IMAGE PROCESSING AS APPLIED TO MEDICAL DIAGNOSTICS

by

KRISTINE A. THOMAS

A THESIS

Presented to the Department of Computer and Information Science
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Master of Science

June 2010

"Image Processing as Applied to Medical Diagnostics," a thesis prepared by Kristine A. Thomas in partial fulfillment of the requirements for the Master of Science degree in the Department of Computer and Information Science. This thesis has been approved and accepted by:

_____

Dr. John Conery, Chair of the Examining Committee

_____JUNE 1, 2010_____

Date

Committee in charge: Dr. John Conery, Chair
                         Dr. Matthew J. Sottile

Accepted by:

_____

Dean of the Graduate School

An Abstract of the Thesis of

Kristine A. Thomas   for the degree of   Master of Science

in the Department of Computer and Information Science

to be taken   June 2010

Title: IMAGE PROCESSING AS APPLIED TO MEDICAL

DIAGNOSTICS

Approved: ———————————————————————————

Dr. John Conery, Chair

Image processing is a powerful tool for increasing the reliability and reproducibility of disease diagnostics. In the hands of pathologists, image processing provides quantitative data from histological images which supplement the qualitative data currently used by specialists. This thesis presents a novel method for analyzing digitized images of hematoxylin and eosin (H&E) stained histology slides to detect and quantify inflammatory polymorphonuclear leukocytes to aid in the grading of acute inflammation of the placenta as an example of the use of image processing in aid of diagnostics.

Methods presented in this thesis include segmentation, a novel threshold selection technique and shape analysis. The most significant contribution is the automated color threshold selection algorithm for H&E stained histology slides which is the only unsupervised method published to date.

CURRICULUM VITAE

NAME OF AUTHOR: Kristine A. Thomas

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

>   University of Oregon
>   Lane Community College
>   Linfield College

DEGREES AWARDED:

>   Master of Science in Computer and Information Science,
>   2010, University of Oregon
>
>   Bachelor of Science in Mathematics, Computer Science minor,
>   2008, University of Oregon

AREAS OF SPECIAL INTEREST:

>   Medical Image Processing
>   Computer Aided Diagnosis

PROFESSIONAL EXPERIENCE:

>   Graduate Research Fellow, Department of Computer and Information
>   Science, University of Oregon, 2008 - 2010
>
>   Undergraduate Research Assistant, Department of Computer and
>   Information Science, University of Oregon, 2007 - 2008

AWARDS AND HONORS:

>   Upsilon Pi Epsilon, Honor Society for the Computing Sciences, 2010
>   Ford Scholar, 2009 – 2010.

PUBLICATIONS:

K.A. Thomas, M.J. Sottile, C.M. Salafia, Unsupervised segmentation for inflammation detection in histopathology images, in *Lecture Notes in Computer Science, Springer Verlag*, 6134:541–549, 2010.

ACKNOWLEDGMENTS

It is a pleasure to thank those who made this thesis possible. I owe my deepest gratitude to my research advisor, Dr. Matthew J. Sottile, whose kind encouragement, remarkable patience, timely advice, brilliant insights and gentle guidance made this thesis a possibility. There are not enough kind words to express my appreciation for his involvement in my education and research.

I wish to thank Dr. John S. Conery for stepping into the position of Chair and believing in my ability to complete my research and produce this thesis.

I want to express my gratitude to Dr. Carolyn M. Salafia, MD, and Placental Analytics, LLC, for their financial support throughout my graduate career as well as the data and expert knowledge that went into this thesis.

I am grateful for Katy J. Kaiser's keen eye and willingness to read and re-read my drafts for spelling and grammatical errors. Without her help, this thesis would have been much more rough around the edges.

Finally, I would like to thank my family and friends for the unbelievable amount of support and love shown me throughout my graduate school journey.

TABLE OF CONTENTS

LIST OF FIGURES

Figure   Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

In an age of mass digitizing of cellular imaging, image processing is clinically relevant in its application to microscopic tissue images. Patients and medical professionals alike have great interest in the development of methods which are reliable and reproducible. Thus, in a field which traditionally has relied on the trained eye of a specialist to make diagnoses from a qualitative perspective, a transition is underway: computer-aided diagnosis is now possible with digitized imaging. To quote Carolyn M. Salafia in [1], "perinatal researchers should, to use a worn out but appropriate cliché, step outside the box and consider alternative approaches to both measurement of histology slides that would yield adequate reliability to allow cross-institutional analysis of the latent construct(s) involved in intraamniotic infection and ultimately to achieve a fuller understanding of the infection-preterm birth pathway." By developing algorithmic methods which are reliable, reproducible and unsupervised, comes the presentation of a powerful tool which aids in the collection of data, assist researchers to further their understanding of the condition, and ultimately help with the critical diagnosis of chorioamnionitis.

1.1 Contributions

This thesis makes two major contributions to the fields of computer science and computer-aided medical diagnostics. First, it introduces a novel method for unsupervised segmentation of Hematoxylin & Eosin (H&E) stained histology slides.

Second, it creates a tool for the quantitative analysis of human inflammatory response through the fusion of the innovative segmentation technique with established image analysis processes. Application of the unsupervised segmentation technique results in the successful analysis of inflammation in histology slides for which quantitative inferences are not possible through current computer-aided diagnostic procedures. Although this thesis applies unsupervised segmentation exclusively to the field of medical diagnostics, this new technique is also a significant contribution to the field of computer science. The following explains the contributions to each field.

1.1.1 Computer Science Contribution

Unsupervised image segmentation aids in the analysis of H&E stained histology slides to determine the severity of tissue inflammation. The technique has clear significance for medical diagnostics, but because it advances image processing tools as well, the unsupervised segmentation technique makes significant contributions to computer science. Automated image segmentation is an open area of research because current methods rely on user input. The users in this context are unlikely to be experts in image processing and are instead experts in the field of medicine. This technique is a ready-to-use tool that doesn't require in-depth training or time-consuming tuning of parameters for each image. The unsupervised and computerized segmentation of cell nuclei produces accurate data about the nuclei which then makes possible accurate and reliable diagnosis of disease.

1.1.2 Medical Diagnostic Contribution

The contributions to medical diagnostics are evident in the need to supplement the qualitative analysis of experienced experts with reliable and reproducible computer-aided analysis. The methods described in this thesis are intended for the identification and quantification of inflammatory polymorphonuclear leukocytes in placental tissue, which are linked to infant

mortality and morbidity. The introduction of quantitative analysis to medical diagnostics makes possible the collection of data that is currently unavailable. These methods extend beyond placental tissue and in fact apply to a wide variety of tissue stained with H&E. Contributions to background and blood removal methods are applicable to research currently being conducted on neuroblastoma [2] and the identification of squamous epithelium in the case of cervical cancer [3]. Contributions to nuclei segmentation apply to research in any area where pathology causes changes to nuclei, like in cases of necrosis, where there is a lack of nuclei, or in cases of neurofibroma, where floret-like giant cells abound [4]. Identification of inflammatory polymorphonuclear leukocytes is relevant to any condition where inflammation is one of the key features in making a diagnosis such as auto-immune and infectious diseases. All of these contributions together yield a tool for computer-aided diagnosis of chorioamnionitis.

1.2   Organization of Thesis

The remainder of this thesis is organized as follows: chapter two gives background information in the areas of human and computer vision and provides information regarding image processing as well as pathology. Chapter three reviews similar research related to medical image processing. Chapter four presents an algorithm that isolates and accounts for inflammatory polymorphonuclear leukocytes in samples of H&E stained tissue, which is crucial to the diagnosis of acute chorioamnionitis. In conclusion, chapter five summarizes this thesis' findings and suggests areas of future research.

CHAPTER II

BACKGROUND

This section provides brief background information from the disciplines which contribute to this thesis: biomedicine and computer science. First presented is the difference between human vision and computer vision both in terms of perception and presentation of image data. Next introduced are topics related to image processing: image data structures, color space, segmentation, thresholding and shape analysis. Finally covered is a brief description of pathology methods, including staining, image acquisition, placental anatomy and characteristics of acute chorioamnionitis.

2.1 Human Perception vs. Computer Vision

Computer vision and human vision appear to have the same function and goal. Both systems interpret multidimensional spatial data for the purpose of information gathering and visual recognition. Due to the complexity of the human visual system and the lack of total understanding of the neural aspects of human perception, even the best computer vision system cannot replicate the human eye. The development of computer vision techniques is focused, then, on improving upon the human vision model rather than replicating it. This section explains basic computer vision systems and how they operate in comparison to the human vision system.

According to Fu and Mui in [5], "the image segmentation problem is one of psychophysical perception and therefore is not susceptible to a purely analytical solution." This view is upheld by many visual illusions which trick the brain into seeing something different than what is actually in the image data, like in the Adelson checkerboard shadow illusion, figure 1(a) [6]. The tile A belongs to the set of darker squares while B belongs to the set of lighter squares. In terms of the intensity value, both tiles are the same shade of gray but they are perceived as different shades because the brain compensates for the shadow. The proof in figure 1(b) superimposes two solid gray lines across the image. The lines are again the same gray value as tile A and tile B. Despite the addition of these two solid lines, the shadow still appears to altar the shade of grey. The easiest way to see that the squares are indeed of uniform color is to remove the context. Only with the removal of the background, in figure 1(c), it becomes clear that both tile A and tile B as well as the solid gray lines are the same shade of gray.

(a)   (b)   (c)
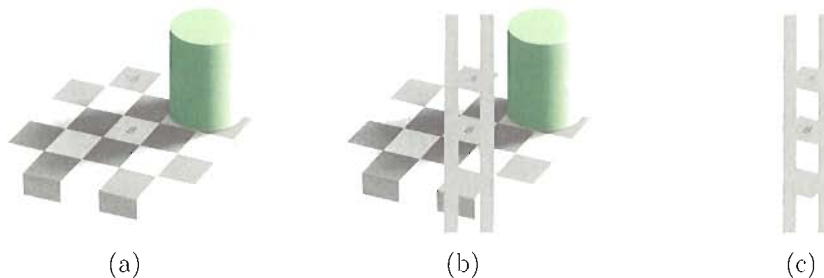
FIGURE 1: Adelson's checker-shadow illusion.

Vision is the ability to focus selectively on objects of interest. The human eye can discern thousands of color shades and about two-dozen shades of gray. Human vision is a complex system that senses and acts on visual stimuli in the form of light. There are nearly 100 million photoreceptors dispersed around the retina [7]. There are two types of photoreceptors: *rods*, which are used for light and dark (black and white) vision; and *cones*, which are used for color vision. There are approximately 10 million cones, most of which fall within 5 degrees of the *fovea*, or the center of gaze [8]. Visual acuity is greatest within 1 degree of the fovea. Within

the entire visual spectrum, rods are more sensitive to light than the cones, but they are entirely absent from the fovea and become more prominent in the periphery [9].

Three types of cones account for the ability to see color over the entire visual spectrum: α sense light toward the blue region of the visual spectrum, β sense toward the green region of the spectrum, and γ sense light toward the red region of the spectrum. α cones are far fewer in number than β and γ cones, but the visual system still perceives blue through neurological compensation.

Visual perception is logarithmic and depends on brightness adaption [7]. In dark conditions images are formed on the rods and in brighter conditions images are formed on the cones. This means that in darker conditions the capacity for perceiving color diminishes and, because the image is sensed further from the fovea, visual acuity diminishes as well.

Physicist Ernst Mach discovered a phenomenon which demonstrates how the human brain compensates for loss of acuity by emphasizing boundaries between regions of differing intensities. The demonstration, called Mach bands, seen in figure 2(a), consists of several stripes represented by constant intensity in increasing or decreasing order. Although each stripe is of a consistent intensity, in the image they appear to change slightly from light to dark. In other words, there appears a faint lighter border to one side and a faint darker border to the other side of each of the internal boundaries. The gradation within each stripe does not exist, but rather is perceived by the human eye: figure 2(b) shows the actual intensity levels and figure 2(c) illustrates of the perceived intensity values. The visual perception of Mach bands at boundaries between intensity levels is thought to assist in the discernment of boundaries to make up for decreased visual acuity. This phenomenon is particularly important in the field of medical diagnosis, as a gradient in intensity may appear in an image where the image data does not support this observation.

Neural signals from sight are transmitted to two areas of the brain for processing: the *associative cortex* processes object information and the *occipital cortex* processes pattern information. When information is missing from an image the brain easily repairs the incomplete image with knowledge of similar patterns
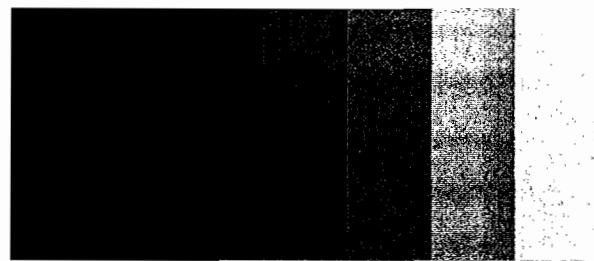
(a)
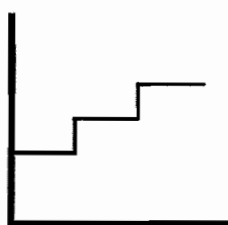
(b) True intensity in mach bands.
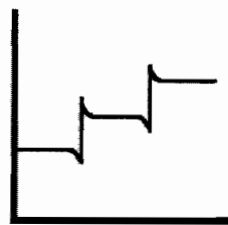
(c) Perceived intensity in mach bands.

