

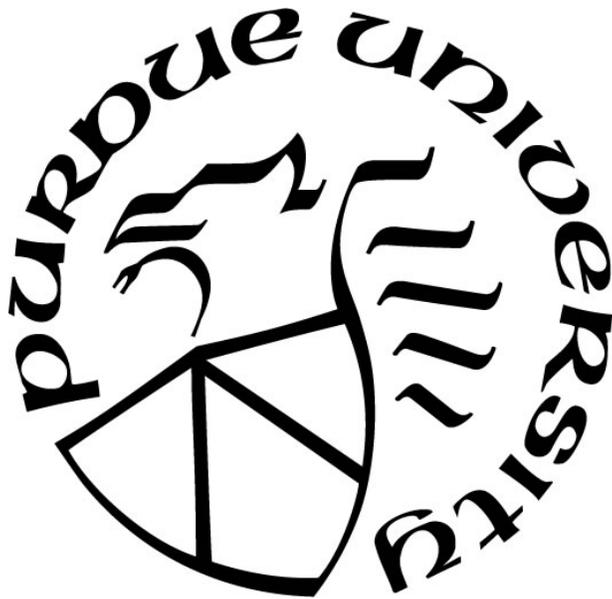
**MITOTIC CELL DETECTION IN H&E STAINED MENINGIOMA
HISTOPATHOLOGY SLIDES**

by
Huiwen Cheng

A Thesis

*Submitted to the Faculty of Purdue University
In Partial Fulfillment of the Requirements for the degree of*

Master of Science



Department of Computer & Information Technology

Indianapolis, Indiana

December 2019

**THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL**

Dr. Gavriil Tsechpenakis, Chair

Department of Computer & Information Science

Dr. Mihran Tuceryan

Department of Computer & Information Science

Dr. Jiang Yu Zheng

Department of Computer & Information Science

Approved by:

Dr. Mihran Tuceryan

Head of the Graduate Program

This thesis is dedicated to my mom Yunxia Zhang, and my dad Jian Cheng for their continuous support and unconditional love.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my advisor Prof. Gavriil Tsechpenakis for his continuous support and guidance. With his immense knowledge and enthusiasm for research, he has inculcated in me the desire to learn and the passion to contribute to society through research. His ideologies, insightful criticisms and words of encouragement have been constant sources of motivation throughout. And also because of him, I started my journey in computer science with greatest support, and will keep being a great journey.

A very special thanks to Dr. Lifan Zeng, who has been my classmates, my best friends for this whole journey. It would be so difficult to finish this whole process from knowing nothing about computer science to finish all projects and found the job at intel, he has been together with me for all those days and nights.

I would like to thank all of my labmates, Dr. Sarun Gulyanon, Dr. Ziyin Wang, and Sepehr Farhand. Who have always been there for me whenever I need help.

I would also like to thank my friend Christopher Goebel, who provided a lot of discussions, help and encouragements during the journey.

Last but not the least, I would like to thank my parents for all their love and support. Their belief in me has always helped me sail through the tough times, I am indebted to them for life.

TABLE OF CONTENTS

LIST OF TABLES	7
LIST OF FIGURES	8
ABSTRACT	10
1. INTRODUCTION.....	15
1.1 Motivation	15
1.2 Problem Statement	16
1.3 Contributions of This Thesis.....	18
2. BACKGROUND.....	19
2.1 Meningioma	19
2.2 Mitotic Cells and Grading.....	20
2.3 Tissue Preparation.....	22
2.4 Automated Histology Image Analysis	24
2.4.1 Image Acquisition.....	24
2.4.2 Image Preprocessing	25
2.4.3 Cell Detection and Segmentation.....	26
2.4.4 Mitosis Detection	28
2.4.5 Feature Extraction and Dimension Reduction	30
2.4.6 Classification.....	31
3. METHODOLOGY	33
3.1 Overall Approach	33
3.2 Data	34
3.2.1 Images	34
3.2.2 Pathologist Annotation.....	37
3.3 Image Preprocessing	38
3.4 Cell Segmentation	40
3.5 K-Means Clustering	48
3.6 Classification.....	50
3.6.1 C4.5 Pruned Decision Tree	51
3.6.2 SMO: Sequential Minimal Optimization	52

