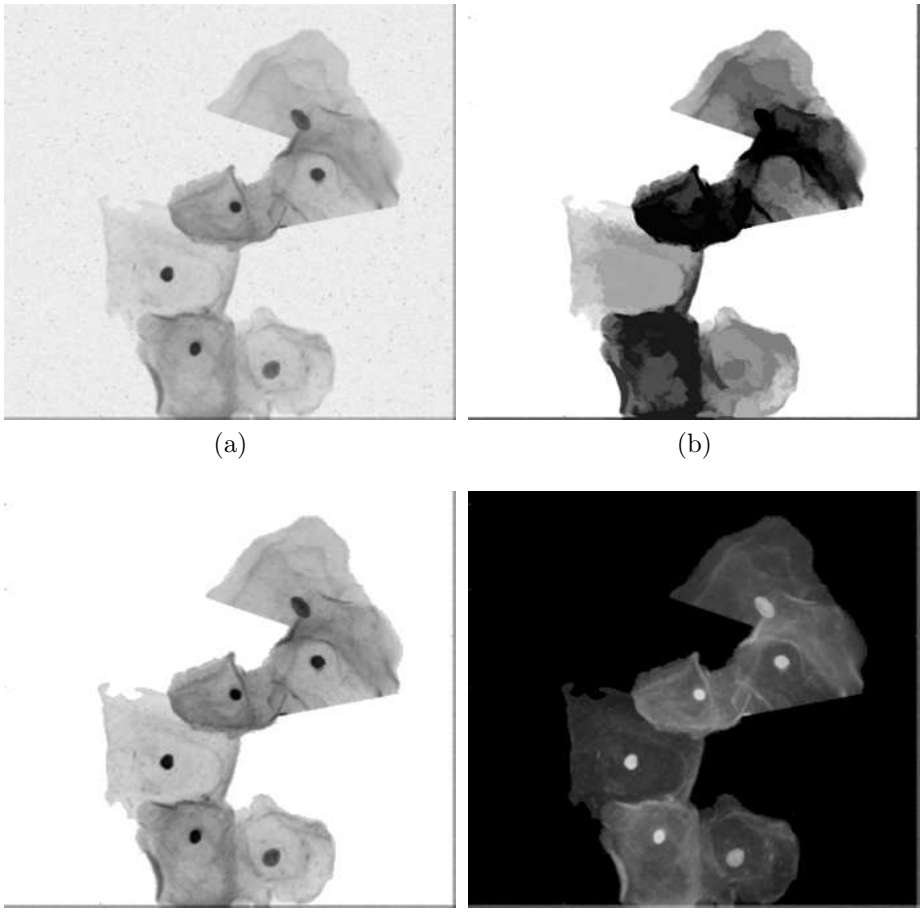


Figure 3.5: Background extraction and finding cell clump, (a) original image, (b) morphologically processed image, (c) foreground cell clump, and (d) complement image of the cell clump.

3.4.2 Implementation of CiscFC

An initial fuzzy partition matrix and complemented foreground image, are required for the proposed circular shape constrained fuzzy clustering (CiscFC). Then the clustering process can be started. The initial fuzzy partition matrix generation process and clustering details of the CiscFC technique, are elaborated in this section.

In the proposed clustering technique, circular shape constrain is imposed on the target cluster containing nucleus candidates. Generally, FCM clustering starts with randomly generated initial fuzzy partition matrix. The partition matrix is stabilized by iterating and minimizing the objective function. In this proposed framework, fuzzy partition matrix needs to be initialized in such a way that target cluster can be identified from the partition matrix and circular shape prior can be

implemented in the target cluster. Therefore, random initial fuzzy partition is not sufficient to start with. In literature, standard FCM (Bezdek 1981) clustering is often chosen to produce desirable initial fuzzy partition matrix for other variants of FCM (Chuang et al. 2006, Cong et al. 2015). It was found empirically that fuzzy partition matrix starts to stabilize after five iterations of standard FCM technique. Hence, the initial fuzzy partition matrix was generated by running standard FCM algorithm for five iterations.

The target cluster was found from the initial partition matrix as the highest mean intensity cluster. The circular shape function finds spatial distances between image pixels and seed points from the target cluster (see Equation (3.2)). Seed points were found using connected component analysis on the target cluster. For this, first target cluster was represented as graph, where target cluster pixels were considered as graph nodes. Graph adjacency was defined with the aid of intensity difference and spatial distances. Graph nodes with intensity difference less than or equal to 8, and maximum Euclidean distance 40, were defined as adjacent. All these thresholds were found with empirical analysis. Strongly connected components were found automatically from the graph using the technique proposed in (Tarjan 1971) (see Section 3.6 for implementation details). Then, member nodes of each connected components were ordered based on intensity values. Finally, top 4% highest intensity member nodes in the connected components were selected as seed points. These seeds were used to enforce circularity in the target cluster using the proposed CiscFC technique.

The proposed clustering (in Section 3.2.2) iterates until convergence. Finally, image segmentation was obtained by assigning image pixels to the cluster with the highest membership values. From all clusters, the highest mean intensity cluster was considered as the segmented nucleus and used for evaluation.

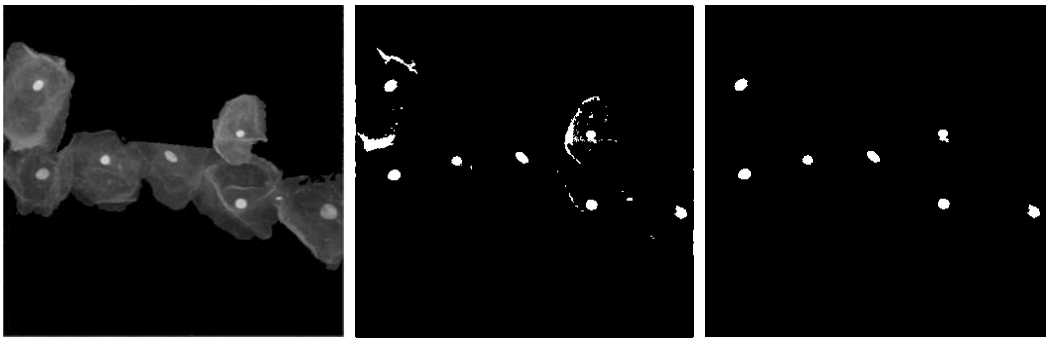
3.4.3 False positive reduction

In this study, nuclei were detected from clustering output using average intensity, area and shape features. From the clustering result, the highest mean intensity cluster was detected as the cluster containing nuclei. In overlapping cervical smear images, some parts of cytoplasm may have similar mean intensity as nucleus and appear in the nucleus cluster. To overcome this issue, further filtering was performed based on area and shape to remove non-nucleus candidates/unwanted regions.

Very small regions (ineligible to be nucleus) were removed from the set of candidates. Then, other non-nucleus candidates were eliminated using shape features as: eccentricity and circularity. Eccentricity was calculated as the ratio of the distance between the foci and major axis length of an ellipse - that has the same second-moment as the region. The eccentricity value of a region ranges in $[0, 1]$. Circularity of a region is the measure of roundness. It was calculated using the area (total number of pixels in that region) and perimeter of the region as:

$$circularity = 4\pi \left(\frac{Area}{Perimeter^2} \right). \tag{3.10}$$

Details of selecting the threshold values for the area, circularity and eccentricity are described in Section 3.8 and these values are reported in Table 3.1. Figure 3.6 shows an example of false positive reduction from the nucleus cluster found after applying proposed CiscFC method.

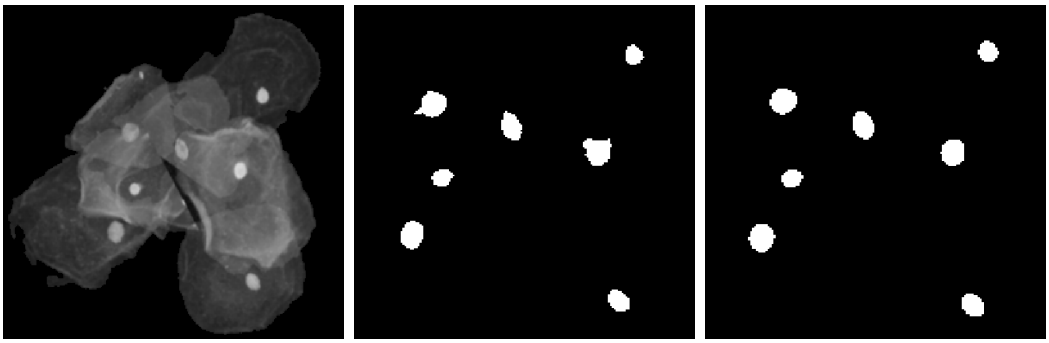


(a) (b) (c)

Figure 3.6: An example of false positive reduction. (a) Complemented foreground image used for clustering (nuclei with high graylevel values), (b) detected nucleus candidates from clustering output, and (c) nucleus cluster after false positive reduction process using area and shape features.

3.4.4 Nucleus shape regularization

Distance regularized level set technique (Li et al. 2010) was applied to improve nucleus boundaries. In this step, nuclei candidates found after false positive reduction stage were used as the zero level set or initial contours. The level set contours are evolved to minimize an energy functional with a distance regularization term, and an external energy term to move the zero level set toward target position. Figure 3.7 provides an example to illustrate the shape regularization process.



(a) (b) (c)

Figure 3.7: An example of nucleus shape regularization. (a) Complemented foreground image, (b) nucleus candidates after false positive reduction, and (c) nucleus candidates after shape regularization.

3.5 Evaluation metrics

Quantitative performance of the proposed framework - in terms of detection and segmentation accuracies, was evaluated using some common measures as: mean Dice Similarity Coefficient (DSC), precision and recall. A nucleus region A was considered correctly detected if:

$$\frac{A \cap B}{A} > \tau \quad \text{and} \quad \frac{A \cap B}{B} > \tau, \tag{3.11}$$

where B is the ground truth nucleus region and τ is the overlap threshold. Following (Gençtav et al. 2012), the value for τ was set to 0.6.

Nucleus detection performance was assessed by calculating object based precision (PR_{obj}) and recall (RC_{obj}) as:

$$PR_{obj} = \frac{\text{No. of correct detections}}{\text{No. of all detections}}, \tag{3.12}$$

$$RC_{obj} = \frac{\text{No. of correct detections}}{\text{No. of all objects in ground truth}}. \tag{3.13}$$

Segmentation accuracy of the proposed framework was assessed using pixel level precision (PR_{pix}) and recall (RC_{pix}), and Dice Similarity Coefficient (DSC) for all correctly detected nucleus regions. These measures were calculated for segmentation output of correctly detected nucleus and the corresponding ground

truth as:

$$PR_{pix} = \frac{\text{No. of correctly detected pixels}}{\text{No. of all detected pixels}} = \frac{TP}{TP + FP}, \tag{3.14}$$

$$RC_{pix} = \frac{\text{No. of correctly detected pixels}}{\text{No. of all pixels in ground truth}} = \frac{TP}{TP + FN}, \tag{3.15}$$

where TP is the number of pixels correctly segmented as nucleus, FP is the number of pixels falsely segmented as nucleus, and FN is the number of pixels falsely detected as background. For two regions A and B, DSC (Dice 1945) is used to measure the similarity between two regions as:

$$DSC(A, B) = \frac{2 \mid A \cap B \mid}{|A| + |B|}, \tag{3.16}$$

resulting in $0 \leq DSC \leq 1$. If the regions A and B have no overlap, DSC value will be 0 and if they are identical, the value of DSC will be 1. This similarity measure is also known as the Zijdenbos similarity index (ZSI) in literature (Gençtav et al. 2012, Chankong et al. 2014, Li et al. 2012, Guan et al. 2015). Mean DSC was calculated across the whole dataset to measure the overall performance of the segmentation technique. The qualitative performance of the proposed approach was assessed by visual inspection of the segmented nuclei boundaries.

An analysis is carried in Section 3.10 to determine the impact of the proposed CSF (detailed in Section 3.2.2) on circularity measure of segmented nuclei. The circularity measure ψ is defined using the region's overlapped area with its equivalent circle, and the area of that region (Giger et al. 1988). A detailed description about this measure is given in Section 4.2. For any segmented region A, the value of ψ lies in the range $[0, 1]$, where higher value indicates that A is more circular.

3.6 Experimental setup

The training dataset of 45 images was used to select background removal technique and determine the parameters required for the proposed framework. Nucleus detection and segmentation performance of the proposed technique was assessed on the test dataset of 900 images. See Section 3.3 for details. The CiscFC algorithm and the proposed framework was implemented in Matlab R2015b and run on Dell R720 Server with Intel Xeon(R) Processor E5-2470 2.30GHz and

32GB RAM. In case of standard FCM clustering, *fcm()* function in Fuzzy logic toolbox of Matlab 2015b package was used. Matlab function *graphconncomp()* was used to generate strongly connected components from target cluster graph and used during CiscFC implementation (see Section 3.4.2). The proposed framework with CiscFC method required on average 3.2 seconds per cell and 13 seconds per image. The Matlab code was not speed optimized.

3.7 Background subtraction technique selection

In the proposed framework, clustering was applied using a fixed number of clusters on background subtracted complemented image. From all the clusters, the brightest cluster was detected as the target cluster containing nucleus. The threshold-based technique (proposed in Section 3.4.1.1) removed most cytoplasm parts along with background, in cases with low contrast between background and cytoplasm (see Figure 3.8a). This affected the final clustering, and in some cases nucleus (or their parts) were assigned to more than one clusters (Figure 3.8b depicts this). Finally, those nucleus fragments were either removed from the target cluster during false positive reduction process (detailed in Section 3.4.3) for not meeting the criteria to be a nucleus candidate (see the example in the top row of Figure 3.8), or resulted in lower segmentation accuracies. Morphological reconstruction based technique (proposed in Section 3.4.1.2) could manage to remove mostly background parts, while leaving nucleus and cytoplasm parts as foreground (refer to Figure 3.8d). This helped the clustering process to accommodate sufficient nucleus pixels in the target cluster, so that candidate nuclei can pass the false positive reduction process (see the bottom row of Figure 3.8).

Two examples from the training set are presented in Figure 3.9 for better visualization of both background subtraction techniques, along with detected final nucleus clusters. From empirical and visual analysis of the training set results for background subtraction, the technique based on morphological reconstruction was found to outperform the thresholding based one. Therefore, morphological reconstruction based background removal technique was used for the proposed framework in this chapter.

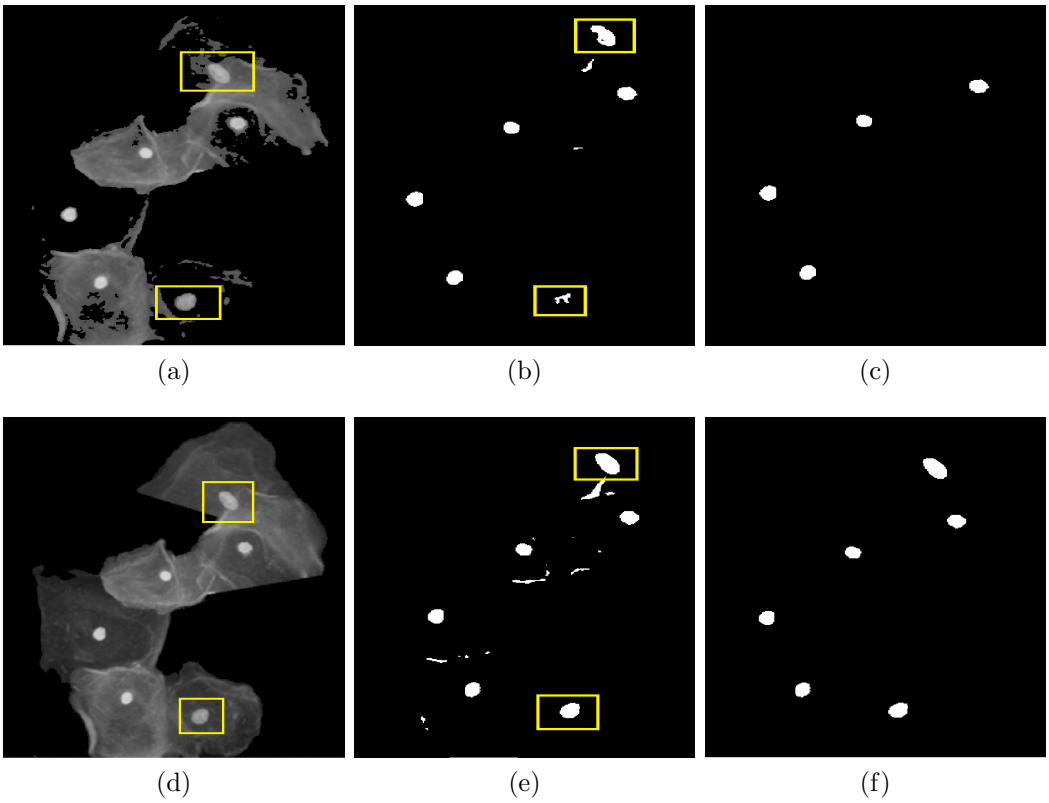


Figure 3.8: Examples of two background removal techniques: top row represents results for thresholding based and bottom row presents morphological reconstruction based, (a), (d) complemented foreground images, (b), (e) brightest clusters as nucleus, and (c), (f) nucleus clusters after false positive reduction. Differences are highlighted with yellow rectangles.

3.8 Parameter tuning

There are some parameters and threshold values required for the proposed framework. Threshold values need to be determined: for area and shape features (Section 3.4.3), to control the range of intensity modification using circular shape shape function (Equation (3.3) in Section 3.2.2), and optimal number of clusters for partitioning the input image. These values were found empirically using the images from the training dataset.

3.8.1 Thresholds for area and shape features

As mentioned in Section 3.4.3, threshold values for area, circularity and eccentricity were found using the training images. First, very small regions - whose area was less than 50 pixels, were discarded from the set of nucleus candidates. In case of eccentricity, nucleus candidates with eccentricity greater than 0.9 were

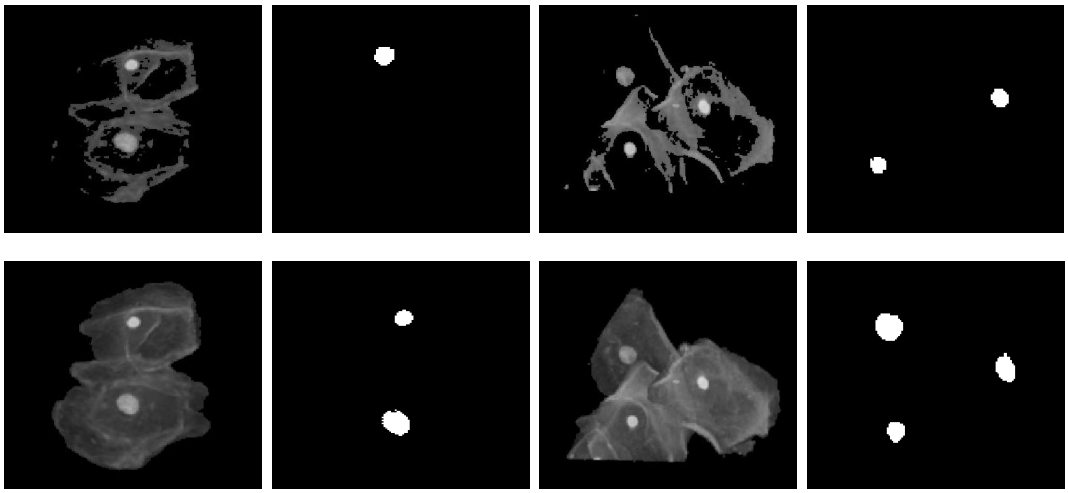


Figure 3.9: Comparison of background subtraction techniques with two example images. Top row presents image foregrounds found with threshold based method and final nucleus clusters, and bottom row shows foregrounds found using morphological reconstruction based technique and corresponding nuclei clusters.

filtered out. For circularity, the suitable range was found empirically as $[0.5, 1.7]$. Nuclei candidates having circularity value outside the range, were not considered and discarded from the set of candidates. Table 3.1 summarizes the findings.

Table 3.1: Parameter values for area and shape features of a candidate nucleus.

Parameter	**Value**
Area	Minimum 50
Circularity	Minimum 0.5 and Maximum 1.7
Eccentricity	Maximum 0.9

3.8.2 Thresholds for CSF

To select a value of intensity modifier δ in Equation (3.3) and modify intensity values using Equation (3.4), two thresholds t_1 and t_2 need to be determined adaptively. Determination of this value is important, as it is related to intensity modification, therefore clustering outcome. Hence, these thresholds have a direct impact on the segmentation accuracy. Two approaches are proposed in this section for finding these thresholds.

3.8.2.1 Otsu multiple thresholding based approach

Initially, the thresholds for implementing CSF were determined adaptively from each complemented foreground image using Otsu multiple thresholding (Otsu 1979). Three classes: nucleus, cytoplasm and background, were generated using multiple thresholds found using Otsu technique. Then the area, circularity, and eccentricity measures were used to discard some false candidates from the target class. Finally, the minimum and maximum radius of candidates in the filtered class, were chosen as the values for t_1 and t_2. Figure 3.10 provides two examples of finding candidates to determine the threshold values using Otsu multiple thresholding.

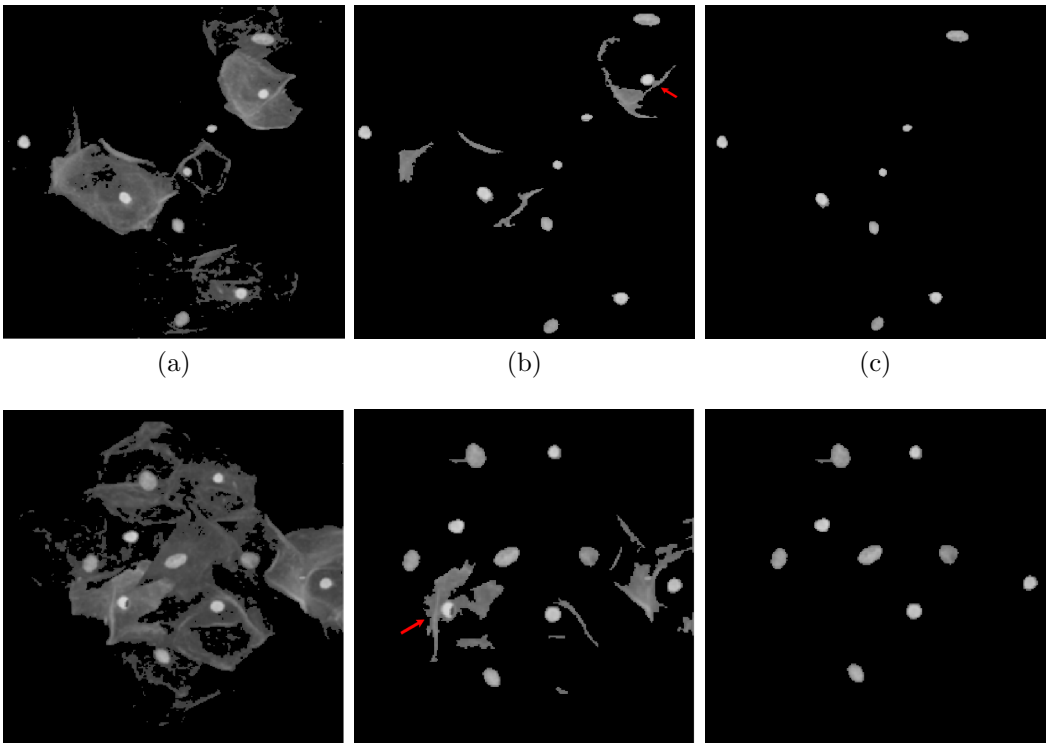


Figure 3.10: Determination of t_1 and t_2 using Otsu multiple thresholding. (a), (d) Complemented foreground images, (b), (e) candidates in nucleus class after thresholding, and (c), (f) filtered candidates.

3.8.2.2 Maximally stable extremal region (MSER) based approach

MSER algorithm (Matas et al. 2004) finds maximally stable connected components from an image. The technique was applied on complemented foreground image to determine the threshold values for t_1 and t_2. Candidates found by the

MSER algorithm were further filtered using area and shape features. The thresholds t_1 and t_2 were selected as the minimum and maximum radius of filtered candidates found by MSER algorithm. For ease of illustration, MSER technique was applied on the same example images from Figure 3.10 and shown in Figure 3.11.

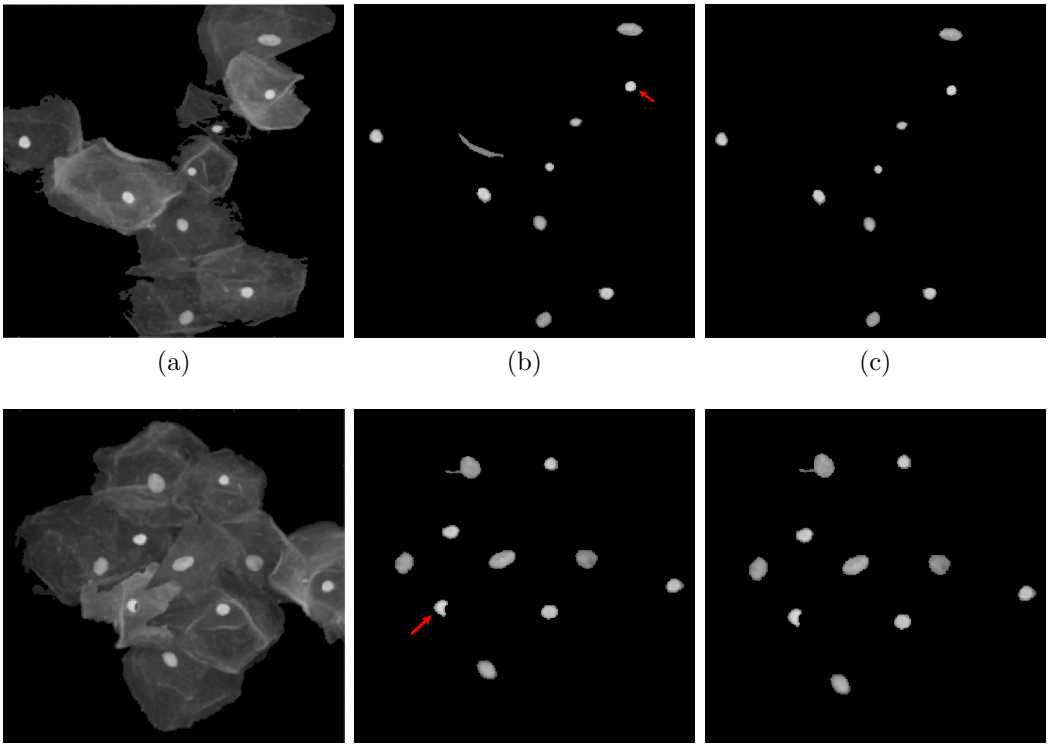


Figure 3.11: MSER technique used to determine t_1 and t_2. (a), (d) Complemented foreground images, (b), (e) candidates after applying MSER algorithm, and (c), (f) area and shape filtered candidates.

The simple thresholding technique described in Section 3.8.2.1 could find candidates successfully to determine the values for t_1 and t_2, if there was sufficient contrast between nucleus and cytoplasm. However in the presence of low contrast and intensity inhomogeneity, this simple technique could not separate nucleus from the cytoplasm parts (see the candidates marked with red arrows in Figure 3.10b and 3.10e). These type of candidates can affect the estimation of t_1 and t_2, therefore they were filtered out using area and shape criteria. As a result, information about those nuclei hiding in the discarded candidates, were not reflected in the values of t_1 and t_2. However, those candidates were present in the MSER technique outcome. Those candidates (marked with red arrow) can be seen in Figure 3.11b and 3.11e.

To overcome this issue, an empirical analysis was carried out using the training dataset. It was found that, while using Otsu multiple thresholding based approach, around 20% candidates were been discarded and has not played any role in determining the values of t_1 and t_2. Only 4% candidates were absent, when calculating the values of thresholds t_1 and t_2 using MSER technique. Therefore in this chapter, MSER algorithm was applied on the complemented foreground image to determine the threshold values required for the intensity modifying function δ in Equation (3.4).

3.8.3 Optimal number of clusters

Optimal number of cluster is one of the required parameters for the proposed framework. Training images were used to determine the optimal number of clusters. Number of dominant peaks in the image histogram was used as the initial cue to find the number of clusters. Using the technique presented in (Silva et al. 2010), dominant peaks were found from complemented foreground training images. In 83% of training images, the number of dominant peaks were 3, 4, or 5. Hence, the proposed framework was analyzed by evaluating all the performance measures with 3, 4 and 5 as number of clusters. In addition, for better understanding of the impact of cluster number on the outcome, proposed framework was further evaluated for 6 and 7 clusters. Table 3.2 shows all the performance measures with number of clusters 3, 4, 5, 6, and 7.

Table 3.2: Finding the optimal number of clusters.

No. of clusters	**DSC**	$\mathbf{PR_{pix}}$	$\mathbf{RC_{pix}}$	$\mathbf{PR_{obj}}$	$\mathbf{RC_{obj}}$
3	**0.942**	0.916	**0.929**	0.926	0.700
4	0.928	0.944	0.921	0.974	0.885
5	0.921	0.964	0.844	0.983	0.878
6	0.908	0.973	0.812	**0.984**	**0.922**
7	0.891	**0.984**	0.770	0.952	0.892

Performance of the proposed framework in terms of DSC is decreased with the increase in number of cluster (see Table 3.2). If the behavior of pixel based precision (PR_{pix}) and recall (RC_{pix}) is observed carefully in Table 3.2, a trade-off can be noticed. When the number of cluster is increased, with the improvement of PR_{pix}, deterioration of RC_{pix} can be seen. Specifically, RC_{pix} shows a sharp declination (from 0.921 to 0.844) for number of cluster 4 to 5. Since, the focus of

this chapter is to improve segmentation accuracy with the incorporation of shape, number of clusters 5, 6, and 7 were not considered further (though the proposed framework achieved highest PR_{obj} and RC_{obj} for cluster 6, however RC_{pix} was only 0.812). With 3 clusters the proposed method achieved the best DSC and RC_{pix}. However RC_{obj} was the lowest (0.700) and 3 clusters were not considered. Therefore, considering the combination of all the evaluation measures, 4 was chosen as the optimal number of cluster for the proposed technique.

3.9 Results

The dataset of 900 images was used for assessing nucleus detection and segmentation performance. Results were compared with standard FCM clustering (Bezdek 1981), baseline and winning techniques (Ushizima et al. 2015, Nosrati & Hamarneh 2015*b*) in ISBI 2014 challenge, and other state-of-the-art methods using the same dataset. To determine the results for the standard FCM clustering, the clustering step in the proposed framework was replaced with standard FCM clustering. Nucleus detection and segmentation results for ISBI 2014 test dataset, were not explicitly reported in short papers for the winning techniques (Ushizima et al. 2015, Nosrati & Hamarneh 2015*b*). However, the baseline and two winning techniques were evaluated and compared, and complete results were reported in (Lu et al. 2017).

Object level comparison of the proposed framework with recent techniques in literature is reported in Table 3.3. Nucleus detection performance of the CiscFC method (in terms of object-based measures) consist of PR_{obj} (0.968) and RC_{obj} (0.882). The highest PR_{obj} (0.994) in (Tareef, Song, Huang, Wang, Feng, Chen & Cai 2017) was achieved on a small subset (90 images) of the ISBI 2014 test set, nevertheless RC_{obj} was low (0.911). Object-based recall RC_{obj} was the highest (0.971) for (Zhang et al. 2019), which utilized a deep learning based technique for nucleus segmentation.