FIGURE 2: Mach band example.

seen in the past. By concentrating on different components of an image, the brain to perceives different images, like in the Kanizsa triangle of figure 3. The Kanizsa triangle can be perceived either as a triangle occluding three circles or three Pac-Man-like figures facing each other [10]. The Pac-Man images are easily recognized with computer vision, but the triangle, which does not exist in the image data and is inferred with the help of the occipital cortex, is not. Structural information theory suggests that the human brain prefers to perceive the triangle because the shape is less complex and therefore has a lower information load than the Pac-Man images [11].

Human vision, which is augmented by a brain that has been specialized to ensure the survival of humankind, may present different perceptions and images than computer vision, which depends upon an algorithm designed for the processing of discrete pixel information contained in image data. The use of experts, or the human eye, makes for subjective grading criteria where inferences are based on experience which may include information not actually present in the image data.
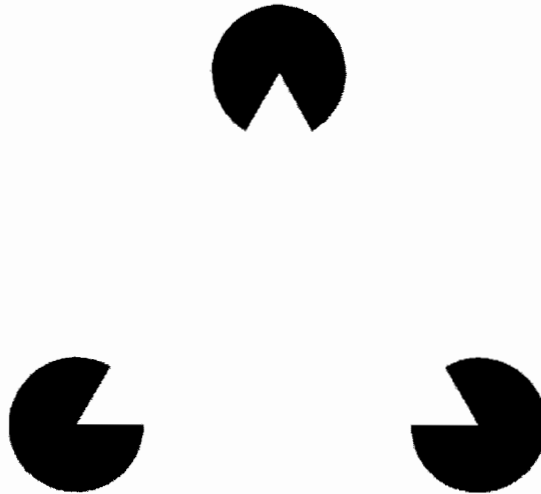
FIGURE 3: Kanizsa triangle.

This may be an advantage or disadvantage depending on the experts and their domain knowledge. In any case, augmenting subjective information with quantitative data from the diagnostic imaging provides the expert with objective information for arriving at a diagnostic conclusion.

2.2 Introduction to Image Processing

Most image processing and computer vision techniques are implemented in software. In general, image processing is employed either to enhance visual appearance or to prepare images for quantitative measurement of features for object recognition [12]. This thesis concentrates on the latter.

2.2.1 Image Data Structures

Computer images are comprised of a set of points or *picture elements*, usually referred to as *pixels*, stored as an array of numbers. Images are spacial data indexed by two spacial coordinates; typically the variables x and y refer to the horizontal and vertical axes of an image. Pixel value represents the color or intensity of each pixel, and the placement of the pixels within the matrix
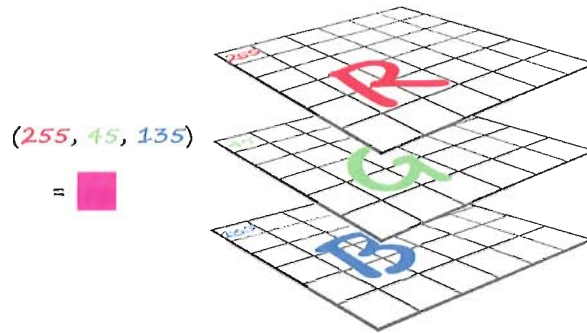
FIGURE 4: RGB color image represented by three matrices.

corresponds to their placement within the image. If more than one value is required to encode pixel information, the image is often represented by a multidimensional matrix. For example, an RGB encoding of an image would contain 3 matrices: one each for red, green and blue intensity as shown in figure 4. In other terms, each pixel represented in the matrix has a value that is encoded as either a scalar (in the case of gray-scale) or a vector (in the case of color). In figure 4, the RGB value $(255, 45, 135)$ produces a pink pixel.

2.2.2 Color Space

Because manipulation of color is possible through additive properties (for example, red and green produce yellow) a wide range of colors is generated from a choice of three primary colors. A red, blue and green (RGB) color space is commonly used in modern displays like televisions, computer monitors and digital cameras. Each primary color has a range of values dependent on bit resolution. Most digital systems store color channels in 8-bit quantities, allowing a range from 0 to 255 to indicate the intensity of a color. In the cases of white, black and shades of gray, white is represented by maximum intensity over all three channels, black is represented by nil intensity over all three channels and shades of gray are represented by equal intensities over all three channels. This is referred to as 24-bit RGB values or *truecolor* and allows for 16,777,216 (256^3) different colors.
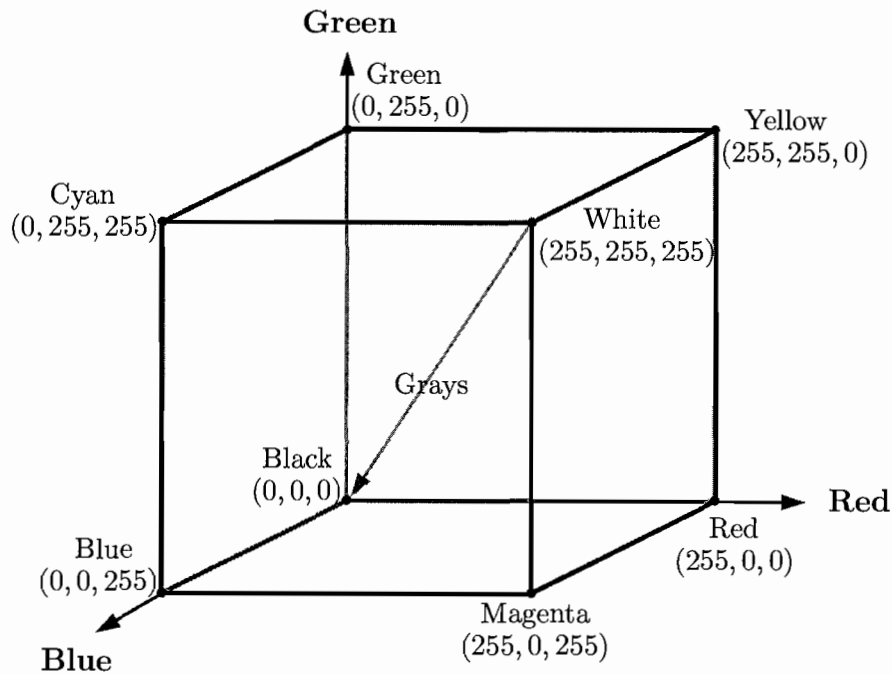
FIGURE 5: RGB color space represented by a cube.

The RGB color space is represented via a three-dimensional cube as seen in figure 5. The coordinates of each point inside the cube represent the values of each primary color. There are three laws of colorimetry: first, the creation of any color is possible with the three primary colors and any given combination of the primary colors is unique; second, any two colors which are equivalent are again equivalent after multiplying or dividing all three components by a scalar; and three, the luminance of a mixture of colors is equal to the sum of the luminance of each color component [13].

2.2.3 Spacial Resolution

Samples of light from a finite number of sensors are stored in digital form to create digital images. Since there is a finite set of sensors, a digital copy of an image is, by definition, degraded. The number of sensors determines the image size (N) and must be sufficiently large to resolve the spacial detail of an image while also being sufficiently small to accommodate efficient memory management and

(a)   (b)   (c)
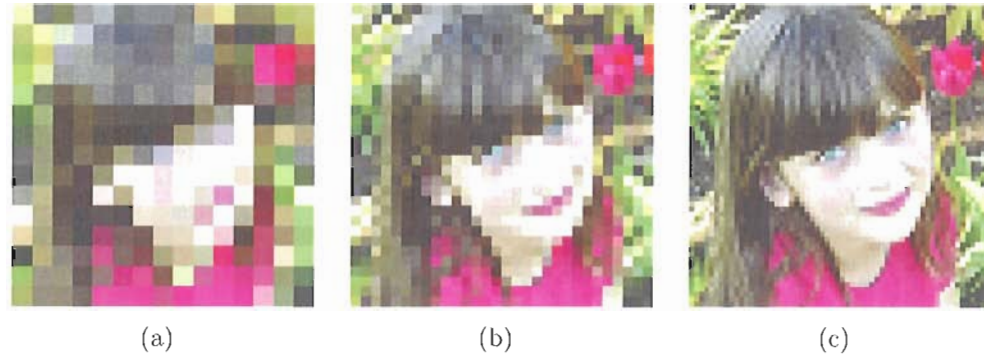
FIGURE 6: Spacial resolution example.

processing time. For example, figure 6 shows the different image resolutions resulting from using different values of N. Figure 6(a) is a 15×15 pixel image which shows only basic structure. The subjects facial features are indecipherable and identification of the pink object in the upper right side of the image is impossible. Figure 6(b), a 30×30 pixel image, begins to show more detail, but identification of the subjects identity or of the type of pink flower in the background is still difficult. Figure 6(c) is a 60×60 pixel image with a much higher level of detail. While the image still suffers from *pixelization*, the subject is identifiable and the type of flower in the background is recognizable.

Determining the appropriate level of spacial resolution is accomplished through *sampling criterion*. For image processing, a sampling of density or of frequency based upon classical signal theory is appropriate. According to the Nyquist sampling theorem, the sampling frequency must be at least twice the highest frequency of the sampled signal to reconstruct a signal from its samples [14]. For image analysis. however, the sampling density should be based upon criteria such as desired accuracy and precision [15].

2.2.4   Segmentation

Image segmentation is a classification problem and therefore is one of the most important and difficult tasks in image processing. Segmentation is the process of partitioning an image into mutually exclusive regions based on pixel
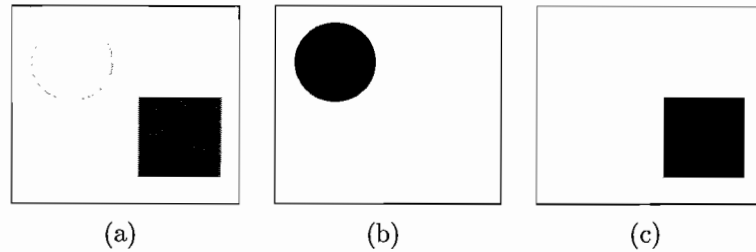
FIGURE 7: Thresholding example.

characteristics. The goal of image segmentation is to assign each pixel to a group or class with similar characteristics. Ideally, classes with similar characteristics correspond to similar objects in an image. In many cases image segmentation remains an unsolved problem [16, 17, 18].

The simplest form of segmentation is *binarization*. Binarization is the process of segmenting an image into 2 classes, such as foreground and background, thus distinguishing the object(s) of interest from the rest of the image. The binary image is referred to as a *mask* when it is used to exclude pixel data from calculations.

Thresholding

Thresholding is an image segmentation technique that classifies portions of the image according to a range of values. Thresholding is based on homogeneity among features. Features are accepted or rejected based on whether or not their values fall within the expected range of values. Figure 7 is an example of grayscale thresholding where figure 7(a) is the original image. Figure 7(b) is the result of thresholding based on a higher intensity range which results in the binarization of the lighter circle from the original image. Similarly, figure 7(c) is the result of thresholding based on a lower intensity range which results in the binarization of the darker square from the original image.

Color thresholding is an extension of grayscale thresholding where the grayscale techniques are extended to each of the color channels based on the desired effect. For example, the color red in RGB space is represented by a high R value

with low G and B values. Therefore, a high threshold range to the R channel and low threshold ranges to the G and B channels effectively thresholds the color red.

Shape Analysis

Shape analysis in image processing is largely dependent on mathematical geometric features. Any shape description scheme should be invariant to changes in translation, scale and rotation. Such features are easily grasped by the human eye and brain but are not easily captured in an algorithm. These features can be described as bent, elongated, circular, convex, sharp, smooth, symmetrical, etc. The method of description is dependent on the particular application. Consistent evaluation criteria do not yet exist for methods of shape description. A common way to explain to an image processing system the variety of shapes in biological preparations is presently not available [19].

2.3   Introduction to Pathology and Histology

Pathology is a study of diseases. Pathologists study the cause and development of disease and determine the difference between healthy tissue and tissue that is diseased or damaged. Histology, the study of microscopic structures in tissue, is a major tool employed by pathologists to determine specific diagnoses. Treatment for many diseases is dependent on a pathologist's ability to make a definitive diagnosis. Diagnosis is often based on a grading scale measuring either the degree of abnormality of the cells or presence of immuno-responsive cells. Many of these grading scales are subject to the expertise of the pathologist and vary by examiner.

2.3.1   Stains

Pathologists observe the cellular structure of tissue to determine if it is infected, mutated, necrotic or otherwise damaged. To accomplish this, pathologists look at biological samples under the microscope. The cellular structure of samples

obtained from biological sources does not contrast well under normal light. Staining creates contrast between objects, which is necessary to observe features of interest such as cell membranes, organelles, nuclei or specific molecules such as nucleic acids, lipids or carbohydrates. Staining method selection is dependent on what is observed within the sample and what microscopy equipment is available. Methods include the utilization of a wide range of stains which are viewed under ultraviolet light, including florescence labels, and various stains that can be viewed on the bright field, such as hematoxylin and eosin (H&E). H&E is the most prevalently used staining methodology in pathology due to its availability and low cost. In the H&E method, the alkaline hematoxylin colors cell nuclei blue while the acidic eosin colors cytoplasm, connective tissues and other organelles different shades of pink.