3.6.3	JRip: Rules Classifieres	54
3.6.4	Libsvm Polynomial	54
3.7	Deep Learning	55
3.7.1	Data Pre-processing and Data Augmentation	55
3.7.2	Transfer Learning with Breast Cancer Training Data.....	56
4.	EXPERIMENTS AND RESULTS	60
4.1	K-Means Results	60
4.1.1	Negative Cluster.....	60
4.1.2	Positive Clusters.....	61
4.1.3	K-means on Breast Cancer Slides.....	63
4.1.4	J48 Results	64
4.1.5	SMO Result.....	65
4.1.6	JRip Result	66
4.1.7	Libsvm Results.....	67
4.2	Deep Learning Results	68
5.	CONCLUSION	70
5.1	Whole Slide Image Process.....	70
5.2	Transfer Learning.....	70
	REFERENCES	72

LIST OF TABLES

Table 1. WHO Grades of Meningioma	16
---	----

LIST OF FIGURES

Figure 1. Meningioma in the brain	19
Figure 2. Different phases of mitosis	21
Figure 3. Histology tissue preparation and image production.....	22
Figure 4. Hematoxylin and eosin (H&E)	23
Figure 5. Computer assisted diagnosis workflow.....	24
Figure 6. Mitotic figures and apoptotic figures.	29
Figure 7. Example digital images with different focus offset values(Caglar Senaras, 2018)	29
Figure 8. Overall Approach.....	34
Figure 9. SVS image view at Aperio ImageScope.....	35
Figure 10. Various layers in the pyramid.	35
Figure 11. Matlab commands for open svx file	36
Figure 12. Matlab ‘imfinfo’ command for image information.....	36
Figure 13. (a) Definitely mitosis cell (b) Definitely not mitosis cell	37
Figure 14. Image preprocess steps for whole slide images:	39
Figure 15. Before normalization and after normalization	40
Figure 16. RGB channel of raw images	41
Figure 17. Green channel before and after adjustment.....	41
Figure 18. Before and after anisotropic diffusion	42
Figure 19. Before and after cell segmentation.....	43
Figure 20. Before and after small region removal, and the display of removed region on raw image	43
Figure 21. Cell boundaries by using red channel	44
Figure 22. Cells before and after splitting	44
Figure 23. Elbow method for k-means	50
Figure 24. VGG16 architecture	57
Figure 25. Model overview before trained on fully-connected layer	58
Figure 26. Fine tune results after each step	59
Figure 27. Epoch 45 - validation accuracy: 0.9812.....	59

Figure 28. Negative clusters from k-means method.....	60
Figure 29. Cluster for positive candidates (red boundaries).....	61
Figure 30. Candidate Clusters on Img2.	62
Figure 31. K-means method on breast cancer slides. Classification Results	63
Figure 32. Classification results with J48 method on img1.	64
Figure 33. SMO result on image1.	65
Figure 34. JRip result on image1, red circles are ground truth.	66
Figure 35. Libsvm Result	67
Figure 36. Prediction results on image1	68
Figure 37. Prediction results on image2	69

ABSTRACT

Author: Cheng, Huiwen. MS

Institution: Purdue University

Degree Received: December 2019

Title: Mitotic cell detection in H&E stained meningioma histopathology slides.