In terms of segmentation accuracy, the proposed technique with CiscFC produced DSC of 0.938, pixel-based precision of 0.927 and pixel-based recall of 0.939. The highest DSC was achieved in (Phoulady et al. 2017), where a nuclei feature (size, average intensity and solidity) based iterative thresholding approach was used. Pixel-based precision was the best for Ushizima *et al.* technique in (Lu et al. 2017) as 0.968, the best recall (0.954) was achieved for a deep learning based technique proposed in (Zhang et al. 2019). Pixel-based evaluation measures and DSC

Table 3.3: Quantitative object based evaluation of the proposed CiscFC technique in comparison with the standard FCM and recent state-of-the-art methods.

Method	**train/test**	$\mathbf{PR_{obj}}$	$\mathbf{RC_{obj}}$
Lu et al. (2015)	45/900	0.730	0.850
Ushizima et al. method in Lu et al. (2017)	45/810	0.959	0.895
Nosrati & Hamarneh method in Lu et al. (2017)	45/810	0.903	0.893
Baseline method in Lu et al. (2017)	45/810	0.977	0.883
Phoulady et al. (2017)	135/810	0.961	0.933
Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen (2017)	45/90	0.990	0.940
Tareef, Song, Huang, Wang, Feng, Chen & Cai (2017)	45/90	**0.994**	0.911
Tareef et al. (2018)	45/900	0.983	0.959
Zhang et al. (2019)	45/900	0.990	**0.971**
Proposed framework with FCM	45/900	0.952	0.852
Proposed framework with CiscFC	45/900	0.968	0.882

Bold values represent the best performance per measure.

for the proposed framework and other recent methods in literature, are reported in Table 3.4.

The influence of shape regularization using level set method (Li et al. 2010) on the proposed framework (with CiscFC and standard FCM) is shown in Table 3.5, in terms of all evaluation measures. The results indicate that shape regularization impact is very limited over the dataset for the proposed framework. While using shape regularization with CiscFC, little improvement was achieved for DSC, RC_{pix} and PR_{obj}, and decline in performance can be seen for PR_{pix} and RC_{obj}. In case of standard FCM, shape regularization has improved only DSC and PR_{pix}, while deteriorating RC_{pix}, PR_{obj} and RC_{obj}. Therefore, this regularization process only helped to refine the nucleus boundaries without much impact on the overall outcome.

The proposed framework with CiscFC was run for 5 consecutive times to identify the impact of different initial partition matrix and seed points on the evaluation measures. These results are reported in Table 3.6. Mean and standard deviation for all performance measures were calculated from the results of all 5 runs and reported in the last row in Table 3.6. The outcome shows that the

Table 3.4: Quantitative comparison of segmentation accuracy of the proposed CiscFC methodology, with the standard FCM and recent state-of-the-art techniques.

Method	**train/test**	**PR_{pix}**	**RC_{pix}**	**DSC**
Lu et al. (2015)	45/900	0.960 (0.060)	0.900 (0.080)	0.920 (0.050)
Ushizima et al. method in Lu et al. (2017)	45/810	**0.968** (0.055)	0.871 (0.069)	0.914 (0.039)
Nosrati & Hamarneh method in Lu et al. (2017)	45/810	0.901 (0.097)	0.916 (0.093)	0.900 (0.053)
Baseline method in Lu et al. (2017)	45/810	0.942 (0.078)	0.912 (0.081)	0.921 (0.049)
Phoulady et al. (2017)	135/810	-	-	**0.947** (-)
Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen (2017)	45/90	0.950 (0.060)	0.930 (0.070)	0.930 (0.040)
Tareef, Song, Huang, Wang, Feng, Chen & Cai (2017)	45/90	0.940 (0.060)	0.950 (0.060)	0.940 (0.040)
Tareef et al. (2018)	45/900	0.906 (0.068)	0.950 (0.051)	0.925 (0.041)
Zhang et al. (2019)	45/900	0.902 (-)	**0.954** (-)	0.931 (-)
Proposed framework with FCM	45/900	0.949 (0.033)	0.875 (0.041)	0.931 (0.040)
Proposed framework with CiscFC	45/900	0.927 (0.095)	0.939 (0.090)	0.938 (0.040)

Bold values represent the best performance per measure.
All results are reported in *mean (std)* format.
'-' indicates unavailable values.

Table 3.5: Evaluation of shape regularization impact on performance measures.

	Proposed framework with CiscFC					**Proposed framework with FCM**				
	DSC	**PR_{pix}**	**RC_{pix}**	**PR_{obj}**	**RC_{obj}**	**DSC**	**PR_{pix}**	**RC_{pix}**	**PR_{obj}**	**RC_{obj}**
Without[a]	0.938	0.927	0.939	0.968	0.882	0.931	0.949	0.875	0.952	0.852
With[b]	0.945	0.909	0.942	0.969	0.866	0.936	0.970	0.871	0.939	0.836
	↑	↓	↑	↑	↓	↑	↑	↓	↓	↓

[a]Without shape regularization.
[b]With shape regularization using level set method (Li et al. 2010).
Last row indicates incline (↑) or decline (↓) of performance with shape regularization.

standard deviations for all evaluation measures over 5 runs are very small, and positively indicates the stability of the proposed framework.

Table 3.6: Identifying stability of the proposed CiscFC framework with respect to different initial partition matrices. Last row shows the mean and standard deviation over 5 runs.

	DSC	$\mathbf{PR_{pix}}$	$\mathbf{RC_{pix}}$	$\mathbf{PR_{obj}}$	$\mathbf{RC_{obj}}$
Run 1	0.938	0.927	0.939	0.968	0.882
Run 2	0.938	0.926	0.928	0.969	0.885
Run 3	0.938	0.927	0.928	0.972	0.876
Run 4	0.938	0.927	0.928	0.969	0.886
Run 5	0.938	0.928	0.938	0.971	0.876
	0.938 (0.000)	0.927 (0.000)	0.932 (0.005)	0.970 (0.001)	0.881 (0.004)

3.10 Discussion

This chapter presented a novel technique to incorporate a circular shape constrain in standard FCM clustering. Performance of the proposed method in terms of: Dice similarity coefficient, pixel based precision and recall, and object based precision and recall, is comparable to the state-of-the art techniques. The proposed CSF included in FCM clustering, helped the algorithm to extract nucleus more precisely, which is noticeable from DSC and higher pixel based recall values. For the test dataset, DSC of more than 83% of correctly detected nuclei is above 0.90 and DSC of only 1% of the correct detections is below 0.80. Figure 3.12a shows the bar graph for DSC. In terms of pixel based measures, for more than 80% of correct detections, the proposed method achieved precision value of at least 0.90. Recall value was at least 0.80 for all and greater than 0.90 for about 95% of the correctly detected nucleus, using the proposed framework. Bar graphs in Figure 3.12b and 3.12c illustrate these pixel level measures.

To better understand the influence of circular shape prior on circularity measure of segmented cervical nuclei, an in depth analysis was performed. Circularity (ψ) of a segmented nucleus was calculated using Equation (4.4). A threshold was found by investigating circularity value distribution of all ground truth nuclei from both training and test images. Circularity distributions are shown in Figure 3.13. By analyzing the histograms presented in Figure 3.13a and 3.13b, it was observed that 99% training and 98% test nuclei, have circularity (ψ) of at least 0.8. Therefore, 0.8 was selected as the circularity threshold for further analysis.

For ISBI test set, histograms of circularity values are reported in Figure 3.14. In case of standard FCM clustering, 93% of correctly detected nuclei have ψ of

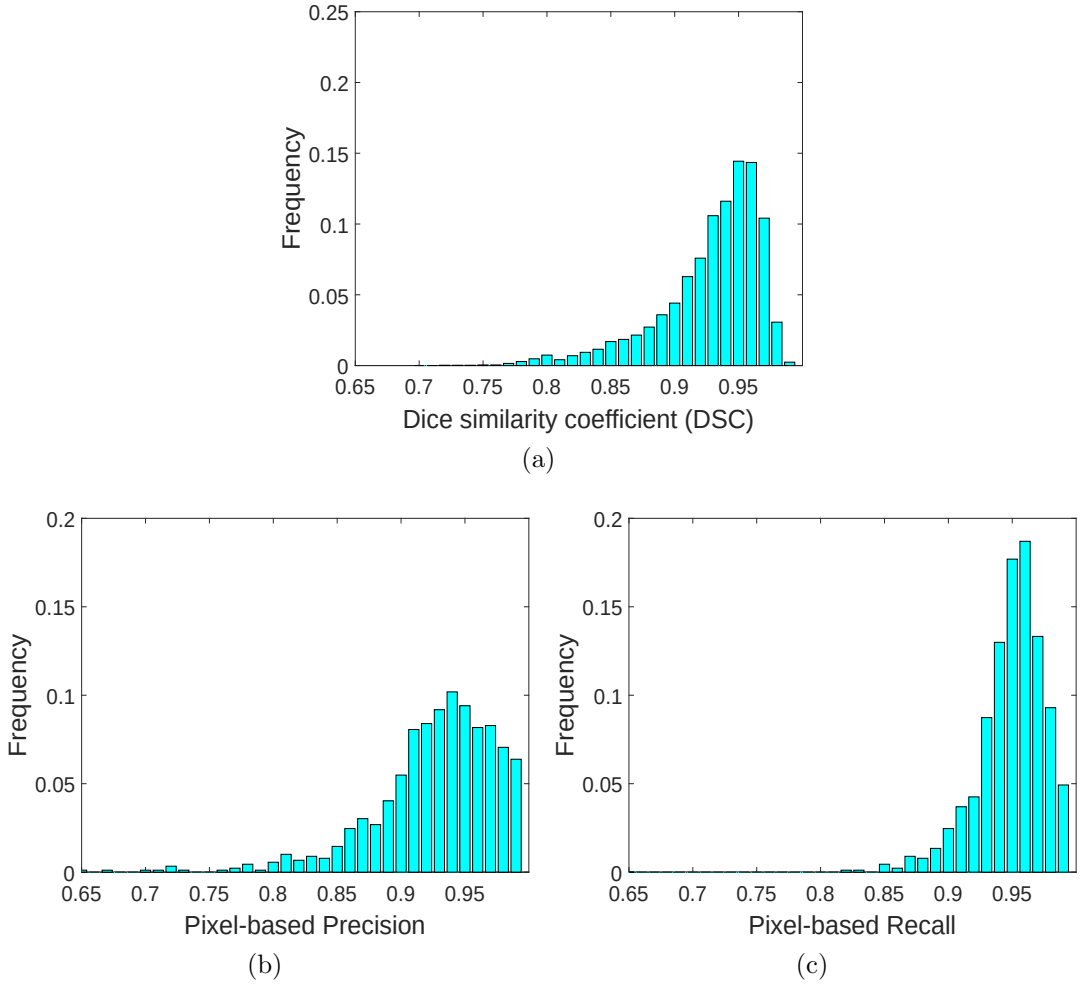
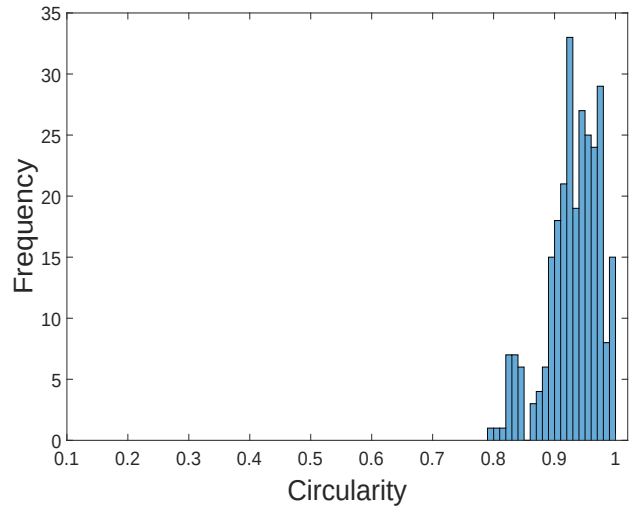


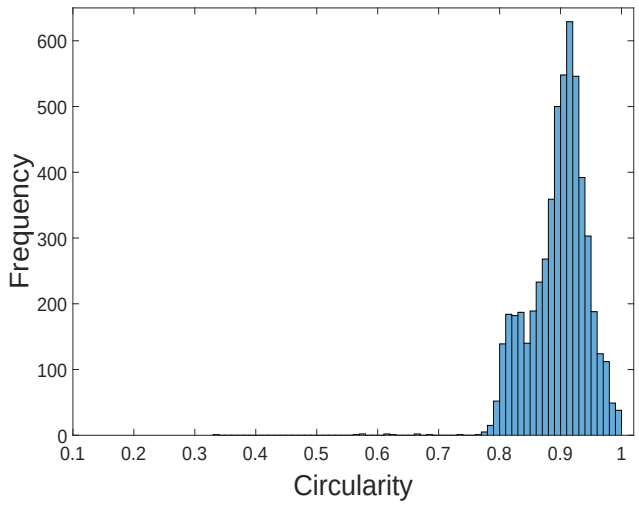
Figure 3.12: Plot of segmentation accuracy measures of the proposed framework for ISBI test dataset: (a) Dice similarity coefficient, (b) pixel based precision and (c) pixel based recall.

at least 0.8 (can be seen from Figure 3.14a). Circularity value histogram for the proposed framework with CiscFC technique is shown in Figure 3.14b. Analysis of this histogram indicates that of 94% correctly detected nuclei have $\psi \geq 0.8$. Therefore, inclusion of circular shape prior has also improved the measure of circularity, while precisely segmenting cervical nuclei.

Visual inspection was performed for qualitative evaluation of the nucleus segmentation results summarized in Figure 3.15. The segmented nuclei found using the proposed CiscFC method and those detected using the standard FCM clustering, can also be visually compared. The proposed framework with the circular shape function (CSF), produced more precise outcome even in challenging situations. Visual results also indicate that the proposed nuclei segmentation approach outperformed the standard FCM algorithm in cases, where proposed approach



(a)

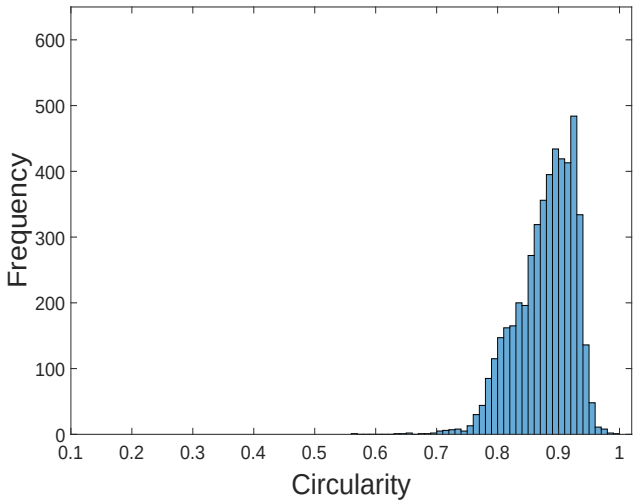


(b)

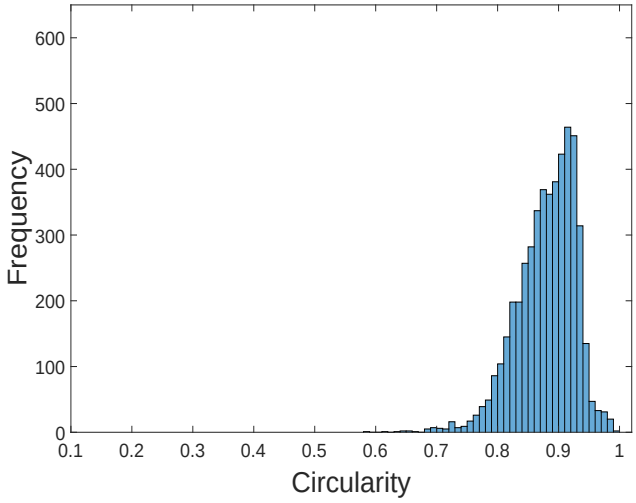
Figure 3.13: Histograms of circularity values of all ground truth nuclei, from ISBI (a) training and (b) test datasets.

could detect and segment nucleus but FCM could not (see Figure 3.15b, 3.15c, 3.15e, 3.15f, and 3.15h, 3.15i).

Finally, it can be concluded from both quantitative and qualitative evaluations that the proposed cervical nuclei segmentation technique detected and segmented nucleus more precisely than the standard FCM clustering. In addition to precise segmentation of nuclei, the proposed technique also improved the circularity measure while segmenting the target region. The proposed method is more resilient in typical difficult situations such as with small average intensity difference



(a)



(b)

Figure 3.14: Circularity value histograms of segmented nucleus using the proposed framework with: (a) standard FCM and (b) CiscFC techniques, from ISBI test dataset.

between the nucleus and cytoplasm - caused by cell overlapping, poor contrast or uneven staining. The outcome of the proposed CiscFC algorithm also matches the performance of recent state-of-the-art methods.

The proposed framework has some limitations that need to be addressed. First, to determine the optimal number of clusters, a training dataset is required. The small training set of 45 images was used to determine the threshold values for area and shape features (the threshold values are reported in Table 3.1). Some

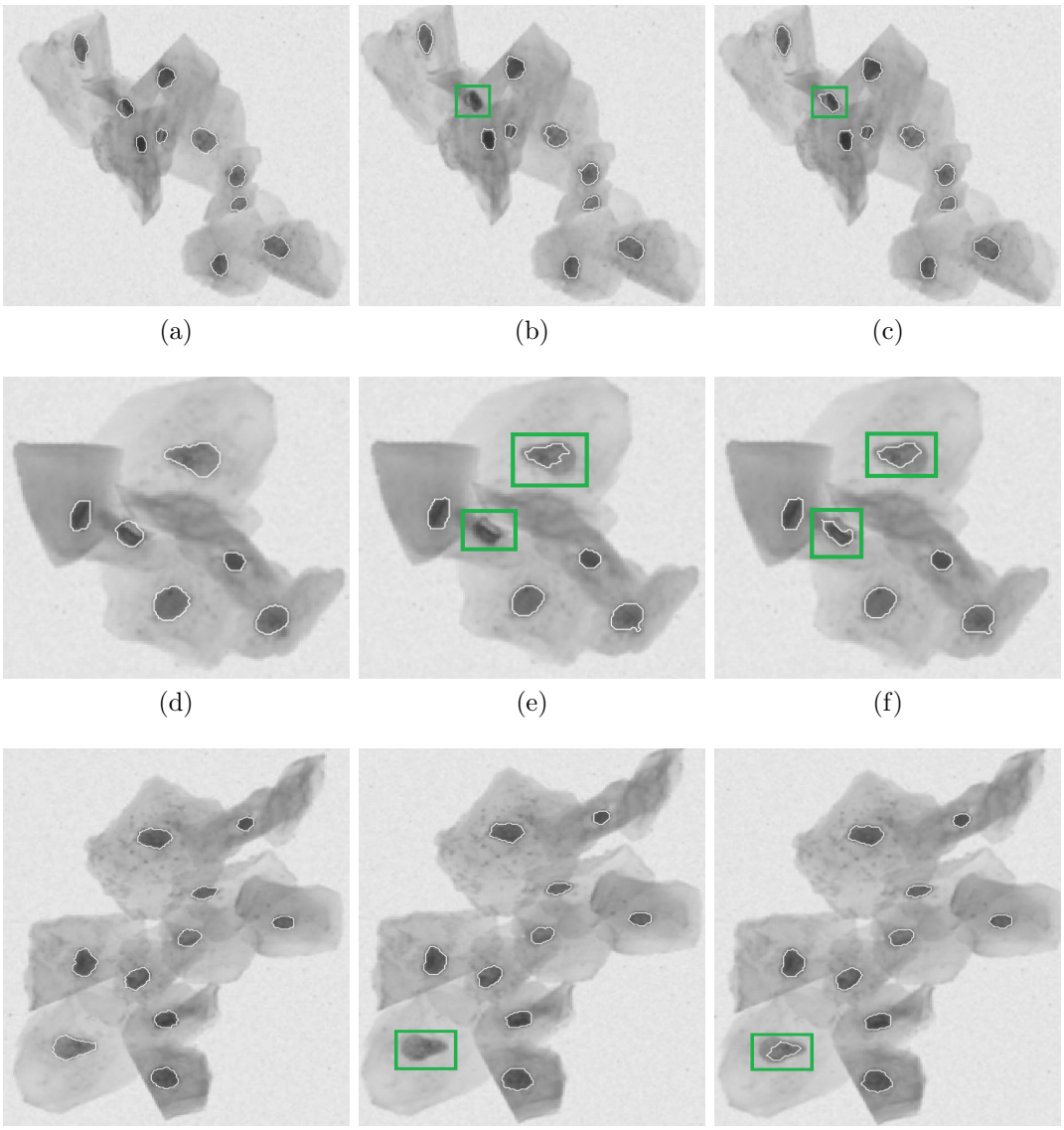


Figure 3.15: Nucleus segmentation examples from the test dataset: first column represents the original images marked with nuclei ground truth, second column presents nuclei segmented using the standard FCM clustering, and the last column shows nuclei segmented using the proposed CiscFC technique. Differences between CiscFC and the standard FCM are marked with green rectangles.

nuclei were removed along with outliers, during the false positive reduction stage using those features. As shown in Figure 3.16, the considered nucleus was present (with eccentricity of 0.92 and circularity of 0.73) in detected nucleus cluster from the output of the proposed method (marked in Figure 3.16b). However, during the outlier removal process that nucleus was removed, since the eccentricity threshold criterion was not been satisfied. Thus it was absent in the final nuclei segmentation output (as can be seen from Figure 3.16c). Another limitation of

(a) (b) (c)

Figure 3.16: An example of missed nucleus detection, (a) nucleus in question (circled in red), (b) detected target nucleus (circled in red) after applying the proposed method and (c) target nucleus removed after false positive reduction.

the proposed method is that the highest average intensity cluster from the clustering output, was detected as the cluster containing nuclei. Thus, the nuclei (with uneven staining and very poor contrast) assigned to other clusters during the clustering process, may not be detected using the proposed framework. Some examples are shown in Figure 3.17, for nucleus detection and segmentation failures due to poor and uneven staining.

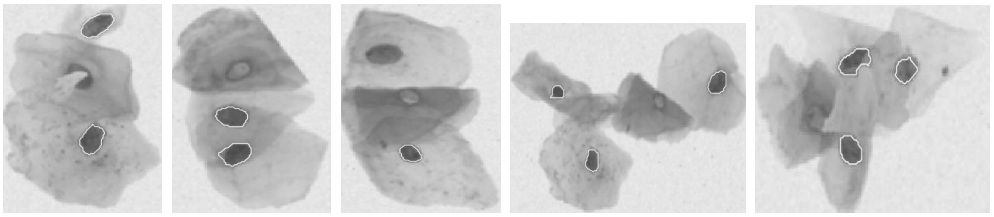


Figure 3.17: Examples of missed nucleus detection and segmentation due to uneven staining and poor contrast.

3.11 Conclusion

This chapter presented a framework for cervical nuclei segmentation in overlapping Pap smear images, and introduced a new circular shape guided fuzzy clustering algorithm. A circular shape function (CSF) was incorporated in the proposed technique, to influence the clustering process and to precisely segment cervical nuclei from foreground images. ISBI 2014 challenge dataset of 900 images was used to evaluate the performance of the proposed framework, both quantitatively and qualitatively. Pixel based and object based measures were used, for quantitative evaluation. Qualitative evaluation was performed by visually inspecting the segmented cervical nuclei. Promising outcome was produced by the

proposed approach, for detection (object based precision 0.968 and recall 0.882) and segmentation (DSC 0.938, pixel based precision 0.927, and recall 0.939) of cervical cell nuclei, compared to the state-of-the-art methods. High segmentation accuracy with small false positive and false negative detection, indicates that the proposed framework is effective for segmenting nuclei from overlapping cervical cytology images.

Chapter 4

Shape prior in graph based segmentation

Graph based segmentation is another useful paradigm in computerized image analysis and machine vision. From graph theoretic segmentation approaches, efficient graph based segmentation (EGS) is well acknowledged and computationally efficient. The EGS technique is based on the concept of minimum spanning trees (MST) in graphs. In this chapter, a novel technique is proposed to incorporate a weighted circular shape prior in efficient graph based segmentation technique. Section 4.1 provides a brief overview of graph based image segmentation techniques and its variants incorporating shape priors in literature. An overview of standard efficient graph based segmentation, and how circular shape prior was defined and incorporated into EGS technique, is described in Section 4.2. Section 4.4 presents the evaluation measures used in this chapter. Experimental set-up and parameter tuning, are described in Section 4.5 and Section 4.6, respectively. Section 4.7 details the experimental results for nucleus detection, segmentation, evaluation of boundary precision, and comparison with other recent state-of-the-art techniques. Discussion and conclusion of this chapter, are provided in Section 4.8 and Section 4.9, respectively.

4.1 Introduction

Mathematical structure of graph is good for presenting and formulating image characteristics. Graph theoretic approaches are widely used for image segmentation, since both local and global image information can be included in the segmentation process (Ma, Bajger, Slavotinek & Bottema 2007, Das et al. 2009). Various

graph based segmentation techniques have been proposed over the years (Wu & Leahy 1993, Boykov & Jolly 2000, Shi & Malik 2000, Boykov & Jolly 2001, Felzenszwalb & Huttenlocher 2004, Boykov & Funka-Lea 2006, Grady & Schwartz 2006, Grady 2006, Veksler 2008, Grosgeorge et al. 2013, Zhang, Kong, Liu, Wang, Chen & Sonka 2017).

In the field of image segmentation and analysis, graph cut based image segmentation (Boykov & Jolly 2001) is one of the most popular techniques. This type of segmentation is based on top-down strategy, where initially whole image is considered as a single segment and then partitioned into sub-segments until a termination criteria is satisfied (Kalinin & Sirota 2015). In order to handle images with noise or other imaging artifacts, new variants of graph cut based segmentation techniques have been developed by including shape priors and adopted to various contexts. A generic star shape prior for convex objects, was defined in (Veksler 2008). This prior helped to lessen the bias of graph cut segmentation towards shorter segmentation boundaries. The prior was defined based on geometric properties of the object and its center. An elliptical shape prior was proposed in (Slabaugh & Unal 2005), and used with graph cut segmentation. The minimum cut was computed from the graph with narrow-band vertices around an initial ellipse. Then, the best fitting ellipse was found iteratively from the minimum cut. Blood vessels and lymph nodes were segmented from pelvic MR images using this modified graph-cut segmentation technique. In (Vu & Manjunath 2008), shape prior was integrated in graph cut for multiple object segmentation. The shape prior was defined in terms of shape distance function. Then, multiple priors were incorporated in the form of a shape energy and used with multiphase graph cut framework. An adaptive way for inclusion of shape prior was proposed in (Wang, Zhang & Ray 2013). Based on the specific needs of different pixels, shape weight was computed adaptively. Similarity in neighborhood pixels was used to determine the shape influence. In (Grosgeorge et al. 2013), PCA was used to develop a statistical shape model, that guided the graph cut segmentation. This modified graph cut segmentation was applied on cardiac MRI images. In (Grosgeorge et al. 2016), a multi-label shape prior guided graph cut segmentation framework was presented. The shape prior was found from probabilistic prior map and integrated into graph cost function. Left and right ventricle, and myocardium, were segmented from cardiac MR images using this technique.

Minimum spanning tree (MST) based graph theoretic approach, is another popular technique of segmentation. Bottom-up strategy is adopted in this seg-

mentation, where initially each image pixel constitutes separate segments, and then segments satisfying specific criteria are merged. Several techniques based on MST are available in literature (Zahn 1971, Urquhart 1982, Felzenszwalb & Huttenlocher 1998, Felzenszwalb & Huttenlocher 2004). The efficient graph based segmentation (EGS) - proposed in (Felzenszwalb & Huttenlocher 2004), is the widely used variant that uses Kruskal's algorithm to find MST in graphs. EGS technique has the ability to extract target object with better segmentation accuracy (Peng et al. 2013). In (Ma, Bajger & Bottema 2007), EGS technique was used to evaluate/investigate the robustness of mammogram segmentation to image distortions such as shifting, rotation, and warping in the field of temporal analysis of screening mammograms. In (Ma, Bajger, Slavotinek & Bottema 2007), pectoral muscle was segmented from mammograms using the EGS algorithm. In (Bajger et al. 2013), standard EGS method was extended to 3D-EGS for segmentation of multiple human organs (lungs, stomach, liver, heart, kidney, spleen, bones, and spinal cord) form CT images. In (Huang et al. 2012), EGS technique was modified by incorporating a new pair-wise region merging predicate, and used to segment breast tumors from ultrasound images. In spite of the popularity of standard EGS algorithm, the technique is sensitive to weak region boundaries - where variability between two regions is small (Peng et al. 2013).

In this chapter, a novel circular shape guided graph based image segmentation technique is proposed. In addition to intensity information, the merging predicate of standard EGS algorithm was redefined to exploit circular shape information. The proposed technique was evaluated by segmenting nuclei from overlapping cervical cytology images.

4.2 Circular shape prior in efficient graph based segmentation (CircEGS)

4.2.1 An overview of EGS method

Let $G = (V, E)$ be a graph, where V is the set of vertices and E is the set of edges. In graph based segmentation approach, a graph is segmented into mutually exclusive components in a way that each component C corresponds to a connected component in a sub-graph $G' = (V, E')$, where $E' \subseteq E$. Edge weight can be defined considering properties associated with the vertices it connects. In EGS technique, segmentation of a graph is performed such that there is enough

similarity within a component and dissimilarity across different components.

Image pixels represent the vertices for graph based image segmentation methods. The standard EGS technique implements bottom-up strategy, therefore segmentation process is performed in conjunction with a region merging process of image pixels. Region merging is performed considering a pairwise region comparison predicate for generating components/regions. This predicate measures and compares the intra/within and inter/between component dissimilarity, where edge weight is defined as the absolute image intensity difference of the corresponding components. Two components are allowed to merge, if the internal difference of at least one component is larger than the difference between two components. The formal definition of the merging predicate for any two components C_1 and C_2 is:

$$D(C_1, C_2) \leq min\left(I(C_1) + \tau(C_1), I(C_2) + \tau(C_2)\right), \tag{4.1}$$

where the difference between two components $D(\cdot, \cdot)$ is calculated as the minimum edge weight connecting the two components. The internal difference for a component $I(\cdot)$ is found as the maximum edge weight in the MST of the component. These differences are adaptively calculated from components' edge weights in the image. The threshold function τ is defined as:

$$\tau(C) = \frac{k}{|C|}, \tag{4.2}$$

where $|\cdot|$ denotes the size of a component, and k is a constant. The value of constant k controls the granularity of the segmentation; larger k imposes preference for larger components. Two components C_1 and C_2 will be merged, if the merging criterion in Equation (4.1) is satisfied. Algorithm 1 presents the pseudo-code for the standard EGS technique (Felzenszwalb & Huttenlocher 2004).

4.2.2 The proposed circular shape prior

The threshold function τ defined in Equation (4.2) is redefined to employ the circular shape constraint. This redefined new threshold is used in the merging predicate in Equation (4.1), so that it is possible to prefer components of a target shape.

The new threshold function is defined as a weighted sum of the original threshold function τ or *area* term and a new circular *shape* term ψ. Given a component

Algorithm 1 Efficient graph based segmentation algorithm

Input: Graph $G = (V, E)$ with n vertices and m edges

Output: Segmented components $S = (C_1, \ldots, C_r)$

1: Sort edges E by increasing edge weights and save the ordered edges as $O = (e_1, \ldots, e_m)$.
2: Start with a segmentation S^0, where each vertex constitutes a single component.
3: Repeat step 4 for $q = 1, \ldots, m$.
4: Construct S^q from S^{q-1}. Let v_i and v_j be two vertices of q-th edge in O. If v_i and v_j are in disjoint components C_i^{q-1} and C_j^{q-1} of S^{q-1}, and difference between these two components is small comparing to the internal difference of both of these components, then S^q is obtained by merging C_i^{q-1} and C_j^{q-1}. Otherwise, $S^q = S^{q-1}$.
5: Return $S = S^m$.