2.3.2 Placenta

During pregnancy the fetus is housed within the placenta. The placenta is created during the first trimester and becomes fully functional in the second trimester. The placenta conducts waste management for the fetus and provides oxygen, nutrition, hormones and antibodies. As shown in figure 8, the placenta is composed of parts from both the fetus and the mother. The placenta is connected to the mother via the uterine wall and to the fetus via the umbilical cord, which inserts into the chorionic plate. The placenta has two fetal membranes. The outer membrane, the chorion, supports the thin inner layer, the amnion. The amnion is in direct contact with the amniotic fluid, which envelopes the fetus. The placenta is the life support system for the fetus during gestation. When the placental tissue is impacted by defect or disease, then, the health of the fetus is at risk.

2.3.3 Acute Chorioamnionitis

Acute chorioamnionitis is the typical pattern of inflammatory changes in the chorionic plate, umbilical cord, and membranes in response to microorganisms in the amniotic fluid [21]. This disorder is the most frequent cause of preterm labor
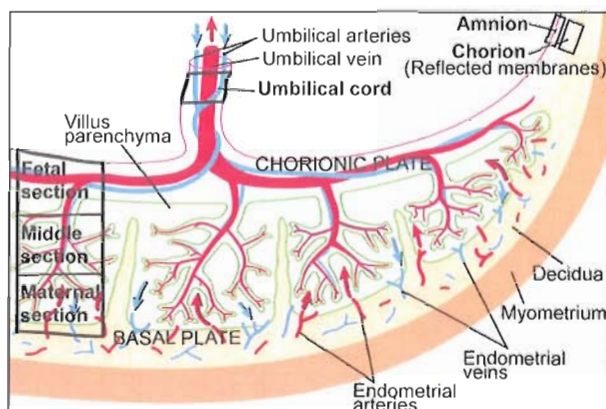
FIGURE 8: Anatomy of the placenta as presented in [20].

and is associated with abnormal fetal heart rates, developmental problems in childhood, and early termination of the pregnancy [21, 22, 23, 24]. Very often acute chorioamnionitis is repeated in successive pregnancies [25].

The *clinical criterion* for diagnosis of chorioamnionitis is based on the presentation of a pregnant woman with premature rupture of membranes, fever, and one or more of the following during labor: maternal tachycardia, fetal tachycardia, maternal leukocytosis (more than 11,000/mm white blood cell count) or foul smelling amniotic fluid. In contrast, the *histologic criterion* for diagnosis of chorioamnionitis is based on the presentation of maternal neutrophils beneath the chorionic plate of the placenta. Although clinical diagnosis correlates poorly with histologic diagnosis, it remains the gold standard [21, 26]. In a study by Redline, et al. [25] of infants presenting with very low birth rate and with histologic chorioamnionitis, 72% (52 out of 72), failed to reach the clinical diagnostic threshold (see table 1). In other words, in 72% of the cases with histologically verified infection, the patient presented without symptoms and therefore was not clinically diagnosed. Conversely, 17% (4 out of 24) of cases diagnosed with clinical chorioamnionitis had no histologic inflammation in the placenta, suggesting that the symptoms were attributable to something other than chorioamnionitis. This indicates that clinical diagnosis, the current "gold standard," is a poor indication of true infection. The accuracy of chorioamnionitis diagnoses is important due to the

Histologic Inflammatory Response	Histologic Chorio-amnionitis	N	Clinical Diagnosis Positive	Clinical Diagnosis Negative	Clinical Histological Agreement
none	negative	47	4 (9%)	43 (91%)	43 (91%)
maternal only	positive	14	1 (7%)	13 (93%)	1 (7%)
maternal and mild-moderate fetal	positive	41	6 (15%)	35 (85%)	6 (15%)
maternal and severe fetal	positive	17	13 (76%)	4 (24%)	13 (76%)
Total:		119	24 (20%)	95 (80%)	63 (53%)

TABLE 1: Clinical vs. histological diagnosis of chorioamnionitis [25].

risk of repetition in successive pregnancies. When responsible for preterm birth in one pregnancy, 46% of successive pregnancies also present with acute chorioamnionitis. If Redline's study is indicative of the very low birth rate deliveries in the population at large, 72% of the patients with chorioamnionitis fail to be diagnosed and have a 46% chance of reoccurrence in the next pregnancy. Therefore, given a pregnancy that ended early due to chorioamnionitis, there is a 32% chance that the mothers next pregnancy will end early due to chorioamnionitis and that the mother will not be warned beforehand. For any mother whose pregnancy has met an untimely end, this is not very reassuring, especially considering that a treatment of antibiotics is an effective preventative measure if the antibiotics are introduced sufficiently early.

Chapter four presents an image processing tool intended to increase the reliability of chorioamnionitis diagnoses. With a reliable tool, the routine examination of placental tissue after premature termination of pregnancy will be

critical in the detection of chorioamnionitis. This will provide the mother and her doctors with the best information to make decisions regarding future pregnancies.

2.4 Conclusion

As digital diagnostic tools are developed to aid in expert analysis it is important to understand the subtleties of human vision versus computer vision. Color space, segmentation, and shape analysis are integral to understanding the tool presented in chapter four. To fully appreciate the impact of its implementation requires an understanding of the physical structure and nature of the placenta and chorioamnionitis. For these reasons, the background presented here will aid in the understanding of the related works presented in chapter three and the contributions presented in chapter four.

CHAPTER III

RELATED WORK

Much of the literature related to the automated analysis of biological tissues is based on supervised techniques that require the user to manually specify parameters such as thresholds and iteration counts. This is detrimental to the acceptance of said techniques by pathologists with limited (if any) familiarity with the meaning of these parameters and their underlying algorithms. Application of image processing techniques to the detection and quantification of inflammatory polymorphonuclear leukocytes (PMLs) is an area that is not well explored or published, however a rich body of literature does exist related to image analysis of stained tissue for diagnostic purposes. The following sections give an overview of current image analysis research in the area of medical diagnostics. A summary of the methods covered in this chapter appear in table III.

3.1   Image Analysis for Diagnostics

Identification of regions of interest (ROIs) is the first step toward the development of computer aided diagnostics (CAD). This is the main focus of Wang, et al. in [3]. The techniques for processing cervical tissue slides covered in [3] include developing a scheme to process the entire slide, detecting edges, extracting features, and classifying tissue as stroma or squamous epithelium. Wang and his team processed a total of twenty tissue slides with mixed results. Some samples were classified correctly with around 95% accuracy; other samples suffered from

Source	Nature of the Algorithm	Application
Deng, Manjunath & Shin, 2001	class map & region growing (JSEG)	general purpose automated segmentation
Hafiane, Bunyak & Palaniappan, 2008	fuzzy clustering & active contours	H&E stained prostate tissue (cancer)
Kong, Sertel, Shimada, Boyer, Saltz & Gurcan, 2009	color clustering based segmentation (EMLDA)	H&E stained neuroblastoma (cancer)
Naik, Doyle, Feldman, Tomaszewski & Maldabhushi, 2007	Bayesian classifier based on color & morphology	H&E stained prostate tissue (cancer)
Otsu, 1979	histogram thresholding	general purpose automated thresholding
Sertel, Kong, Shimada, Catalyurek, Saltz & Gurcan, 2009	modified k-nearest & neighbor classifier based on texture features	H&E stained neuroblastoma (cancer)
Tosun, Kandemir, Sokmensuer, & Dunduz-Demir, 2009	k-means, region growing & merging	H&E stained colon tissue (cancer)
Wang, Turner, Crooks, Diamond & Hamilton, 2007	edge detection & texture based region segmentation	H&E stained cervical tissue (cancer)
Zheng, Chang & Gur, 1995	region growing & shape based segmentation	mammogram breast tissue

TABLE 2: Summary of related publications.

misclassifications due to morphologically similar features, like red blood cells co-occurring with squamous epithelium within the tissue slide. Because they only classified slide regions as background, stroma and squamous epithelium without a category for red blood cells, the red blood cells were classified as squamous epithelium. The image processing technique presented in section 4.2.1 resolves this dilemma by removing red blood cells from the image data prior to feature extraction and classification.

Hafiane's research in [27] focuses on cell nuclei counting on H&E stained slides of prostate tissue using fuzzy clustering and active contours. The goal of their technique was to count the number of nuclei within a known diagnosis rather than to infer a diagnosis from spacial distribution of nuclei. Measuring the success of their algorithm proved difficult because "there was some degree of inconsistency in the quality of the ground truth across experts in identifying indistinct nuclei, so that some of the false positives detected by the algorithm may indeed be correct." The researchers make no mention of how many experts were used and how the experts compare amongst themselves. Hafiane, et al. state that further interaction with experts is necessary to accurately quantify the performance of their method. Gaining consensus among experts who can then verify a reliable and reproducible method is a common problem within the domain of medical diagnosis and also was experienced by [28, 29, 30, 31, 32].

In [33], Naik, et al. presents a method for classifying Gleason grades among prostate cancer cases for 3 classes (benign, Gleason grade three, and Gleason grade four) with the intention of developing an automated system that reduces inter- and intra-observer variability. Their methods include segmentation, morphological feature extraction, and manifold learning. According to [33], variability in appearance across the grades gives rise to difficulty with automated segmentation of gland structures. They quote an inter-observer variability up to 48% for Gleason grades three and four. Naik and his team claim to achieve accuracies of 95% when distinguishing between grades three and four, but they do not specify how those results are verified.

In [34], Tosun, et al. made the observation that when colon tissue deviates from a homogeneous organization, it is considered cancerous and the grade is determined by the severity of the deviation. Because there was no considerable difference in the color distribution between healthy and unhealthy tissue samples their methods relied heavily on morphological features. The proposed algorithm was run on sixteen randomly chosen cases and compared to images which were manually segmented by one specialist. The specialist's segmentation included a region marked as positive for cancer, negative for cancer and a generous region which could be classified in either direction. They reported an accuracy rate of almost 95% when counting all segmentation lines falling in the neutral area as a correct segmentation.

3.2   Semi-Automatic Segmentation Techniques

Researchers have proposed analytical methods that increase automation and reduce the burden on the user. Semi-automation is achieved either by building a self-adjusting algorithm (a difficult undertaking) or by changing adjustable values to fixed values. This automation can be achieved in one of two ways. The ideal way to automate an method is by building the algorithm to be self adjusting so it uses the proper values depending on the input image. The easier way to automate a method is by changing adjustable values to fixed values. Hard coding a parameter is only appropriate when the input is invariant to that value. For example, at a particular magnification and resolution all red blood cells should be about the same size. Hard coding the size of the cell would be appropriate if it can be assumed that all input images are acquired at the same magnification and using the same resolution. If that assumption cannot be made, the cell size is more appropriately set as a variable.

Sertel et al. attempted to automate the removal of white background pixels by using a hardcoded threshold and classified as the white background all pixels with intensity level greater than two hundred in all three 8-bit RGB color channels [2]. Although Sertel et al. used H&E stained tissue slides, I was unable to reproduce their result using this fixed value. Upon implementation of this

(a) Original image.   (b) Pixels misclassified as 'white.'
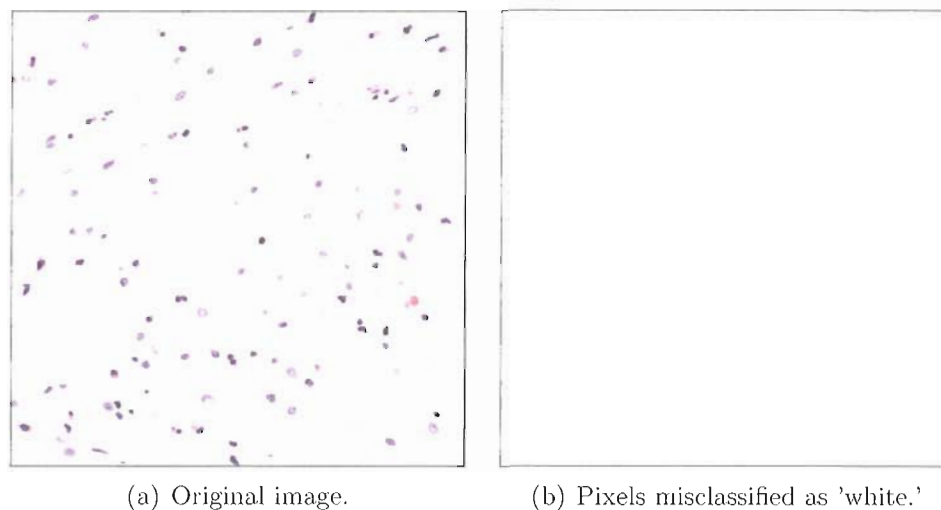
FIGURE 9: Sertel et al. fixed threshold for white background detection.

technique, the threshold of two hundred was too low for some instances of H&E stained tissue and oversegmentation resulted. Figure 9 illustrates the problem. The original tissue image in figure 9(a) does not contain any white background, however all pixels shown in figure 9(b) are classified as white using Sertel's thresholding value. This image contains no glass background yet the majority of the pixels are misclassified as such. This poses a significant problem for researchers identifying the density of a particular cell within the tissue—misclassifying tissue as background falsely inflates the density ratio. An alternate approach to background detection is a fully automated method which responds to image conditions. This approach is presented in Section 4.2.1.