Major Professor: Gavriil Tsechpenakis

Meningioma represent more than one-third of all primary central nervous system (CNS) tumors, and it can be classified into three grades according to WHO (World Health Organization) in terms of clinical aggressiveness and risk of recurrence. A key component of meningioma grades is the mitotic count, which is defined as quantifying the number of cells in the process of dividing (i.e., undergoing mitosis) at a specific point in time. Currently, mitosis counting is done manually by a pathologist looking at 10 consecutive high-power fields (HPF) on a glass slide under a microscope, which is an extremely laborious and time-consuming process. The goal of this thesis is to investigate the use of computerized methods to automate the detection of mitotic nuclei with limited labeled data. We built computational methods to detect and quantify the histological features of mitotic cells on a whole slides image which mimic the exact process of pathologist workflow. Since we do not have enough training data from meningioma slide, we learned the mitotic cell features through public available breast cancer datasets, and predicted on meningioma slide for accuracy. We use either handcrafted features that capture certain morphological, statistical, or textural attributes of mitoses or features learned with convolutional neural networks (CNN). Hand crafted features are inspired by the domain knowledge, while the data-driven VGG16 models tend to be domain agnostic and attempt to learn additional feature bases that cannot be represented through any of the handcrafted features. Our work on detection of mitotic cells shows 100% recall, 9% precision and 0.17 F1 score. The detection using VGG16 performs with 71% recall, 73% precision, and 0.77 F1 score. Finally, this research of automated image analysis could drastically increase diagnostic efficiency and reduce inter-observer variability and errors in pathology diagnosis, which would allow fewer pathologists to serve more patients while maintaining diagnostic accuracy and precision. And all these methodologies will increasingly transform practice of pathology, allowing it to mature toward a quantitative science.

1. INTRODUCTION

1.1 Motivation

Meningioma represent more than one-third of all primary central nervous system (CNS) tumors, and the prediction of recurrence and, for malignant variants, prediction of survival is the most important prognostic question. The mitotic index (MI), which is defined as the sum of mitotic figures per 10 consecutive high-power fields (HPF), has been shown to be one of the most reliable predictors of the likelihood of meningioma recurrence. According to current World Health Organization (WHO) classification of meningioma, there are 3 grades: benign (WHO grade I), atypical (WHO grade II) and anaplastic (WHO grade III). However, meningioma classification is still based on ‘gold standard’ pathology diagnosis on the standard H&E-stained slides, which is a challenging and time-consuming task that can lead to variability in measurements and high inter-observer error. Therefore, there is clearly a need for quantitative image-based assessment of digital pathology slides. Computer-aided diagnosis has become more and more important to help increase efficiency of routine work, to improve accuracy of tasks in which some grading is involved and help to yield relevant information for diagnosis and prognosis.

The approaches to computer aided diagnosis of histopathology images involve segmentation of the image into relevant structures, feature extraction, classification and quantification. In addition to the numerous potential challenges of pathological image analysis, one of the key challenges for digital pathology is the limited data sets that can be used for data mining, and the limited efforts that pathologist can put to provide annotated datasets and to validate the results. Therefore, there is a clear need to develop and deploy sophisticated image processing and machine learning methods with limited data sets for automatic cancer cell detection and classification.

1.2 Problem Statement

According to WHO grading criteria, different grades are defined in table 1.

Table 1. WHO Grades of Meningioma

Grade	Mitoses/10 HPF	5-year Recurrence
Grade I (benign)	<4	20%
Grade II (atypical)	≥ 4 and < 20	40%
Grade III (anaplastic)	≥ 20	

It usually takes hours to manually count the number of mitotic cells in H&E stained slides by a well-trained pathologist, and it is really a challenge to differentiate Grade I and Grad II since it is critique to find 4~20 mitotic cells among large number of cell nuclei. Recently, the developments of imaging technology have led to the availability of high resolution slide scanners which can scan and digitize histopathology slides at magnifications of 40X up-to 80X microscope resolution. The availability of this technology has led to a new area of quantitative analysis of pathology images. Image analysis tools for automated grading of cancer stage have been developed for neuroblastoma(Gurcan et al., 2007; Sertel et al., 2009), breast cancer(Karacali & Tozeren, 2007; Schnorrenberg, Pattichis, Kyriacou, & Schizas, 1997), follicular lymphoma(Belkacem-Boussaid, Pennell, Lozanski, Shana'ah, & Gurcan, 2010; Fauzi et al., 2015), and prostate cancer (Liu, Tian, Zhang, & Fei, 2016), along with other types of cancer. The approaches to computer aided diagnosis of histopathology images involve segmentation of the image into relevant structures, feature extraction, classification and quantification. However, Mitosis detection is very hard. First, a large number of cells can be observed in histology images stained with Hematosin & Eosin, among which are many dark-purple spots corresponding to cell nuclei, but only a subset of them is in a mitotic phase and must be detected. Therefore, the challenging of pathological image analysis lies