C, the new threshold function $\bar{\tau}$ is defined as:

$$\bar{\tau}(C) = \lambda \cdot \tau(C) + (1 - \lambda) \cdot \psi(C), \tag{4.3}$$

where weight parameter $0 \leq \lambda \leq 1$ controls the relative importance between the *area* and *shape* terms. Smaller λ makes the shape term more important and vice versa.

Circular *shape* term ψ reflects the circularity measure of a component. Following (Giger et al. 1988), an equivalent circle is used to measure the circularity criterion. The circle with equivalent area of the considered component is called an equivalent circle. The overlapped area of the component with its equivalent circle and the total area of that component, is used to define ψ. This is schematically demonstrated in Figure 4.1.

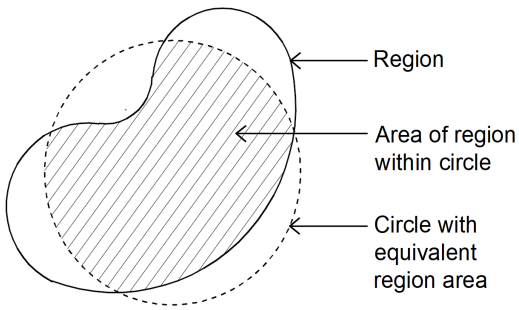


Figure 4.1: Schematic illustration of circularity definition.

Circularity measure ψ for a component C is defined as:

$$\psi(C) = \frac{|C \cap EQ_c|}{|C|}, \tag{4.4}$$

where EQ_c is the equivalent circle, centered at the component centroid with radius $r = \sqrt{|C|/\pi}$. Note that, EQ_c has the equivalent area as C and if C is a circle, then $\psi(C) = 1$. Hence, for any component C, the value of circularity measure ψ lies in the range $[0, 1]$. Higher value of ψ indicates that the component C is more circular. Two components C_1 and C_2 are merged, if the following merging predicate is satisfied:

$$\bar{D}(C_1, C_2) \leq min\left(I(C_1) + \bar{\tau}(C_1), I(C_2) + \bar{\tau}(C_2)\right). \tag{4.5}$$

The proposed CircEGS segmentation technique enforces more circular nucleus in the segmentation output. In case of images with weak edges between regions caused by low contrast, object occlusion, overlapping, noise, the proposed CircEGS method helps to preclude circular components from merging with other non-circular components.

4.3 Dataset

The segmentation technique proposed in this chapter was evaluated using the ISBI 2014 challenge dataset (described in Section 3.3). This dataset contains 945 grayscale synthetic cervical smear images of 512 $\times$ 512 pixels, subdivided into the training set of 45 images - used for parameter tuning, and a test set of 900 images - used for evaluation.

4.4 Evaluation metrics

After the segmentation process is completed, the region with the highest overlap with the ground truth nucleus (measured by Dice similarity coefficient (DSC) using Equation (3.16)) - was considered as the segmented nucleus and used for evaluation of the nucleus segmentation performance. Quantitative evaluation measures defined in Section 3.5 were used in this study to evaluate the proposed CircEGS technique. The measures are: object-based precision (PR_{obj}), object-based recall (RC_{obj}), pixel based precision (PR_{pix}), pixel based recall (RC_{pix}),

DSC, and circularity (ψ). In addition to these, Hausdorff distance (H) and border error with tolerance (H_t metric) were also used to evaluate the segmentation boundaries. Performance evaluation measures were calculated using all correctly detected nucleus regions that satisfy Equation (3.11).

Hausdorff distance is the maximum distance of a set to the nearest point in the other set (Huttenlocher et al. 1993). This distance can be used to measure the proximity of two sets. Let $A = \{a_1, a_2, \ldots, a_p\}$ and $B = \{b_1, b_2, \ldots, a_q\}$ be two finite sets with p and q number of elements, respectively. The Hausdorff distance h from set A to B is defined as:

$$h(A, B) = \max_{a \in A} (\min_{b \in B} \|a - b\|), \tag{4.6}$$

where $\|\cdot\|$ is some underlying norm on the points of the sets A and B. Since, the Hausdorff distance is asymmetric (usually $h(A, B) \neq h(B, A)$), a general definition of Hausdorff distance H between two sets A and B is:

$$H(A, B) = \max (h(A, B),\, h(B, A)). \tag{4.7}$$

Note that $H(A, B) = H(B, A)$. In this study, H was used as a measure of distance (in pixels) between the segmented nucleus and its ground truth, to indicate how close the boundaries are. Lower value for H is preferable. Since, H is computed using the estimated and ground truth ROI boundaries, it is sensitive to noise and outliers. In medical image segmentation, usually ground truth boundaries are delineated by human observer and there exist variability (Bajger et al. 2013). Hence, solely it may not be the best indication of segmentation accuracy for techniques that do not aim to minimize this measure.

The border positioning error or H_t metric (Kubassova et al. 2006), considers the uncertainty in ground truth boundaries by using a tolerance. It is calculated as the mean deviation of boundary pixels from the ground truth over the entire border length of the segmented region. The H_t metric is the percentage of pixel-wise closeness between two region boundaries with tolerance t pixels. For two boundaries A and B, H_t metric is defined as:

$$H_t(A, B) = \frac{1}{2} \left(\frac{N_{A_t}}{N_A} + \frac{N_{B_t}}{N_B} \right), \tag{4.8}$$

where N_A and N_B are the number of pixels in region boundaries A and B, t is the tolerance in pixels, and N_{A_t} and N_{B_t} are the number of pixels correctly identified with tolerance t for the boundaries A and B. For a given boundary,

the H_t metric increases monotonically with t and converges to 1. This metric was used to measure the boundary precision of the segmented nucleus compared to the ground truth.

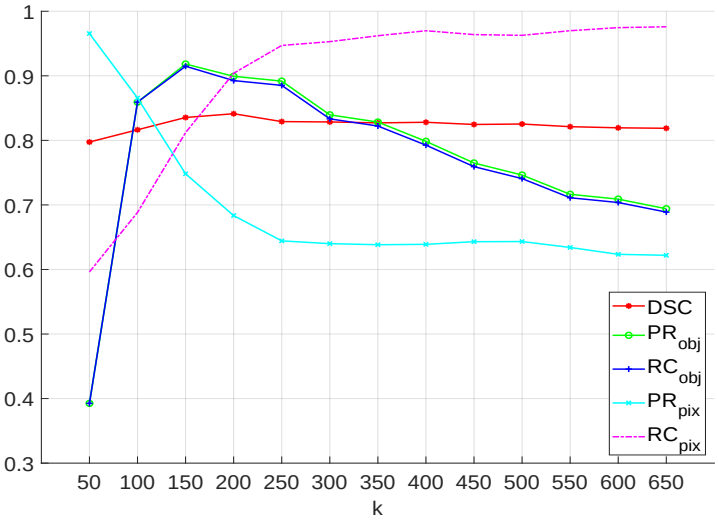
4.5 Experimental setup

Parameters required for the EGS and the proposed CircEGS algorithms, were tuned using 45 training images. Test dataset of 900 images was used to evaluate detection and segmentation performance of the proposed framework. Both EGS and CircEGS techniques were implemented in-house using Java programming language. A PC with Intel Core i7-4770 3.40 GHz processor and 8 GB RAM was used to run both segmentation techniques. After segmenting cervical smear images with standard EGS and proposed CircEGS, all performance evaluations were implemented in Matlab R2017b.

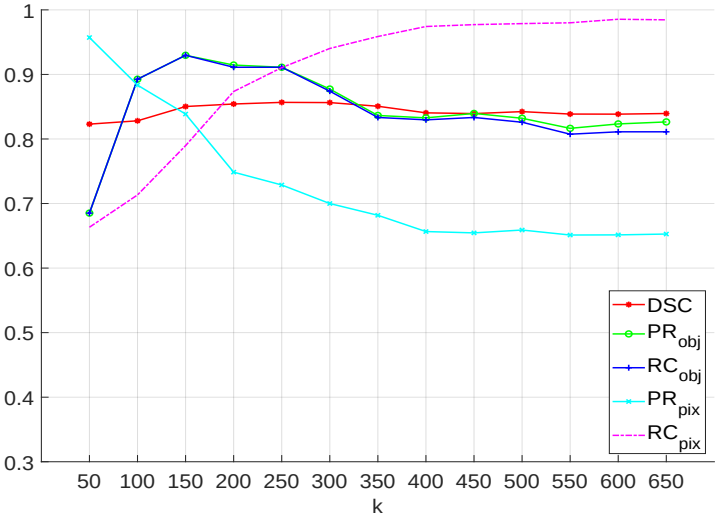
4.6 Parameter tuning

The constant k (see Equation (4.2)) need to be determined empirically, for both EGS and CircEGS techniques. Parameter k controls the degree of similarity between components, therefore the final number of components (and their average size) in the segmented output (Ma, Bajger, Slavotinek & Bottema 2007, Bajger et al. 2013). Size of nucleus varies from 100 to 600, in 45 images of ISBI training set. Hence, training images were segmented using EGS and CircEGS techniques with k values in the range $[50, 650]$ with step of 50.

Detection and segmentation evaluation measures (PR_{obj}, RC_{obj}, DSC, PR_{pix}, and RC_{pix}) for the training set, were found for EGS and CircEGS segmented nucleus regions and plotted in graphs (shown in Figure 4.2). DSC is an overlap index - calculated from correct detections only. It can be observed from the graphs that DSC is rather stable over the whole range of k values for both EGS and CircEGS methods. The trade-off between PR_{pix} and RC_{pix} can be clearly seen from the graphs for EGS and CircEGS techniques, and both plots intersect around $k = 150$. The sharp decline of PR_{pix} and incline of RC_{pix} stopped at $k = 250$, and both plots relatively stabilized after $k = 250$. A quick improvement of PR_{obj} and RC_{obj} is visible in Figure 4.2a and 4.2b, for $k = 50$ to $k = 100$, reached maximum for $k = 150$, and then started to decline. Since larger k value produces larger components, the increase in k value impacted object based evaluation measures



(a)



(b)

Figure 4.2: Plots of evaluation measures for k in the range $[50, 650]$ using ISBI training set using, (a) EGS and (b) CircEGS techniques.

calculated from all segmented regions and number correct detections from the segmentation outcome. It is worth mentioning that, for CircEGS technique both PR_{obj} and RC_{obj} remained above 0.8 even with large k values, whereas for EGS technique both object based measure declined below 0.7 for $k = 650$. It can be concluded after the analysis of the graphs in Figure 4.2a and 4.2b that both detection and segmentation performance of EGS and CircEGS are comparable for k in the range $[100, 250]$. For ISBI dataset, $k = 150$ was selected and used for

both EGS and CircEGS techniques in this study.

Weight parameter λ for CircEGS method in Equation (4.3), controls the relative importance of area or circular shape constrain. The value of λ was found empirically as 0.65, using the training images. Table 4.1 summarizes these findings.

Table 4.1: Empirical values for parameter k and weight parameter λ.

Parameter	**Value**
k	150
λ	0.65

4.7 Results

Nucleus detection and segmentation performance of the proposed CircEGS technique was assessed using 900 images from the test dataset. Before segmentation process started, first morphological reconstruction based background removal technique (see Section 3.4.1.2 for details) was applied on the original images, to produce the foreground images. Then, foreground images were rescaled by 0.5 to reduce processing time. Finally, a Gaussian filter with standard deviation 0.8 was used to slightly smooth the images. These pre-processed foreground images were used for segmentation using both standard EGS and the proposed CircEGS techniques.

All results were compared with the standard EGS method (Felzenszwalb & Huttenlocher 2004), the baseline and winning techniques (Ushizima et al. 2015, Nosrati & Hamarneh 2015*b*) in ISBI 2014 challenge, and other state-of-the-art methods using the same dataset. Nucleus detection and segmentation results for ISBI 2014 test dataset, were not explicitly reported in short papers for the winning techniques (Ushizima et al. 2015, Nosrati & Hamarneh 2015*b*). However, the baseline and two winning techniques were evaluated and compared, and complete results were reported in (Lu et al. 2017).

4.7.1 Quantitative evaluation

For the ISBI dataset, DSC, pixel and object-based precision and recall, Hausdorff distance, circularity, and H_t metric were computed with $k = 150$, for both EGS

and CircEGS techniques.

Object level comparison of the proposed algorithm with standard EGS, CiscFC technique proposed in Chapter 3 and recent techniques in literature, is reported in Table 4.2. The proposed CircEGS outperformed standard EGS method in terms of both object based precision (0.868/0.830) and recall (0.861/0.811), however these results were lower than the precision and recall values achieved for CiscFC. In (Tareef, Song, Huang, Wang, Feng, Chen & Cai 2017), the highest object-based precision (0.994) was achieved on a small subset (90 images) of the ISBI 2014 test set, nevertheless object based recall was low (0.911). The highest value for RC_{obj} (0.971) was achieved in (Zhang et al. 2019), where a deep learning based technique was used.

Table 4.2: Object level evaluation of the proposed CircEGS technique, in comparison with the standard EGS and recent state-of-the-art methods.

Method	**train/test**	$\mathbf{PR_{obj}}$	$\mathbf{RC_{obj}}$
Lu et al. (2015)	45/900	0.730	0.850
Ushizima et al. method in Lu et al. (2017)	45/810	0.959	0.895
Nosrati & Hamarneh method in Lu et al. (2017)	45/810	0.903	0.893
Baseline method in Lu et al. (2017)	45/810	0.977	0.883
Phoulady et al. (2017)	135/810	0.961	0.933
Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen (2017)	45/90	0.990	0.940
Tareef, Song, Huang, Wang, Feng, Chen & Cai (2017)	45/90	**0.994**	0.911
Tareef et al. (2018)	45/900	0.983	0.959
Zhang et al. (2019)	45/900	0.990	**0.971**
CiscFC method in Chapter 3	45/900	0.968	0.882
Standard EGS	45/900	0.830	0.811
Proposed CircEGS	45/900	0.868	0.861

Bold values represent the best performance per measure.

In terms of segmentation accuracy, the proposed CircEGS produced DSC of 0.897, pixel-based precision of 0.782 and pixel-based recall of 0.955. Segmentation results of standard EGS technique was lower (DSC - 0.839, PR_{pix} - 0.724, and RC_{pix} - 0.866), compared to the proposed method. Evaluation measures to determine segmentation accuracies for the proposed CircEGS, standard EGS,

CiscFC technique proposed in Chapter 3, and other recent methods in literature, are reported in Table 4.3. CiscFC technique outperformed the proposed CircEGS algorithm, in terms of all segmentation accuracy evaluation metrics. The highest PR_{pix} (0.968) was found for the technique proposed by Ushizima *et al.* in (Lu et al. 2017). The proposed CircEGS technique achieved the best value for RC_{pix} (0.954), and the highest DSC (0.947) was achieved in (Phoulady et al. 2017).

Hausdorff distance (H) gives an indication of the largest distance between the ground truth and segmented nucleus regions (Bajger et al. 2013), therefore lower value of H is preferable. None of the state-of-the-art methods reported this measure. However, H is calculated for the proposed technique and reported in Table 4.4. The CircEGS method achieved better result for H than the standard EGS technique (1.628/2.019).

Table 4.3: Pixel level evaluation of the proposed CircEGS technique.

Method	**train/test**	$\mathbf{PR_{pix}}$	$\mathbf{RC_{pix}}$	**DSC**
Lu et al. (2015)	45/900	0.960 (0.060)	0.900 (0.080)	0.920 (0.050)
Ushizima et al. method in Lu et al. (2017)	45/810	**0.968** (0.055)	0.871 (0.069)	0.914 (0.039)
Nosrati & Hamarneh method in Lu et al. (2017)	45/810	0.901 (0.097)	0.916 (0.093)	0.900 (0.053)
Baseline method in Lu et al. (2017)	45/810	0.942 (0.078)	0.912 (0.081)	0.921 (0.049)
Phoulady et al. (2017)	135/810	-	-	**0.947** (-)
Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen (2017)	45/90	0.950 (0.060)	0.930 (0.070)	0.930 (0.040)
Tareef, Song, Huang, Wang, Feng, Chen & Cai (2017)	45/90	0.940 (0.060)	0.950 (0.060)	0.940 (0.040)
Tareef et al. (2018)	45/900	0.906 (0.068)	0.950 (0.051)	0.925 (0.041)
Zhang et al. (2019)	45/900	0.902 (-)	0.954 (-)	0.931 (-)
CiscFC method in Chapter 3	45/900	0.927 (0.095)	0.939 (0.090)	0.938 (0.040)
Standard EGS	45/900	0.724 (0.116)	0.866 (0.095)	0.839 (0.067)
Proposed CircEGS	45/900	0.782 (0.103)	**0.955** (0.091)	0.897 (0.064)

Bold values represent the best performance per measure.
All results are reported in *mean (std)* format.
'-' indicates unavailable values.

Table 4.4: Hausdorff distance evaluation.

Method	**H**
Standard EGS	2.019 (0.539)
Proposed CircEGS	**1.628** (0.458)

Bold values represent the best performance per measure.
All results are reported in *mean (std)* format.

4.7.2 Border precision evaluation

H_t metric was employed to evaluate the border precision of the segmented nuclei, where t is the precision tolerance in pixels. Figure 4.3 shows the plot of H_t measure with tolerance t in the range [0, 10] pixels for EGS and CircEGS techniques. It can be observed from Figure 4.3 that for the whole range of t, the H_t metric

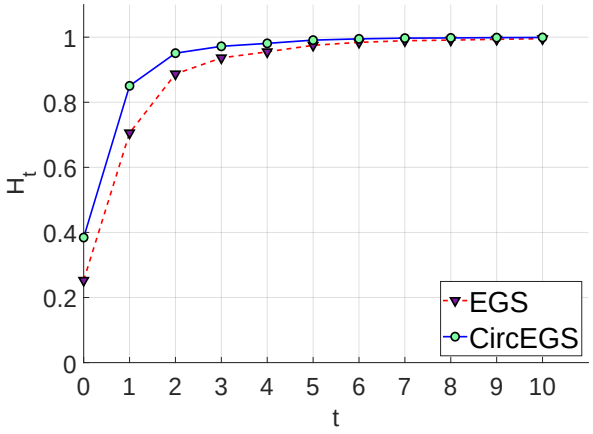


Figure 4.3: Plot of H_t metric for EGS and CircEGS techniques.

for CircEGS is better than standard EGS method. H_t measures for both EGS and CircEGS techniques are stabilized for t over 4 pixels.

Table 4.5 presents the values of H_t metric for tolerance thresholds 1, 2, ..., 5 pixels. With 1 and 2 pixels tolerances, H_1 and H_2 values using CircEGS and EGS for the ISBI test set are 0.850/0.706 and 0.951/0.887, respectively. Overall, the H_t values for CircEGS are higher, particularly for low values of t.

Table 4.5: H_t metric with tolerance threshold t set to 1, 2, ..., 5 (pixels).

Method	$\mathbf{H_1}$	$\mathbf{H_2}$	$\mathbf{H_3}$	$\mathbf{H_4}$	$\mathbf{H_5}$
Standard EGS	0.706	0.887	0.936	0.955	0.975
Proposed CircEGS	**0.850**	**0.951**	**0.972**	**0.981**	**0.991**

Bold values represent the best performance per measure.

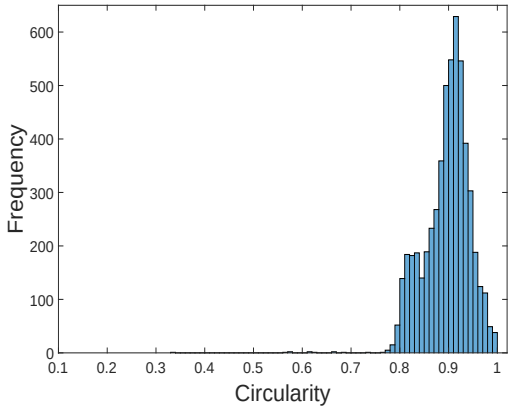
4.8 Discussion

This chapter presented a novel technique to incorporate a circular shape prior in efficient graph based segmentation. Superior DSC, pixel-based precision and recall, and object-based precision and recall, indicate that the proposed technique with circular shape prior, can segment nucleus more accurately than the standard one. The Hausdorff distance for CircEGS was lower (which is preferable) and H_t metric analysis also indicates that the proposed CircEGS produced better nucleus boundaries. Overall nucleus segmentation performance of the proposed CircEGS technique has improved compared to the standard EGS method. Performance of the proposed new technique is also comparable with the state-of-the art techniques.

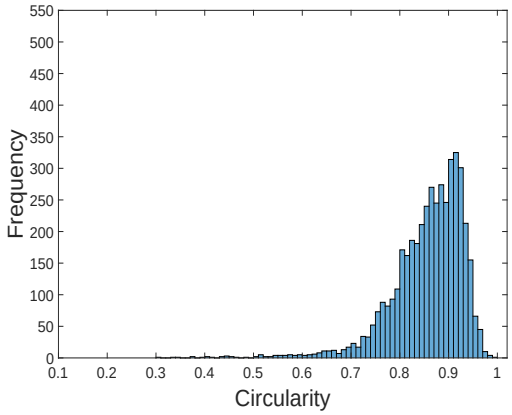
An in-depth analysis of circularity (ψ) was performed, to better understand the influence of shape constrain on the nuclei segmentation outcome. By analyzing circularity distribution of ground truth nuclei from ISBI train and test images (see Section 3.10 and Figure 3.13 for details), circularity threshold was selected as 0.8 for further analysis.

Figure 4.4 represents the circularity distribution of segmented nuclei using standard EGS and proposed CircEGS techniques. In case of EGS segmentation, 80.98% segmentation outputs were correctly detected, and among these, 82.8% (3620 out of 4373) detections had $\psi \geq 0.8$. For CircEGS segmentation, 86% of nucleus were correctly detected, where circularity of 96.8% (4496 out of 4644) were at least 0.8. From the histograms of Figure 4.4 one can see that the histogram for CircEGS outcome (Figure 4.4c) resembles the ground truth histogram (Figure 4.4a) better, than that for EGS method (Figure 4.4b).

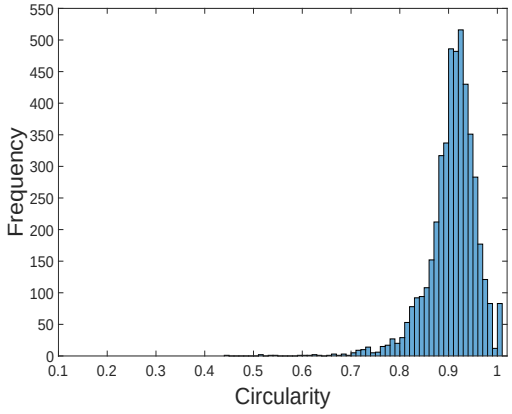
The influence of circularity on nucleus segmentation evaluation measures (DSC, PR_{pix}, RC_{pix}, and H) can be observed from Table 4.6, where comparison is presented between EGS and the proposed CircEGS techniques, on the basis of ψ below and above 0.8. When the circularity is above 0.8, elevated results can be observed for all segmentation measures, for both EGS and CircEGS techniques.



(a)



(b)



(c)

Figure 4.4: Circularity histograms of: (a) ground truth, and nuclei segmented using (b) standard EGS and (c) CircEGS techniques.

Table 4.6: Influence of circularity on nucleus segmentation measures, in terms of DSC, pixel based precision (PR_{pix}) and recall (RC_{pix}), and Hausdorff distance (H).

	EGS		**CircEGS**	
Measures	$\psi < 0.8$	$\psi \geq 0.8$	$\psi < 0.8$	$\psi \geq 0.8$
DSC	0.796	0.848	0.797	**0.900**
PR_{pix}	0.668	0.774	0.592	**0.811**
RC_{pix}	0.854	0.886	0.960	**0.962**
H	2.471	1.924	2.827	**1.589**

Bold values represent the best performance per measure.

Therefore, circularity measure and segmentation performance for cervical nucleus are related. Overall, the inclusion of circular shape prior in EGS method has improved both segmentation and detection performance.

Examples of nuclei boundaries segmented by the standard EGS (Felzenszwalb & Huttenlocher 2004) and the proposed CircEGS method, are presented in Figure 4.5 for qualitative evaluation and visual comparison. Visual inspection indicates that the proposed CircEGS approach outperforms the standard EGS technique (see Figure 4.5c, 4.5f, 4.5i). In cases like Figure 4.5l, the proposed algorithm could detect and segment some nucleus, where the standard EGS technique failed. Finally, from both quantitative and qualitative evaluations, it can be concluded that CircEGS detected and segmented nuclei more precisely compared to the standard EGS method.

The proposed technique has some limitations that need to be addressed. In this study, the value of parameter λ (see Equation (4.3) in Section 4.2.2), which controls the relative importance between area term τ and circular shape prior ψ, and parameter k - controlling the granularity of segmentation (see Equation (4.2)), were found using a small set of 45 images as 0.65 and 150, respectively. These values of λ and k, were used to evaluate the proposed CircEGS technique with 900 images from the test set. These values did not perform well, where nuclei are in close proximity (see Figure 4.6h, 4.6l), in the presence of poor contrast between nucleus and cytoplasm due to cell overlapping (see Figure 4.6i, 4.6j) and uneven staining (see Figure 4.6g, 4.6k).

4.9 Conclusion

This chapter presented a novel graph based segmentation method with shape prior, to segment nuclei from overlapping cervical cytology images. A new weighted circular shape guidance was incorporated in the merging predicate of the standard EGS method. The proposed algorithm used both intensity and shape information while merging regions; therefore it produced nucleus segmentation boundaries with better circularity than the standard technique. Experimentation of the proposed method was conducted on ISBI 2014 challenge dataset, and performance was evaluated qualitatively and quantitatively in terms of: Dice similarity coefficient, pixel and object-based precision and recall, circularity, Hausdorff distance, and H_t metric. Experimental results and visual inspection indicate that the proposed CircEGS can produce better segmentation of cervical nuclei, when compared with the standard EGS method.

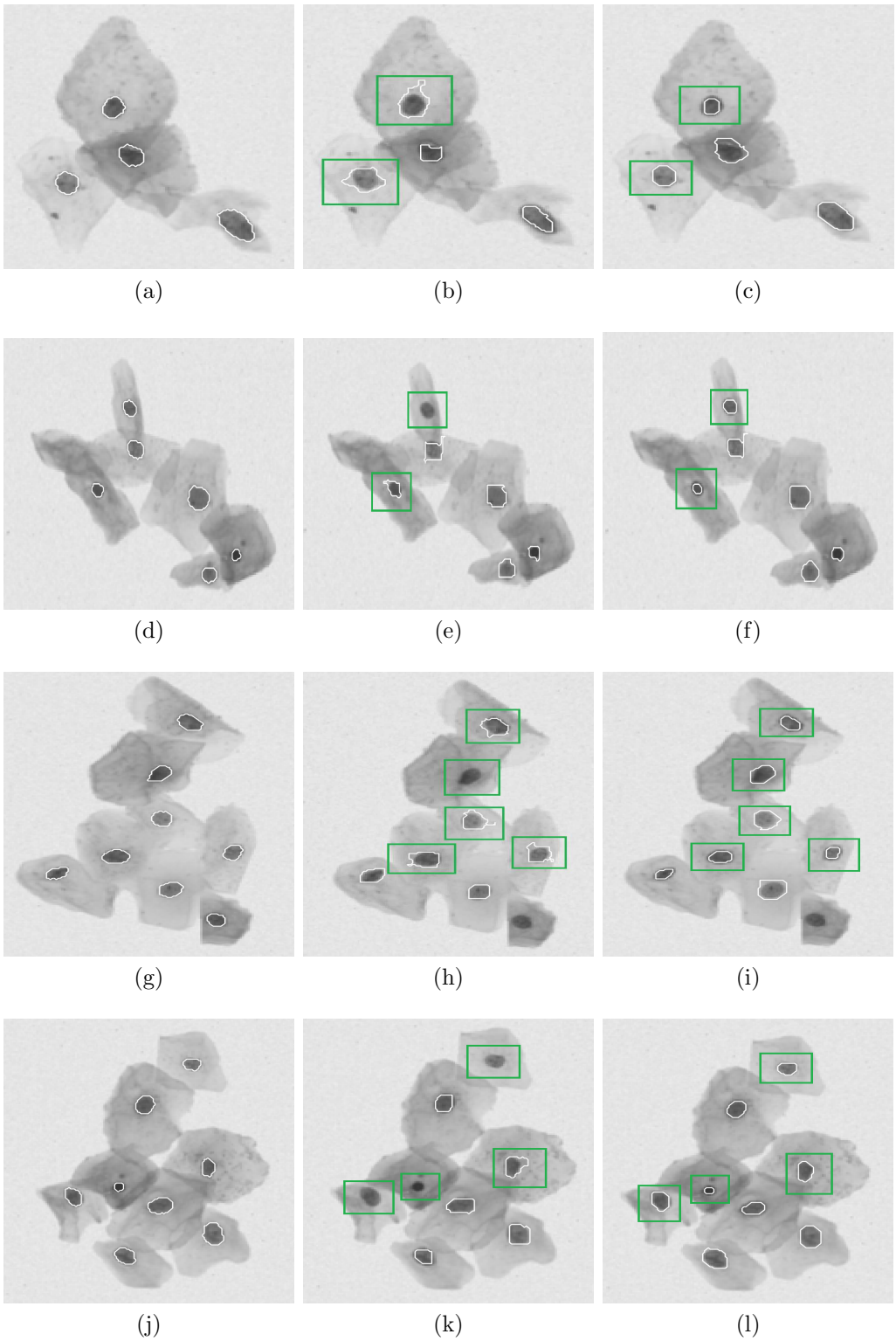


Figure 4.5: Nucleus segmentation examples. First column presents original images marked with nucleus ground truth, second column shows the segmented boundaries using the EGS method, and third column contains the segmentation boundaries using CircEGS technique. Differences between CircEGS and EGS techniques are marked with green rectangles.

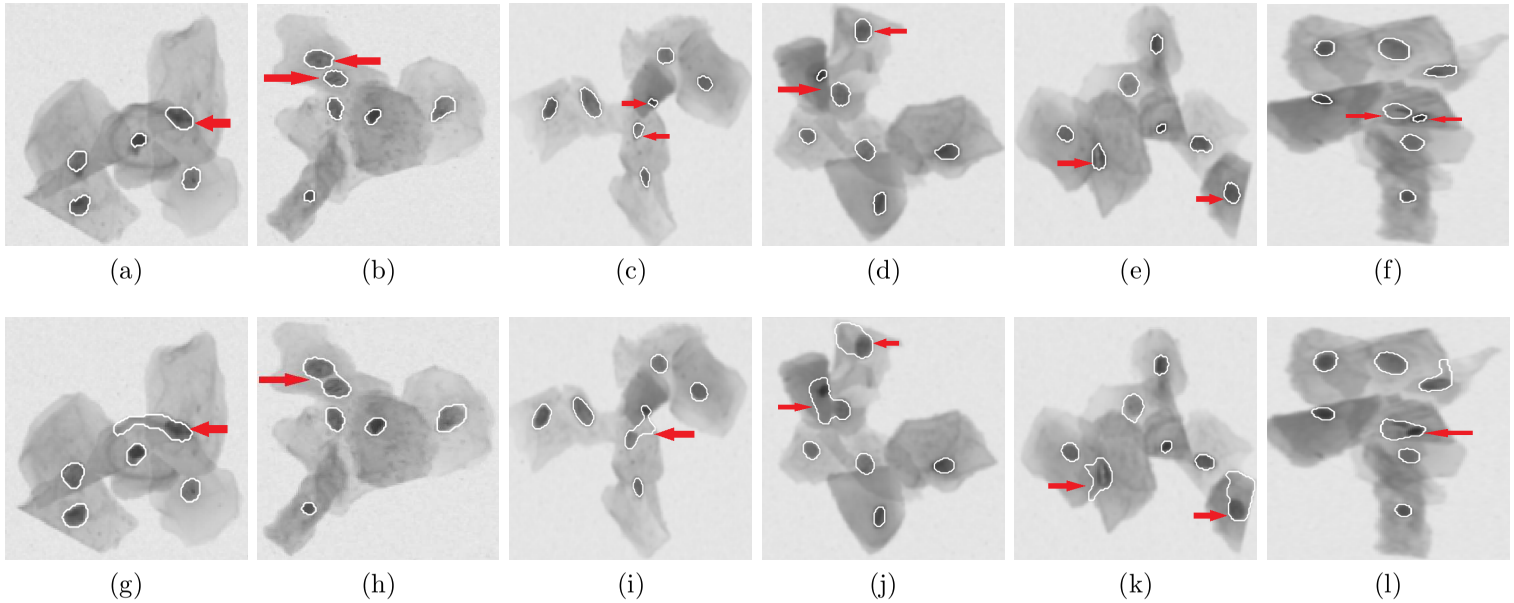


Figure 4.6: Examples of nucleus segmentation failures: top row shows nucleus ground truth boundaries, and bottom row presents corresponding CircEGS technique segmented nucleus boundaries. Failed segmentation cases and nucleus in ground truth, are indicated with red arrows.