Classification of tissue types is achieved by Naik, et al. in [33] with the use of a color classifier and thresholding to remove background noise. A Gleason grade is a numerical value used to describe prostate cancer malignancy where a high value corresponds with a poor prognosis.

The process begins with the detection of tissue types based on their color properties. A Bayesian classifier is trained to detect if a pixel value is likely lumen, cytoplasm or nuclei. In any given digital image, each pixel $c \in C$,where C is a two-dimensional grid of image pixels, has a value assigned by the function f which

represents red, green, blue, hue, saturation, and intensity. A training set is selected for each tissue type and is used to generate the probability density function $p(c, f(c)|w_v)$, where w_v is the pixel class with $v \in Lumen, Nucleus, Cytoplasm$. Then for each image, the probability that $c \in w_v$ is

$$P(w_v|c, f(c)) = \frac{P(w_v)p(c, f(c)|w_v)}{\sum_{v \in \{L,N,S\}} P(w_v)p(c, f(c)|w_v)}, \qquad (3.1)$$

by Bayes theorem. The output for each pixel can be reconstructed into an intensity image with values between (0,1). Bright intensity corresponds to a high likelihood that the pixel belongs to the class. A threshold is applied to define the object and remove background noise and finally the true tissue is determined based on neighborhood criterion.

Like in [33], the first step to defining the objects in [34] is accomplished by applying k-means to color intensities to distinguish between the three tissue types. Purple regions correspond to epithelial and lymphoid cells, pink regions represent connective tissue, and white regions signify luminal structures and connective tissue components. Tosun's technique goes on to use uniformity measures for each pixel in a three-step region growing method:

- *Seed Determination*: Seeds are formed by connecting pixels using four connectivity (the pixels directly above, below, to the right and to the left) for which all twelve uniformity measures are smaller than their corresponding thresholds. The threshold is the mean plus one standard deviation of the associated measure computed over all the pixels.

- *Seed Growing*: The threshold value is recalculated for all pixels that are not already assigned to the existing seeds. The pixels that fall under the threshold are added either to the neighboring seed or, if a neighbor does not exist, to a new seed. Tosun et al report over-segmentation of the image at this point in the process.

- *Region Merge*: To adjust for over-segmentation, regions are merged based on two criteria: (1) the percentage of total area for the same type of objects in

the region and (2) the percentage of the combined areas of the different objects which correspond to the same cluster in the region.

The proposed algorithm was run on sixteen randomly chosen cases and compared to images which were manually segmented by one specialist (whose results were set as the gold standard) and images which were segmented by a separate process proposed by Deng and Manjunath [35]. As previously mentioned, Tosun et al. claimed an accuracy rate of nearly 95% even though, in many of the cases, the algorithm included some of the healthy tissue in sections marked as cancerous tissue and cancerous tissue in sections marked as healthy tissue. The accuracy rate of 95% was achieved through the introduction of a third category to the gold standard. Instead of classifying tissue as either healthy or cancerous, the specialist classified tissue as healthy, cancerous and a third category which could be correctly segmented as either healthy or cancerous. This is a huge advantage for achieving such a high accuracy rating since the categorization for much of the image would be deemed correct regardless of the algorithmic result.

Otsu's method [36] has been modified many times since its publication in 1979 [37, 38, 39, 40] but in its original form it remains a mainstay of the histogram thresholding techniques. Otsu's method for segmentation does an exhaustive search of every thresholding level and creates a record of the between-class variance. The optimal threshold is defined as the value with the maximum variance between two classes or with the minimum variance within two classes which yields the same threshold result [36]. Defining the optimal threshold depends on the existence of a clear foreground and background on opposite sides of the gray-level histogram, also referred to as the bimodal distribution of gray levels. Minimization of the intra-class variance can be used for multilevel thresholding as shown in [39, 40]. When converted to multilevel thresholding, Otsu's binarized, fully automated algorithm becomes parameterized. The fully automated version of Otsu's method is a poor fit for H&E stained tissue due to variance among slidesthe presence of blood, the visibility of the glass background, or the type of tissue being segmented, for example. Unfortunately, though, multilevel thresholding is flawed for our purpose also, as the

number of levels necessary to achieve optimal segmentation is unclear. As with the iterative threshold selection algorithm presented in Section 4.2.2, preprocessing as presented in Section 4.2.1 and postprocessing as presented in Section 4.2.3 is required to fully automate Otsu's method for the purpose of H&E segmentation. We compare Otsu's method to other segmentation methods in Section 4.3.

3.3 Shape Analysis

Kong et al. differentiate between pathological components of H&E stained tissue [41] using only texture and color. While they recognize nuclear size and shape as an integral part of the process used by pathologists in their visual grading process, they contend that morphological features cannot be expressed quantitatively. Because diagnostics rely heavily on morphological components, this methodology must be refuted. [34, 33, 42] and Section 4.2.3, all demonstrate the feasibility of incorporating morphological features into a quantitative system. While consistent evaluation criteria do not exist for methods of shape description, the development of such a system continues to be an active area of research rooted in differential geometry [19].

In the diagnosis of prostate cancer, a gleason grade is assigned according to the arrangement of nuclei and the shape of the gland structure. Naik et al. employs two methods of shape analysis to assign Gleason grades: First, a size constraint removes noise. Then, a Bayesian classifier analyzes morphological features to assign a grade. Morphological feature extraction is based on eight features calculated across the exterior lumen boundary and the interior nuclei boundary. The eight features are as follows:

- *Area*, a measure of the space enclosed by the boundary B;

- *Area overlap ratio*, a measure of shape irregularity, is the area enclosed by B divided by the area of the smallest circle enclosing B;

- *Distance ratio*, the ratio of the average distance to the maximum difference of the distance from the centroid of B to the points lying on B;

- *Standard deviation* of the Euclidean distance from the centroid of B to the points lying on B;

- *Variance* of the Euclidean distance from the centroid of B to the points lying on B;

- *Perimeter ratio*, the ratio of the estimated length of B (via linear interpolation) to the true length of B (the actual number of pixels);

- *Compactness*, the quotient of the true length of B squared and the area enclosed by B; and

- *Smoothness*, the measure of a boundary pixel, c, and the Euclidian distance of its clock-wise neighbor, c_{i+1}, and its counter-clock-wise neighbor, c_{i-1}, to the centroid, c_g, multiplied by the pixel's Euclidean distance from the centroid, $d(c_g, c_i)$, divided by two and summed over all the boundary pixels.

$$Smoothness = \sum_{c_i \in B} \left| d(c_i, c_g)\big(d\left(c_{i-1}, c_g\right) + d\left(c_{i+1}, c_g\right)\big)/2 \right| \qquad (3.2)$$

These eight features determine which Gleason grade to assign to the analyzed tissue. For example, Gleason grade four is made up of smaller but more uniform lumen areas which lack glandular structure, whereas Gleason grade one is made up of benign epithelium of larger varied sized glands closely resembling healthy tissue. As the tissue strays from it's healthy formation, the worse the patient's prognosis.

Tosun et al. proposed an object-oriented approach which treats sections of the image representing portions of the same tissue type as a single object as opposed to looking at individual pixels. Shape analysis based on object size is used to group various tissue components [34]. Each of the three object types (epithelial tissue, connective tissue and luminal structures) are processed twice to determine circular structures within the tissue. The circular structures are then categorized into two object groups based on size. The homogeneity criterion is based on the distribution of the objects in size and space and is determined by two metrics:

- *Object Size Uniformity*: For each object type, the standard deviation of the area of the object indicates object size uniformity. A completely uniform collection results in a standard deviation of zero.

- *Object Spacial Distribution Uniformity*: For each object type, first calculate the sum of the position vectors of every object with reference to the image centroid. Next, calculate the magnitude of the resulting vector. The magnitude indicates spacial distribution uniformity where a perfectly uniform spacial distribution results in a sum and magnitude of zero.

Due to the uniform nature of healthy tissue, uniformity metrics distinguish between cancerous and benign colon tissue.

This literature review is a testament to the limitations of current research and the need for further study. Before image analysis is successfully established as an invaluable tool in the field of medical diagnostics, the challenges impeding its integration must be resolved. These challenges include algorithms which require user input beyond the scope of their primary users, improper preprocessing of samples, thresholding techniques which implement hard values inappropriately, results which are difficult to verify, and, generally speaking, processes which are unnecessarily complex. The tool presented in chapter four addresses these concerns. It is not dependent on user input, it accounts for context in its preprocessing techniques, it introduces a novel thresholding solution, it produces results which are easily verified, and, most importantly, is user-friendly and practical.

CHAPTER IV

AUTOMATED DETECTION OF INFLAMMATORY RESPONSE

Presently, image processing for automated (or even semi-automated) medical image processing is an open and active area of research. There are no known publications related to image processing for the purposes of identifying inflammatory PML nuclei. Medical diagnosis of placental inflammation is also an open area of research, and attempts at standardized scoring of intra-amniotic infection is a source of controversy [22, 23]. Salafia et al. argue that the current qualitative approach needs to be improved upon via a reliable quantitative automated tool [1]. The technique presented in this thesis answers that call and creates a universal and uniform system of diagnosis for the presence of inflammation through the identification and quantification of inflammatory PML nuclei.

4.1 Image Dataset

The image set for developing and verifying this algorithm is made up of four-hundred and fifty tissue slides: forty-five present variations in H&E stain age and four-hundred and five present variations in H&E stain brand. Samples from placental tissue are soaked in paraffin and cut at a thickness of fiveμm before being placed on a glass slide. Stain age is varied on three sequential tissue slices from fifteen cases which are stained at zero, two, and four days old. Stain brand is varied

with Myelo IHC and with eight unique combinations of four brands of Hematoxylin (Harris-R, Gills-R, Harris-P and Gills-P) and two types of Eosin (Eosin and Eosin Phloxin). This rotation of nine stain combinations is performed three times over twenty-seven sequential tissue slices from fifteen tissue samples for a total of four-hundred and five slides.

Varying brand and stain age in this way determines whether variations that are expected in practical laboratory and hospital environments have a negative impact on the performance and reliability of the algorithm. It is not practical to assume that staining conditions can be exactly replicated across a diverse set of lab environments; therefore, in order to be viable, a robust diagnostic algorithm must account for these variables.

4.2   Methods

The following sections introduce three components of quantitative identification of inflammatory PML nuclei: background detection via image segmentation, cell nuclei detection via a novel iterative threshold selection algorithm, and isolation of particular nuclei of interest via shape analysis.

4.2.1   Background Detection (Image Segmentation)

The iterative thresholding technique requires that an image consist of only H&E stained tissue. Removal of extraneous features, like white background and red blood cells, normalizes the image for input into the iterative thresholding technique. Extraneous features are removed through a method that utilizes color ratios and is therefore neutral to the number of bits for color representation. Consider the raw image in figure 11(a) in the following discussion.
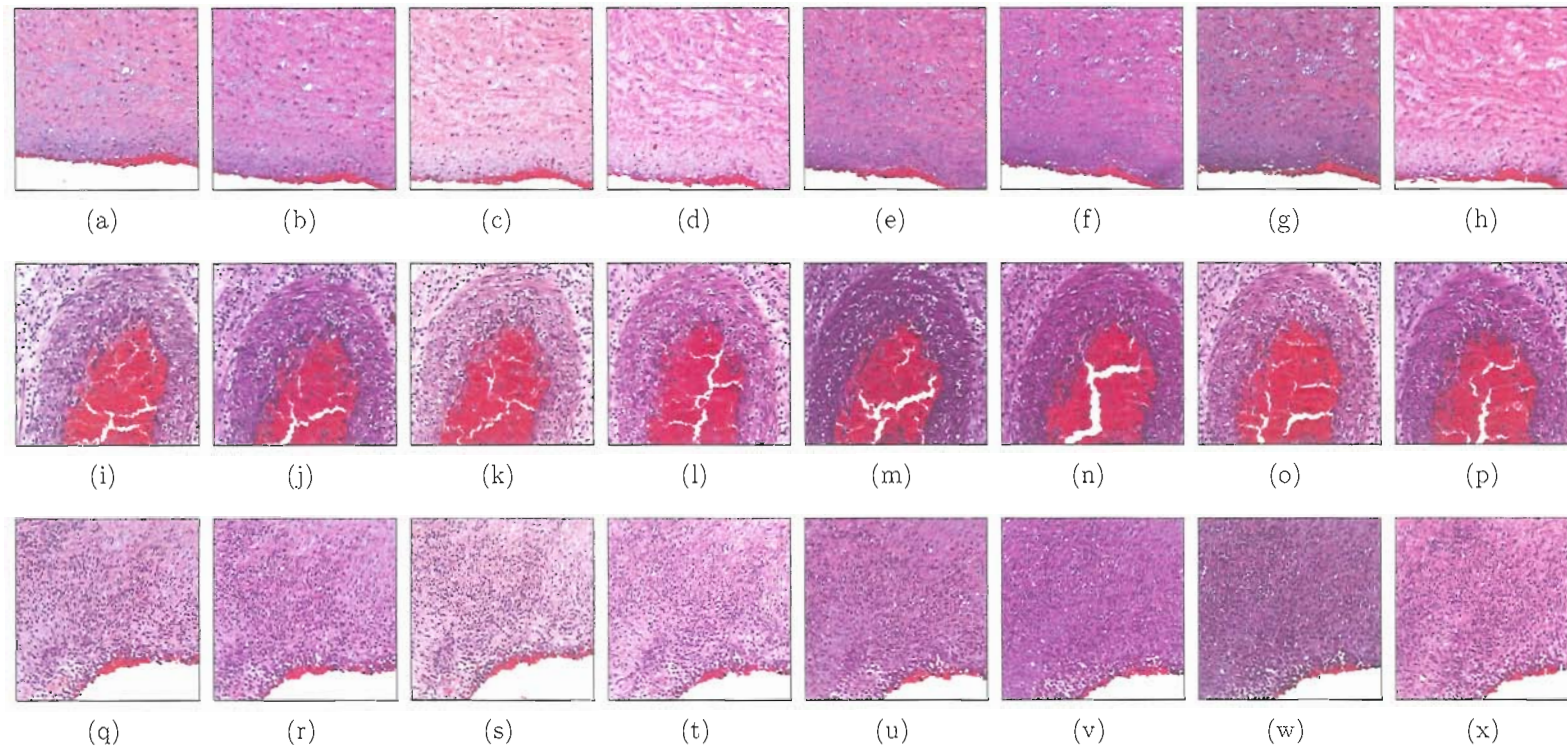
FIGURE 10: Hematoxylin & Eosin stain combinations on sequential tissue slices. From left to right the Harris-R Hematoxylin & Eosin, Harris-R Hematoxylin & Eosin-Phloxin, Gills-R Hematoxylin & Eosin, Gills-R Hematoxylin & Eosin-Phloxin, Harris-P Hematoxylin & Eosin, Harris-P Hematoxylin & Eosin-Phloxin, Gills-P Hematoxylin & Eosin, and Gills-P Hematoxylin & Eosin-Phloxin. From top to bottom, the tissues represent no inflammation, moderate inflammation and severe inflammation.