on automated analysis of targets 1) the large number of cell nuclei among background; and the big size of high-resolution digitized pathology images 2) the variability in size, shape, appearance, and texture of the individual cell nuclei. 3) numerous cell candidates with similar appearance as the candidate cells, which is quite complicate to discern candidate cells amongst similar-looking cells with image artifacts. Recently, the development of computerized systems for automated mitosis detection has become an active area of research with the goal of developing decision support systems to retrieve pathologists work load. Contests have been held to encourage research in this topic, including the 2012 International Conference on Pattern Recognition (ICPR12) Mitosis Detection Contest (Roux et al., 2013), the Assessment of Mitosis Detection Algorithms 2013 Challenge (AMIDA13) (Veta et al., 2015) and the 2014 ICPR Mitosis Detection Challenge (MITOS-ATYPIA-14) (Roux, 2014). However, all these contest are for breast cancer, and there is no public available mitotic datasets for meningioma. Meanwhile the collection and annotation of mitotic cells on scanned slides can be very costly and time consuming.

So the goal of this thesis is to determine if it is feasible to develop sophisticated image processing and machine learning methods with limited data sets for automatic mitotic cell detection and classification.

Many researches have attempted to solve the problem. Most of approaches fall into one of the following categories: 1) efficient augmentation of label data, 2) utilization of weak label or unlabeled information, or 3) utilization of existing models/parameters for other tasks.

The formal definition of the problem is as follows: Given a patient P , and a H&E stained image I_{HE} , we will develop a model that takes I_{HE} as input, and output predictions for the mitotic cells M_P

which should include all possible candidates, therefore the pathologist can decide further based on marked candidates.

1.3 Contributions of This Thesis

This thesis provides the following contributions:

1. This work is the first to identify mitotic cells in meningioma using domain inspired features (hand crafted) from other cell types which allow for explicit modeling of the kinds of features that pathologists look for when identifying mitoses.
2. Developed and deployed sophisticated image processing and machine learning methods with limited data sets for automatic cancer cell detection and classification.

Compared the hand crafted features and domain agnostic features (deep learning method) for sensitivity of mitosis detection.

2. BACKGROUND

2.1 Meningioma

Meningiomas are the most common adult primary tumor of the central nervous system accounting for about 30 percent of all brain tumors (Buerki et al., 2018; N. Wang & Osswald, 2018), with an annual incidence of 5/1000,000. Meningiomas arise from the membranes surrounding the brain and spinal cord (Figure 1). Meningiomas are usually slow-growing tumors, and most meningiomas are grade I benign tumors, but up to 15% are atypical (grade II) and 2% anaplastic (grade III). However, some benign meningiomas follow a more aggressive progression with multiple recurrences, whereas some atypical meningiomas and malignant meningiomas can have a relative benign progression with long progression-free survival and overall survival (Rockhill, Mrugala, & Chamberlain, 2007).

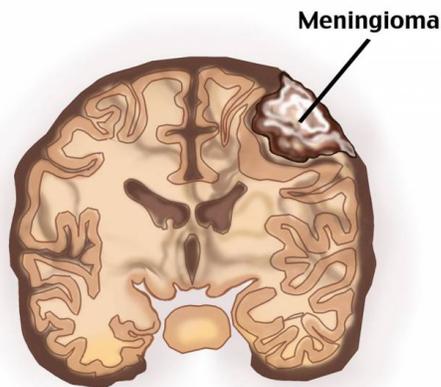


Figure 1. Meningioma in the brain

Actually, there are two important factors that determine the prognosis in meningiomas patients: the extent of resection and the tumor's histological grade. Higher grade meningiomas are more likely not to receive a clean resection, and even when they do, they have higher possibility of

recurrence. Once a tumor recurs, it is more likely to do so again, ultimately leading to a loss of local control and rarely, metastasis. Counterintuitively, it has been realized that high-grade tumor does not always correlate with high recurrence/progression, hence, much work is being performed and needed, such as advanced imaging to discover other features or phenomena that contribute to tumor growth and recurrence.