Chapter 5

Superpixel merging with gradient guidance

Region merging is one of the basic approaches in the field of image segmentation. In pixel level representation of image, region merging based segmentation techniques are affected by noise. To overcome this limitation, a superpixel based region merging technique is proposed to find nucleus accurately from cervical cell images. In this chapter, a novel region merging criteria with pairwise contrast and gradient boundary is introduced and a superpixel merging framework is presented. Section 5.3 presents step-wise detail of the proposed methodology of superpixel merging, and also introduces superpixel merging criteria. Section 5.4, Section 5.5, and Section 5.6, presents evaluation metrics, experimental set-up, and tuning of required threshold values for parameters, respectively. Experimental results and comparison with the recent state-of-the-art cervical nucleus segmentation techniques, are reported in Section 5.7. Section 5.8 provides a formal discussion and use of an alternative superpixel generation technique, for the proposed framework. Summary of this chapter is provided in Section 5.9.

5.1 Introduction

Accurate detection and segmentation of cell nuclei is the most important step towards Malignancy Associated Change (MAC) analysis and computer-aided analysis of cervical cells. Researchers used a wide range of methods to find nucleus accurately. While segmenting cell images, many of the methods used cluster of pixels or superpixels to represent the underlying image instead of pixel-grid

based presentation, to reduce the effect of noise in segmentation outcome (Ushizima et al. 2015, Nosrati & Hamarneh 2015*b*, Lu et al. 2015, Lee & Kim 2016, Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017).

SLIC (Achanta et al. 2012) is one of the most popular techniques to generate regular and compact superpixels, and also been used to segment cervical nuclei in (Lee & Kim 2016, Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017). Preferred average size of superpixels need to be known before running the SLIC method. However, size of cervical cell nuclei varies over a wide range. Depending on the grade of abnormality, nucleus enlargement can be even six times of the normal size of nucleus (Nayar & Wilbur 2015). Cell overlapping and uneven staining in cervical images cause intensity inhomogeneity and poor contrast. Hence, there is strong possibility of over-segmentation or under-segmentation of nucleus with a fixed preferred size of SLIC superpixel, and nuclei segmentation task become more challenging in the aforementioned situations. In (Lee & Kim 2016), a local thresholding technique (Phansalkar et al. 2011) was used to segment nuclei candidates from over-segmented SLIC superpixels. However, four fixed weighting variables used in local thresholding technique, make it difficult to cope with the challenge of varying nucleus size. In (Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen 2017), the difficulty imposed by over-segmentation of SLIC superpixels was mentioned, and they tried to handle it using several pre-processing techniques and a feature vector composed of shape, texture and boundary clues for superpixel classification, however under-segmentation problem was not considered.

In this chapter, a superpixel merging framework with a novel region merging criteria is proposed to segment nuclei from overlapping cervical cytology images. SLIC technique is used to over-segment images in a way that nucleus is divided into tiny superpixel regions, thus there is no issue with under-segmentation of nucleus. Then, the region merging process is initiated from detected nuclei markers and guided by pairwise regional contrast and image gradient boundary analysis.

5.2 Dataset

ISBI 2014 challenge dataset (described in Section 3.3) is used to evaluate the proposed segmentation framework. There are 945 cervical smear images of 512×512 pixels, and all are 8 bit grayscale images. The training set of 45 images is used for tuning parameter values and the test set of 900 images is used for evaluation.

Training images had 270 and test images contained 5400 nucleus in total.

5.3 Methodology

The proposed nucleus segmentation framework has five key steps: nuclei marker detection, gradient boundary generation, superpixel generation, superpixel merging and shape regularization. Figure 5.1 presents a flowchart of the proposed methodology. In nuclei marker detection step, markers or seed points for nuclei

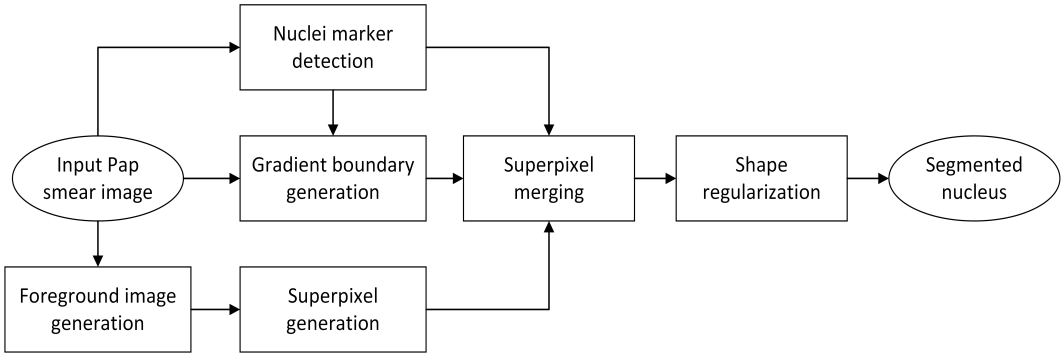


Figure 5.1: The flowchart of the proposed nucleus segmentation framework.

were obtained from original images. Gradient boundaries were generated using nuclei markers and gradient magnitude of Pap smear images. Superpixel images were generated using appropriate superpixel generation technique on foreground images. Then, superpixels were merged to generate segmented nucleus, using nuclei markers, pairwise regional contrast, and gradient guidance. Finally, contours of the segmented nucleus were regularized or smoothed using a level set technique. All the key steps of the proposed framework are detailed in this section.

5.3.1 Nuclei marker detection

Detection of nuclei markers from original grayscale images, is the first step of the proposed superpixel merging framework. Nuclei markers were detected by combining bit plane slicing and maximally stable extremal regions (MSER) (Matas et al. 2004) techniques.

In n bit per pixel binary representation of images, pixel values ranges in $[0, 2^n]$ and contain n number of 1-bit planes. Plane 1 contains the lowest order bit of all pixels in the image, similarly plane n contains the highest order bit of all pixels. Therefore, a 8-bit grayscale image consists of 2^8 or 256 color levels - from 0 (black)

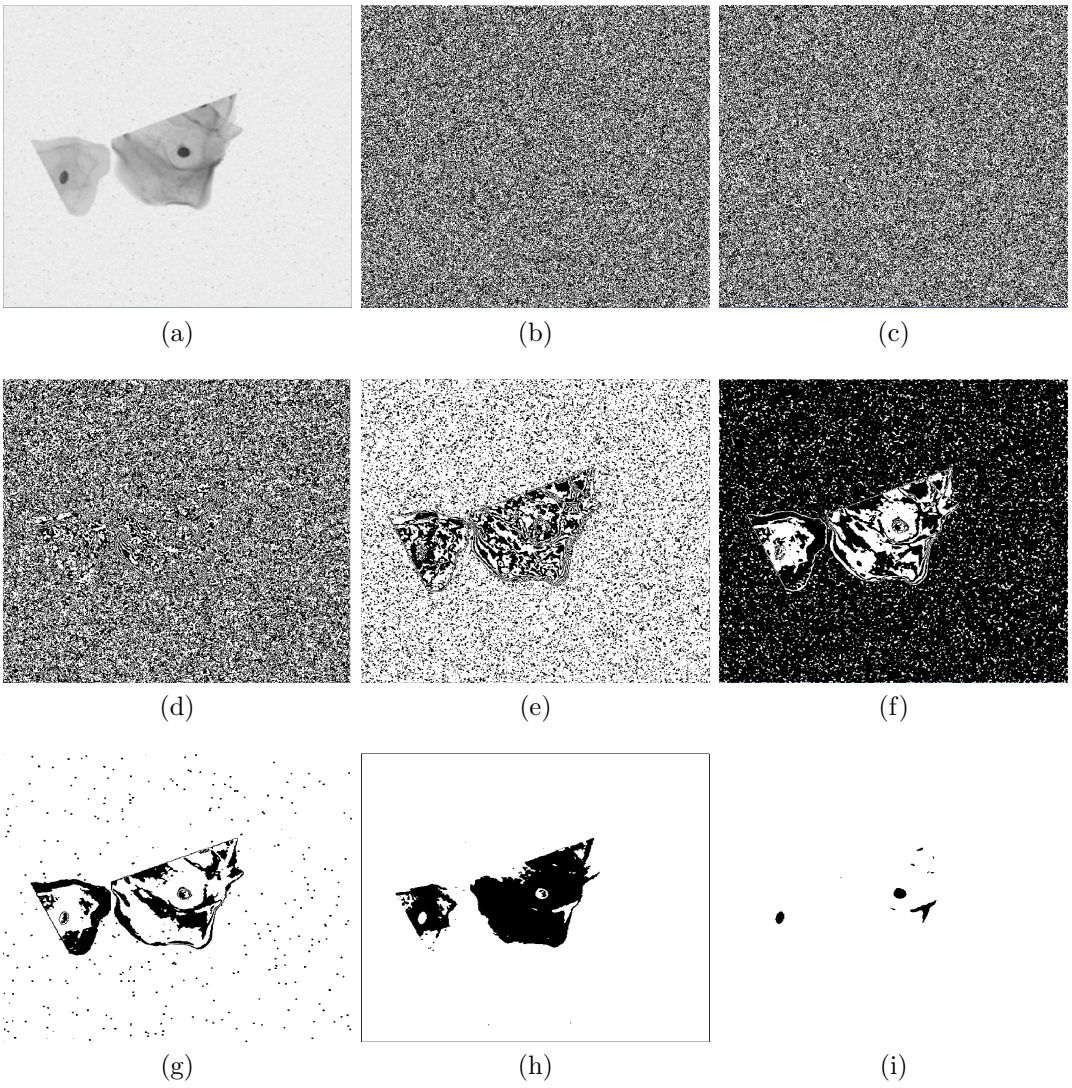


Figure 5.2: Visual representation of (a) a 8 bit Pap smear image, and its 8 bit planes as: (b) bit 1 plane, (c) bit 2 plane, (d) bit 3 plane, (e) bit 4 plane, (f) bit 5 plane, (g) bit 6 plane, (h) bit 7 plane, and (i) bit 8 plane.

to 255 (white), and each pixel can be represented with eight 1-bit planes (Shih 2010). Grayscale images can be sliced from bit 1 plane or Least Significant Bit (LSB) plane to bit 8 plane or Most Significant Bit (MSB) plane. Most image information resides in more significant bits, and less significant bits contain noise and some finer details (Shih 2010). Figure 5.2 presents an example of bit plane slicing of a cervical cell image. In this study, MSB planes were considered to contain candidates for nuclei markers. MSER technique finds maximally stable connected components from image (Matas et al. 2004). Outputs from MSER algorithm were also considered as marker candidates.

After finding initial marker candidates, three morphological operations: re-

moving small regions, filling holes and opening, were applied on both outputs of MSB plane and MSER techniques to remove some of the false findings. Then initial candidates were filtered using area, circularity, eccentricity, solidity, and maximum intensity thresholds. These threshold values were found empirically and are reported in Section 5.6. Finally, primary nucleus markers were found by combining filtered initial outputs from MSB plane and MSER algorithm.

Primary marker candidates were further filtered to remove false detections using a technique from (Phoulady et al. 2017). Candidate regions were dilated to find the outer boundary. Regions were accepted, if the mean intensity difference between the outer boundary and the region was greater than a threshold d (see Section 5.6 for the value of d). Centroids of these final candidates were used as nuclei markers during superpixel region merging process. The whole process of nucleus marker detection is summarized with an example in Figure 5.3.

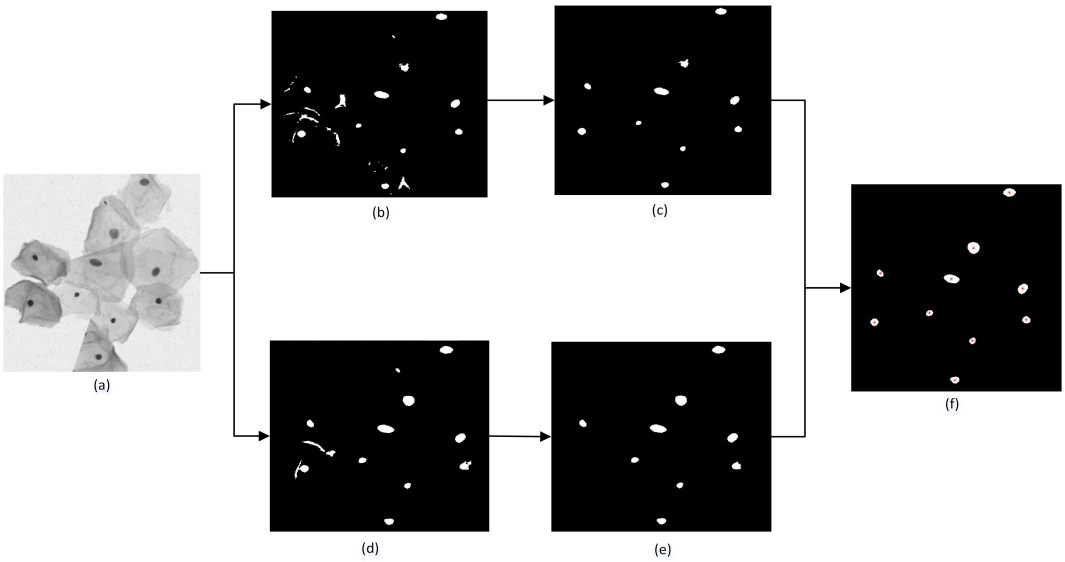


Figure 5.3: An example of nucleus marker detection process: (a) original image, (b) MSB plane with marker candidates, (c) MSB plane after applying morphological operation and filtering, (d) MSER output with nuclei marker candidates, (e) MSER output after applying morphological operation and filtering, (f) centroids of final nuclei candidates - found by combining outputs in (c) and (e), are marked with red stars and used as nuclei markers in the proposed framework.

5.3.2 Superpixel merging criteria

5.3.2.1 Pairwise regional contrast

A measure of pairwise contrast between two regions is introduced in this study, to control the merging process of the proposed framework. Pairwise regional contrast considers region size in addition to similarity between candidate regions intensity levels. Let R_i and R_j be two candidate regions, then pairwise contrast C is defined as:

$$C = abs\left(\mu(R_i) - \mu(R_j)\right) \cdot k, \tag{5.1}$$

where $\mu(.)$ refers to the mean intensity of the corresponding region, abs(.) returns the absolute value, and the contrast coefficient k is defined as:

$$k = \frac{|R_i| \cdot |R_j|}{|R_i| + |R_j|}, \tag{5.2}$$

where $|.|$ indicates the number of pixels in the corresponding region. This contrast criteria can be used to control superpixel merging in the proposed framework, by finding an appropriate threshold for the pairwise contrast value.

5.3.2.2 Gradient boundary

Image gradient boundaries are found from gradient magnitude of grayscale Wiener filtered images. Considering θ° equal angle interval, N intensity radial profiles are drawn from nucleus markers on gradient magnitude image. Maximum intensity points (closest to the nucleus markers) on the radial profiles, are selected as the gradient boundary points. Gradient boundary generation process for a selected nucleus marker is illustrated in Figure 5.4, which shows a nucleus marker, intensity profiles, and gradient boundary points. Finally, gradient boundary G_b is found by connecting N gradient boundary points. These boundaries provide guidance during the proposed superpixel merging process to find nuclei contours.

5.3.3 Superpixel generation

Before superpixel generation process starts, foreground image is generated. Foreground image is found using the morphological reconstruction based technique - described in Section 3.4.1.2. Then, the foreground image is over-segmented to

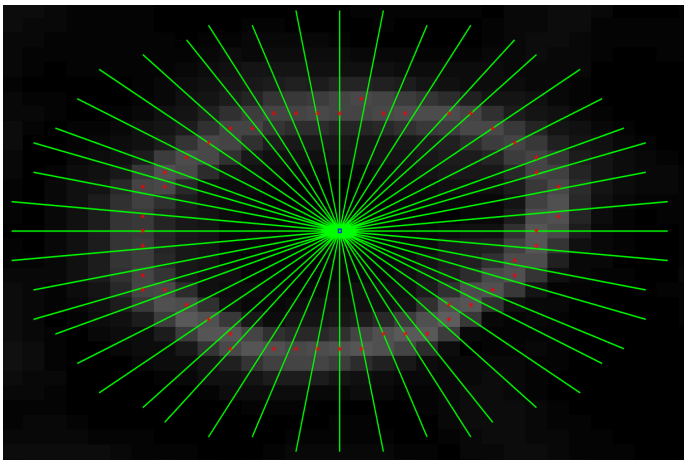


Figure 5.4: An example of gradient boundary generation, where blue square in the center indicates nucleus marker, green lines represent intensity profiles, and red stars stand for gradient boundary points.

generate the superpixel image. The regions in the superpixel image will be called superpixels in this study.

5.3.3.1 Simple linear iterative clustering (SLIC) superpixels

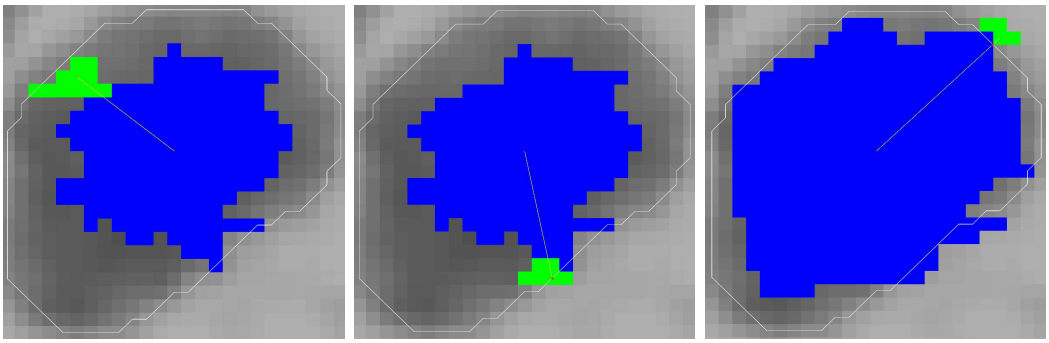
This technique is an adaptation of the popular k-means clustering for superpixel generation (Achanta et al. 2012), with low computational complexity and some control over superpixel size and compactness. The desired number of superpixels is the only parameter required for the SLIC algorithm to generate approximately equal sized superpixels. Section 2.1.2.3 provides an overview of SLIC technique. In this chapter, foreground images are over-segmented using SLIC, so that approximate average superpixel size remains fixed.

5.3.4 A novel superpixel merging technique with pairwise regional contrast and gradient boundary

Let the parent superpixel be the superpixel that contains the nucleus marker. The algorithm for nucleus segmentation is an iterative approach. It starts from the parent superpixel, which is dilated repeatedly to identify neighboring superpixels, which are then tested for merging condition.

For all mergings of neighboring superpixels with the parent, pairwise regional contrast criterion need to be satisfied first. Superpixels are considered for merging, if their pairwise regional contrast C does not exceed a certain threshold value.

Then, the location of neighbors with respect to the gradient boundary is determined. The neighboring superpixel will merge with the parent when the neighbor superpixel centroid is located inside (see Figure 5.5a) or on top (see Figure 5.5b) of the gradient boundary. If the neighboring superpixel centroid is outside of the gradient boundary (see Figure 5.5c), the distance between these two is checked. Maximum Euclidean distance allowed for superpixel merging is 2 pixels (found empirically from the training set, see Section 5.6 for details). The newly merged



(a) (b) (c)

Figure 5.5: Test cases for a neighbor superpixel during the proposed merging process. The centroid of a neighbor superpixel (in green) is: (a) inside, (b) on and (c) outside, the gradient boundary (in white).

superpixel becomes the parent superpixel for the next dilation. This dilation and merging process of the algorithm iterates until there are no more superpixels left that satisfy the merging criteria. Note that due to finite number of superpixels, the algorithm always converges. The pseudo-code for this process is presented in Algorithm 2.

This merging algorithm is applied for each nucleus marker in an image. When the merging process is finished for all the detected markers, merged superpixels are returned as the binary masks for the segmented nucleus candidates. Figure 5.6 illustrates the superpixel merging procedure for an individual nucleus.

5.3.5 Shape regularization

The proposed framework utilizes over-segmented superpixel images. Hence, segmented nucleus boundaries are jagged by nature (example can be seen in Figure 5.6f). To smooth the boundaries, distance regularized level set (Li et al. 2010) was applied. Nucleus boundaries found using the proposed technique were used as the initial level set contours.

Algorithm 2 Superpixel merging algorithm

Input: Superpixel image, gradient boundary G_b, pairwise contrast threshold C_t, and a nucleus marker

Output: Binary mask of segmented nucleus

1: Locate the parent superpixel in the input superpixel image that contains the nucleus marker
2: Dilate the parent superpixel using disk structuring element of size 1 and identify the neighboring superpixels
3: Test each neighbor superpixel for merging:
   a) Draw a line L to connect parent and neighbor superpixel centroids
   b) Find intersection point of L and gradient boundary G_b
   c) If there is no intersection point and pairwise contrast $C \leq C_t$, merge the neighbor with the parent superpixel
   d) If there is an intersection point and $C \leq C_t$, then merge with the parent superpixel if:
      i) neighbor superpixel centroid is located on the gradient boundary G_b, or
      ii) neighbor superpixel centroid is not further than 2 pixels (Euclidean distance), from G_b
4: Repeat steps 2 and 3 until all relevant neighbor superpixels are merged
5: Return merged superpixel as nucleus binary mask

5.4 Evaluation metrics

The proposed framework was evaluated quantitatively using the notations and definitions in Section 3.5. All segmented nucleus were checked using ground truths for being correct detections or not. Object based precision (PR_{obj}) and recall (RC_{obj}), pixel based precision (PR_{pix}) and recall (RC_{pix}), and Dice similarity coefficient (DSC), were used for evaluation.

5.5 Experimental setup

Parameters and threshold values required for the proposed superpixel merging framework were found using 45 images from the training set. Detection and segmentation performance of the proposed framework was evaluated using 900 images from the test dataset. The framework was implemented in Matlab R2017b, running on a PC with Intel Core i7-4770 3.40 GHz processor and 8 GB RAM. SLIC superpixel was generated using the Matlab function *superpixels()* from the Image Processing toolbox. Average running time for the proposed framework with SLIC superpixel was approximately 3.3 seconds per cell and 18 seconds per

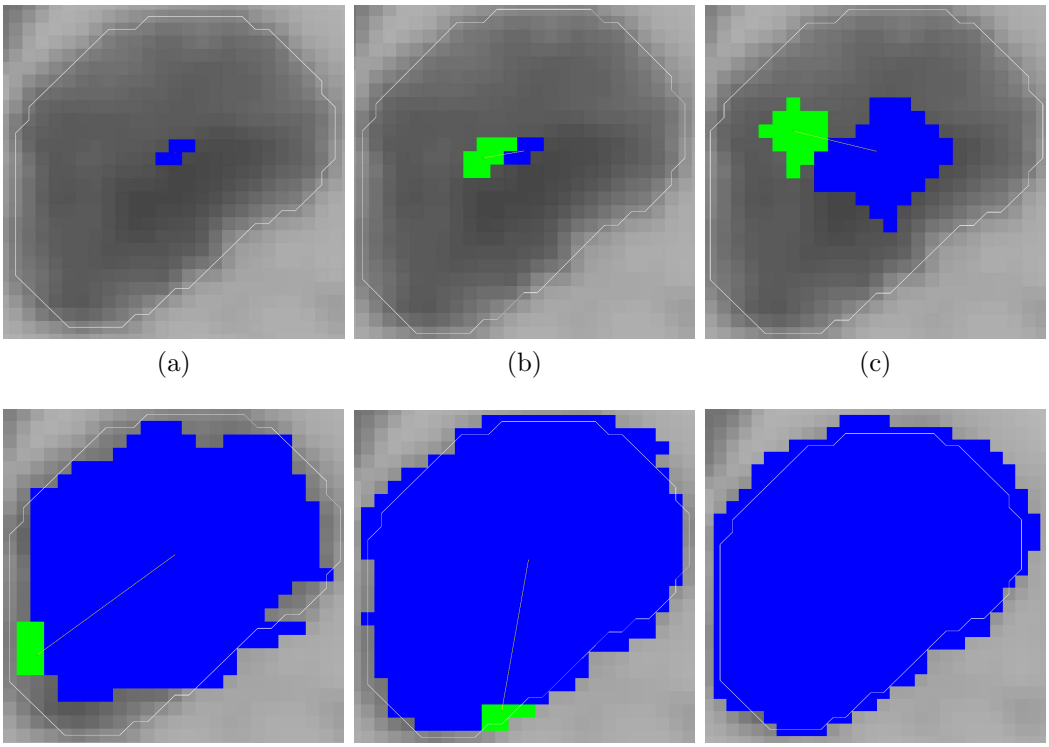


Figure 5.6: Superpixel merging process illustration. (a) Parent superpixel (in blue) containing nucleus marker, (b) - (e) show neighbor superpixels (in green) tested for merging with the parent, and (f) is the final merged superpixel as segmented nucleus.

image. The code used to implement the framework was not optimized.

5.6 Finding thresholds for parameters

In the proposed framework, some parameters require threshold values to be found empirically. Training images were used to determine these values.

While finding nuclei markers in Section 5.3.1, thresholds values are required to filter initial candidates using area, circularity, eccentricity, solidity, and maximum intensity. Very small regions (area less than 65 pixels) were discarded. Candidates with eccentricity greater than 0.9, solidity greater than 0.88, and mean intensity of region greater than 160, were also filtered out. To be candidate for nuclei markers, suitable range of circularity was found as $[0.5, 1.7]$. Mean intensity difference d between inner region and outer boundary was found as 15; that is candidates with d less than 15 were not considered. Table 5.1 summarizes these findings.

Table 5.1: Threshold values to detect candidate nuclei markers.

Parameter	**Value**
Area	Minimum 65
Circularity	Minimum 0.5 and Maximum 1.7
Eccentricity	Maximum 0.9
Solidity	Maximum 0.88
Mean intensity of region	Maximum 160
Mean intensity difference d	Minimum 15

Gradient boundary G_b was generated considering $\theta = 7.5°$ angles and $N = 49$ intensity radial profiles (see Section 5.3.2.2). Foreground images were over-segmented using SLIC technique in a way, such that average superpixel size W is not bigger that 9 pixels.

Threshold for pairwise regional contrast: In the proposed framework, regions are allowed to merge if their pairwise contrast do not exceed a certain threshold value. Note that setting $k = 1$ (see Equation (5.2)) would allow merging of any two regions with sufficiently small average intensity difference regardless of their size (too generous condition, as confirmed empirically). For n merging steps, k can be approximated as follows:

$$k = \frac{|R_i| \cdot |R_j|}{|R_i| + |R_j|} \approx \frac{W \cdot W}{W + W} \leq \frac{2W \cdot W}{2W + W} \leq \frac{3W \cdot W}{3W + W} \leq \ldots \leq \frac{nW \cdot W}{nW + W} = \frac{n\,W^2}{(n+1)\,W} = \frac{n}{n+1}\,W < W. \tag{5.3}$$

This gives an upper estimation for the contrast coefficient which depends only on the segmentation parameter W (average size of superpixels for SLIC technique). Summarizing, for any segmented image with average superpixel size W, if $k = 1$, all superpixels could be merged, provided their intensities are sufficiently similar, and when $k = W$ none of them could be merged despite very small intensity difference for the two superpixels. The correct/optimal threshold value for a data set lies between these two extremes and can now be estimated from Equation (5.1) as follows. Assuming grayscale images, mean intensity difference $abs(\mu(R_i) - \mu(R_j)) < 256$ for any two superpixels R_i and R_j. Then, the maximum range to search for the threshold is $[256, W \cdot 256)$. In this study, for SLIC superpixels $W = 9$, so the search range for the threshold was between $[256, 2304)$. An

analysis was performed on the training set by plotting DSC over the threshold range, and presented in Figure 5.7. The actual value of the threshold for this study was selected as $C_t = 450$, by analyzing the graph.

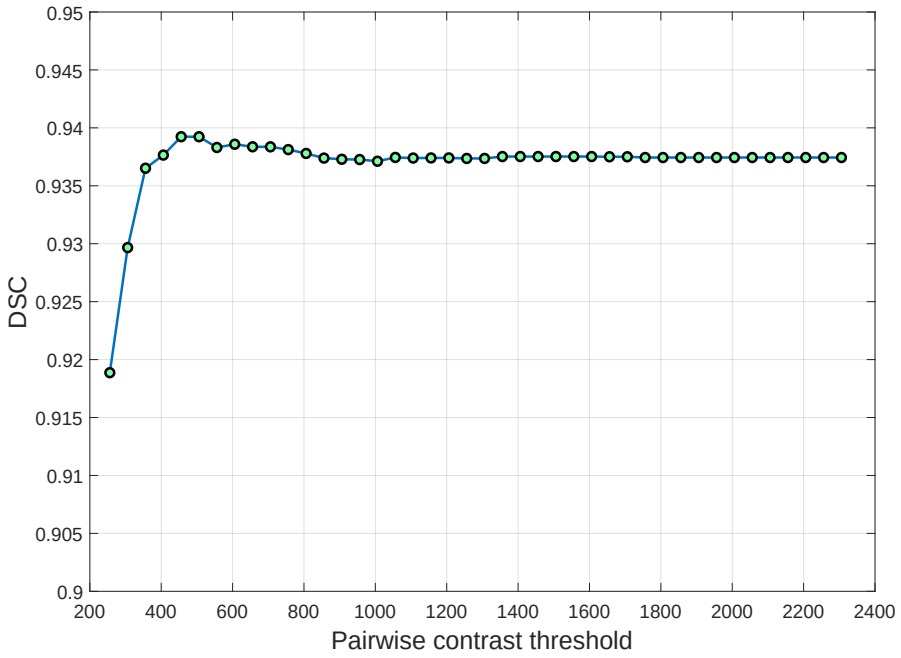


Figure 5.7: Plot of Dice similarity coefficient for pairwise regional contrast threshold in the range $[256, 2306]$ with step of 50, using the proposed superpixel merging framework on ISBI training set.

Maximum distance from gradient boundary: In the proposed superpixel merging algorithm, if the centroid of a neighbor superpixel is located outside the gradient boundary, then the Euclidean distance between the centroid and the boundary is checked. For merging purpose, a maximum allowed distance need to be defined. After finding the threshold value for pairwise regional contrast $C_t = 450$, an analysis was carried out to find the impact of Euclidean distance on nuclei segmentation performance (particularly on DSC), using the training images. This is presented as a graph in Figure 5.8. By inspecting the graph, it was found that the maximum Euclidean distance 2 pixels is the best choice for $C_t = 450$.