(a) Original image.   (b) Red blood cell mask.   (c) Preprocessed image after removal of red blood cells.
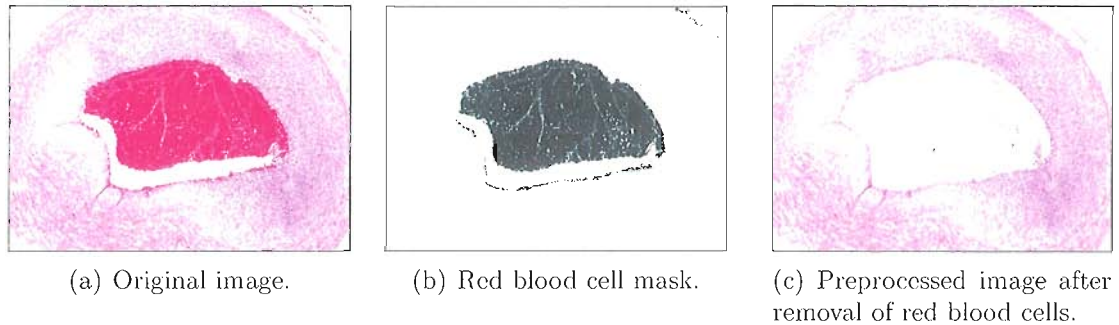
FIGURE 11: Preprocessing for removal of red blood cells.

Red Blood Cell Detection

Removal of red blood cells is accomplished with a simple threshold. Within an image of H&E stained tissue, any pixel with a red to blue ratio greater than 1.25 is too red to be connective tissue or nuclei and therefore is masked out and ignored. The pixels in figure 11(a) identified as meeting this criteria are masked in black in figure 11(b). The result after removal of the red blood cells is shown in 11(c).

White Background Detection

The iterative thresholding technique requires that all white background be removed from the image. In slide images, the background is a relative white but is not necessarily the maximum intensity RGB color value for white. To account for this discrepancy, the threshold is based on a percentage of the maximum value of each channel. This eliminates the need for manual adjustments when a slide is very dark or very light, when it is overcast with excess dye, or when the scanner is incorrectly calibrated to the background. To compute the color channel thresholds that identify the white background of a given slide, first identify the maximum intensity in each color channel: r_{max}, g_{max}, and b_{max}. Then, given these values, calculate individual thresholds for each channel as follows:

$$t_{red} = 0.90 \cdot r_{max}$$
$$t_{green} = 0.90 \cdot g_{max}$$
$$t_{blue} = 0.90 \cdot b_{max}$$

Instead of exclusively identifying pixels that are at maximum value (a rare occurrence in this context), the multipliers in the above equations detect pixels that are very bright but are not necessarily at maximum value. This accounts for variations in the imaging process like lighting and background. These image-specific thresholds automate the identification of pixels which constitute white background.

$$isWhite(x) = (x_{red} > t_{red}) \wedge (x_{green} > t_{green}) \wedge (x_{blue} > t_{blue}) \tag{4.1}$$

In cases where an image does not contain any white pixels, there exists enough variability in the color channels to result in image specific thresholds which are sufficiently high to exclude all pixels in the image. Because every pixel fails to meet the criteria for thresholding in all three channels, at least one clause in $isWhite()$ is found to be false. This renders $isWhite()$ to be false and thus validates the method for images which do not contain white background.

Figure 4.2.1 is an illustration of the automated white thresholding method. All three original images resulted in different RGB values for the isWhite threshold. The threshold for figure 12(a) was [242 179 188], for figure 12(b) was [238 218 225], and for figure 12(c) was [242 234 234]. The resulting white masks are shown as black pixels in figures 12(d), 12(e), and 12(f).

4.2.2   Cell Nuclei Detection (Iterative Threshold)

Once the red blood cells are removed from the image and the white background is excluded, the remaining pixels correspond exclusively to tissue or nuclei; a threshold must be established to differentiate the two. Due to lighting conditions and to shifts in the mean color intensities related to stain aging and brand variations, the precise value of this threshold is a function of the image
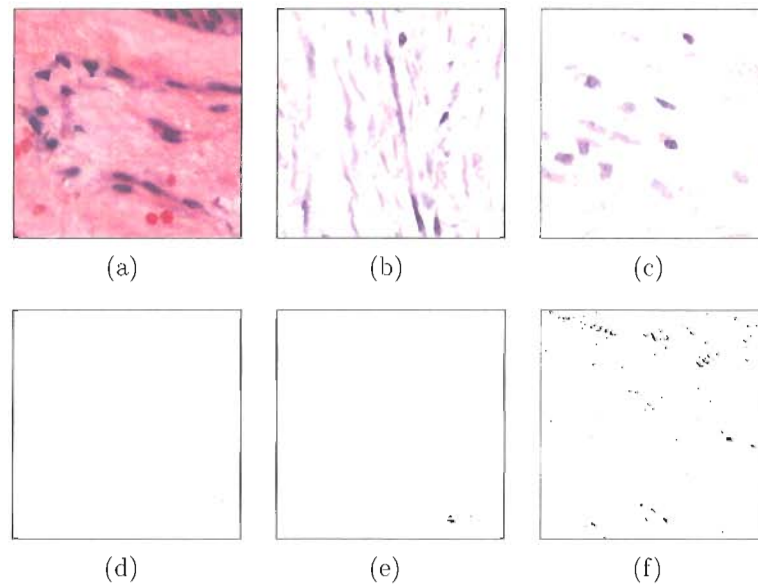
FIGURE 12: Results of isWhite() mask for images with no white pixels.

contrast and the red channel distribution. The following method, the iterative technique, eliminates the need for technicians to manually specify the optimal threshold value by incrementally differentiating tissue from nuclei:

Algorithm 4.2.1. *Iterative Threshold Selection.*

    Input: *RGB pixel values.*

    Output: *Threshold.*

    $threshold \leftarrow meanIntensity(Pixels)/2$

    repeat

        for all *pixels* $p \in Pixels$

            if $p_{red} < threshold$

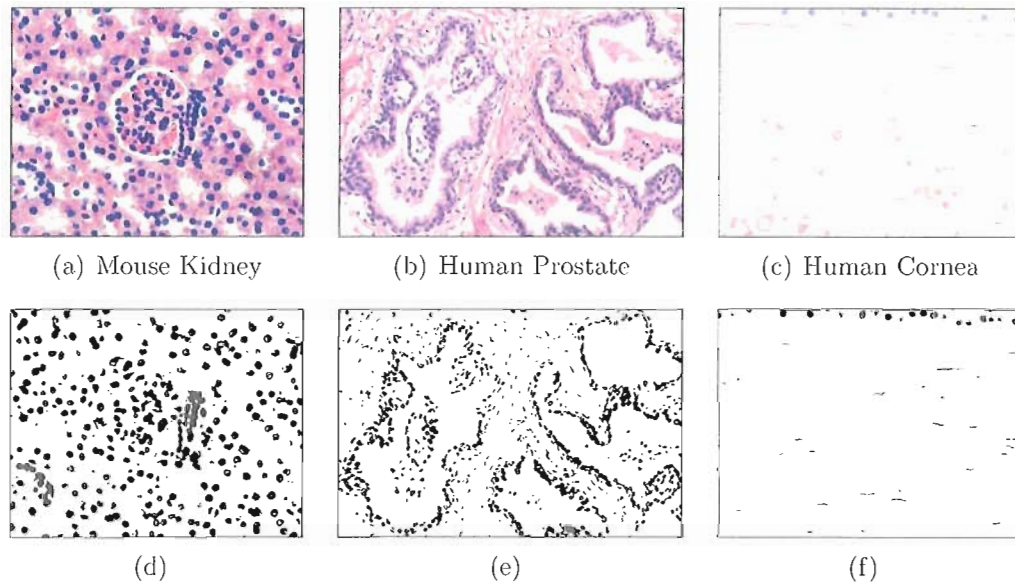   *remove p from Pixels*

                $threshold \leftarrow meanIntensity(Pixels)/2$

    until *nothing is removed from Pixels*

    return *threshold*

The resulting threshold value distinguishes pixels belonging to tissue from those belonging to nuclei based on a comparison of the red channel value. Tissue,

FIGURE 13: Nuclei detection in various H&E stained tissue.

which appears in shades of pink, has a much higher red channel value than nuclei, which appear purple or blue in color. The red channel is therefore a reliable determiner in the identification of tissue and nuclei.

$$isNuclei(x) = x_{red} < threshold \qquad (4.2)$$

Due to the inconsistent receptivity of some nuclei to dye or to the overlay of some nuclei with tissue, the final threshold is found to be slightly low (dark) for segmenting all nuclei after the threshold selection algorithm is applied to the image set. A ten percent correction (twenty-five for a slide scanner with 8-bit encoding) brings some of the outlier nuclei into view.

The methods covered up to this point (detection of red blood cells, white background, and nuclei) are designed for use with placental tissue. As figure 13 illustrates, the success of these methods are not limited to placental tissue and can be applied to other tissue types as well. The algorithm has a broad application base due to its self-adjusting nature and does not require user-defined tuning for the extension of its application to other diagnostic domains.

4.2.3 Cell Nuclei of Interest Detection (Size and Shape Filter)

It is now essential to determine which cell nuclei are the nuclei of interest. That is, which nuclei are inflammatory PML nuclei. Identification of nuclei of interest is possible through shape analysis due to their morphologically distinct size and form.
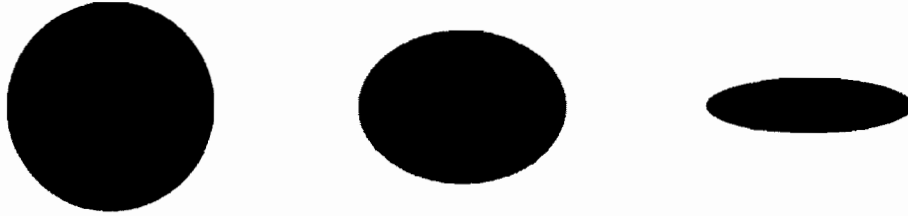
Filtration by Size

PML nuclei consistently take on a characteristic size within an image and therefore are easily differentiated from other, non-PML nuclei. From a biological standpoint, every cell type has a normal size range. Filtering out nuclei that are outside of this range eliminates a large number of non-PML nuclei or other artifacts that have a similar response to the H&E stain. The pixel count of the candidate nuclei defines this filter.

Inflammatory PML nuclei differ from chorion and amnion epithelium in size (see figure 15(c)). Because size analysis of nuclei and epithelium depends on resolution and magnification, empirical values for the range of normal pixel group size for these two features do not apply. The size threshold is easily calculated with the following formula, however:

$$pixel\ group\ size = resolution\ (pixels\ per\ \mu m^2) \times cell\ size\ (\mu m^2) \qquad (4.3)$$

The normal range of cell size is the basis for the threshold range. Following the computation of the number of pixels of each connected component, the connected components are eliminated as noise if the number of pixels is below the pixel group size interval and are eliminated as epithelial cells or larger artifacts if the number of pixels is above the pixel group size.

**FIGURE 14**: Examples of shapes with eccentricity 0.00, 0.70 and 0.95.

Filtration by Shape

Just as inflammatory PML nuclei are distinguished from non-PML nuclei by their size, inflammatory PML nuclei are differentiated from other cell features by shape. PML nuclei have a characteristic circular (or near circular shape) that is distinct from similarly sized cell features (like connective tissue fibroblasts as seen in figure 15(b)) which take on a more elliptical and elongated shape. Eccentricity, a measure of shape, is computed for each feature to perform filtration by shape. A circle has an eccentricity of zero, and increasingly flattened oval shapes have eccentricities that approach one. For the purposes of identifying inflammatory PML nuclei, features with an eccentricity of .95 or above are identified as fibroblasts (which are more elliptical in shape). Matlab's built in function was used to calculate eccentricity based on each connected group of pixels. Example shapes with varying eccentricity are shown in figure 14.