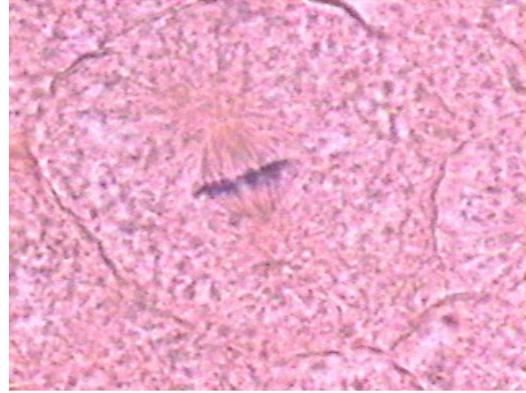
2.2 Mitotic Cells and Grading

The World Health Organization (WHO) classification for meningiomas is based solely on histopathological characterizations of mitotic rate, cellular features of atypia, and local invasion. About 80% are WHO grade I (also referred to as BM), 17% are WHO grade II (AM), and 2% are WHO grade III (anaplastic meningioma/MM)(Ostrom QT, 2017). And the key component of meningioma grad is the mitotic count (MI- mitotic index), which involves quantifying the number of cells in the process of dividing (i.e., undergoing mitosis).

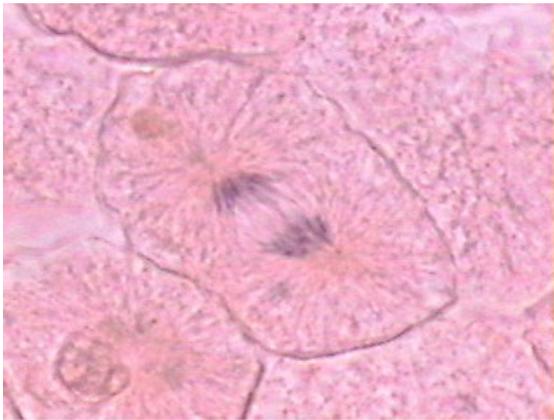
Mitosis is the process that cells reproduce to two identical daughter cells, which is the necessary step for cancer cell to form tumors and spread through the body. Mitotic figures evolve over a continuum spanning five distinct phases during which a cell nucleus undergoes various transformations: prophase, prometaphase, metaphase, anaphase and telophase. Each phase is associated with a unique shape and texture, during which chromosomes (stained in purple) is the main marker for differentiation between different phases. (Figure 2)



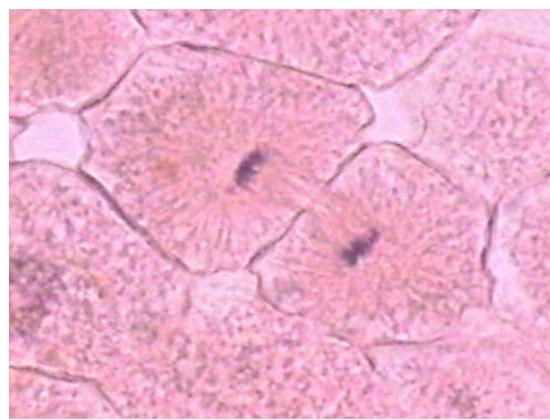
(a) **Prophase** is a stage of mitosis in which the chromatin condenses (it becomes shorter and fatter) into a highly ordered structure called a chromosome in which the chromatin becomes visible.



(b) **Metaphase** is a stage of mitosis in the eukaryotic cell cycle in which condensed & highly coiled chromosomes, carrying genetic information, align in the middle of the cell before being separated into each of the two daughter cells.



(c) **Anaphase** is the stage of mitosis when chromosomes separate in an eukaryotic cell. Each chromatid moves to opposite poles of the cell, the opposite ends of the mitotic spindle, near the micro-tubule organizing centers.