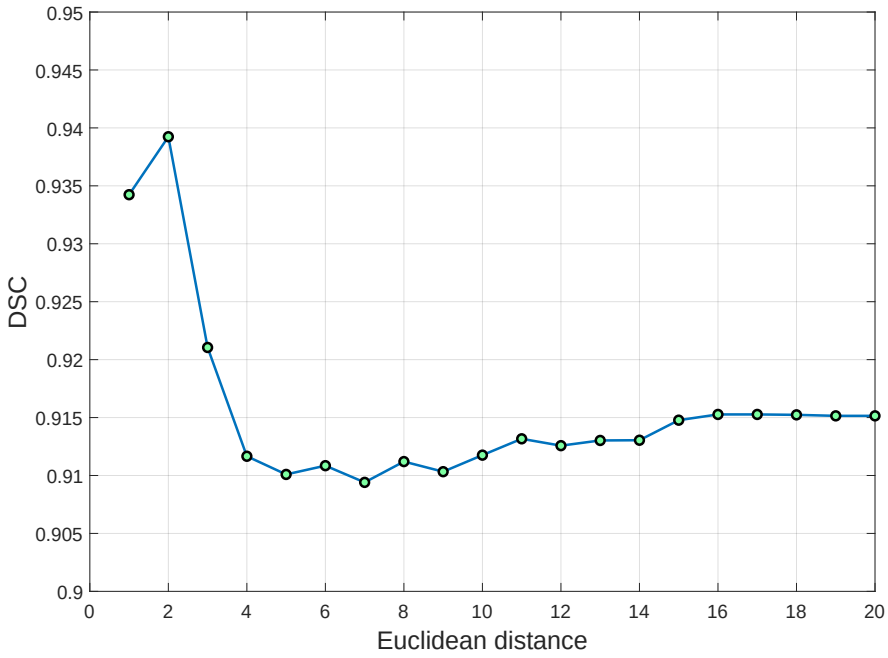


Figure 5.8: Plot of Dice similarity coefficient for Euclidean distance in the range $[1, 20]$, using the proposed superpixel merging framework on ISBI training set.

5.7 Results

The test dataset of 900 images was used to evaluate the proposed superpixel merging framework, in terms of nucleus detection and segmentation performance. Results were compared with the baseline and winning techniques (Ushizima et al. 2015, Nosrati & Hamarneh 2015*b*) in ISBI 2014 challenge (complete results, as reported in (Lu et al. 2017) were used), and other recent state-of-the-art methods that used the same public dataset.

Performance of the proposed framework for detecting nucleus from cervical smear images, was evaluated in terms of object based measures. Nucleus detection results, and comparison with CiscFC (presented in Chapter 3), CircEGS (presented in Chapter 4) and recent techniques are reported in Table 5.2. The proposed framework's performance exceeded the one of CiscFC and CircEGS techniques. Object based precision and recall values for the proposed method, are also comparable to the state-of-the-art methods. The highest object based precision PR_{obj} was 0.994 in (Tareef, Song, Huang, Wang, Feng, Chen & Cai 2017) where only a small subset of 90 test images was used, and the proposed approach achieved $PR_{obj} = 0.980$ on the whole 900 image set. The proposed

approach achieved the object based recall value of $RC_{obj} = 0.944$. While, the highest RC_{obj} (0.971) was achieved in (Zhang et al. 2019) using a deep learning based technique.

Table 5.2: Object based evaluation of the proposed superpixel merging framework and comparison with the state-of-the-art methods.

Method	**train/test**	$\mathbf{PR_{obj}}$	$\mathbf{RC_{obj}}$
Lu et al. (2015)	45/900	0.730	0.850
Ushizima et al. method in Lu et al. (2017)	45/810	0.959	0.895
Nosrati & Hamarneh method in Lu et al. (2017)	45/810	0.903	0.893
Baseline method in Lu et al. (2017)	45/810	0.977	0.883
Phoulady et al. (2017)	135/810	0.961	0.933
Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen (2017)	45/90	0.990	0.940
Tareef, Song, Huang, Wang, Feng, Chen & Cai (2017)	45/90	**0.994**	0.911
Tareef et al. (2018)	45/900	0.983	0.959
Zhang et al. (2019)	45/900	0.990	**0.971**
CiscFC method in Chapter 3	45/900	0.968	0.882
CircEGS method in Chapter 4	45/900	0.868	0.861
Proposed framework with SLIC	45/900	0.980	0.944

Bold values represent the best performance per measure.

Evaluation measures to determine segmentation accuracies for the proposed framework and comparison with CiscFC (proposed in Chapter 3), CircEGS (proposed in Chapter 4) and other recent techniques, are reported in Table 5.3. The proposed technique scored the highest DSC (0.948) and pixel based recall (0.966). This indicates that the proposed superpixel merging framework missed less nucleus pixels and nucleus segmentation is more accurate compared to CiscFC, CircEGS and other state-of-the-art techniques. Pixel-based precision of the proposed framework is lower than CiscFC technique (0.908/0.927), nevertheless this measure has significantly improved compared to CircEGS technique (0.908/0.782). The highest pixel-based precision value (0.968) was achieved for the technique proposed by Ushizima et al. reported in (Lu et al. 2017), however their RC_{pix} (0.871) was the lowest of all the methods in comparison.

Table 5.3: Evaluation of segmentation accuracy measures for the proposed superpixel merging framework, in comparison to the state-of-the-art techniques.

Method	**train/test**	$\mathbf{PR_{pix}}$	$\mathbf{RC_{pix}}$	**DSC**
Lu et al. (2015)	45/900	0.960 (0.060)	0.900 (0.080)	0.920 (0.050)
Ushizima et al. method in Lu et al. (2017)	45/810	**0.968** (0.055)	0.871 (0.069)	0.914 (0.039)
Nosrati & Hamarneh method in Lu et al. (2017)	45/810	0.901 (0.097)	0.916 (0.093)	0.900 (0.053)
Baseline method in Lu et al. (2017)	45/810	0.942 (0.078)	0.912 (0.081)	0.921 (0.049)
Phoulady et al. (2017)	135/810	-	-	0.947 (-)
Tareef, Song, Cai, Huang, Chang, Wang, Fulham, Feng & Chen (2017)	45/90	0.950 (0.060)	0.930 (0.070)	0.930 (0.040)
Tareef, Song, Huang, Wang, Feng, Chen & Cai (2017)	45/90	0.940 (0.060)	0.950 (0.060)	0.940 (0.040)
Tareef et al. (2018)	45/900	0.906 (0.068)	0.950 (0.051)	0.925 (0.041)
Zhang et al. (2019)	45/900	0.902 (-)	0.954 (-)	0.931 (-)
CiscFC method in Chapter 3	45/900	0.927 (0.095)	0.939 (0.090)	0.938 (0.040)
CircEGS method in Chapter 4	45/900	0.782 (0.103)	0.955 (0.091)	0.897 (0.064)
Proposed framework with SLIC	45/900	0.908 (0.031)	**0.966** (0.026)	**0.948** (0.033)

Bold values represent the best performance per measure.
All results are reported in *mean (std)* format.
'-' indicates unavailable values.

5.8 Discussion

This chapter presented a superpixel merging framework to address the challenging task of cervical cytology nuclei segmentation. A novel superpixel merging criteria with pairwise regional contrast and gradient boundary was introduced. SLIC technique was used to generate superpixels in this study. The segmentation and detection results indicate competitive performance of the proposed method. The proposed approach helped to overcome the problem of defining an optimal region size for generating SLIC superpixels, so that nuclei regions over a dataset are neither over-segmented nor under-segmented.

SLIC is a popular superpixel generation technique due to its speed, memory efficiency, and segmentation performance. However, approximate equal sized su-

perpixels criteria and post-processing step to reassign disjoint pixels to nearby superpixels, may compromise homogeneity of superpixel regions. Consequently, this may affect the accuracy of nucleus segmentation. To overcome the homogeneity issue, other superpixel generation techniques can be used that considers region homogeneity while segmenting image, for example the Statistical Region Merging (SRM) segmentation technique (Nock & Nielsen 2004). SRM algorithm is based on probability and statistical theory, and generates statistically homogenous superpixels. The proposed merging framework utilizing SRM superpixels is explored in next section.

5.8.1 SRM superpixels

SRM segmentation technique is based on a merging predicate and an order to perform region merging (see Section 2.1.2.2 for an overview of SRM segmentation technique). The parameter Q embodied in the merging predicate definition controls the coarseness of segmentation, with bigger Q value resulting in a finer segmentation. SRM algorithm does not produce equal sized superpixels and the number of superpixels can not be controlled. Here the scale of segmentation can be controlled by tuning the parameter Q. However, finding the best or optimal Q value for superpixel generation to suit the problem at hand is a challenge (Bajger et al. 2010, Bajger et al. 2013). Pairwise regional contrast threshold C_t was found by analyzing superpixel size (see Section 5.6 and Equation (5.3)). Since, size of SRM superpixels can not be approximated beforehand, therefore merging technique from Section 5.3 is not directly applicable on SRM segmentation.

To overcome this issue, SRM superpixel image partition was generated in a novel way by fusing two: coarse and fine, SRM outputs. In this strategy, superpixel merging starts from a big piece of candidate SRM superpixel found using a small value of parameter Q, call it Q_s. Then, the segmentation is updated iteratively by merging smaller SRM superpixels generated using a bigger value of Q, call it Q_b. The process completes when all relevant superpixels are merged.

Finding $\mathbf{Q_s}$ value for SRM: The Q_s needs to be selected carefully for generating coarse SRM superpixels, such that small nucleus candidate superpixels can still be detected. Assuming that the merging predicate of SRM technique is true for two regions R and R' from image I, the predicate can be solved for Q_s,

similarly as in (Bajger et al. 2010):

$$Q_s \leq \frac{g^2}{2T_0^2}\left(\frac{1}{|R|}+\frac{1}{|R'|}\right)\ln\frac{2}{\delta}, \tag{5.4}$$

where g is the number of image intensity levels, $T_0 = |\bar{R} - \bar{R}'|$, with $\bar{R}$ indicating the average intensity across the region R. The value for δ is selected small so that it guarantees high probability of merging relevant regions in the SRM technique (see Nock & Nielsen (2004) for details). An upper bound for Q_s value can be estimated from Equation (5.4), so that two regions with average intensity difference T_0 or smaller would merge.

In the dataset used in this study (detailed in Section 5.2), the minimum size of nucleus for training images was 124 pixels, image size $|I| = 512 \times 512$ and $g = 256$. Following (Nock & Nielsen 2004), $\delta = \frac{1}{6|I|^2}$ was used, and the value of Q_s was selected to limit the superpixel size to 124 or smaller (otherwise some nuclei could be under-segmented). If $|R \cup R'| \approx 124$ and $T_0 = 47$ (found empirically from training images), merging of regions R and R' can be prevented with $Q_s \approx 16$ (found by solving Equation (5.4)). Therefore, the size of nucleus candidate 90 pixels is acceptable as the starting point. To illustrate the merging process, once the size of one region R reaches 90 pixels, a graph was drawn to analyze the behavior of Q_s by plotting Q_s as a function of the size of the other region R' (see Figure 5.9a). By analyzing the graph, $Q_s = 15$ was found as the optimal value to generate coarse SRM superpixels, with sizes not exceeding $125 (= 90 + 35)$ pixels.

Finding Q_b value for SRM: The next step is to select a value for Q_b to produce finer or over-segmented SRM superpixels. In the proposed framework, this fine superpixel image was used to refine the segmentation by merging superpixels, considering pairwise regional contrast criterion. In Section 5.6, the threshold value suitable for pairwise regional contrast technique $C_t = 450$ was found by analyzing SLIC superpixel size. For SLIC technique, the value for C_t was found assuming nucleus size to be on average 9 pixels and not bigger than 36 pixels. However, SRM technique does not guarantee any bound on the size of superpixels. Hence, the value of Q_b to generate over-segmented image was selected by analyzing average and maximum SRM superpixel size, so that the previously validated threshold value C_t can still be used. Applying these findings to the current setting for SRM superpixel generation, and by analyzing Figure 5.9b and 5.9c, it can be seen that the corresponding Q values for these superpixel sizes fall between $[25000, 35000]$. Therefore, $Q_b = 30000$ was selected.

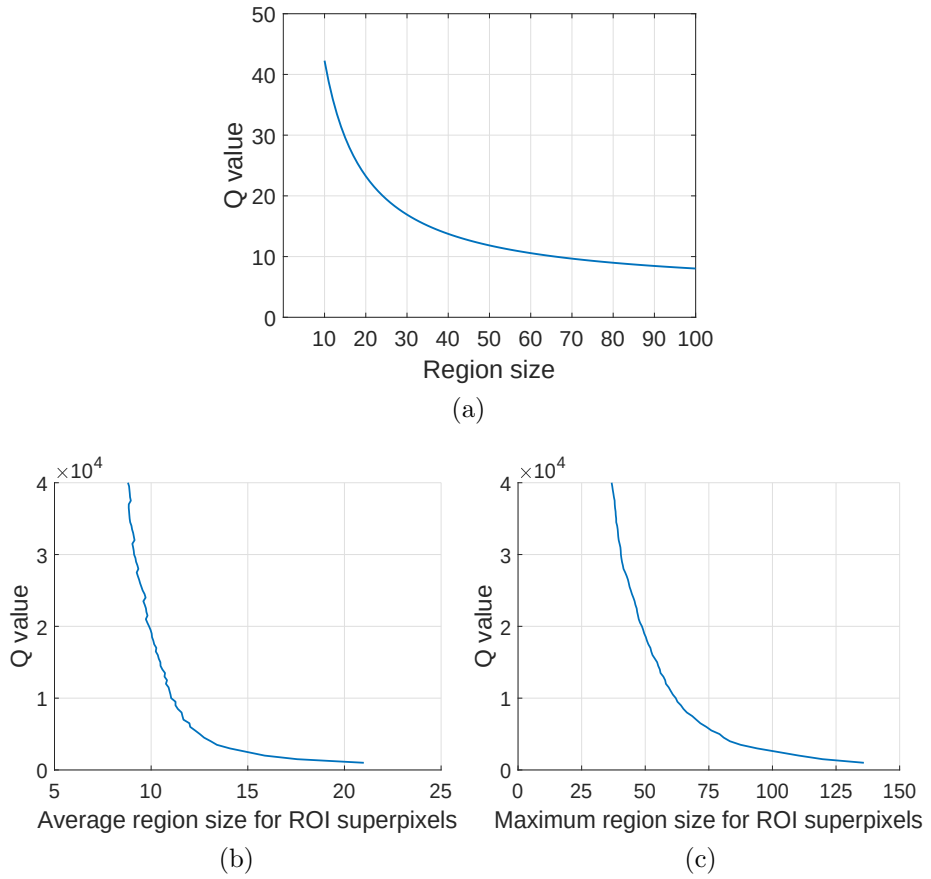


Figure 5.9: Estimation of Q values for SRM segmentation. (a) Finding Q_s for coarse SRM segmentation, and finding Q_b from (b) average and (c) maximum region size analysis for finer SRM segmentation.

Summarizing, the final fused SRM superpixel image was found in three steps. First, coarse SRM segmentation of the input image was generated with $Q_s = 15$, and all superpixels not containing nuclei markers were removed. In second step, a finer SRM segmentation of the same input image was found using $Q_b = 30000$. Finally, in the last step, the coarse superpixel partition found in the first step was fused with the finer partition found in the second step, to get the final segmentation to be used in superpixel merging process.

Implementation and comparison: A Java in-house implementation of the original SRM algorithm was used to generate SRM superpixels. Average computational time for the superpixel merging framework with SRM superpixels was approximately 4.5 seconds per cell and 25 seconds per image, which is little higher than with SLIC superpixel generation. The proposed framework was compared for SLIC and SRM superpixel generation techniques and the results are reported in Table 5.4.

Table 5.4: Comparison of nucleus segmentation and detection evaluation for the proposed superpixel merging framework with SLIC and SRM techniques.

Method	**DSC**	$\mathbf{PR_{pix}}$	$\mathbf{RC_{pix}}$	$\mathbf{PR_{obj}}$	$\mathbf{RC_{obj}}$
Proposed framework with SLIC	0.948 (0.033)	0.908 (0.031)	**0.966** (0.026)	0.980	0.944
Proposed framework with SRM	**0.956** (0.028)	**0.930** (0.024)	0.962 (0.017)	**0.987**	0.944

Bold values represent the best performance per measure.

In terms of nucleus segmentation, the proposed framework with SRM superpixels yielded higher DSC than with SLIC superpixels (0.956/0.948). This DSC value (0.956) is also the highest while compared to other state-of-the-art techniques in Table 5.3. Pixel based precision value PR_{pix} for the proposed framework with SRM, is higher than in SLIC technique (0.930/0.908), and also comparable with other methods from literature. RC_{pix} for SRM is slightly less than that of SLIC (0.962/0.966), and the second highest compared to recent techniques (see Table 5.3).

From the object level perspective, improvement in true positive detection and low occurrences of false positive detections can be observed from Table 5.4. The proposed framework with SRM superpixel segmented 5168 objects, among these 5099 were correctly detected nuclei. The framework with SLIC superpixel found 5077 correctly detected nucleus out of 5163 segmented candidates. More specifically, object level precision PR_{obj} of the proposed framework with SRM was higher than SLIC technique as (0.987/0.980). In addition to that, higher rate of true positive detection was coupled with a low false negative rate of detection 0.056 for SRM superpixels. Object based recall RC_{obj} was same for both superpixel generation techniques as 0.944. While compared to nuclei detection performance of the state-of-the-art techniques (in Table 5.2), results achieved for the proposed framework with SRM technique were comparable.

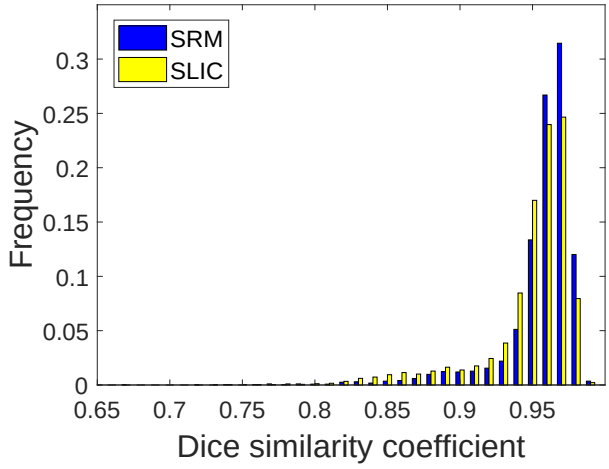
To better understand the impact of the two different superpixel based segmentation techniques on the accuracy of segmentation, an in-depth analysis was performed. Figure 5.10 presents the bar graphs for DSC, PR_{pix} and RC_{pix} using the proposed framework with SRM and SLIC superpixels. For SRM superpixel, around 84% of correct detections have DSC of 0.95 (see Figure 5.10a), and it is 73.8% for SLIC superpixel. In case of pixel level measures for SRM superpixels, PR_{pix} is above or equal to 0.90 for 92% correctly detected nucleus, and precision value stands between 0.8 and 0.9 for only 7.7% correct detections (see

Figure 5.10b). While using SLIC superpixel, 74% of correctly detected nuclei have PR_{pix} at least 0.9, and 25.4% have $0.8 < PR_{pix} < 0.9$. Although proposed framework with SRM superpixel has shown superiority in terms of DSC and PR_{pix}, for SLIC method RC_{pix} was slightly higher. The pixel based recall is 0.98 or above for 47.1% nucleus using SLIC, and it is 18.8% for SRM method generated superpixels. However, presence of $RC_{pix} \geq 0.95$ in 88% cases (see Figure 5.10c) for SRM superpixels resulted in a very small difference in RC_{pix} for SLIC and SRM superpixels. While producing superpixels, SRM segmentation algorithm retains statistical homogeneity. SLIC technique generates superpixels of equal average size, compromising homogeneity in the superpixel regions itself. This fact helped the proposed framework with SRM method to improve segmentation accuracies in terms of DSC and PR_{pix} over SLIC technique.

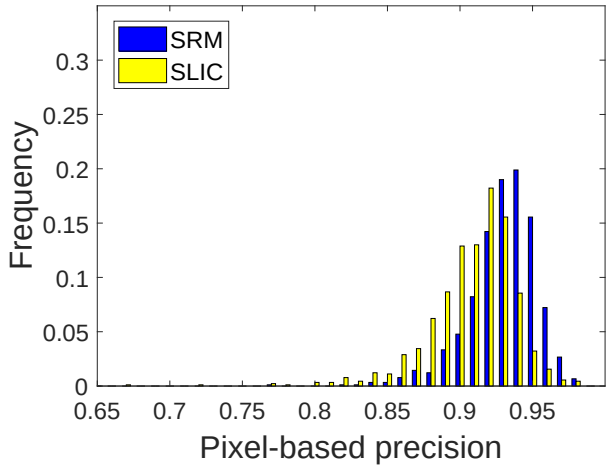
Nucleus segmentation results can be qualitatively assessed by visual inspection. Example images from the test dataset, with ground truth and segmented boundaries using the proposed methodology with SLIC and SRM superpixels, are displayed in Figure 5.11 for qualitative analysis. It can be seen that the proposed superpixel merging framework with SRM superpixel produced precise estimation of nucleus boundaries. The proposed technique with SRM, segmented nucleus more accurately even in cases where mean intensity of cytoplasm was very similar to nucleus due to cell overlapping.

Limitation: The proposed framework with SRM superpixels still has some limitations, despite of superior segmentation results. A training dataset is required to find the threshold value for pairwise regional contrast and parameter Q to control the scale of segmentation while generating superpixels. Estimation for the range of pairwise regional contrast threshold is related to approximate superpixel size (see Section 5.6, Equation (5.3)). However, superpixel size can not be directly approximated for SRM technique. Hence, the merging framework proposed in Section 5.3 can not be directly implemented with SRM superpixels.

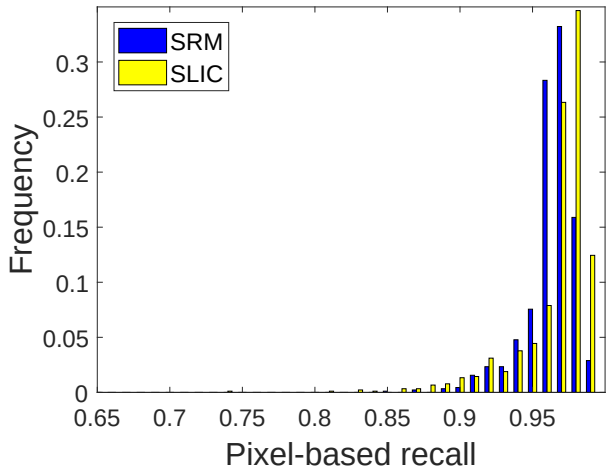
Examples of some failed cases in terms of correctly segmenting nucleus boundaries are shown in Figure 5.12. The main reason of failures of the proposed approach is missing nucleus markers (examples shown in Figure 5.12, where segmented nucleus boundaries are absent). The presence of a marker is critical for the framework because the merging process starts from the superpixel containing nucleus maker. This limitation applies for both SLIC and SRM superpixel generation techniques. Another reason of failure of the proposed approach with SRM superpixels, is the superpixel generation process. The value of Q_s for SRM



(a)



(b)



(c)

Figure 5.10: Plot of (a) Dice similarity coefficient (DSC), (b) pixel-based precision (PR_{pix}) and (c) pixel-based recall (RC_{pix}) - for the proposed framework using SRM and SLIC methods.

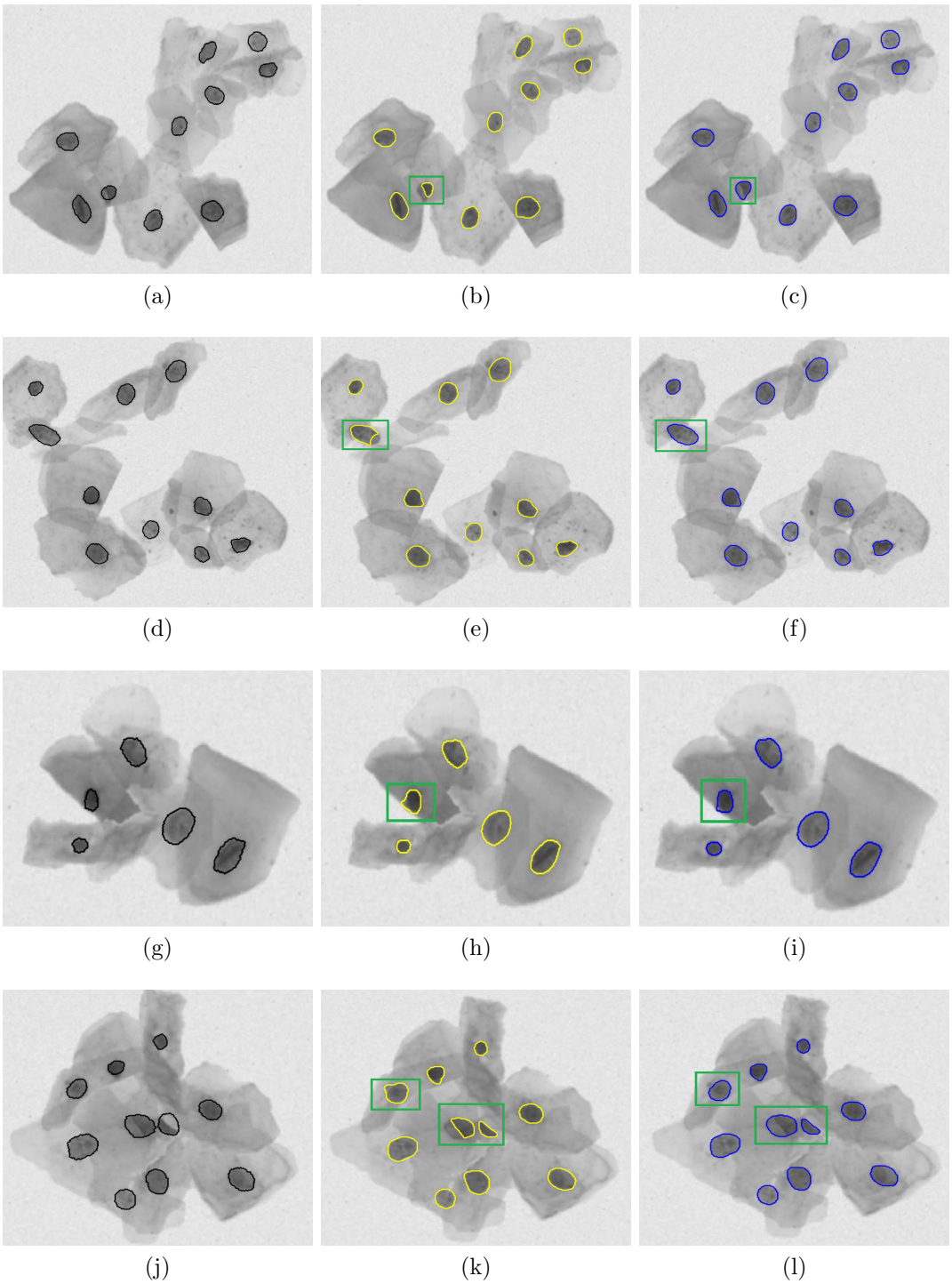


Figure 5.11: Visual comparison of ground truth (in black, left column), and segmented boundaries found by the proposed superpixel merging framework with SLIC (in yellow, middle column) and SRM (in blue, right column) superpixels. Differences in segmented boundaries found with SLIC and SRM superpixel generation techniques are highlighted with green rectangles.

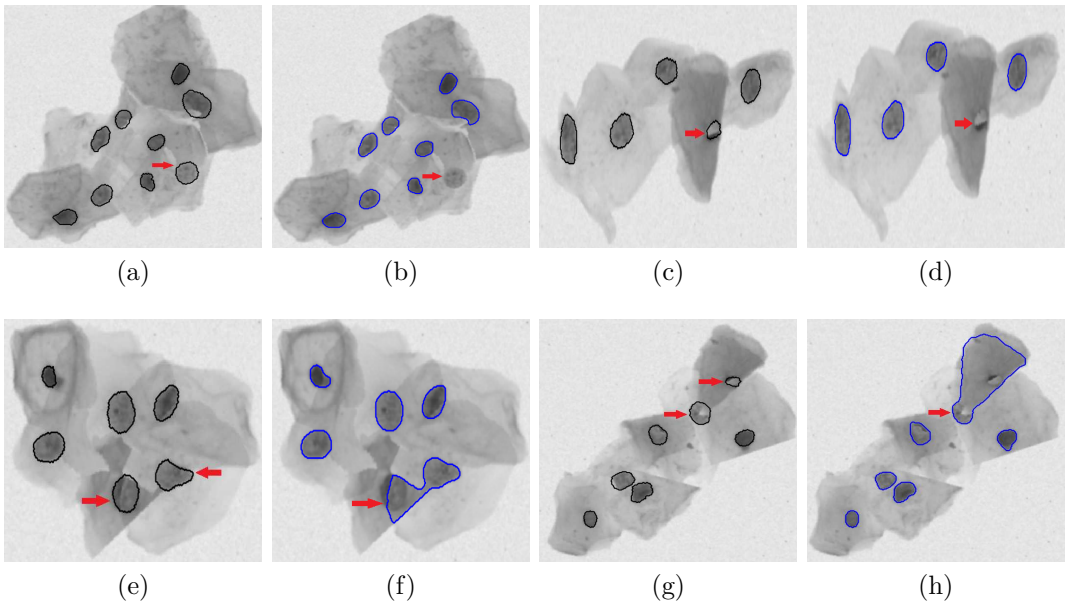


Figure 5.12: Examples of failure cases with proposed framework with SRM superpixels. Top row presents failures due to missed marker detection, and bottom row shows failed cases for superpixel generation. Ground truth boundaries are in black, segmented boundaries are in blue, and differences are indicated with red arrows.

superpixel generation, estimated from Equation (5.4) and analyzing Figure 5.9a, could not always separate candidate nucleus superpixel from the surrounding cytoplasm because of poor contrast due to cell overlapping. Hence, the final fused superpixel image was not appropriate for the proposed approach in those cases and resulted in failure (see bottom row of Figure 5.12).

5.9 Summary and conclusion

This study stands as the first stage for computer-aided analysis of cervical cells. This chapter proposed a superpixel merging framework to address the challenging problem of nucleus segmentation from overlapping cervical cytology images. A novel region merging criteria with pairwise contrast and gradient boundary was introduced in this chapter. Superpixels were generated using SLIC and SRM segmentation techniques. Due to the formulation of the algorithm, SLIC technique may compromise homogeneity in superpixels. Therefore, to alleviate this limitation, SRM segmentation technique was explored for superpixel generation and used with the proposed framework. A novel approach to generate SRM superpixel image was also introduced, to compensate the problem of finding the

optimal value to control the scale of segmentation for SRM.

The performance of the proposed framework was analyzed both qualitatively and quantitatively, using 900 test images from ISBI 2014 dataset. Segmentation and detection accuracies of the proposed framework utilizing SRM superpixels were superior, compared to SLIC technique in terms of DSC, PR_{pix}, PR_{obj}. Also, similar results obtained for RC_{pix} and RC_{obj}. However, average superpixel size can not be approximated for SRM segmentation method, hence similar analysis used for SLIC technique to find the threshold value range for pairwise regional contrast is not applicable for SRM algorithm. The threshold value found for SLIC superpixels was re-used while utilizing SRM approach. Therefore, SLIC superpixels are more appropriate for the proposed superpixel merging framework, and will be used later on. The proposed approach helped to overcome the difficulty to define an optimal region size for SLIC method. The highest DSC and pixel based recall RC_{pix} values indicate that the proposed approach with SLIC superpixels, can produce more accurate segmentation of nuclei than the recently published state-of-the-art techniques.

Chapter 6

Abnormality detection in cervical cells

In Chapters 3, 4 and 5, three segmentation techniques were presented to precisely segment cervical nuclei from overlapping Pap smear images. In this chapter, a framework is proposed to detect abnormality in cervical cells using the techniques developed in previous chapters. Section 6.1 provides an introduction and Section 6.2 describes the image dataset used for evaluation. The methodology used in this chapter is described in Section 6.3. Section 6.4 presents the evaluation metrics used. Experimental results and discussion are presented in Section 6.5, followed by conclusion in Section 6.6.