According to Matlab documentation, the eccentricity of a region is calculated for the ellipse with the same second moments as the region. Eccentricity is a measure of how much a conic section deviates from circular. Thus, eccentricity is defined as e, $0 \le e \le 1$, where $e = 0$ is a perfect circle and as e approaches one the shape is increasingly elliptical.

4.3   Analysis

The method presented in chapter four, henceforth referred to as the Thomas method, is evaluated through comparison with currently available automated
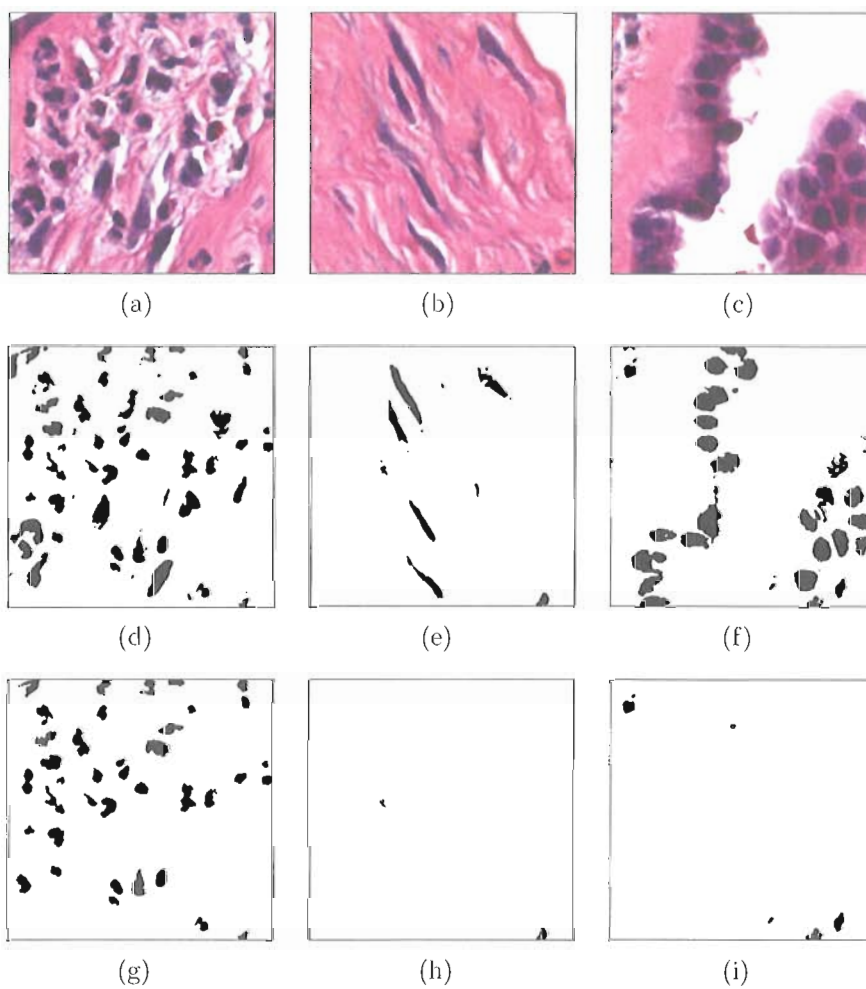
FIGURE 15: Neutrophil, fibroblast and epithelium segmentation. Case study with inflammatory PML nuclei shown in 15(a), fibroblasts in 15(b) and epithelium in 15(c). The mask after thresholding shown in 15(d), 15(e) and 15(f) respectively. Finally 15(g), 15(h) and 15(i) show the results after size and shape filtering.

segmentation methods, namely the Otsu method and the JSEG method. The three methods are compared on the basis of processing time, relative error in nuclei quantification and agreement with pathologist diagnosis. Processing time, measured in seconds, is the time required for execution of the method. Relative error is the change in positive pixel counts across sequential tissue slices with a variety of stain combinations. Pathologist diagnosis is the qualitative analysis resulting in the identification of inflammation as absent, moderate or severe. These metrics are assessed across sequential tissue slices processed with a variety of stain combinations, as shown in figure 10 and outlined in section 4.1. The results of the segmentation methods are shown in figures 16, 17 and 18, which represent cells without inflammation, cells with moderate inflammation, and cells with severe inflammation, respectively.

4.3.1 Processing Time

The Thomas, Otsu, and JSEG methods were all processed on an Intel® Core™2 Duo T9300 2.5 GHz processor with 4GB RAM and 32-bit operating system using MatLab R2008b Student Edition. Execution time in table 3 is the average execution time in seconds for processing 700×700 pixel photomicrographs representing each of the staining combinations over the three grades of inflammation (none, moderate and severe). On all counts, the Thomas method was slightly faster than the Otsu method and much faster than the JSEG. The Thomas method performs size and shape analysis for each detected nuclei, and its execution time, therefore, is a function of the number of nuclei present. As such, computation time varies directly with the rate of infection. The Otsu method is not dependent on nuclei count and, as such, its computation time, while slightly longer, is comparable to the Thomas method. The JSEG method performed approximately 300-400 times slower than the Otsu method and the Thomas method: where the Otsu and the Thomas methods execute in a fraction of a second, the JSEG methods average execution time is nearly a minute.
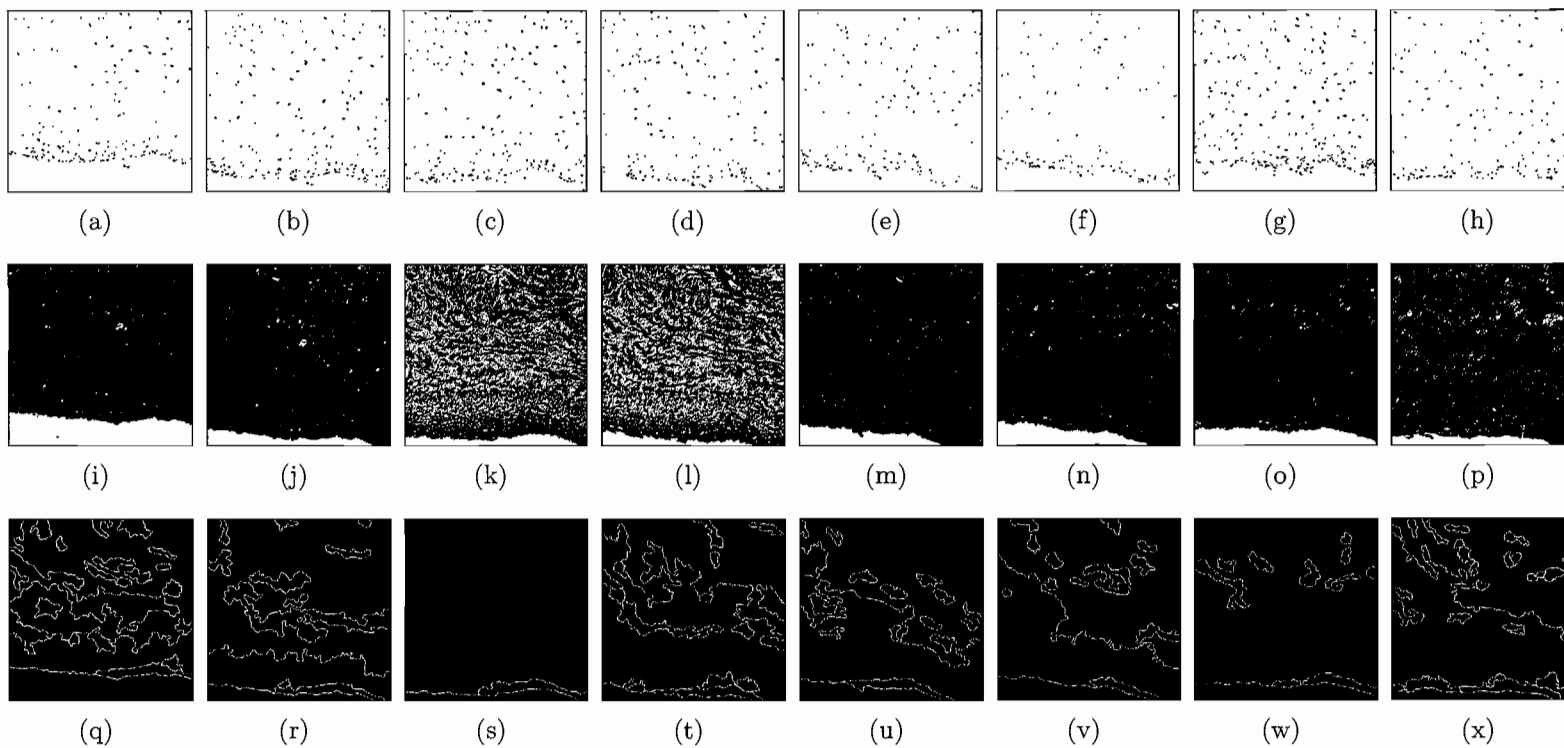
FIGURE 16: From top to bottom, sequential tissue slices without inflammation are segmented by the Thomas, Otsu, and JSEG methods. The corresponding original images are shown in figures 10(a)–10(h).
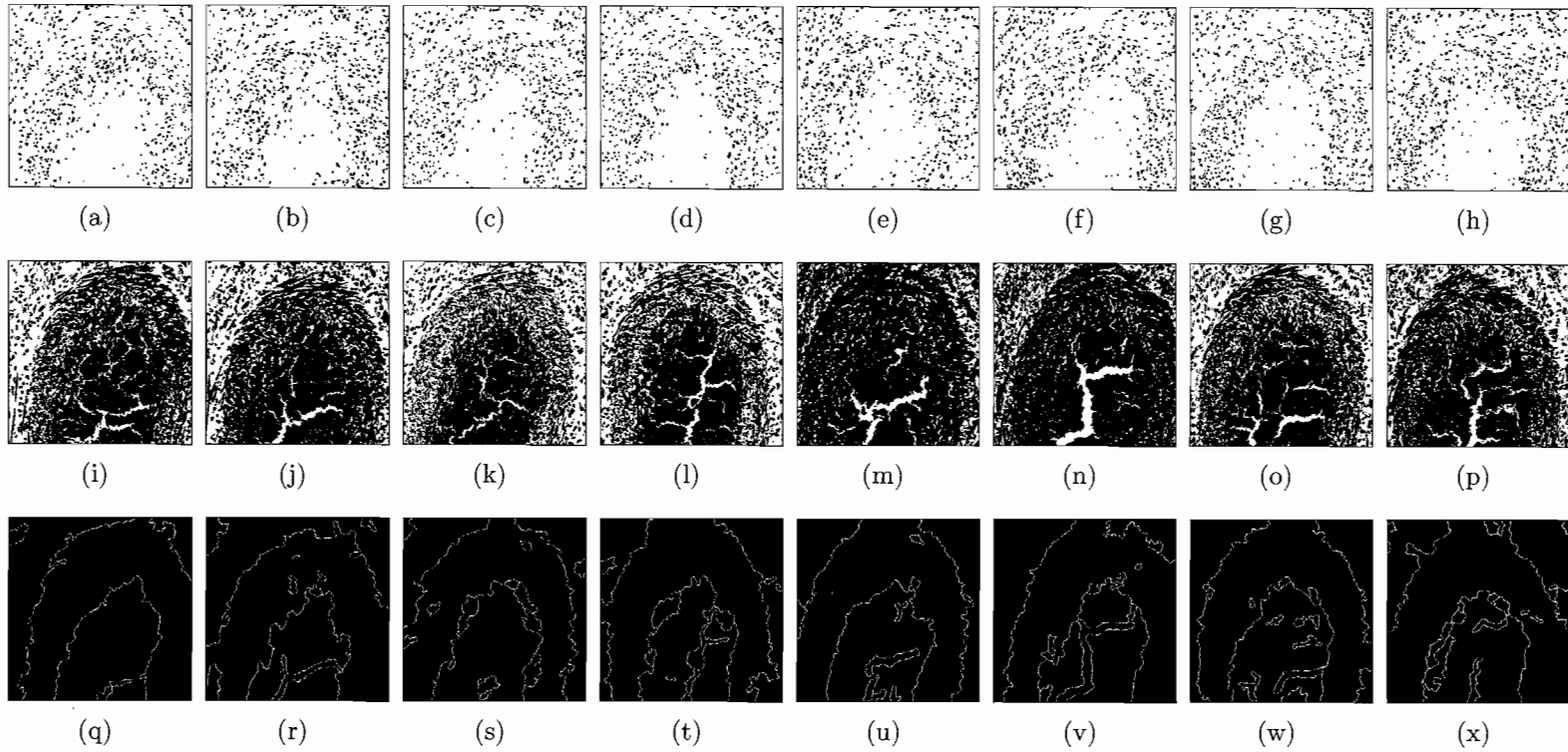
FIGURE 17: From top to bottom, sequential tissue slices with moderate inflammation are segmented by Thomas, Otsu and JSEG. The corresponding original images are shown in figures 10(i)–10(p).
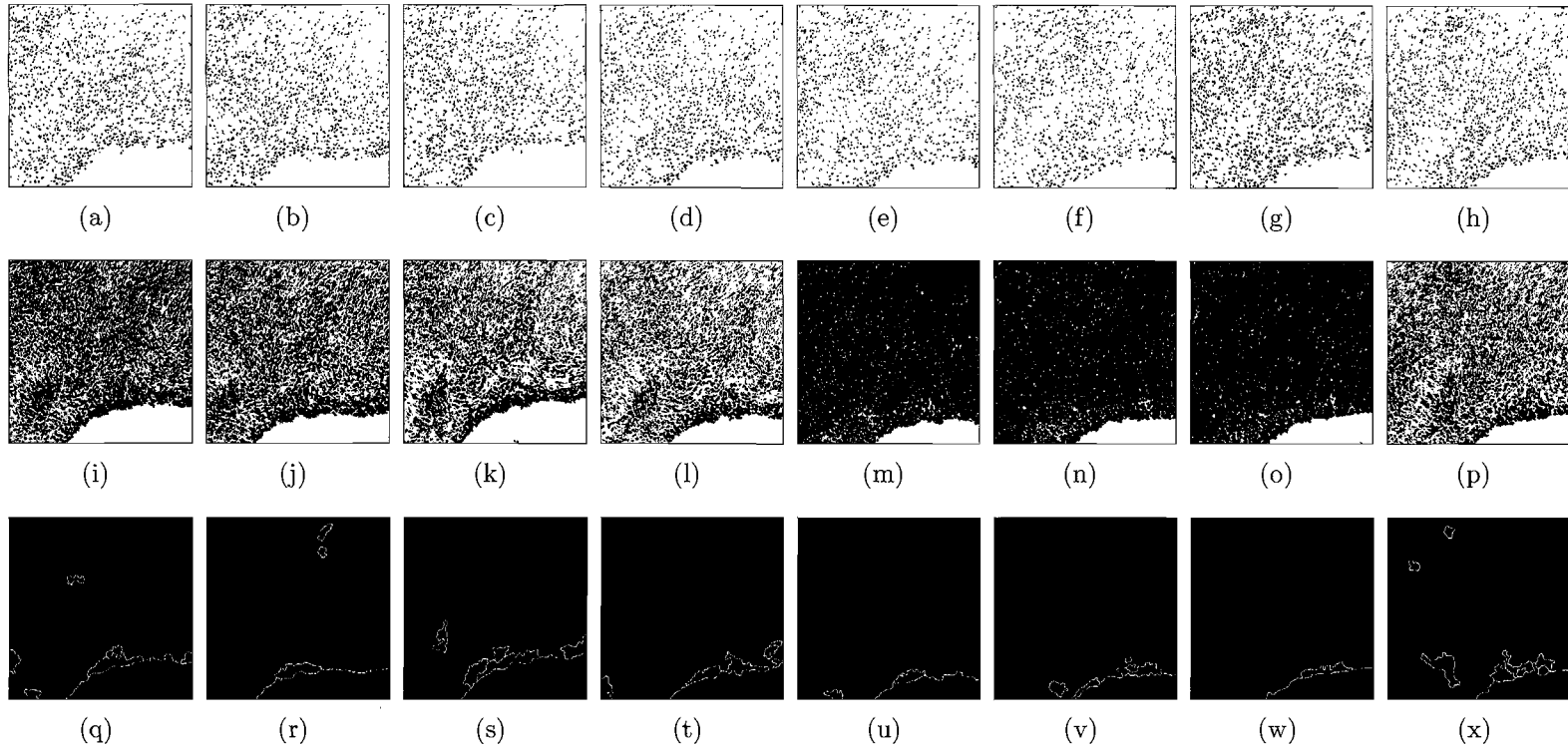
FIGURE 18: From top to bottom, sequential tissue slices with severe inflammation are segmented by Thomas, Otsu and JSEG. The corresponding original images are shown in figures 10(q)–10(x).