(d) **Telophase** is a stage of mitosis in a eukaryotic cell in which the effects of prophase and prometaphase events are reversed. Two daughter nuclei form in the cell. The nuclear envelopes of the daughter cells are formed from the fragments of the nuclear envelope of the parent cell. As the nuclear envelope forms around each pair of chromatids, the nucleoli reappear.

Figure 2. Different phases of mitosis

Specifically, prophase may be detected as nonmitotic cells, while later stages like anaphase and telophase of the mitotic cells may appear to split into two dark-blue spots which will count as one. There are also some abnormal mitoses with tripode, tetrapod formation, which all add the complication of mitosis detection. Meanwhile, scanned images from a single slide may not show all mitotic figures on the plane of focus, making their recognition more difficult due to areas being out of focus.

2.3 Tissue Preparation

The normal tissue preparation, which involves chemically and physically stabilizing the tissue are as follows. (Figure 3)(Lei He, 2010)

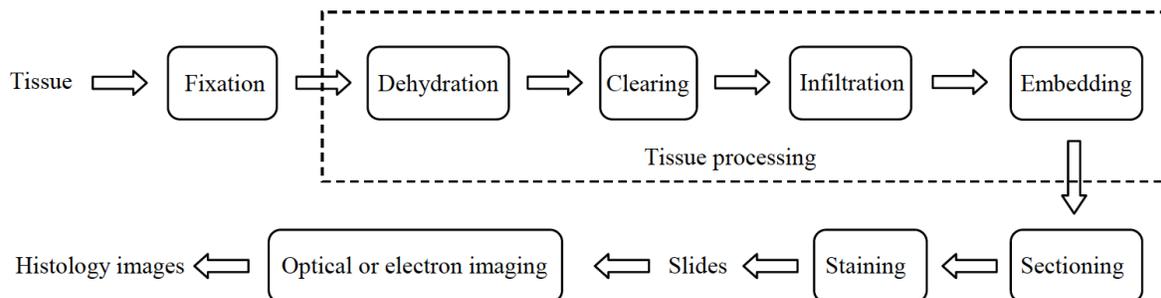


Figure 3. Histology tissue preparation and image production.

The first step of the tissue preparation is formalin fixation, it is to immerse the tissue into a fixative solution, which is helping to stop cells from breaking down. In general, tissue is fixed for a few hours (small biopsies) to about 24 hours (large biopsies). Fixation is critical because poorly fixed tissue leads to poor tissue sectioning and poor microscopic morphology.

After fixation, the tissue is physically stabilized by one of several methods (freeze drying, microwave, chemical) with the end goal of preserving the cellular morphology. Then the tissue is embedded in the paraffin. From the paraffin blocks, sections with a thickness of 3-5 μm are cut using a microtome (a high precision cutting instrument) and mounted on glass slides.

At this point, the structures of interest in the tissue, in most instances the nuclei and cytoplasm, are nearly invisible under a light microscope and must be stained to create contrast. The most widely used stains for both diagnostic and research histology are hematoxylin and eosin (H&E, Figure 4). In spite of the fact that this staining protocol has been in use for around a century, the diagnostic and prognostic procedure for all patients still almost always starts by staining the sections with H&E. The reason that H&E staining of tissue has persisted for decades is that these stains attach to almost every cellular component, providing excellent contrast between cellular constituents. Hematoxylin binds to DNA and dyes the nuclei with blue/purple color, and eosin binds to proteins and dyes other structures (cytoplasm, stroma, etc.) with pink color.

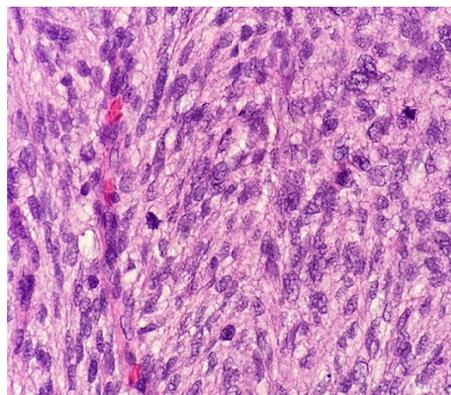


Figure 4. Hematoxylin and eosin (H&E)