6.1 Introduction

Cervix tissues undergo dysplastic changes before true development of cancer and these changes refer to the abnormalities in cell morphology and structure. Dysplastic changes in cell are accompanied with different levels of nucleus and cytoplasmic changes and structural disorders. Therefore, morphological features - extracted from nucleus and cytoplasm, play prominent role in diagnostic decisions, and are crucial for computerized cell image analysis. Due to cell overlapping in Pap smear images, cytoplasm boundaries have fuzzy appearance in cell clusters (Plissiti & Nikou 2012*a*). Therefore, correct delineation of cytoplasm boundaries in Pap smear images and extraction of accurate morphological features from cytoplasm, are still challenging and an active field of research (Chankong et al. 2014, Dong et al. 2019, Li et al. 2012). However, nucleus is the most distinguishable region in Pap smear images even with high degree of cell overlap.

A substantial progress has already been achieved in the direction of precise segmentation of cervical nuclei (Bora et al. 2017, Liu et al. 2018). In addition, three novel segmentation techniques were proposed in Chapters 3, 4 and 5, to segment cervical nuclei from overlapping cervical cell images. Based on this scenario, detecting abnormality in Pap smear images - using the features extracted solely from nucleus, seems quite appropriate.

A new cervical cell classification framework - exclusively based on nuclei features, is proposed in this chapter. The proposed framework consists of nucleus segmentation, feature generation and selection, and classification. The three successful segmentation techniques proposed in Chapters 3, 4 and 5, with shape and prior guidance are used for nucleus segmentation. Features extracted from segmented nuclei are used in the process of cervical cell classification to detect abnormality.

6.2 Dataset

The ISBI 2014 dataset used in Chapters 3, 4 and 5 (see Section 3.3 for details) to evaluate the proposed segmentation techniques, is a synthetic dataset. The challenge and complexity encountered when processing synthetic data is different from that of screening real microscopic Pap smear slides. Hence, algorithms developed and validated on synthetic data should always be verified on real data, to identify the strength and robustness of the new techniques in real world situation. The ISBI 2014 dataset contains ground truths for nucleus and cytoplasm, however cell classification labels are not available for this dataset. Therefore, publicly available Pap smear benchmark database, or Herlev Pap smear database, presented in (Jantzen et al. 2005, Jantzen & Dounias 2006), was used in this chapter to verify the segmentation techniques developed in Chapters 3, 4 and 5 and to evaluate the proposed abnormality detection or classification framework. This database is located at `http://mde-lab.aegean.gr/index.php/downloads` for download. Pap smear specimens were prepared using conventional Pap staining technique at Herlev University Hospital, Denmark, and images in this dataset were collected using digital camera and microscope.

There are 917 labeled images in this dataset, acquired at a magnification of 0.201 μm/pixel. Cervical cells were manually classified as normal and abnormal by skilled cyto-technicians and doctors. The normal class contains 242 normal cell images with cell types: superficial squamous (74), intermediate squamous

(70) and columnar (98). Abnormal class is populated with 675 abnormal cell images (mild dysplasia - 182, moderate dysplasia - 146, severe dysplasia - 197, and carcinoma in situ/CIS - 150). Two cyto-technicians examined each cell and difficult samples were further examined by a doctor, for maximizing the confidence of classification. In case of disagreement between cyto-technicians and doctors, the samples were discarded. All images in Herlev database are in RGB format and associated with ground truths. Figure 6.1 shows examples from normal and abnormal classes in Herlev dataset.

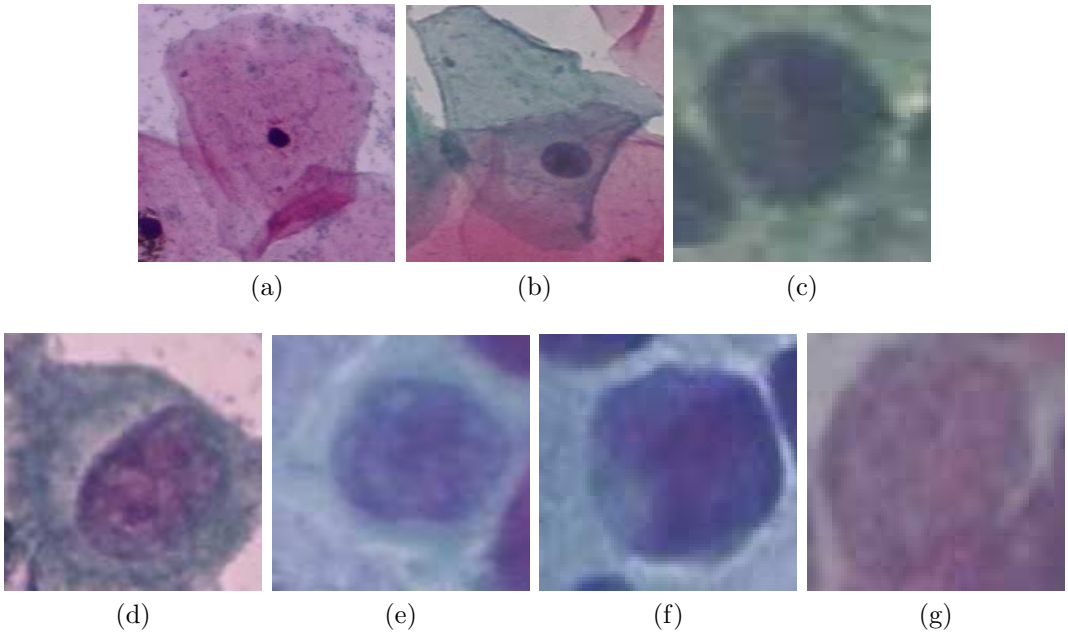


Figure 6.1: Example of cervical cell images from Herlev dataset: (a)-(c) normal and (d)-(g) abnormal.

6.3 Methodology

In the proposed framework, nuclei were segmented from cervical cell images using three segmentation techniques: CircEGS (proposed in Chapter 4), CiscFC (proposed in Chapter 3), and superpixel merging framework (proposed in Chapter 5). The first two techniques are guided by circular shape prior and the last one utilizes gradient boundary information, while segmenting nucleus. After segmentation, some morphological, intensity and texture features were calculated solely from cell nuclei. Then, fuzzy entropy measure was used adaptively to rank features in the feature set, and top ranking features were selected to form the feature vector for classification. Several classifiers were employed to discriminate abnormal

and normal cervical cells. Figure 6.2 shows a block diagram or flowchart of the proposed methodology.

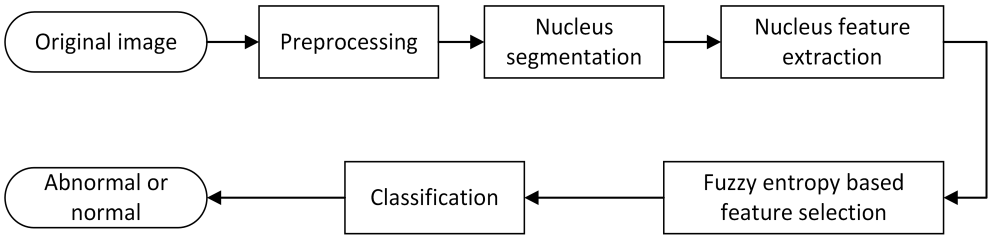


Figure 6.2: Block diagram of the proposed cervical cell classification framework.

6.3.1 Nucleus segmentation

Images in Herlev dataset are different from the overlapping cervical cell images in ISBI 2014 challenge dataset (see Section 3.3). The Herlev dataset contains RGB images, however the segmentation techniques proposed in Chapters 3, 4 and 5, are designed for grayscale images. These images also contain noise and other imaging artifacts. Therefore, it is required to convert the RGB images to grayscale and apply some pre-processing to denoise the images before segmentation.

Preprocessing: Before segmenting nucleus, all images were transformed from RGB to grayscale images using the same technique as used in (Li et al. 2012, Zhang, Kong, Liu, Wang, Chen & Sonka 2017). First, original RGB images were converted to CIELAB color space. CIELAB color space is composed of L* dimension for luminance or lightness, a* stands for red/green, and b* is for blue/yellow dimensions. Since, L* dimension separates lightness from color, it was chosen to generate grayscale images. The L* dimension was extracted from CIELAB space image and values were normalized to [0, 255] to produce the grayscale image. Then, a 7×7 median filter was applied on the grayscale image to remove noise and smooth the image like in (Chankong et al. 2014, Gautam et al. 2017). Finally, contrast-limited adaptive histogram equalization (CLAHE) was applied to improve grayscale image contrast. These pre-processed grayscale images were used in the segmentation process.

Segmentation: The first nucleus segmentation technique is the circular shape guided graph based segmentation - CircEGS, proposed in Chapter 4. In addition to intensity, a weighted shape prior utilizing circularity of nuclei was used in this

algorithm. Circular shape constrained fuzzy clustering or CiscFC, is the second method used for segmenting nucleus (presented in Chapter 3). Here, a shape function was integrated into fuzzy clustering process for precise segmentation of nuclei boundaries. The third segmentation technique used here, is a gradient guided superpixel merging based framework (proposed in Chapter 5). This technique will be abbreviated as SPmerg in this chapter. A novel merging criterion with pairwise regional contrast and gradient boundary guidance, was used in SPmerg algorithm to merge SLIC and SRM superpixels. SPmerg technique utilizing SLIC superpixels will be used in this chapter for Herlev dataset.

Herlev dataset contains single cell images, hence the background part of the cell is very small, especially in abnormal cell images. Therefore, the background subtraction steps from the original frameworks proposed in Chapters 3, 4 and 5, were excluded while segmenting nuclei from Herlev dataset. The segmentation techniques were not retrained for this datset and the default parameter and threshold values found in Chapters 3, 4 and 5 were used. From segmented regions, the region with the highest overlap with the ground truth nucleus was considered as the segmented nucleus, as in (Gençtav et al. 2012, Li et al. 2012). A segmented nucleus region is called correctly detected against ground truth region, if the criteria in Equation (3.11) is satisfied (see Section 3.5 for details). Correctly detected nuclei were used in further steps of this proposed framework for detecting abnormality in cervical cells.

6.3.2 Feature extraction and normalization

Dysplastic changes in cervical cells are visible through morphological changes (size, shape, texture, and intensity) in cell nuclei (Gençtav et al. 2012, Bora et al. 2017). A set of 13 features grouped as intensity/shape features and texture features from gray level co-occurrence matrix, were extracted from correctly detected nucleus in each image. The 10 intensity/shape related features form a subset of features used in (Chankong et al. 2014, Jantzen & Dounias 2006). These features are:

- area
- brightness or mean intensity
- major axis length (length of the major axis of an ellipse that fully encloses the segmented nucleus)

- minor axis length (length of the minor axis of an ellipse that fully encloses the nucleus region)
- elongation (ratio between minor axis length and major axis length of nucleus)
- roundness (ratio between segmented nucleus area and the area of the circle corresponding to nucleus major axis length)
- perimeter
- compactness (ratio between squared perimeter and area of the segmented nucleus)
- count of local maxima in 3×3 windows.
- count of local minima in 3×3 windows.

The 3 texture features used in this study are: homogeneity, contrast, and energy. These features were calculated from gray level co-occurrence matrix of the segmented nucleus using the definitions from (Theodoridis & Koutroumbas 2008).

If feature values vary in different ranges, there is a chance that features with large values will have a stronger influence on the classifiers (Theodoridis & Koutroumbas 2008). Therefore, normalization of feature values is an important step before using those values for classification. With normalization process, it is possible to restrict the feature values to stay in a certain range. In this study, all feature values were normalized to zero mean and unit variance (see (Theodoridis & Koutroumbas 2008) for details). If there are N values for a specific feature x, where $\bar{x}$ is the mean and σ is the standard deviation of feature x, then values of feature x were normalized to $\hat{x}$ as follows:

$$\hat{x} = \frac{x_i - \bar{x}}{\sigma}, \text{where } i = 1, 2, \ldots, N. \tag{6.1}$$

6.3.3 Feature selection

The basic idea of this step is to filter out some features, whose distributions are relatively random. To determine the importance of features in a feature set, fuzzy entropy and similarity classifier based feature selection technique - presented in (Luukka 2011), is used in this study. This technique is based on the concept of fuzziness measure from (Luca & Termini 1972). Fuzzy entropy values were

calculated considering a similarity measure between the ideal vectors and sample vectors for classification. Using this measure, one can rank and order the features in a feature set. Smaller entropy value of a feature indicates that the feature is more informative. Therefore, all features were ranked based on the fuzzy entropy value, and feature vectors were subsequently formed considering this ranking for classification purpose.

6.3.4 Classifiers

Five popular classifiers from literature named as: Fisher linear discriminant analysis (LDA), k-nearest neighbor (KNN), support vector machine (SVM) with two different kernels, and Ensemble, were used in this chapter to evaluate two-class (normal or abnormal) classification performance. These classifiers are briefly described in Section 2.1.5. In case of KNN, 7 nearest neighbors ($k = 7$) were used, as in (Chankong et al. 2014). While using SVM classifier, linear and radial basis function (RBF) kernels were employed. Ensemble method builds predictive models by combining multiple models, to improve accuracy of decisions. In ensemble methodology, bagging (bootstrap aggregating) is the most popular method. In this study, ensemble with bagging technique was used. It is worth to mention that all the conventional classifiers used in this chapter require inexpensive computational resources, compared to deep-learning or feature learning based approaches.

In Herlev dataset, no designated training and test sets are available. This dataset contains only 917 images, and is divided in an imbalanced way into normal (with 3 cell types) and abnormal (with 4 cell types) classes of cell images. Manual division of this dataset into train set (for training and validation) and test set (for evaluation), is not straightforward and may impose bias. Therefore, all classifiers were used with 10 independent runs of k-fold cross validation ($k = 5$). In k fold cross validation (Leisch et al. 1998), the input dataset is randomly partitioned in k equal sized subsets/folds. The process is repeated for k times and each time - the classifier is trained with $k - 1$ subsets and tested with remaining 1 subset of that fold.

6.4 Evaluation measures

Correctly detected nuclei regions were used to determine segmentation and classification accuracies. Cervical cell nucleus detection and segmentation were eval-

uated using the notations and definitions presented in Section 3.5. Nucleus detection performance was assessed using object level precision (PR_{obj}) and recall (RC_{obj}). Dice similarity coefficient (DSC), pixel level precision (PR_{pix}) and recall (RC_{pix}), were used to find accuracy of segmented nuclei boundaries.

For ease of comparison with literature, performance of the classifiers were assessed using five well-known measures from literature that were used to evaluate cervical cell classification. The measures are as: accuracy, sensitivity, specificity, area under the receiver operating characteristics or ROC curve (AUC), and Spearman rank-order correlation coefficient (ρ). While evaluating classification performance in this study, true positive (tp) indicates the number of abnormal cells classified as abnormal, true negative (tn) stands for number of normal cells classified as normal, false positive (fp) is the count of normal cells classified as abnormal, and false negative (fn) stands for the number of abnormal cells classified as normal.

Confusion matrix is a popular choice to visualize the performance of a classifier. It represents the summary of prediction results in a tabular form, for a specific classification problem (Ting 2017). For the two class classification problem (abnormal/normal) of this study, the confusion matrix will look like Table 6.1.

Table 6.1: An example of confusion matrix for a two class problem.

		Predicted	
		Normal	Abnormal
Actual	Normal	tn	fp
	Abnormal	fn	tp

Accuracy is calculated as the ratio of correct predictions over the total number of instances (Chankong et al. 2014). Accuracy for a C class classification can be defined as:

$$\begin{aligned} \text{Accuracy} &= \frac{\sum_{i=1}^{C} \text{No. of correct predictions in class } i}{\sum_{i=1}^{C} \text{No. of instances in class } i} \\ &= \frac{tp + tn}{tp + fp + tn + fn}. \end{aligned} \tag{6.2}$$

Sensitivity or true positive rate is found from the ratio of true positive instances (correctly predicted as abnormal cell) and total number of positive in-

stances (total number abnormal cells) as (Chankong et al. 2014):

$$\text{Sensitivity} = \frac{tp}{tp + fn}. \tag{6.3}$$

Specificity or true negative rate is defined as the ratio of correctly detected true negatives and total number of negative instances as (Chankong et al. 2014):

$$\text{Specificity} = \frac{tn}{tn + fp}. \tag{6.4}$$

AUC is a popular classification performance measure, calculated from ROC curve (Bradley 1997). ROC curve can be found by plotting classifier's false positive rate along x axis and true positive rate or sensitivity along y axis. Note, false positive rate is equal to $(1 -$ true negative rate$)$ or $(1 -$ specificity$)$. Therefore, ROC curve is also called sensitivity vs $(1 -$ specificity$)$ plot. The area under the ROC curve provides a measure of predictive power of the classifier.

Spearman rank-order correlation coefficient (ρ) represents the correlation between ground truth label (X) and predicted label (Y) of a classifier. This is defined as in (Gençtav et al. 2012):

$$\rho = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\left\{\sum_{i=1}^{N}(X_i - \bar{X})^2 \sum_{i=1}^{N}(Y_i - \bar{Y})^2\right\}^{1/2}}, \tag{6.5}$$

where N is the total number of instances, and $\bar{X}$ and $\bar{Y}$ are the mean values of ground truth and predicated labels, respectively. Values for ρ range in $[-1, +1]$, where -1 indicates perfect negative correlation, $+1$ indicates perfect positive correlation, and 0 indicates no correlation between the labels. If both labels are highly correlated, the value of ρ will be close to $+1$ (Gençtav et al. 2012, Chankong et al. 2014). In summary, higher values (close or equal to 1) are preferable for all evaluation measures used for nucleus detection, segmentation and classification.

6.5 Experimental results and discussion

In this chapter, the proposed framework to detect abnormality from nuclei features was evaluated using the Herlev dataset. Evaluation of nuclei segmentation and cell classification results, and comparison of these results with the state-of-the-art methods are presented in this section.

Nuclei segmentation: Shape and gradient boundary guided segmentation techniques were used to segment nuclei, to overcome the difficulties associated with poor contrast and intensity inhomogeneity. Table 6.2 presents the results of nucleus segmentation (calculated from correctly detected nuclei) and detection evaluations. The values for nucleus segmentation and detection measures for all three segmentation techniques were close to or above 0.9. This performance is very satisfactory, since the parameters for the segmentation techniques proposed in Chapters 3, 4 and 5 were not re-tuned for Herlev dataset, and previously established on a different dataset. This indicates that the segmentation techniques developed in Chapters 3, 4 and 5 are robust and applicable directly to other datasets.

The best segmentation performance, in terms of average DSC, PR_{pix} and RC_{pix}, was achieved for CiscFC technique as 0.94, 0.94 and 0.92, respectively. Average nuclei detection performance with CircEGS was the best with PR_{obj} (0.91) and RC_{obj} (0.94). From 917 cervical cell nuclei, CircEGS, CiscFC and SPmerg (using SLIC superpixels) segmentation techniques could correctly segment 859, 814 and 799 nuclei, respectively. Examples are presented in Figure 6.3 to compare ground truth nuclei boundaries with boundaries found using all three segmentation techniques. The segmentation using SPmerg technique with SLIC

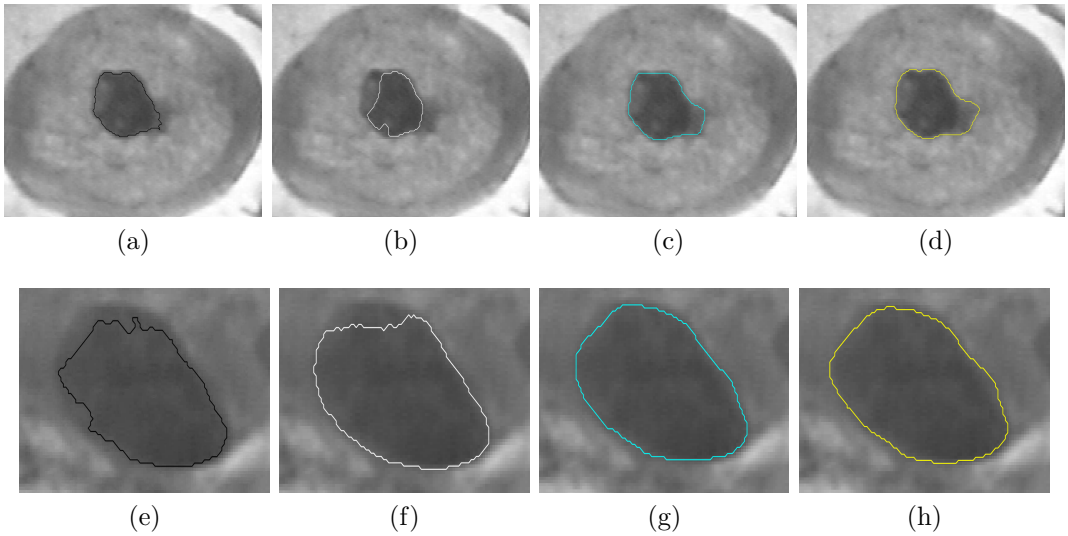


Figure 6.3: Comparison of ground truth (in black, first column), and segmented nuclei boundaries found using: CircEGS (in white, second column), CiscFC (in cyan, third column), and SPmerg (in yellow, last column) techniques.

superpixels and guided by gradient boundaries, did not perform well for abnormal cells compared to other two methods, because approximating gradient boundary

was very challenging for Herlev dataset.

Table 6.2: Evaluation of nucleus segmentation and detection for Herlev dataset.

			Performance measures				
			Segmentation			**Detection**	
Method	**Class**	**Cell type**	**DSC**	**PR_{pix}**	**RC_{pix}**	**PR_{obj}**	**RC_{obj}**
CircEGS	Normal	Super.	0.94 (0.05)	0.87 (0.13)	0.98 (0.04)	0.97	0.99
		Inter.	0.93 (0.07)	0.94 (0.07)	0.92 (0.12)	0.93	0.97
		Column.	0.90 (0.08)	0.91 (0.09)	0.85 (0.15)	0.90	0.93
	Abormal	Mild dys.	0.90 (0.08)	0.95 (0.08)	0.83 (0.15)	0.86	0.90
		Mod. dys.	0.91 (0.07)	0.93 (0.09)	0.87 (0.13)	0.89	0.92
		Sev. dys.	0.92 (0.07)	0.92 (0.12)	0.91 (0.12)	0.92	0.93
		CIS	0.91 (0.07)	0.91 (0.12)	0.90 (0.11)	0.95	0.97
Average with CircEGS			0.91 (0.07)	0.92 (0.11)	0.88 (0.13)	**0.91**	**0.94**
CiscFC	Normal	Super.	0.95 (0.05)	0.93 (0.11)	0.96 (0.06)	0.91	0.91
		Inter.	0.96 (0.03)	0.96 (0.07)	0.95 (0.06)	0.91	0.91
		Column.	0.93 (0.06)	0.94 (0.09)	0.89 (0.12)	0.92	0.92
	Abormal	Mild dys.	0.96 (0.05)	0.93 (0.11)	0.96 (0.05)	0.92	0.92
		Mod. dys.	0.95 (0.05)	0.93 (0.11)	0.95 (0.08)	0.95	0.95
		Sev. dys.	0.93 (0.06)	0.96 (0.08)	0.87 (0.13)	0.81	0.81
		CIS	0.93 (0.05)	0.95 (0.09)	0.90 (0.10)	0.85	0.85
Average with CiscFC			**0.94** (0.05)	**0.94** (0.10)	**0.92** (0.10)	0.89	0.89
SPmerg	Normal	Super.	0.90 (0.05)	0.83 (0.08)	1.00 (0.00)	0.88	0.85
		Inter.	0.92 (0.04)	0.85 (0.07)	0.99 (0.02)	0.93	0.90
		Column.	0.87 (0.09)	0.81 (0.13)	0.95 (0.09)	0.79	0.80
	Abormal	Mild dys.	0.86 (0.10)	0.93 (0.08)	0.83 (0.17)	0.79	0.81
		Mod. dys.	0.87 (0.10)	0.92 (0.08)	0.85 (0.16)	0.83	0.83
		Sev. dys.	0.89 (0.08)	0.91 (0.09)	0.89 (0.14)	0.90	0.92
		CIS	0.88 (0.09)	0.90 (0.10)	0.87 (0.14)	0.91	0.92
Average with SPmerg			0.88 (0.09)	0.89 (0.10)	0.89 (0.14)	0.86	0.86

Bold values represent the best average performance of segmentation techniques.
Segmentation results are reported in *mean (std)* format.
Detection results are reported as *mean* values.

CiscFC was the best performing method in terms of segmentation accuracy and was compared with the recent state-of-the-art techniques in literature that also used the Herlev dataset for evaluation. The comparison is presented in Table 6.3. The best values for DSC, PR_{pix} and RC_{pix} were obtained in (Liu et al. 2018), where nuclei were segmented using mask regional convolutional neural network (Mask-RCNN) and local fully connected conditional random field (LFC-CRF) techniques. However, this result differs by only around 1% (DSC) from the proposed techniques outcome. The highest DSC (0.95) was also found in (Dong et al. 2019) using adaptive gradient vector flow (AGVF) snake method. Nu-

cleus segmentation results found using CiscFC algorithm were higher than that achieved using graph-based fully convolutional network (FCN-G) (Zhang, Sonka, Lu, Summers & Yao 2017). Overall, nucleus segmentation performance achieved from this study is comparable with the recent state-of-the-art methods.

Table 6.3: Comparison of nucleus segmentation performance of the proposed framework, with the state-of-the-art methods using Herlev dataset.

Method	**DSC**	$\mathbf{PR_{pix}}$	$\mathbf{RC_{pix}}$
Gençtav et al. (2012)	0.89	0.88	0.93
Li et al. (2012)	0.94	-	-
Chankong et al. (2014)	0.80	0.85	0.83
Zhang, Sonka, Lu, Summers & Yao (2017)	0.92	-	-
Gautam et al. (2017)	0.86	0.85	0.89
Zhang, Kong, Liu, Wang, Chen & Sonka (2017)	-	0.90	0.95
Liu et al. (2018)	**0.95**	**0.96**	0.96
Dong et al. (2019)	**0.95**	-	-
Proposed with CiscFC	0.94	0.94	0.92

Bold values represent the best performance per measure.
'-' represents values not available.

Cell classification: Fuzzy entropy based hierarchical order/ranking was found adaptively from the training set in each fold of a run. This ordering was used to select the top ranking features for the test set, and form the feature vector for classification task in that fold. Rankings of 13 features in 5-folds of a run are presented in Table 6.4. Following this adaptive way, thirteen sets of feature vectors were generated, based on the ranking; with top 1, 2, 3, ..., 13 features, respectively. The effectiveness of these 13 feature vectors was evaluated using LDA, KNN, SVM-linear, Ensemble-bag, and SVM-RBF classifiers, and all classification evaluation metrics were computed.

AUC score is widely used to assess the predictive power of classifiers. In Figure 6.4, AUC scores found for all classifiers are plotted against the aforementioned thirteen sets of features. By analyzing AUC scores over number of ranked features/feature sets, it was found that the classification performance stabilizes for the feature set with 10 best features, for all classifiers and nuclei segmentation

Table 6.4: Fuzzy entropy based hierarchical ordering/ranking of 13 features of 5-folds in a run.

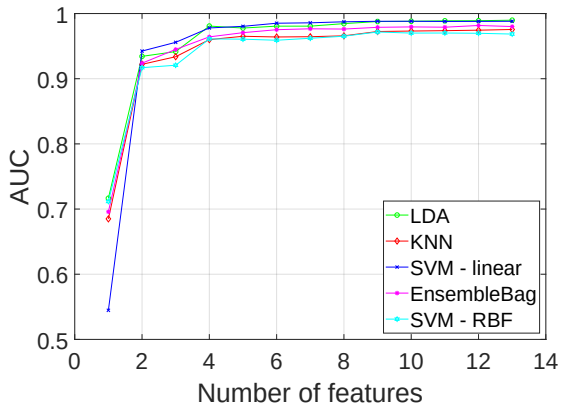
Rank	**Fold 1**	**Fold 2**	**Fold 3**	**Fold 4**	**Fold 5**
1	compactness	compactness	compactness	local minima	compactness
2	local minima	local minima	local maxima	compactness	local minima
3	local maxima	local maxima	local minima	local maxima	local maxima
4	perimeter	perimeter	perimeter	perimeter	perimeter
5	area	area	area	area	major axis
6	brightness	major axis	major axis	major axis	area
7	major axis	brightness	brightness	brightness	brightness
8	minor axis	minor axis	contrast	minor axis	contrast
9	contrast	contrast	minor axis	contrast	minor axis
10	roundness	roundness	homogeneity	roundness	roundness
11	elongation	elongation	roundness	homogeneity	elongation
12	homogeneity	homogeneity	elongation	elongation	homogeneity
13	energy	energy	energy	energy	energy

Rankings are presented in descending order, where rank 1 indicates the top and rank 13 is the lowest rank.

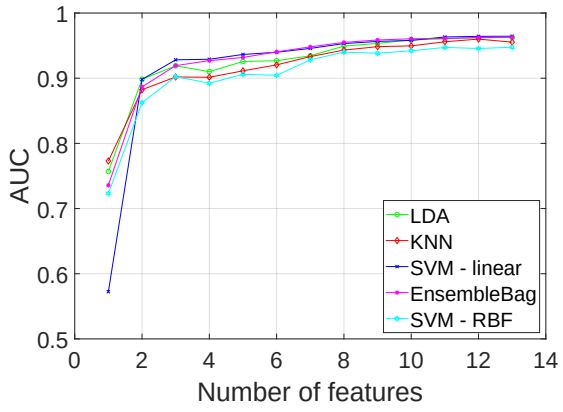
techniques. Thus, classification performance outcome of the proposed framework with the 10 best features were reported in this study.

Table 6.5 presents two class (normal and abnormal cells) classification results with 10 features. Classification performance for CircEGS, CiscFC and SPmerg nucleus segmentation techniques, was the best with SVM-linear, Ensemble-bag and LDA, respectively. A direct relation between segmentation/detection and classification performance can be observed from Table 6.2 and Table 6.5. Segmentation accuracy was the best for CiscFC technique, however nuclei detection accuracy of CircEGS technique was the highest with second best segmentation accuracy. It can be seen from Table 6.5 that cell classification performance was the best for CircEGS algorithm.

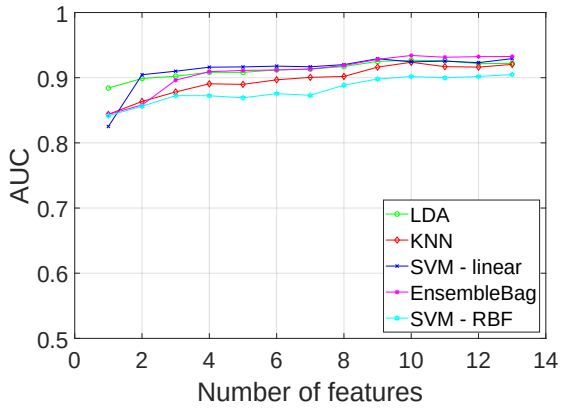
Confusion matrices of the best classification performance for three nuclei segmentation techniques, are presented in Table 6.6. These matrices were generated as the summation of the confusion matrices found for each fold in a run. It can be summarized from the confusion matrices that, CircEGS segmented nuclei with SVM-linear classifier outperformed others. A snapshot of classification measures



(a)



(b)



(c)

Figure 6.4: Plot of AUC over number of ranked features with LDA, KNN, SVM-linear, Ensemble-Bag, and SVM-RBF classifiers, for Herlev dataset segmented with: (a) CircEGS, (b) CiscFC and (c) SPmerg techniques.

Table 6.5: Two-class classification performance of Herlev dataset - segmented with CircEGS, CiscFC and SPmerg methods, with 10 independent runs of 5-fold cross validation and 10 nuclei features.