	Thomas	Otsu	JSEG
None	0.119	0.141	61.865
Moderate	0.123	0.141	42.035
Severe	0.133	0.151	57.010

TABLE 3: Execution time of Thomas, Otsu and JSEG methods. Execution time measured in seconds for 700x700 pixel photomicrographs by level of inflammation.

4.3.2 Reliability of Quantification

Although each slice of tissue contains different nuclei (and therefore no two slices of tissue are exactly alike), each tissue slice represents approximately the same amount of inflammation because each is sequential to the next. That is, the tissue slices are only micrometers apart in the original tissue sample. The specific variability tolerated for relative error is based on the expected biological change in nuclei counts across sequential tissue slices. Domain specialists must determine the tolerance of this error in accordance with their expertise. While some natural variation is expected, ideally the different staining methods do not exert large influences on relative error. The robustness of the algorithm is analyzed with the goal of overcoming stain variation and achieving a relative error close to the natural biological variability.

The JSEG method does not perform well for the task of nuclei segmentation. The results produced by the JSEG method fail to identify nuclei and the method does not appear suitable for this application. So while the JSEG method is an automated segmentation technique and its results appear in figures 16, 17, and 18, its validation metrics were not calculated.

Output produced by the Thomas method is quantified by *positive pixel groups* (which correspond to nuclei) per $100\mu m^2$ of tissue; output produced by the Otsu method is quantified by *positive pixel count*. For the sake of comparison and normalization, output for the Otsu method is scaled to the amount of tissue in the image to convert the output metric for the Otsu method to positive pixel count per

$100\mu m^2$ of tissue. The quantification of the positive output is henceforth referred to as P.

The change in the quantification metric due to stain variation is as follows:

$$\delta_{stain} = P_{max} - P_{min}, \tag{4.4}$$

where P_{max} and P_{min} are the highest and lowest quantification result across the stain variations.

Quantification of the Thomas method, δ_{stain}, is defined as the change in nuclei count. This metric does not extend to the Otsu Method, however, and another calculation is necessary to obtain an equivalent metric. Assuming the Otsu method gives a correct segmentation result, a rough estimate for the change in nuclei count is calculated as follows:

$$Otsu \ \Delta_{nuclei} = \frac{\delta_{stain}}{nuclei \ size} \tag{4.5}$$

$$Thomas \ \Delta_{nuclei} = \delta_{stain} \tag{4.6}$$

The percent of classification change due to stain variations can be viewed as the relative error. Relative area, ε, is calculated as follows:

$$\varepsilon = \frac{\Delta_{nuclei}}{N_{tissue}}, \tag{4.7}$$

where N_{tissue} is the number of pixels within the image that represent tissue.

Relative error is a measure of precision rather than accuracy. A method is precise if its results are consistent. It is important to show the method's precision does not suffer as a result of stain variations. Accuracy, which is measured according to a methods correlation with expert diagnosis, is addressed in subsection 4.3.3.

Validating the precision of a method is dependant on the tolerance, α, to be determined by a domain expert as the number of expected change in nuclei count per $100\mu m^2$ of tissue. Validation, γ, is determined as follows:

Method	Inflammation Level	$\mathbf{P_{min}}$	$\mathbf{P_{max}}$	$\delta_{\mathbf{stain}}$	$\mathbf{\Delta_{nuclei}}$	ε	ρ
Thomas	none	9	19	10	10	.03%	
	moderate	61	71	10	10	.03%	.9923
	severe	88	105	17	17	.05%	
Otsu	none	273315	462399	189084	378	1.02%	
	moderate	269206	401883	132677	265	.85%	-.4635
	severe	187587	430602	243015	486	1.32%	

TABLE 4: Analysis metrics for the Thomas and the Otsu methods.

$$\gamma = \begin{cases} 1 & \varepsilon < \alpha \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

See table 4 for a comparison of the Thomas and Otsu methods over the previously described metrics.

4.3.3 Correlation with Expert Diagnosis

A lower relative error, ε, is desirable provided that agreement with pathologist diagnosis does not suffer. A consistent algorithm, no matter how precise, is not helpful if it is inaccurate. A low relative error in this case reflects the precision of the algorithm while correlation with expert diagnosis reflects the accuracy.

The output from the Thomas method agrees with clinical diagnosis with a correlation coefficient, ρ, of 0.9923. In contrast, output for the Otsu method is scattered and results in a correlation coefficient of -0.4635. These results are illustrated in the graphs of algorithm output presented in figure 19 and 20.

The results clearly show that the Otsu method is not suitable for the quantification of cell nuclei. In contrast, the Thomas method not only is stable across stain variations but also correlates strongly with expert diagnosis and makes clear distinctions between the levels of inflammation.
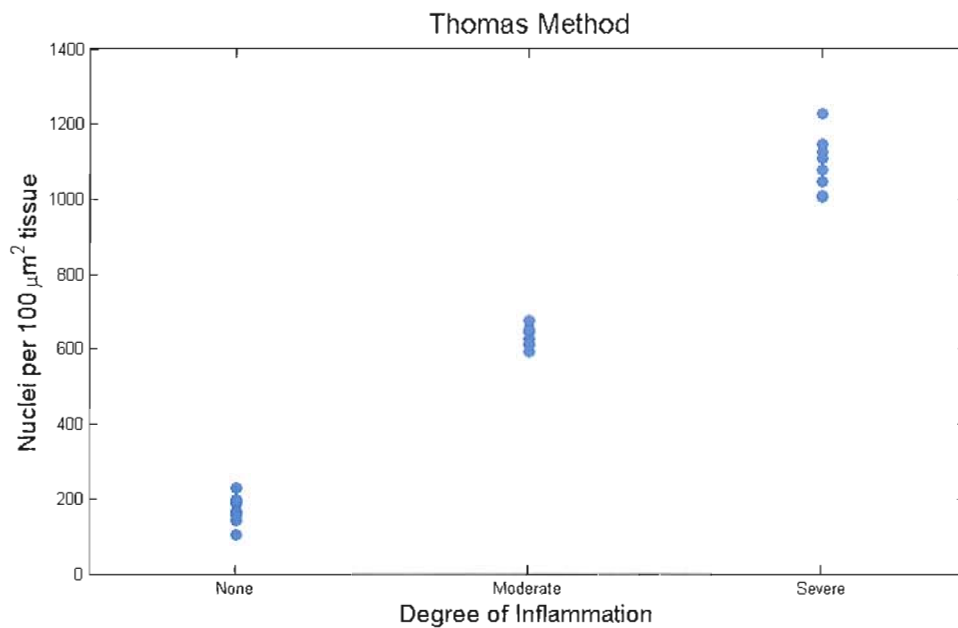
FIGURE 19: Graph of Thomas method results vs. degree of inflammation. Nuclei detected by Thomas method compared to the degree of inflammation.
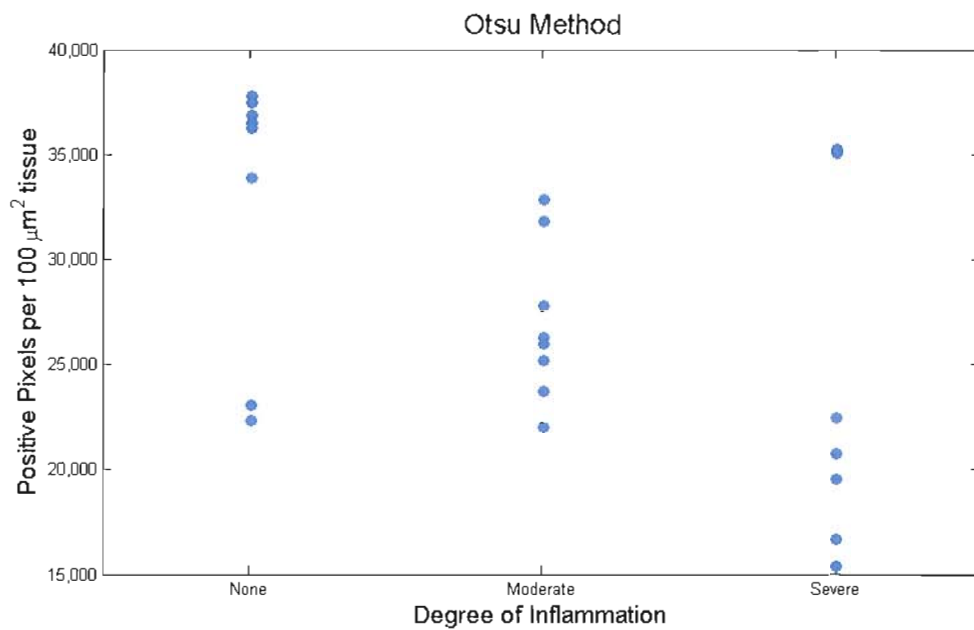
FIGURE 20: Graph of Otsu method results vs. degree of inflammation. Positive pixels detected by Otsu method compared to the degree of inflammation.

CHAPTER V

CONCLUSION

The Thomas Method presented herein effectively isolates and identifies nuclei which match the size and shape constraints of inflammatory PMLs in H&E stained placental tissue samples that present varying stages of acute chorioamnionitis. The method utilizes segmentation, adaptive thresholding, and shape analysis techniques. The analysis shows this method to be not only robust, but also unaffected by over-staining or by under-staining. A strong correlation exists between the method's results and expert pathologists diagnoses of inflammation as absent, moderate or severe; the method is likewise unsupervised and undemanding of user input. As such, the Thomas method is a reliable diagnostic tool with ready access for users with little or no image processing knowledge.