2.4 Automated Histology Image Analysis

After the tissue preparation and image production through different imaging technologies, the resulting digital histology images are ready for analysis by pathologists or CAD systems. Computer assisted disease diagnosis (CAD) has been conducted for various cancer detection and grading applications, including prostate, breast, renal cell carcinoma, pediatric tumor neuroblastoma and lung cancer grading. A typical CAD system for histology image analysis is shown in Fig. 5. This system consists of conventional image processing and analysis tools, including preprocessing, image segmentation, feature extraction, feature dimension reduction, feature-based classification, and postprocessing. The sequential order of these functional modules may be changed or omitted in practical applications.

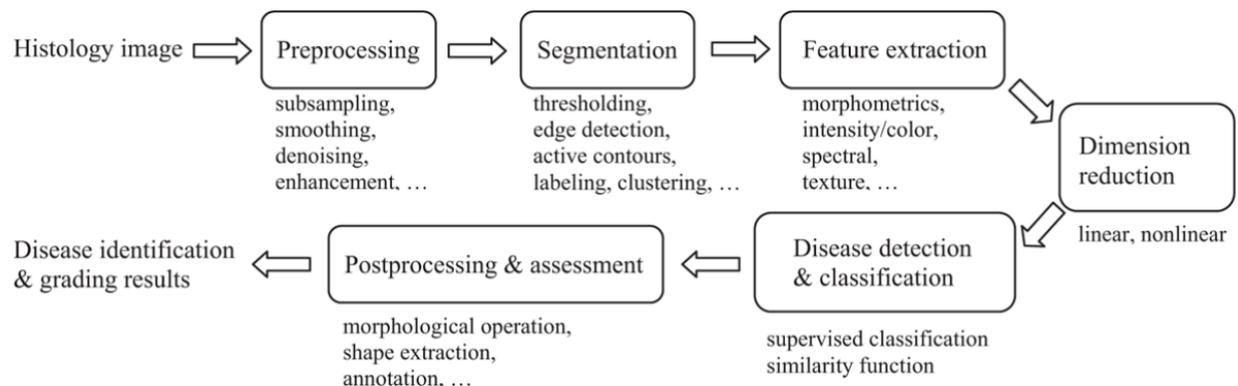


Figure 5. Computer assisted diagnosis workflow.

2.4.1 Image Acquisition

Image acquisition is the prerequisite in image analysis. Traditional method is to use a microscope-mounted digital camera that captures individual field images. Currently, a whole-slide-image (WSI) scanner can scan through an entire tissue slice to obtain a high-resolution whole-slide-image for pathology archiving, education, and image analysis. Recently, The U.S. Food and Drug

Administration just permitted marketing of the Philips IntelliSite Pathology Solution (PIPS), the first whole slide imaging (WSI) system that allows for review and interpretation of digital surgical pathology slides prepared from biopsied tissue, which opened a new era of digital histology. Whole slide scanners capture images of tissue sections tile by tile or in a line-scanning fashion, and the images are usually in gigabytes.

Scanning can occur at multiple magnifications. Scanning at 20X magnification is usually acceptable for standard viewing and interpretation, including routine image analysis of hematoxylin-eosin (H&E) and IHC slides. For other applications, such as digitization of in situ hybridization slides, images should be acquired at 40X to resolve information that may be separated by distances less than about 0.5 μm . And for mitotic detection, the pathologists selected 10 high power fields (HPF) at 40X magnification. An HPF has a size of 512 x 512 μm^2 (an area of 0.262 mm^2), which is the equivalent of a microscope field diameter of 0.58 mm.

When performing image analysis of histopathology images, it is of interest to separate the histological stains that dye different tissue components. Multiple resolutions are stored for a single whole slide image for streamlined image loading. For example, a sample whole slide image acquired at 40x by the Aperio Scanscope whole slide scanner is accompanied by the same image down sampled at 10x, 2.5x, and 1.25x, as well as a thumbnail image that represents the entire tissue fit within a ~1-megapixel frame.