		Classifiers				
Method	**Measures**	**LDA**	**KNN**	**SVM-lin.**	**Ensemble-Bag**	**SVM-RBF**
CircEGS	Accuracy	**0.96**	0.95	**0.96**	**0.96**	0.94
	AUC	**0.99**	0.97	**0.99**	0.98	0.97
	Sensitivity	**0.99**	**0.99**	0.98	0.98	**0.99**
	Specificity	0.88	0.84	**0.91**	**0.91**	0.81
	ρ	0.89	0.87	**0.91**	0.90	0.85
CiscFC	Accuracy	0.92	0.92	0.92	**0.93**	0.92
	AUC	**0.96**	0.95	**0.96**	**0.96**	0.94
	Sensitivity	0.97	0.97	0.96	0.97	**0.98**
	Specificity	0.80	0.80	**0.82**	**0.82**	0.77
	ρ	0.80	0.80	0.80	**0.81**	0.80
SPmerg	Accuracy	**0.91**	**0.91**	**0.91**	**0.91**	**0.91**
	AUC	**0.93**	0.92	0.92	**0.93**	0.90
	Sensitivity	0.97	0.97	0.97	0.96	**0.99**
	Specificity	**0.75**	0.73	0.74	**0.75**	0.69
	ρ	**0.77**	0.75	0.76	0.76	0.76

Bold values represent the best performance for each measure per segmentation technique.

Table 6.6: Confusion matrices for the best classifiers in 2 class classification problem of Herlev dataset with 10 features, accumulated over 5-folds in a run.

		Predicted					
		CircEGS + SVM-linear		CiscFC + Ensemble-bag		SPmerg + LDA	
		Normal	Abnormal	Normal	Abnormal	Normal	Abnormal
Actual	Normal	213	8	178	19	153	16
	Abnormal	19	619	43	574	50	580

found in 5 folds of an iteration using CircEGS segmentation and SVM-linear classification, is presented in Table 6.7. From these results, the highest AUC score

Table 6.7: Classification results of 5-folds in a run, found for CircEGS segmentation and SVM linear classification with 10 nuclei features.

Measures	**Fold 1**	**Fold 2**	**Fold 3**	**Fold 4**	**Fold 5**
Accuracy	0.97	0.97	0.95	0.97	0.97
AUC	0.99	0.99	0.98	0.99	0.99
Sensitivity	0.98	0.99	0.98	0.99	0.98
Specificity	0.94	0.89	0.87	0.89	0.94
ρ	0.93	0.91	0.88	0.93	0.93

was in fold 1 (0.994382...) and the lowest for fold 3 (0.979304...) (the values are without rounding up to two digits after decimal point). ROC curves for these two folds are represented in Figure 6.5

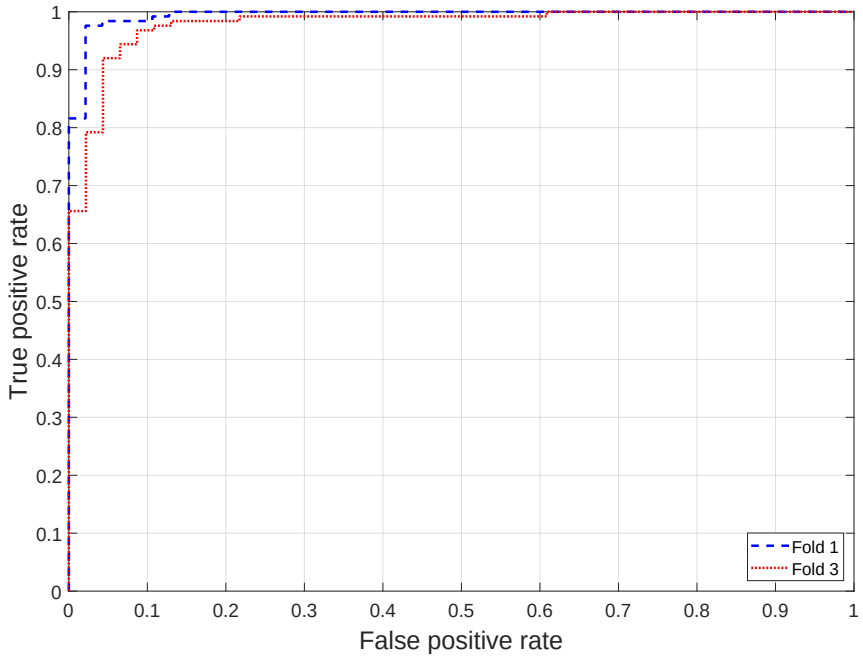


Figure 6.5: ROC curves illustrating the best and lowest AUC scores found in two folds of classification in a run, with CircEGS segmented nuclei and SVM linear classifier.

A comparison of the results found for CircEGS + SVM-linear (average of 10 iterations/runs of 5 folds) with the state-of-the art methods classification performance on Herlev dataset, is presented in Table 6.8. The best classification performance (in terms of accuracy, sensitivity, specificity, and ρ) was found with artificial neural network (ANN) classifier in (Chankong et al. 2014). Same accuracy level (0.99) was also achieved in (Jith et al. 2018) using deep neural network based on AlexNet. Though, many of the studies did not report AUC, however the proposed technique achieved higher AUC score (0.99) than GoogLeNet-5C (Lin et al. 2019). In addition, accuracy, sensitivity, specificity values for the proposed methodology were also higher than the deep learning based technique GoogLeNet-5C in (Lin et al. 2019).

The outcome of the proposed methodology is very promising for both cervical nuclei segmentation and abnormality detection, in comparison with recent literature using the Herlev dataset for evaluation. Classification performance indicates

Table 6.8: Comparison of Herlev dataset classification performance of the proposed approach with state-of-the-art techniques.

Method	**Accuracy**	**Sensitivity**	**Specificity**	ρ	**AUC**
ANN (Chankong et al. 2014)	**0.99**	**0.99**	**0.96**	**0.97**	-
Ensemble (Bora et al. 2017)	0.96	0.98	0.89	-	-
DeepCerv (Jith et al. 2018)	**0.99**	-	-	-	-
GoogLeNet-5C (Lin et al. 2019)	0.94	0.97	0.90	-	0.98
Proposed (CircEGS + SVM-lin.)	0.96	0.98	0.91	0.91	**0.99**

'-' represents values not available.

that features extracted only from nucleus, are adequate enough to identify abnormality in cervical cells. This finding is important, since precise segmentation of nucleus is more convenient and meaningful, than finding accurately overlapping and fuzzy cytoplasm boundaries. The framework was implemented in Matlab R2018b, running on a PC with Intel Core i7-4770 3.40 GHz processor and 8 GB RAM. Matlab function *crossvalind* from Bioinformatics Toolbox was used for k fold cross validation. The code was not optimized.

6.6 Conclusion

A novel nuclei feature based cervical cell classification framework is proposed in this chapter. Nucleus segmentation plays the vital role in computer-aided screening and diagnostic decision for cervical cytology images, therefore nucleus detection and boundary delineation accuracies are important. Three segmentation techniques with circular shape prior (CiscFC, CircEGS) and gradient boundary guidance (SPmerg) proposed in Chapters 3, 4 and 5, were exploited for segmenting nucleus. Thirteen morphological features were extracted from segmented nucleus and a fuzzy entropy based features selection/ranking process was applied to reduce the dimensionality of the feature vector. From AUC score analysis, the feature set with the 10 most discriminative features was found to achieve optimal performance for detecting abnormality in cervical cytology images using KNN, LDA, SVM-linear, SVM-RBF, and Ensemble-bag classifiers. The experimental

results on the Herlev dataset demonstrate that the proposed framework achieved promising results for nucleus segmentation and cervical cell classification, while compared to the state-of-the-art methods. This indicates that it is possible to characterize cervical cells as normal or abnormal using only nuclei features. The proposed abnormality detection system can help to improve efficiency and reduce observer bias, and decrease the rate of cervical cancer by assisting the cytotechnologists to identify abnormal cells in Pap smear images.

Chapter 7

Conclusion

Image segmentation is the crucial step in the field of computer-aided image analysis. The quality of ROI segmentation has a great influence on the performance of the whole application. In medical images, precise segmentation is precluded by noise, intensity inhomogeneity, poor contrast, object occlusion, and other imaging artifacts. Incorporation of prior knowledge about the desired object into segmentation frameworks has proven useful to improve the accuracy and plausibility of segmentation outcome. Prior guidance can reduce ambiguity in segmentation by ruling out inconsistent regions.

In this thesis, three prior guided novel segmentation frameworks were proposed to bridge the gap between experts' knowledge and computer vision techniques. Two circular shape constrained pixel grid level and one gradient guided superpixel level segmentation approaches were presented for precise segmentation of cervical nuclei. Finally, a cervical cell abnormality detection framework was proposed exclusively based on features extracted from nuclei.

7.1 Thesis summary and contribution

This thesis started with the motivation, a brief review of cervical cancer, Pap smear screening, and automated screening systems. Major approaches of nuclei segmentation in computer-aided screening and usefulness of prior knowledge in segmentation techniques were reviewed. Research objectives and outline of this thesis were also discussed in Chapter 1.

In Chapter 2, a literature review and technical background of this thesis was presented. The major steps involved in computer-aided cervical cytology screen-

ing and some popular techniques for those steps, were outlined. In addition, some commonly used priors and their incorporation in segmentation frameworks, and cervical nuclei segmentation techniques or related works from literature, were presented in this chapter.

In Chapter 3, a pixel level circular shape constrained clustering based segmentation framework (CiscFC) was presented, to segment cervical nuclei from overlapping Pap smear images. This is the first contribution of this thesis. A novel circular shape function (CSF) was proposed and incorporated into Fuzzy c-means (FCM) clustering technique. The CSF imposed a circular shape constrain on the target cluster and influenced the cluster formation process. Spatial location of a pixel was considered during the computation of CSF, thus the proposed algorithm could differentiate the pixels with similar intensity values using their spatial distance. The proposed CiscFC approach was evaluated using ISBI 2014 challenge dataset. The accuracy of nuclei boundary delineation was improved for the proposed algorithm, while compared to the standard FCM technique. Cervical nuclei segmentation performance of the CiscFC framework was also comparable with recent state-of-the-art techniques.

In Chapter 4, the second contribution of this thesis, a circular shape guided graph based segmentation (CircEGS) technique was presented on pixel grid level. A novel adaptive circular shape prior was proposed and incorporated in the merging predicate of efficient graph based segmentation (EGS) technique, to consider the circularity measure of regions during the merging process. Specifically, the threshold function of the merging predicate was redefined to employ the shape prior. The redefined threshold function was designed to balance the relative importance between area and shape term. ISBI 2014 challenge dataset was employed to evaluate cervical nuclei segmentation performance of the proposed CircEGS technique. The proposed algorithm outperformed the standard EGS technique, in terms of segmentation accuracy, quality of segmented boundary and border precision. In addition, circularity of segmented nuclei also improved for CircEGS. While compared to the state-of-the-art techniques, CircEGS performed favorably with others.

The proposed pixel level segmentation frameworks delineated cervical nuclei boundaries very precisely, and in some aspects exceeded the performance of other existing algorithms. However, effect of noise is more on the segmentation outcome for the frameworks - designed on pixel grid level representation of an image. Superpixels are the group of pixels with similar characteristics, thus they are less sensitive to noise and also have the ability to capture textural informa-

tion from regions. In literature, many ROI segmentation frameworks exist that are composed of superpixel tessellation and clustering/grouping of superpixels. In Chapter 5, the third contribution of this thesis, a novel superpixel merging framework (SPmerg) with a novel region merging criterion was proposed. The region merging criterion was composed of pairwise regional contrast and gradient boundary information. The value of pairwise regional contrast was thresholded to control the superpixel merging process, along with gradient guidance. SLIC and SRM segmentation techniques were employed to generate image superpixels. The proposed SPmerg framework was evaluated on ISBI 2014 challenge dataset, and achieved the best segmentation results compared to the techniques proposed in Chapter 3 and 4. The accuracy of nuclei boundary delineation using SPmerg framework also outperformed the state-of-the-art techniques found in literature.

Finally in Chapter 6, a new cervical cell classification framework was proposed to detect abnormality in cervical cells exclusively based on nuclei features. Cervical nuclei were segmented using the prior guided segmentation techniques developed in Chapters 3, 4 and 5. Herlev Pap smear database was used to evaluate this classification framework, since classification labels are not available for ISBI 2014 dataset. Parameter and threshold values found for ISBI 2014 dataset in Chapters 3, 4 and 5 were reused for Herlev dataset; the segmentation techniques were not retrained for the new dataset. Nucleus segmentation results for Herlev dataset were compared with related publications in literature. Thirteen intensity, shape and texture features were extracted from segmented nuclei, and ranked using a Fuzzy entropy and similarity classifier based feature selection technique. Results demonstrated that the classification performance of the proposed framework with only 10 nuclei features were very promising and comparable with the state-of-the-art methods. This is a positive indication towards the development of solely nuclei feature based characterization/classification systems for cervical cytological image analysis, in case of accurate segmentation of nuclei.

7.2 Future work

Despite of the high potential of the frameworks proposed in this thesis, in terms of detection, segmentation and classification of cervical nuclei, there are still some observations that lead towards future directions. These observations are discussed below.

- In Chapter 3, the highest mean intensity cluster was selected as the target

cluster and CSF was employed on this cluster. As a result, nuclei assigned to other clusters during the clustering process could not be detected using the proposed CiscFC framework, due to the presence of uneven staining, cell overlapping or poor contrast in the image. In future, the proposed framework may benefit from appropriate local enhancement process, which would improve the contrast between nucleus and cytoplasm and compensate the effect of uneven staining. Future work may also investigate the behavior of the proposed CiscFC method on detection of elliptical and other varied shaped nuclei.

- In Chapter 4, weight parameter to control the relative importance between area and shape term, and the parameter to control the scale of segmentation, were found from a small training set and kept fixed for evaluation using ISBI 2014 test set and Herlev database. Instead of fixed values, adaptive calculation of these values may show further improvement in segmentation accuracy of CircEGS technique.

- In Chapter 5, detection of nuclei markers are crucial for SPmerg algorithm, but the marker detection technique used in this chapter was not perfect. In the proposed framework, failure in nuclei marker detection directly impacted the detection and segmentation performance. Nucleus segmentation and detection accuracies can be potentially improved by replacing, the MSB and MSER based nuclei marker detection technique with a more efficient marker detector. In addition, other superpixel generation algorithms can be explored to determine their applicability with SPmerg framework, other than SLIC and SRM techniques.

- In Chapter 6, a set of 13 features (10 morphological and intensity, and 3 basic texture features) were used for classification. Capability of some other features, such as local binary pattern or radial features, could be explored in future. Ten most discriminatory features were selected as optimal subset, by analyzing the AUC scores. However, a classifier could be over-fitted to the experimental dataset in presence of more features for the small sized dataset used in this study. In such scenario, there is a chance that performance of classifier may worsen for a new independent dataset. Classification or abnormality detection performance could be explored and analyzed for less features, to decrease the risk of over-fitting. Five conventional classifiers (LDA, KNN, SVM with linear and RBF kernels, and Ensemble-bagging) were used, since they require inexpensive computational resources. In fu-

ture, deep learning based techniques could be employed for the classification task. Additionally, the classification task can also be extended to identify different stages of abnormality.

- The two datasets used in this thesis were generated using well preserved cervical cells, thus the developed frameworks has not faced the challenges of real world slides which may contain red-blood cells, degenerated cells, or some other non-cellular artifacts. To incorporate the proposed segmentation and classification frameworks in computerized cervical screening system, it would have to be tested on real world cervical whole slide images.

- The computational performance of the techniques developed in this thesis would have to be improved to be applicable in real-time screening systems. For example, SLIC superpixel merging based technique (presented in Chapter 5) takes 3.3 seconds to process one cell. Processing of one cervical smear sample with 20,000 cells would require around 24 hours. Therefore, the developed techniques would require speed optimization.

- Finally, the proposed frameworks could be adapted and applied to different modalities of medical image analysis applications such as: histopathology, fluorescence microscopy or other cytology images.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. & Süsstrunk, S. (2010), SLIC superpixels, Technical Report 149300, EPFL.

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. & Süsstrunk, S. (2012), 'SLIC superpixels compared to state-of-the-art superpixel methods', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(11), 2274–2282.

Adhikari, S. K., Sing, J. K., Basu, D. K. & Nasipuri, M. (2015), 'Conditional spatial fuzzy c-means clustering algorithm for segmentation of MRI images', *Applied Soft Computing* **34**, 758–769.

Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A. & Moriarty, T. (2002), 'A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data', *IEEE Transactions on Medical Imaging* **21**(3), 193–199.

American Cancer Society (2018), *Global Cancer Facts & Figures*, 4th edn, Atlanta: American Cancer Society.

American Cancer Society (2019), *Cancer Prevention & Early Detection Facts & Figures 2019-2020*, Atlanta: American Cancer Society.

Araújo, F. H., Silva, R. R., Ushizima, D. M., Rezende, M. T., Carneiro, C. M., Bianchi, A. G. C. & Medeiros, F. N. (2019), 'Deep learning for cell image segmentation and ranking', *Computerized Medical Imaging and Graphics* **72**, 13–21.

Australian Institute of Health and Welfare (2019), *Cervical screening in Australia 2019*, Cancer series no. 123. Cat. no. CAN 124. Canberra: AIHW.

Australian Institute of Health and Welfare & Cancer Australia (2012), *Gynaecological cancers in Australia: an overview*, Cancer series no. 70. Cat. no. CAN 66. Canberra: AIHW.

Bai, J., Miri, M. S., Liu, Y., Saha, P., Garvin, M. & Wu, X. (2014), Graph-based optimal multi-surface segmentation with a star-shaped prior: Application to the segmentation of the optic disc and cup, *in* '2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)', pp. 525–528.

Bai, X., Sun, C. & Zhou, F. (2009), 'Splitting touching cells based on concave points and ellipse fitting', *Pattern Recognition* **42**(11), 2434–2446.

Bajger, M., Lee, G. & Caon, M. (2013), '3D segmentation for multi-organs in CT images', *ELCVIA Electronic Letters on Computer Vision and Image Analysis* **12**(2), 13–27.

Bajger, M., Ma, F., Williams, S. & Bottema, M. (2010), Mammographic mass detection with statistical region merging, *in* '2010 International Conference on Digital Image Computing: Techniques and Applications (DICTA)', pp. 27–32.

Bamford, P. & Lovell, B. (1998), 'Unsupervised cell nucleus segmentation with active contours', *Signal Processing* **71**(2), 203–213.

Barrett, W. A. & Mortensen, E. N. (1997), 'Interactive live-wire boundary extraction', *Medical Image Analysis* **1**(4), 331–341.

BD FocalPoint GS imaging system (2019), `https://www.bd.com/scripts/anz/productsdrilldown.asp?CatID=221&SubID=1821&siteID=7&d=&s=anz&sTitle=BD+%2D+Australia%2FNew+Zealand&metaTitle=Tripath&dc=anz&dcTitle=BD+%2D+Australia%2FNew+Zealand#4168`. [Online; Accessed: 2019-11-10].

BD PrepStain Slide Processor (2019), `https://www.bd.com/scripts/anz/productsdrilldown.asp?CatID=221&SubID=1821&siteID=7&d=&s=anz&sTitle=BD+%2D+Australia%2FNew+Zealand&metaTitle=Tripath&dc=anz&dcTitle=BD+%2D+Australia%2FNew+Zealand#4167`. [Online; Accessed: 2019-11-10].

BD SurePath (2019), `https://www.bd.com/scripts/anz/productsdrilldown.asp?CatID=221&SubID=1821&siteID=7&d=&s=anz&sTitle=BD+%2D+Australia%2FNew+Zealand&metaTitle=Tripath&dc=anz&dcTitle=BD+%2D+Australia%2FNew+Zealand#4166`. [Online; Accessed: 2019-11-10].

Béliz-Osorio, N., Crespo, J., García-Rojo, M., Muñoz, A. & Azpiazu, J. (2011), Cytology imaging segmentation using the locally constrained watershed transform, *in* 'Mathematical Morphology and Its Applications to Image and Signal Processing', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 429–438.

Ben-Zadok, N., Riklin-Raviv, T. & Kiryati, N. (2009), Interactive level set segmentation for image-guided therapy, *in* '2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro', pp. 1079–1082.

Bengtsson, E. (2003), Computerized cell image analysis: Past, present, and future, *in* J. Bigun & T. Gustavsson, eds, 'Image Analysis', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 395–407.

Bengtsson, E., Eriksson, O., Holmquist, J., Jarkrans, T., Nordin, B. & Stenkvist, B. (1981), 'Segmentation of cervical cells: Detection of overlapping cell nuclei', *Computer Graphics and Image Processing* **16**(4), 382–394.

Bengtsson, E., Eriksson, O., Holmquist, J., Nordin, B. & Stenkvist, B. (1979), 'High resolution segmentation of cervical cells', *The Journal of Histochemistry and Cytochemistry* **27**(1), 621–628.

Bengtsson, E. & Malm, P. (2014), 'Screening for cervical cancer using automated analysis of Pap-smears', *Computational and Mathematical Methods in Medicine* **2014**, 1–12. Article ID 842037.

Bergmeir, C., Silvente, M. G. & Bentez, J. M. (2012), 'Segmentation of cervical cell nuclei in high-resolution microscopic images: A new algorithm and a web-based software framework', *Computer Methods and Programs in Biomedicine* **107**(3), 497–512.

Bezdek, J. C. (1981), *Pattern Recognition with Fuzzy Objective Function Algorithms*, Springer US, Plenum Press, New York, USA.

Bishop J, W., Bigner S, H., Colgan T, J., Husain, M., Howell L, P., McIntosh K, M., Taylor D, A. & Sadeghi M, H. (1998), 'Multicenter masked evaluation of AutoCyte PREP thin layers with matched conventional smears', *Acta Cytologica* **42**(1), 189–197.

Bora, K., Chowdhury, M., Mahanta, L. B., Kundu, M. K. & Das, A. K. (2017), 'Automated classification of Pap smear images to detect cervical dysplasia', *Computer Methods and Programs in Biomedicine* **138**, 31–47.

Borst, H., Abmayr, W. & Gais, P. (1979), 'A thresholding method for automatic cell image segmentation', *The Journal of Histochemistry and Cytochemistry* **27**(1), 180–187.

Boykov, Y. & Funka-Lea, G. (2006), 'Graph cuts and efficient N-D image segmentation', *International Journal of Computer Vision* **70**(2), 109–131.

Boykov, Y. & Jolly, M.-P. (2000), Interactive organ segmentation using graph cuts, *in* 'Medical Image Computing and Computer-Assisted Intervention - MICCAI', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 276–286.

Boykov, Y., Veksler, O. & Zabih, R. (1999), Fast approximate energy minimization via graph cuts, *in* 'Proceedings of the Seventh IEEE International Conference on Computer Vision', Vol. 1, pp. 377–384.

Boykov, Y., Veksler, O. & Zabih, R. (2001), 'Fast approximate energy minimization via graph cuts', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(11), 1222–1239.

Boykov, Y. Y. & Jolly, M.-P. (2001), Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images, *in* 'International Conference on Computer Vision, Vancouver, Canada', Vol. 1, pp. 105–112.

Bradley, A. P. (1997), 'The use of the area under the ROC curve in the evaluation of machine learning algorithms', *Pattern Recognition* **30**(7), 1145–1159.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. & Jemal, A. (2018), 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA: A Cancer Journal for Clinicians* **68**(6), 394–424.

Breiman, L. (1996), 'Bagging predictors', *Machine Learning* **24**(2), 123–140.

Buades, A., Coll, B. & Morel, J. (2005), A non-local algorithm for image denoising, *in* '2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)', Vol. 2, pp. 60–65.

Cahn, R. L., Poulsen, R. S., & Toussaint, G. (1977), 'Segmentation of cervical cell images', *The Journal of Histochemistry and Cytochemistry* **25**(7), 681–688.

Cai, W., Chen, S. & Zhang, D. (2007), 'Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation', *Pattern Recognition* **40**(3), 825–838.

Carothers, A., McGoogan, E., Vooijs, P., Bird, C., Colquhoun, M., Eason, P., McKie, M., Nieuwenhuis, F., Pitt, P. & Rutovitz, D. (1994), 'A collaborative trial of a semi-automatic system for slide preparation and screening in cervical cytopathology', *Analytical cellular pathology : the journal of the European Society for Analytical Cellular Pathology* **7**(4), 261–274.

Caselles, V., Kimmel, R. & Sapiro, G. (1995), Geodesic active contours, *in* 'Proceedings of IEEE International Conference on Computer Vision', pp. 694–699.

Catarino, R., Petignat, P., Dongui, G. & Vassilakos, P. (2015), 'Cervical cancer screening in developing countries at a crossroad: Emerging technologies and policy choices', *World Journal of Clinical Oncology* **6**(6), 281–290.

Cervical cancer (2019), `https://cervical-cancer.canceraustralia.gov.au/`. [Online; Accessed: 2019-11-10].

Cervical cancer statistics (2019), `https://www.wcrf.org/dietandcancer/cancer-trends/cervical-cancer-statistics`. [Online; Accessed: 2019-11-10].

Chankong, T., Theera-Umpon, N. & Auephanwiriyakul, S. (2014), 'Automatic cervical cell segmentation and classification in Pap smears', *Computer Methods and Programs in Biomedicine* **113**(2), 539–556.

Chen, S. & Zhang, D. (2004), 'Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure', *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **34**(4), 1907–1916.

Cheng, J. & Rajapakse, J. C. (2009), 'Segmentation of clustered nuclei with shape markers and marking function', *IEEE Transactions on Biomedical Engineering* **56**(3), 741–748.

Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J. & Chen, T.-J. (2006), 'Fuzzy c-means clustering with spatial information for image segmentation', *Computerized Medical Imaging and Graphics* **30**(1), 9–15.

Cloppet, F. & Boucher, A. (2010), 'Segmentation of complex nucleus configurations in biological images', *Pattern Recognition Letters* **31**(8), 755–761. Award winning papers from the 19th International Conference on Pattern Recognition (ICPR).

Cong, W., Song, J., Wang, L., Liang, H. & Li, J. (2015), 'A fuzzy c-means clustering scheme incorporating non-local spatial constraint for brain magnetic resonance image segmentation', *Journal of Medical Imaging and Health Informatics* **5**(8), 1821–1825.

Cootes, T. F., Edwards, G. J. & Taylor, C. J. (2001), 'Active appearance models', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(6), 681–685.

Cootes, T. F., Taylor, C. J., Cooper, D. H. & Graham, J. (1992), Training models of shape from sets of examples, *in* 'British Machine Vision Conference (BMVC)', Springer London, London, pp. 9–18.

Cootes, T., Taylor, C., Cooper, D. & Graham, J. (1995), 'Active shape models-their training and application', *Computer Vision and Image Understanding* **61**(1), 38–59.

Cremers, D., Fluck, O., Rousson, M. & Aharon, S. (2007), A probabilistic level set formulation for interactive organ segmentation, *in* 'Proceedings of the SPIE', Vol. 6512, pp. 1–9.

Crowell, E. F., Bazin, C., Saunier, F., Brixtel, R., Caillot, Y., Lesner, B., Toutain, M., Ferreri, C., Garcia, I., Mathieu, M.-C., Vaussanvin, J., Depardon, J. & Renouf, A. (2019), 'Cytoprocessor: A new cervical cancer screening system for remote diagnosis', *Acta Cytologica* **63**, 215–223.

CytoProcessor (2019), `https://www.datexim.com/en/cytoprocessor-3/`. [Online; Accessed: 2019-11-10].

Dalal, N. & Triggs, B. (2005), Histograms of oriented gradients for human detection, *in* 'IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)', Vol. 1, pp. 886–893.

Das, P., Veksler, O., Zavadsky, V. & Boykov, Y. (2009), 'Semiautomatic segmentation with compact shape prior', *Image and Vision Computing* **27**(1), 206–219. Canadian Robotic Vision 2005 and 2006.

Dice, L. R. (1945), 'Measures of the amount of ecologic association between species', *Ecology* **26**(3), 297–302.

Dietterich, T. G. (2000), Ensemble methods in machine learning, *in* 'Multiple Classifier Systems', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–15.

Dong, N., Zhao, L. & Wu, A. (2019), 'Cervical cell recognition based on AGVF-snake algorithm', *International Journal of Computer Assisted Radiology and Surgery* .

Dunn, J. C. (1973), 'A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters', *Journal of Cybernetics* **3**(3), 32–57.

Elsheikh, T. M., Austin, R. M., Chhieng, D. F., Miller, F. S., Moriarty, A. T. & Renshaw, A. A. (2013), 'American society of cytopathology workload recommendations for automated Pap test screening: Developed by the productivity and quality assurance in the era of automated screening task force', *Diagnostic Cytopathology* **41**(2), 174–178.

Erhardt, R., Reinhardt, E., Schlipf, W. & Bloss, W. (1980), 'FAZYTAN: a system for fast automated cell segmentation, cell image analysis and feature extraction based on TV-image pickup and parallel processing', *Analytical and quantitative cytology* **2**(1), 25–40.

Falcão, A. X., Udupa, J. K., Samarasekera, S., Sharma, S., Hirsch, B. E. & de A. Lotufo, R. (1998), 'User-steered image segmentation paradigms: Live wire and live lane', *Graphical Models and Image Processing* **60**(4), 233–260.

Felzenszwalb, P. F. & Huttenlocher, D. P. (1998), Image segmentation using local variation, *in* 'Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231)', pp. 98–104.

Felzenszwalb, P. F. & Huttenlocher, D. P. (2004), 'Efficient graph-based image segmentation', *International Journal of Computer Vision* **59**(2), 167–181.

Fisher, R. A. (1936), 'The use of multiple measurements in taxonomic problems', *Annals of Eugenics* **7**(2), 179–188.

Freedman, D. & Zhang, T. (2005), Interactive graph cut based segmentation with shape priors, *in* '2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)', Vol. 1, pp. 755–762.

Fukunaga, K. & Narendra, P. M. (1975), 'A branch and bound algorithm for computing k-nearest neighbors', *IEEE Transactions on Computers* **C-24**(7), 750–753.

Galloway, M. M. (1975), 'Texture analysis using gray level run lengths', *Computer Graphics and Image Processing* **4**(2), 172–179.

Gautam, S., Gupta, K., Bhavsar, A. & Sao, A. K. (2017), Unsupervised segmentation of cervical cell nuclei via adaptive clustering, *in* 'Medical Image Understanding and Analysis', Springer International Publishing, Cham, pp. 815–826.

Gençtav, A., Aksoy, S. & Őnder, S. (2012), 'Unsupervised segmentation and classification of cervical cell images', *Pattern Recognition* **45**(12), 4151–4168.

Giger, M. L., Doi, K. & MacMahon, H. (1988), 'Image feature analysis and computer-aided diagnosis in digital radiography. 3. automated detection of nodules in peripheral lung fields', *Medical Physics* **15**(2), 158–166.

Gonzalez, R. & Woods, R. E. (2008), *Digital Image Processing*, Pearson Prentice Hall.

Grady, L. (2006), 'Random walks for image segmentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(11), 1768–1783.

Grady, L. & Schwartz, E. L. (2006), 'Isoperimetric graph partitioning for image segmentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(3), 469–475.

Grosgeorge, D., Petitjean, C., Dacher, J.-N. & Ruan, S. (2013), 'Graph cut segmentation with a statistical shape model in cardiac MRI', *Computer Vision and Image Understanding* **117**(9), 1027–1035.

Grosgeorge, D., Petitjean, C. & Ruan, S. (2016), 'Multilabel statistical shape prior for image segmentation', *IET Image Processing* **10**, 710–716.

Guan, T., Zhou, D. & Liu, Y. (2015), 'Accurate segmentation of partially overlapping cervical cells based on dynamic sparse contour searching and GVF snake model', *IEEE Journal of Biomedical and Health Informatics* **19**(4), 1494–1504.