The diagnostic and technological gains of the Thomas method extend to other areas of pathology research, as well. The Thomas methods isolation of cell nuclei through its innovative iterative thresholding technique makes possible advances in the diagnosis of cancer, inflammation and necrosis. Other features of the method are easily adapted to enhance already existing methods for segmentation of red blood cell regions, shape analysis and classification through eccentricity, and the replacement of hard valued thresholding systems. Because the method applies to all biological tissue samples with H&E staining, its applications are limitless.

While this work establishes the Thomas method as stable over stain variation and on track with expert grading of inflammatory levels, the need for more evaluation and validation is apparent. Three expert pathologists with Placental

Analytics, LLC will test the algorithm over a large set of photomicrographs later this year. Showing correlation between the Thomas method and the experts grading system over a larger tissue set is another step toward establishing the method as a reliable and repeatable diagnostic tool.

APPENDIX

MATLAB IMPLEMENTATION OF THOMAS METHOD

The following code is an implementation of the algorithm presented in chapter four. It was written and executed using MATLAB 2008b Student Edition.

```
function [IDX, threshold] = thomas_IT(I, areaThresh, eccThresh)
%Iterative threshold algorithm for H&E stained tissue.
%   IDX = thomas_IT(I) segments an RGB image I by means of iterative
%   threshold selection.  It is specifically developed for working
%   with H&E stained histology slides. Thomas_IT returns a 2D array
%   IDX containing the indices of nuclei stained by H&E.
%
%   Area and eccentricity filters are optional and are useful for
%   filtering out fibroblasts and epithelium.
%
%   Notes:
%   ------
%   (1) Size and eccentricity thresholds are optional. Excluding them
%       results in nuclei segmentation for H&E stained tissue.
%   (2) Size constraint is dependent on resolution.
%   (3) lowering the R/B ratio for blood filtering can be done if the
%       tissue is lighter in color or has a blue overcast.  This can
%       eliminate clotted blood which does not stain as dark as fresh
%       blood.
%
%   Example:
%   -------
%   I = imread(R0009.tif)
%   [IDX, threshold] = Thomas_IT(I)
```

```
%    figure, imshow(IDX,[]);
%
%    Reference: K.A. Thomas, M.J. Sottile, C.M. Salafia. Unsupervised
%    Segmentation for In?ammation Detection in Histopathology Images.
%    ICISP 2010, LNCS 6134, pp.541-549, 2010.
%
%    Code prepared by Kristine A. Thomas 5-14-2010


%% Method Parameters
if nargin<3
    % default: don't segment based on eccentricity
    eccThresh = -1;
    if nargin<2
        %default: don't segment based on size
        areaThresh=-1;
    end
end

backgroundThresh = 0.9; % percent of max value to consider white
redBlueRatio = 1.25;

%% Identify indices of red blood cells
bloodIDX = (double(I(:,:,1))./double(I(:,:,3))) >= redBlueRatio;

%% Identify indices of white background
t_red = double(max(max(I(:,:,1))))* backgroundThresh;
t_green = double(max(max(I(:,:,2))))* backgroundThresh;
t_blue = double(max(max(I(:,:,3))))* backgroundThresh;
whiteIDX = (double(I(:,:,1))> t_red)   ...
           & (double(I(:,:,2))> t_blue)...
           & (double(I(:,:,3))> t_green);

%% Identify indicies of tissue pixels
tissueIDX = (~bloodIDX) & (~whiteIDX);

%% Iterative Threshold Selection
%create grayscale image based on the average intesity
grayI = double(mean(I,3));
%red channel image
I_red = double(I(:,:,1));
```

```matlab
% pixel lists used in the iterative threshold selection method
redPixelList = I_red(tissueIDX);
grayPixelList = grayI(tissueIDX);

numPixels = sum(tissueIDX);
numPixels_init = 0;

% initial threshold
threshold = mean(grayPixelList)/2;

% iterates as long as pixels are being removed from the set
while(numPixels_init ~= numPixels)
    %record pixel count before iteration
    numPixels_init = numPixels;

    % identify nuclei based on red channel
    nucleiIDX = redPixelList < threshold;

    % remove pixels from list that are likely nuclei
    grayPixelList = grayPixelList(~nucleiIDX);
    redPixelList = redPixelList(~nucleiIDX);

    % recalculate threshold based on average intensity
    % of remaining pixels
    threshold = mean(grayPixelList)/2;

    % new pixel count
    numPixels = sum(nucleiIDX);
end

% adjustment for outlier data
threshold = threshold + 25;

% memory management
clear grayPixelList;
clear redPixelList;

%% Segment based on threshold
nucleiIDX = uint8(((I(:,:,1)<threshold) & tissueIDX));
IDX = nucleiIDX;

%% Segment based on area
```

```matlab
if (areaThresh > 0)
    ImLabel = bwlabel(IDX);
    stats = regionprops(ImLabel,'Area');

    % segmentation based on area
    upperThresh= (areaThresh*15);
    areaIDX = find([stats.Area]>areaThresh ...
                & [stats.Area]< upperThresh);
    IDX = uint8(ismember(ImLabel, areaIDX));
end

%% Segment based on eccentricity
if (eccThresh >= 0)
    ImLabel = bwlabel(IDX);
    stats = regionprops(ImLabel,'Eccentricity');

    % segmentation based on eccentricity
    eccIDX = find([stats.Eccentricity]<eccThresh);
    IDX = uint8(ismember(ImLabel, eccIDX));
end

end
```

BIBLIOGRAPHY

[1] C.M. Salafia, D. Misra, and J.N.V. Miles. Methodologic issues in the study of the relationship between histologic indicators of intraamniotic infection and clinical outcomes. *Placenta*, 30(11), 2009.

[2] O. Sertel, J. Kong, H. Shimada, U. V. Catalyurek, J. H. Saltz, and M. N. Gurcan. Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development. *Pattern Recognition*, 42(6), 2009.

[3] Y. Wang, R. Turner, D. Crookes, J. Diamond, and P. Hamilton. Investigation of methodologies for the segmentation of squamous epithelium from cervical histological virtual slides. In *IMVIP '07: Proceedings of the International Machine Vision and Image Processing Conference*, pages 83–90, Washington, DC, USA, 2007. IEEE Computer Society.

[4] S. Shaktawat and D. Golka. Floret-like multinucleated giant cells in neurofibroma. *Diagnostic Pathology*, 2(1):47, 2007.

[5] K.S. Fu and J.K. Mui. A survey on image segmentation. *Pattern Recognition*, 13(1):3–16, 1981.

[6] E.H. Adelson. *The New Cognitive Neurosciences*, chapter Lightness Perception and Lightness Illusions, pages 339–351. MIT Press, 2nd edition, 2000.

[7] R. Wodnicki, G.W. Roberts, and M.D. Levine. A foveated image sensor in standard cmos technology. In *In Custom Integrated Circuits Conference*, pages 357–360, 1995.

[8] B.R. Conway. *Neural Mechanisms of Color Vision*. Kluwer Academic Publishers, 2002.

[9] S. Polyak. *Vertebrate Visual System*. University of Chicago Press, 1957.

[10] G. Kanizsa. *Organization in Vision*. Twayne Publishers, Boston, 1979.

[11] H. Buffart, E. Leeuwenberg, and F. Restle Coding theory of visual pattern completion. *Journal of Experimental Psychology: Human Perception and Performance*, 7:241–274, 1981.

[12] J.C. Russ. *The Image Processing Handbook.* CRC Press, 2007.

[13] H.D. Cheng, X.H. Jiang, Y. Sun, and J.L. Wang. Color image segmentation: advances and prospects. *Pattern Recognition*, 34(12):2259–2281, December 2001.

[14] M. Nixon and A.S. Aguado. *Feature Extraction & Image Processing, Second Edition.* Academic Press, 2008.

[15] I.T. Young. Sampling density and quantitative microscopy. *Analytical and Quantitative Cytology and Histology*, 10:269–275, 1988.

[16] T. Adamek, N. O'Connor, and N. Murphy. Region-based segmentation of images using syntactic visual features. In *Workshop on Image Analysis for Multimedia Interactive Services,(WIAMIS), Montreux, Switzerland*, 2005.

[17] N. Bonnet. Artificial intelligence and pattern recognition techniques in microscope image processing and analysis. *Advances in Imaging and Electron Physics*, 114:1–77, 2000.

[18] J. Freixenet, X. Muñoz, D. Raba, J. Martí, and X. Cufí. Yet another survey on image segmentation: Region and boundary information integration. *Lecture Notes in Computer Science*, pages 408–422, 2002.

[19] P. Van Osta, J.M. Geusebroek, K. Ver Donck, L. Bols, and J. Geysen. The principles of scale space applied to structure and colour in light microscopy, September 13 2002.

[20] R. Sood, J.L. Zehnder, M.L. Druzin, and P.O. Brown. Gene expression patterns in human placenta. 103(14):5478–5483, 2006.

[21] F.T. Kraus M.D., R.W. Redline M.D., D.J. Gersell Ph.D., D.M. Nelson and J.M. Dicke M.D. *Placental Pathology.* American Registry of Pathology, 2004.

[22] R.L. Naeye. *Disorders of the placenta, fetus, and neonate : diagnosis and clinical significance.* Mosby Year Book, St. Louis, 1992.

[23] F. Arias, A. Victoria, K. Cho, and F. Kraus. Placental histology and clinical characteristics of patients with preterm premature rupture of membranes. *Obstet Gynecol*, 89(2):265–71, 1997.

[24] C.M. Salafia, H.E. Mangam, C.A. Weigl, G.J. Foye, and L. Silberman. Abnormal fetal heart rate patterns and placental inflammation. *Am J Obstet Gynecol*, 160(1):140–7, 1989.

[25] R. Redline, D. Wilson-Costello, E. Borawski, A.A. Fanaroff, and M. Hack. The relationship between placental and other perinatal risk factors for neurologic impairment in very low birth weight children. *Pediatric Research*, 47(6):721 – 726, 2000.

[26] C. Hameed, N. Tejani, U.L. Verma, and F. Archbald. Silent chorioamnionitis as a cause of preterm labor refractory to tocolytic therapy. *Am J Obstet Gynecol*, 149(7):726–30, 1984.

[27] A. Hafiane, F. Bunyak, and K. Palaniappan. Clustering initiated multiphase active contours and robust separation of nuclei groups for tissue segmentation. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.

[28] O. Sertel, J. Kong, H. Shimada, U. Catalyurek, J.H. Saltz, and M. Gurcan. Computer-aided prognosis of neuroblastoma: classification of stromal development on whole-slide images. In *Proceedings of SPIE*, volume 6915, page 69150P.1–69150P.10, 2008.

[29] J. Kong, O. Sertel, H. Shimada, K. Boyer, J. Saltz, and M. Gurcan. Computer-aided grading of neuroblastic differentiation: Multi-resolution and multi-classifier approach. In *IEEE International Conference on Image Processing, ICIP*, volume 5, pages 525–528, 2007.

[30] C. Pilette, M.C. Rousselet, P. Bedossa, D. Chappard, F. Oberti, H. Rifflet, M.Y. Maïga, Y. Gallois, and P. Calòs. Histopathological evaluation of liver fibrosis: quantitative image analysis vs semi-quantitative scores: Comparison with serum markers. *Journal of hepatology*, 28(3):439 – 446, 1998.

[31] M. Volk, M. Strotzer, M. Lenhart, C. Manke, W.R. Nitz, J. Seitz, S. Feuerbach, and J. Link. Time-Resolved Contrast-Enhanced MR Angiography of Renal Artery Stenosis: Diagnostic Accuracy and Interobserver Variability. *Am. J. Roentgenol.*, 174(6):1583–1588, 2000.

[32] L. Hollink, S. Little, and J. Hunter. Evaluating the application of semantic inferencing rules to image annotation. January 01 2005.

[33] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi. Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information. *Proceedings of MIAAB*, 2007.

[34] A.B. Tosun, M. Kandemir, C. Sokmensuer, and C. Gunduz-Demir. Object-oriented texture analysis for the unsupervised segmentation of biopsy images for cancer detection. *Pattern Recognition*, 42(6), 2009.

[35] Y. Deng and BS Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 800–810, 2001.

[36] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, January 1979.

[37] X. Lang, F. Zhu, Y. Hao, and J. Ou. Integral image based fast algorithm for two-dimensional otsu thresholding. *Image and Signal Processing, 2008. CISP '08. Congress on*, 3:677–681, 2008.

[38] B.F. Buxton, H. Abdallahi, D. Fernandez-Reyes, and W. Jarra. Development of an extension of the otsu algorithm for multidimensional image segmentation of thin-film blood slides. *International Conference on Computing: Theory and Applications*, 0:552–562, 2007.

[39] P.S. Liao, T.S. Chen, and P.C. Chung. A fast algorithm for multilevel thresholding. *J. Inf. Sci. Eng*, 17(5):713–727, 2001.

[40] D. Y. Huang and C. H. Wang. Optimal multi-level thresholding using a two-stage otsu optimization approach. *Pattern Recognition Letters*, 30(3):275–284, February 2009.

[41] J. Kong, O. Sertel, H. Shimada, K.L. Boyer, J.H. Saltz, and M.N. Gurcan. Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation. *Pattern Recognition*, 42(6), 2009.

[42] B. Zheng, Y.H. Chang, and D. Gur. Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. *Acad Radiol*, 2(11):959–66, 1995.