Haddad, C. W., Drukker, K., Gullett, R., Carroll, T. J., Christoforidis, G. A. & Giger, M. L. (2018), 'Fuzzy c-means segmentation of major vessels in angiographic images of stroke', *Journal of medical imaging (Bellingham, Wash.)* **5**(1), 14501–14511.

Hamarneh, G. & Li, X. (2009), 'Watershed segmentation using prior shape and appearance knowledge', *Image and Vision Computing* **27**(1), 59–68. Canadian Robotic Vision 2005 and 2006.

Haralick, R. M., Shanmugam, K. & Dinstein, I. (1973), 'Textural features for image classification', *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**(6), 610–621.

Harandi, N. M., Sadri, S., Moghaddam, N. A. & Amirfattahi, R. (2010), 'An automated method for segmentation of epithelial cervical cells in images of ThinPrep', *Journal of Medical Systems* **34**(6), 1043–1058.

Hayakawa, T., Prasath, V. B. S., Kawanaka, H., Aronow, B. J. & Tsuruoka, S. (2019), 'Computational nuclei segmentation methods in digital pathology: A survey', *Archives of Computational Methods in Engineering* .

Heimann, T. & Meinzer, H.-P. (2009), 'Statistical shape models for 3D medical image segmentation: A review', *Medical Image Analysis* **13**(4), 543–563.

Holmquist, J., Bengtsson, E., Olsen, B., Stenkvist, B. & Noguchi, Y. (1976), 'Analysis of grey-level histograms as a method of classifying cells from the uterine cervix', *Computers in Biology and Medicine* **6**(3), 213–223.

Holmquist, J., Imasoto, Y., Bengtsson, E., Olsen, B. & Stenkvist, B. (1976), 'A microspectrophotometric study of papanicolaou-stained cervical cells as an aid in computerized image processing.', *Journal of Histochemistry & Cytochemistry* **24**(12), 1218–1224. PMID: 63509.

Howell, L. P., Davis, R. L., Belk, T. I., Agdigos, R. & Lowe, J. (1998), 'The autocyte preparation system for gynecologic cytology', *Acta Cytologica* **43**, 171–177.

Hua, J., Xiong, Z., Lowey, J., Suh, E. & Dougherty, E. R. (2004), 'Optimal number of features as a function of sample size for various classification rules', *Bioinformatics* **21**(8), 1509–1515.

Huang, Q.-H., Lee, S.-Y., Liu, L.-Z., Lu, M.-H., Jin, L.-W. & Li, A.-H. (2012), 'A robust graph-based segmentation method for breast tumors in ultrasound images', *Ultrasonics* **52**(2), 266–275.

Husham, A., Hazim Alkawaz, M., Saba, T., Rehman, A. & Saleh Alghamdi, J. (2016), 'Automated nuclei segmentation of malignant using level sets', *Microscopy Research and Technique* **79**(10), 993–997.

Hutchinson, M. L., Cassin, C. M. & Ball, H. G. (1991), 'The efficacy of an automated preparation device for cervical cytology', *American Journal of Clinical Pathology* **96**(3), 300–305.

Huttenlocher, D. P., Klanderman, G. A. & Rucklidge, W. J. (1993), 'Comparing images using the Hausdorff distance', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(9), 850–863.

Ibragimov, B., Likar, B., Pernu, F. & Vrtovec, T. (2014), 'Shape representation for efficient landmark-based segmentation in 3-D', *IEEE Transactions on Medical Imaging* **33**(4), 861–874.

Irshad, H., Veillard, A., Roux, L. & Racoceanu, D. (2014), 'Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential', *IEEE Reviews in Biomedical Engineering* **7**, 97–114.

Isa, N. A. M. (2005), 'Automated edge detection technique for Pap smear images using moving k-means clustering and modified seed based region growing algorithm', *International Journal of the Computer, the Internet and Management* **13**(3), 45–59.

Jantzen, J. & Dounias, G. (2006), Analysis of Pap-smear image data, *in* 'Proc. of the Nature-Inspired Smart Information Systems', NiSIS.

Jantzen, J., Norup, J., Dounias, G. & Bjerregaard, B. (2005), Pap-smear benchmark data for pattern classification, *in* 'Proc. of the Nature-Inspired Smart Information Systems', NiSIS, pp. 1–9.

Jith, O. U. N., Harinarayanan, K. K., Gautam, S., Bhavsar, A. & Sao, A. K. (2018), DeepCerv: Deep neural network for segmentation free robust cervical cell classification, *in* 'Computational Pathology and Ophthalmic Medical Image Analysis', Springer Int. Publishing, Cham, pp. 86–94.

Jung, C. & Kim, C. (2010), 'Segmenting clustered nuclei using H-minima transform-based marker extraction and contour parameterization', *IEEE Transactions on Biomedical Engineering* **57**(10), 2600–2604.

Jung, C., Kim, C., Chae, S. W. & Oh, S. (2010), 'Unsupervised segmentation of overlapped nuclei using Bayesian classification', *IEEE Transactions on Biomedical Engineering* **57**(12), 2825–2832.

Kale, A. & Aksoy, S. (2010), Segmentation of cervical cell images, *in* '20th International Conference on Pattern Recognition', pp. 2399–2402.

Kalinin, P. & Sirota, A. (2015), 'A graph based approach to hierarchical image over-segmentation', *Computer Vision and Image Understanding* **130**, 80–86.

Kaur, S. & Sahambi, J. (2016), 'Curvelet initialized level set cell segmentation for touching cells in low contrast images', *Computerized Medical Imaging and Graphics* **49**, 46–57.

Koss, L. G., Lin, E., Schreiber, K., Elgert, P. & Mango, L. (1994), 'Evaluation of the PAPNET cytologic screening system for quality control of cervical smears', *American Journal of Clinical Pathology* **101**(2), 220229.

Kostrykin, L., Schnörr, C. & Rohr, K. (2019), 'Globally optimal segmentation of cell nuclei in fluorescence microscopy images using shape and intensity information', *Medical Image Analysis* **58**, 101536.

Kubassova, O., Boyle, R. D. & Radjenovic, A. (2006), A novel method for quantitative evaluation of segmentation outputs for dynamic contrast-enhanced MRI data in RA studies., *in* 'Proceedings of the Joint Disease Workshop, 9th International Conference on Medical Image Computing and Computer Assisted Intervention', Vol. 1, pp. 72–79.

Lee, H. & Kim, J. (2016), Segmentation of overlapping cervical cells in microscopic images with superpixel partitioning and cell-wise contour refinement, *in* '2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)', pp. 1367–1373.

Leisch, F., Jain, L. C. & Hornik, K. (1998), 'Cross-validation with active pattern selection for neural-network classifiers', *IEEE Transactions on Neural Networks* **9**(1), 35–41.

Lempitsky, V., Kohli, P., Rother, C. & Sharp, T. (2009), Image segmentation with a bounding box prior, *in* '2009 IEEE 12th International Conference on Computer Vision', pp. 277–284.

Leung, S.-H., Wang, S.-L. & Laus, W.-H. (2004), 'Lip image segmentation using fuzzy clustering incorporating an elliptic shape function', *IEEE Transactions on Image Processing* **13**(1), 51–62.

Leventon, M. E., Grimson, W. E. L. & Faugeras, O. (2000), Statistical shape influence in geodesic active contours, *in* 'Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)', Vol. 1, pp. 316–323.

Li, B. N., Chui, C. K., Chang, S. & Ong, S. (2011), 'Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation', *Computers in Biology and Medicine* **41**(1), 1–10.

Li, C., Xu, C., Gui, C. & Fox, M. D. (2010), 'Distance regularized level set evolution and its application to image segmentation', *IEEE Transactions on Image Processing* **19**(12), 3243–3254.

Li, K., Lu, Z., Liu, W. & Yin, J. (2012), 'Cytoplasm and nucleus segmentation in cervical smear images using radiating GVF snake', *Pattern Recognition* **45**(4), 1255–1264.

Lin, H., Hu, Y., Chen, S., Yao, J. & Zhang, L. (2019), 'Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks', *IEEE Access* **7**, 71541–71549.

Litvin, A. & Karl, W. C. (2005), Coupled shape distribution-based segmentation of multiple objects, *in* 'Information Processing in Medical Imaging', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 345–356.

Liu, Y., Zhang, P., Song, Q., Li, A., Zhang, P. & Gui, Z. (2018), 'Automatic segmentation of cervical nuclei based on deep learning and a conditional random field', *IEEE Access* **6**, 53709–53721.

Lu, Z., Carneiro, G. & Bradley, A. P. (2013), Automated nucleus and cytoplasm segmentation of overlapping cervical cells, *in* 'Medical Image Computing and Computer-Assisted Intervention - MICCAI', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 452–460.

Lu, Z., Carneiro, G. & Bradley, A. P. (2015), 'An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells', *IEEE Transactions on Image Processing* **24**(4), 1261–1272.

Lu, Z., Carneiro, G., Bradley, A. P., Ushizima, D., Nosrati, M. S., Bianchi, A. G. C., Carneiro, C. M. & Hamarneh, G. (2017), 'Evaluation of three algorithms for the segmentation of overlapping cervical cells', *IEEE Journal of Biomedical and Health Informatics* **21**(2), 441–450.

Luca, A. D. & Termini, S. (1972), 'A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory', *Information and Control* **20**(4), 301–312.

Luukka, P. (2011), 'Feature selection using fuzzy entropy measures with similarity classifier', *Expert Systems with Applications* **38**(4), 4600–4607.

Ma, F., Bajger, M. & Bottema, M. J. (2007), Robustness of two methods for segmenting salient features in screening mammograms, *in* '9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications (DICTA)', pp. 112–117.

Ma, F., Bajger, M., Slavotinek, J. P. & Bottema, M. J. (2007), 'Two graph theory based methods for identifying the pectoral muscle in mammograms', *Pattern Recognition* **40**(9), 2592–2602.

MacQueen, J. (1967), Some methods for classification and analysis of multivariate observations, *in* 'Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics', University of California Press, Berkeley, Calif., pp. 281–297.

Malladi, R., Sethian, J. A. & Vemuri, B. C. (1995), 'Shape modeling with front propagation: a level set approach', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(2), 158–175.

Malm, P. (2013), Image Analysis in Support of Computer-Assisted Cervical Cancer Screening, PhD thesis, Department of Information Technology, Division of Visual Information and Interaction, Box 337, Uppsala University, SE-751 05 Uppsala, Sweden.

Malm, P., Balakrishnan, B. N., Sujathan, V. K., Kumar, R. & Bengtsson, E. (2013), 'Debris removal in Pap-smear images', *Computer Methods and Programs in Biomedicine* **111**(1), 128–138.

Malpica, N., de Solrzano, C. O., Vaquero, J. J., Santos, A., Vallcorba, I., Garca-Sagredo, J. M. & del Pozo, F. (1997), 'Applying watershed algorithms to the segmentation of clustered nuclei', *Cytometry* **28**(4), 289–297.

Mariarputham, E. J. & Stephen, A. (2015), 'Nominated texture based cervical cancer classification', *Computational and Mathematical Methods in Medicine* **2015**, 1–10. Article ID 586928.

Matas, J., Chum, O., Urban, M. & Pajdla, T. (2004), 'Robust wide-baseline stereo from maximally stable extremal regions', *Image and Vision Computing* **22**(10), 761–767. British Machine Vision Computing 2002.

Molnar, C., Kato, Z. & Jermyn, I. H. (2015), A new model for the segmentation of multiple, overlapping, near-circular objects, *in* '2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)', pp. 1–5.

Moshavegh, R., Ehteshami Bejnordi, B., Mehnert, A., Sujathan, K., Malm, P. & Bengtsson, E. (2012), Automated segmentation of free-lying cell nuclei in Pap smears for malignancy-associated change analysis, *in* '2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society', pp. 5372–5375.

Nayar, R. & Wilbur, D. (2015), *The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*, Springer International Publishing.

Nock, R. & Nielsen, F. (2004), 'Statistical region merging', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(11), 1452–1458.

Noordam, J. C., van den Broek, W. H. A. M. & Buydens, L. M. C. (2000), Geometrically guided fuzzy c-means clustering for multivariate image segmentation, *in* 'Proceedings 15th International Conference on Pattern Recognition. ICPR-2000', Vol. 1, pp. 462–465.

Nosrati, M. S. & Hamarneh, G. (2015*a*), Segmentation of overlapping cervical cells: A variational method with star-shape prior, *in* '2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)', pp. 186–189.

Nosrati, M. S. & Hamarneh, G. (2015*b*), 'A variational approach for overlapping cell segmentation'. [Online; Accessed: 2019-11-10].
URL: *http://cs.adelaide.edu.au/∼carneiro/isbi14_challenge/results_release.html*

Nosrati, M. S. & Hamarneh, G. (2016), 'Incorporating prior knowledge in medical image segmentation: a survey', *arXiv e-prints* p. arXiv:1607.01092.

Ojala, T., Pietikainen, M. & Maenpaa, T. (2002), 'Multiresolution gray-scale and rotation invariant texture classification with local binary patterns', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(7), 971–987.

Oliveira, P. H. C., Moreira, G., Ushizima, D. M., Carneiro, C. M., Medeiros, F. N. S., de Araujo, F. H. D., e Silva, R. R. V. & Bianchi, A. G. C. (2017), A multi-objective approach for calibration and detection of cervical cells nuclei, *in* '2017 IEEE Congress on Evolutionary Computation (CEC)', pp. 2321–2327.

Oprisescu, S., Rădulescu, T., Sultana, A., Rasche, C. & Ciuc, M. (2015), Automatic Pap smear nuclei detection using mean-shift and region growing, *in* 'International Symposium on Signals, Circuits and Systems (ISSCS)', pp. 1–4.

Osher, S. & Sethian, J. A. (1988), 'Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations', *Journal of Computational Physics* **79**(1), 12–49.

Otsu, N. (1979), 'A threshold selection method from gray-level histograms', *IEEE Transactions on Systems, Man, and Cybernetics* **9**(1), 62–66.

Pai, P.-Y., Chang, C.-C. & Chan, Y.-K. (2012), 'Nucleus and cytoplast contour detector from a cervical smear image', *Expert Systems with Applications* **39**(1), 154–161.

Papanicolaou, G. N. (1942), 'A new procedure for staining vaginal smears', *Science* **95**(2469), 438–439.

Paragios, N. (2003), User-aided boundary delineation through the propagation of implicit representations, *in* 'Medical Image Computing and Computer-Assisted Intervention - MICCAI', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 678–686.

Paragios, N., Rousson, M. & Ramesh, V. (2002), Matching distance functions: A shape-to-area variational approach for global-to-local registration, *in* 'Computer Vision - ECCV', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 775–789.

Patten Jr., S. F., Lee, J. S. & Nelson, A. C. (1996), 'NeoPath AutoPap 300 automatic Pap screener system', *Acta Cytologica* **40**, 45–52.

Peng, B., Zhang, L. & Zhang, D. (2013), 'A survey of graph theoretical approaches to image segmentation', *Pattern Recognition* **46**(3), 1020–1038.

Phansalkar, N., More, S., Sabale, A. & Joshi, M. (2011), Adaptive local thresholding for detection of nuclei in diversity stained cytology images, *in* '2011 International Conference on Communications and Signal Processing', pp. 218–220.

Phoulady, H. A., Goldgof, D. B., Hall, L. O. & Mouton, P. R. (2016), A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images, *in* '2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)', pp. 201–204.

Phoulady, H. A., Goldgof, D., Hall, L. O. & Mouton, P. R. (2017), 'A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images', *Computerized Medical Imaging and Graphics* **59**, 38–49.

Phoulady, H. A., Zhou, M., Goldgof, D. B., Hall, L. O. & Mouton, P. R. (2016), Automatic quantification and classification of cervical cancer via adaptive nucleus shape modeling, *in* '2016 IEEE International Conference on Image Processing (ICIP)', pp. 2658–2662.

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., Romeny, B. T. H., Zimmerman, J. B. & Zuiderveld, K. (1987), 'Adaptive histogram equalization and its variations', *Computer Vision, Graphics, and Image Processing* **39**(3), 355–368.

Pizer, S. M., Fletcher, P. T., Joshi, S., Thall, A., Chen, J. Z., Fridman, Y., Fritsch, D. S., Gash, A. G., Glotzer, J. M., Jiroutek, M. R., Lu, C., Muller, K. E., Tracton, G., Yushkevich, P. & Chaney, E. L. (2003), 'Deformable M-Reps for 3D medical image segmentation', *International Journal of Computer Vision* **55**(2), 85–106.

Pizer, S. M., Fritsch, D. S., Yushkevich, P. A., Johnson, V. E. & Chaney, E. L. (1999), 'Segmentation, registration, and measurement of shape variation via image object shape', *IEEE Transactions on Medical Imaging* **18**(10), 851–865.

Plissiti, M. E., Charchanti, A., Krikoni, O. & Fotiadis, D. I. (2006), Automated segmentation of cell nuclei in Pap smear images, *in* 'Proceedings of IEEE International Special Topic Conference on Information Technology in Biomedicine'.

Plissiti, M. E. & Nikou, C. (2012*a*), Cervical cell classification based exclusively on nucleus features, *in* 'Image Analysis and Recognition', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 483–490.

Plissiti, M. E. & Nikou, C. (2012*b*), 'Overlapping cell nuclei segmentation using a spatially adaptive active physical model', *IEEE Transactions on Image Processing* **21**(11), 4568–4580.

Plissiti, M. E., Nikou, C. & Charchanti, A. (2011*a*), 'Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering', *IEEE Transactions on Information Technology in Biomedicine* **15**(2), 233–241.

Plissiti, M. E., Nikou, C. & Charchanti, A. (2011*b*), 'Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images', *Pattern Recognition Letters* **32**(6), 838–853.

Plissiti, M. E., Vrigkas, M. & Nikou, C. (2015), Segmentation of cell clusters in Pap smear images using intensity variation between superpixels, *in* '2015 International Conference on Systems, Signals and Image Processing (IWSSIP)', pp. 184–187.

Pratt, W. K. (1972), 'Generalized Wiener filtering computation techniques', *IEEE Transactions on Computers* **C-21**(7), 636–641.

Raudys, S. J. & Jain, A. K. (1991), 'Small sample size effects in statistical pattern recognition: recommendations for practitioners', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(3), 252–264.

Reagan, J. W., Seidemann, I. L. & Saracusa, Y. (1953), 'The cellular morphology of carcinoma in situ and dysplasia or atypical hyperplasia of the uterine cervix', *Cancer* **6**(2), 224–235.

Riana, D., Plissiti, M. E., Nikou, C., Widyantoro, D. H., Mengko, T. L. R. & Kalsoem, O. (2015), 'Inflammatory cell extraction and nuclei detection in Pap smear images', *Int. J. E-Health Med. Commun.* **6**(2), 27–43.

Richart, R. M. (1967), 'Natural history of cervical intraepithelial neoplasia', *Clinical Obstetrics and Gynecology* **10**(4), 748–784. Symposium on Dysplasia and Carcinoma in Situ of the Cervix.

Ridler, T. W. & Calvard, S. (1978), 'Picture thresholding using an iterative selection method', *IEEE Transactions on Systems, Man, and Cybernetics* **8**(8), 630–632.

Rokach, L. (2005), *Ensemble Methods for Classifiers*, Springer US, Boston, MA, pp. 957–980.

Rother, C., Kolmogorov, V. & Blake, A. (2004), 'Grabcut: Interactive foreground extraction using iterated graph cuts', *ACM Transactions on Graphics* **23**(3), 309–314.

Roy, K., Bhattacharjee, D. & Nasipuri, M. (2020), Automated segmentation of cervical cells using mser algorithm and gradient embedded cost function-based level-set method, *in* M. Gupta, D. Konar, S. Bhattacharyya & S. Biswas, eds, 'Computer Vision and Machine Intelligence in Medical Image Analysis', Springer Singapore, Singapore, pp. 91–99.

Russell, J., Crothers, B., Kaplan, K. & Zahn, C. (2005), 'Current cervical screening technology considerations: Liquid-based cytology and automated screening', *Clinical Obstetrics and Gynecology* **1**(48), 108–119.

Sarwar, A., Sheikh, A. A., Manhas, J. & Sharma, V. (2019), 'Segmentation of cervical cells for automated screening of cervical cancer: a review', *Artificial Intelligence Review* .

Schindler, K. & Suter, D. (2008), 'Object detection by global contour shape', *Pattern Recognition* **41**(12), 3736–3748.

Sellors, J. W. & Sankaranarayanan, R. (2003), *Colposcopy and Treatment of Cervical Intraepithelial Neoplasia: A Beginners Manual*, International Agency for Research on Cancer.

Shannon, C. E. (1948), 'A mathematical theory of communication', *Bell System Technical Journal* **27**(3), 379–423.

Shi, J. & Malik, J. (2000), 'Normalized cuts and image segmentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(8), 888–905.

Shih, F. Y. (2010), *Image Processing and Pattern Recognition*, Wiley-IEEE Press.

Silva, D. V. S. X. D., Fernando, W. A. C., Kodikaraarachchi, H., Worrall, S. T. & Kondoz, A. M. (2010), 'Adaptive sharpening of depth maps for 3D-TV', *Electronics Letters* **46**(23), 1546–1548.

Silva, R. R., Araujo, F. H., Ushizima, D. M., Bianchi, A. G., Carneiro, C. M. & Medeiros, F. N. (2019), 'Radial feature descriptors for cell classification and recommendation', *Journal of Visual Communication and Image Representation* **62**, 105–116.

Singaraju, D., Grady, L. & Vidal, R. (2008), Interactive image segmentation via minimization of quadratic energies on directed graphs, *in* '2008 IEEE Conference on Computer Vision and Pattern Recognition', pp. 1–8.

Slabaugh, G. & Unal, G. (2005), Graph cuts segmentation using an elliptical shape prior, *in* 'IEEE International Conference on Image Processing 2005', Vol. 2, pp. 1222–1225.

Smith, J. H. (2011), 'Cytology, liquid-based cytology and automation', *Best Practice & Research Clinical Obstetrics & Gynaecology* **25**(5), 585–596. Colposcopy and Cervical Pathology.

Smith, R. A., Andrews, K. S., Brooks, D., Fedewa, S. A., Manassaram-Baptiste, D., Saslow, D. & Wender, R. C. (2019), 'Cancer screening in the united states, 2019: A review of current american cancer society guidelines and current issues in cancer screening', *CA: A Cancer Journal for Clinicians* **69**(3), 184–210.

Song, Y., Zhang, L., Chen, S., Ni, D., Lei, B. & Wang, T. (2015), 'Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning', *IEEE Transactions on Biomedical Engineering* **62**(10), 2421–2433.

Sonka, M., Hlavac, V. & Boyle, R. (1993), *Image Processing, Analysis and Machine Vision*, 1st edn, Chapman & Hall.

Spencer, C. C. & Bostrom, R. C. (1962), 'Performance of the Cytoanalyzer in recent clinical trials', *JNCI: Journal of the National Cancer Institute* **29**(2), 267–276.

Stegmann, M. B. & Gomez, D. D. (2002), 'A brief introduction to statistical shape analysis'. Images, annotations and data reports are placed in the enclosed zip-file.
URL: *http://www2.imm.dtu.dk/pubdb/p.php?403*

Sun, T. & Neuvo, Y. (1994), 'Detail-preserving median based filters in image processing', *Pattern Recognition Letters* **15**(4), 341–347.

Tareef, A., Song, Y., Cai, W., Huang, H., Chang, H., Wang, Y., Fulham, M., Feng, D. & Chen, M. (2017), 'Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation', *Neurocomputing* **221**, 94–107.

Tareef, A., Song, Y., Huang, H., Feng, D., Chen, M., Wang, Y. & Cai, W. (2018), 'Multi-pass fast watershed for accurate segmentation of overlapping cervical cells', *IEEE Transactions on Medical Imaging* **37**(9), 2044–2059.

Tareef, A., Song, Y., Huang, H., Wang, Y., Feng, D., Chen, M. & Cai, W. (2017), 'Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling', *Neurocomputing* **248**, 28–40.

Tareef, A., Song, Y., Lee, M., Feng, D. D., Chen, M. & Cai, W. (2015), Morphological filtering and hierarchical deformation for partially overlapping cell segmentation, *in* '2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)', pp. 1–7.

Tarjan, R. (1971), Depth-first search and linear graph algorithms, *in* '12th Annual Symposium on Switching and Automata Theory (swat 1971)', pp. 114–121.

Theodoridis, S. & Koutroumbas, K. (2008), *Pattern Recognition*, 4th edn, Academic Press, Inc., Orlando, FL, USA.

ThinPrep Imaging System (2019), `http://www.hologic.ca/products/clinical-diagnostics-and-blood-screening/instrument-systems/thinprep-imaging-system`. [Online; Accessed: 2019-11-10].

Ting, K. M. (2017), *Confusion Matrix*, Springer US, Boston, MA, pp. 260–260.

Tolles, W. E. & Bostrom, R. C. (1956), 'Automatic screening of cytological smears for cancer: The instrumentation', *Annals of the New York Academy of Sciences* **63**(6), 1211–1218.

Tomasi, C. & Manduchi, R. (1998), Bilateral filtering for gray and color images, *in* 'Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)', pp. 839–846.

Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E. & Willsky, A. (2003), 'A shape-based approach to the segmentation of medical imagery using level sets', *IEEE Transactions on Medical Imaging* **22**(2), 137–154.

Tsai, M.-H., Chan, Y.-K., Lin, Z.-Z., Yang-Mao, S.-F. & Huang, P.-C. (2008), 'Nucleus and cytoplast contour detector of cervical smear image', *Pattern Recognition Letters* **29**(9), 1441–1453.

Tucker, J. (1976), 'Cerviscan: An image analysis system for experiments in automatic cervical smear prescreening', *Computers and Biomedical Research* **9**(2), 93–107.

Tucker, J. & Husain, O. (1981), 'Trials with the cerviscan experimental prescreening device on polylysine-prepared slides', *Analytical and quantitative cytology* **3**(2), 117–120.

Urquhart, R. (1982), 'Graph theoretical clustering based on limited neighbourhood sets', *Pattern Recognition* **15**(3), 173–187.

Ushizima, D. M., Bianchi, A. G. C. & Carneiro, C. M. (2015), 'Segmentation of subcellular compartments combining superpixel representation with Voronoi diagrams'. [Online; Accessed: 2019-11-10].
URL: *http://cs.adelaide.edu.au/∼carneiro/isbi14_challenge/results_release.html*

van Driel-Kulker, A. & Ploem-Zaaijer, J. (1989), 'Image cytometry in automated cervical screening', *Analytical cellular pathology : the journal of the European Society for Analytical Cellular Pathology* **1**(1), 63–77.

Vapnik, V. & Lerner, A. (1963), 'Pattern recognition using generalized portrait method', *Automation and Remote Control* **24**(3), 774–780.

Vedaldi, A. & Soatto, S. (2008), Quick shift and kernel methods for mode seeking, *in* D. Forsyth, P. Torr & A. Zisserman, eds, 'Computer Vision - ECCV', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 705–718.

Veksler, O. (2008), Star shape prior for graph-cut image segmentation, *in* 'Computer Vision - ECCV', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 454–467.

Verma, H., Agrawal, R. & Sharan, A. (2016), 'An improved intuitionistic fuzzy c-means clustering algorithm incorporating local information for brain image segmentation', *Applied Soft Computing* **46**, 543–557.

Vincent, L. (1993), 'Morphological grayscale reconstruction in image analysis: applications and efficient algorithms', *IEEE Transactions on Image Processing* **2**(2), 176–201.

Vu, N. & Manjunath, B. S. (2008), Shape prior segmentation of multiple objects with graph cuts, *in* '2008 IEEE Conference on Computer Vision and Pattern Recognition', pp. 1–8.

Wählby, C., Sintorn, I.-M., Erlandsson, F., Borgefors, G. & Bengtsson, E. (2004), 'Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections', *Journal of Microscopy* **215**(1), 67–76.

Wang, H., Zhang, H. & Ray, N. (2013), 'Adaptive shape prior in graph cut image segmentation', *Pattern Recognition* **46**(5), 1409–1414.

Wang, J., Kong, J., Lu, Y., Qi, M. & Zhang, B. (2008), 'A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints', *Computerized Medical Imaging and Graphics* **32**(8), 685–698.

Wang, Z., Song, Q., Soh, Y. C. & Sim, K. (2013), 'An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation', *Computer Vision and Image Understanding* **117**(10), 1412–1420.

Watanabe, S. (1974), 'An automated apparatus for cancer prescreening: CYBEST', *Computer Graphics and Image Processing* **3**(4), 350–358.

WHO — Cervical cancer (2019), `http://www.who.int/cancer/prevention/diagnosis-screening/cervical-cancer/en/`. [Online; Accessed: 2019-11-10].

Win, K. Y., Choomchuay, S., Hamamoto, K. & Raveesunthornkiat, M. (2018), 'Detection and classification of overlapping cell nuclei in cytology effusion images using a double-strategy random forest', *Applied Sciences* **8**(9).

Wu, H. S., Gil, J. & Barba, J. (1998), 'Optimal segmentation of cell images', *IEE Proceedings - Vision, Image and Signal Processing* **145**(1), 50–56.

Wu, Z. & Leahy, R. (1993), 'An optimal graph theoretic approach to data clustering: theory and its application to image segmentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(11), 1101–1113.

Xing, F. & Yang, L. (2016), 'Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review', *IEEE Reviews in Biomedical Engineering* **9**, 234–263.

Yang-Mao, S. F., Chan, Y. K. & Chu, Y. P. (2008), 'Edge enhancement nucleus and cytoplast contour detector of cervical smear images', *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **38**(2), 353–366.

Zahn, C. T. (1971), 'Graph-theoretical methods for detecting and describing Gestalt clusters', *IEEE Transactions on Computers* **C-20**(1), 68–86.

Zahniser, D. J. & Hurley, A. A. (1996), 'Automated slide preparation system for the clinical laboratory', *Cytometry* **26**(1), 60–64.

Zahniser, D. J., Oud, P. S., Raaijmakers, M. C., Vooys, G. P. & de Walle, R. T. V. (1979), 'BioPEPR: a system for the automatic prescreening of cervical smears', *Journal of Histochemistry and cytochemistry* **27**(1), 635–641.

Zhang, J., Liu, Z., Du, B., He, J., Li, G. & Chen, D. (2019), 'Binary tree-like network with two-path fusion attention feature for cervical cell nucleus segmentation', *Computers in Biology and Medicine* **108**, 223–233.

Zhang, L., Chen, S., Wang, T., Chen, Y., Liu, S. & Li, M. (2011), A practical segmentation method for automated screening of cervical cytology, *in* '2011 International Conference on Intelligent Computation and Bio-Medical Instrumentation', pp. 140–143.

Zhang, L., Kong, H., Chin, C. T., Liu, S., Chen, Z., Wang, T. & Chen, S. (2014), 'Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts', *Computerized Medical Imaging and Graphics* **38**(5), 369–380.

Zhang, L., Kong, H., Liu, S., Wang, T., Chen, S. & Sonka, M. (2017), 'Graph-based segmentation of abnormal nuclei in cervical cytology', *Computerized Medical Imaging and Graphics* **56**, 38–48.

Zhang, L., Kong, H., Ting Chin, C., Liu, S., Fan, X., Wang, T. & Chen, S. (2014), 'Automation-assisted cervical cancer screening in manual liquid-based cytology with hematoxylin and eosin staining', *Cytometry Part A* **85**(3), 214–230.

Zhang, L., Sonka, M., Lu, L., Summers, R. M. & Yao, J. (2017), Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei, *in* '2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI)', pp. 406–409.

Zhao, L., Li, K., Wang, M., Yin, J., Zhu, E., Wu, C., Wang, S. & Zhu, C. (2016), 'Automatic cytoplasm and nuclei segmentation for color cervical smear image using an efficient gap-search MRF', *Computers in Biology and Medicine* **71**, 46–56.

Zhu-Jacquot, J. & Zabih, R. (2007), Graph cuts segmentation with statistical shape priors for medical images, *in* '2007 Third International IEEE Conference on Signal-Image Technologies and Internet-Based System', pp. 631–635.