2.4.2 Image Preprocessing

Image quality acquired by scanner varies due to many factors including the innate meningioma. Therefore, appropriate preprocessing methods could reduce variations to some degree(Bhattacharyya, 2011), such as color normalization which is to minimize staining

variations(Khan, Rajpoot, Treanor, & Magee, 2014), spatial filtering which is to highlight major image structure, denoising which is to reduce image noise, and enhancement which is to optimize contrast between objects of interest and background(He, Long, Antani, & Thoma, 2012). Major categories of image denoising methods include partial differential equation (PDE)-based anisotropic diffusion(Malik, 1990) variational methods(Rudin, Osher, & Fatemi, 1992), robust statistics(Black, 1998), and wavelet thresholding(Donoho, 1995). Traditional enhancement techniques include adaptive filters(Gonzalez, 2008) and inverse (backward) anisotropic diffusions(Gilboa, Sochen, & Zeevi, 2002). Particularly, intensity centering and histogram equalization were normally used to normalize a diverse set of pathology images.

2.4.3 Cell Detection and Segmentation

Detection and segmentation of cell nuclei as region of interest from the background is the crucial and fundamental steps for further cancer identification and classification. And the measurements of the nucleus features, especially mitotic cell features that are distinct among different types of cells are crucial for meningioma diagnosis and prognosis. However, nuclei segmentation is a very challenging problem, especially for H&E stained slides, due to the variability of tissue appearance caused by imperfections in the staining process.

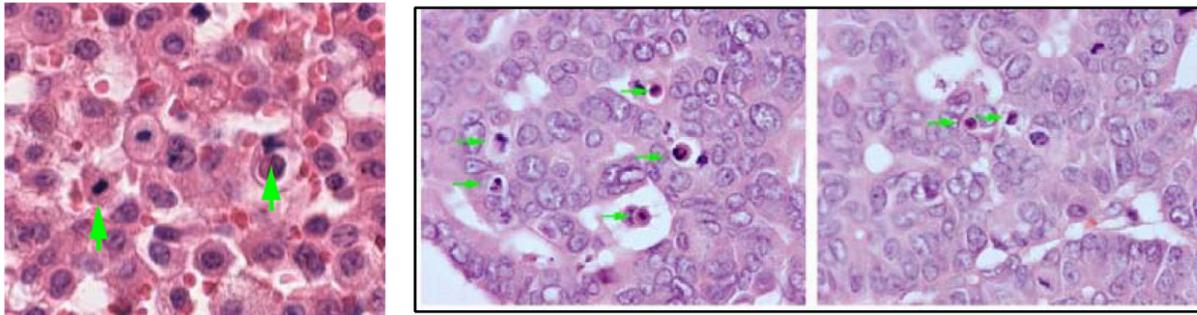
A large variety of approaches for segmentation of nuclei in histopathology images have been proposed by researchers. And the most difficult part of nuclei segmentation in H&E stained histology images is to detect individual nuclei, especially when they are clustered together and overlap on top of each other. Easy segmentation methods have so far been proposed for H&E images including threshold and clustering, and complex algorithm consist of intensity-based (Arteta, Lempitsky, Noble, & Zisserman, 2012), water shed based (Ali & Madabhushi, 2012;

Cheng & Rajapakse, 2009; Li, Zhou, Ma, & Wong, 2010), voting-based (Parvin et al., 2007), mathematical morphology– based (Esmaeilsabzali, Sakaki, Dechev, Burke, & Park, 2012; Yan et al., 2013), texture based (Omar S. Al-Kadi, 2017), color-based Laplacian of Gaussian (LoG)-filter based (Al-Kofahi, Lassoued, Lee, & Roysam, 2010), gradient-based (Esmaeilsabzali et al., 2012), region growth and Markov random field (MRF) (Fatakdawala et al., 2010), H-minima transform– based (Jung & Kim, 2010), Gaussian mixture model (Vink, Van Leeuwen, Van Deurzen, & De Haan, 2013), and deep learning (Xu et al., 2016) approaches. Other popular models consist of active contour model and level set model (Hoque et al., 2001; Qi, Xing, Foran, & Yang, 2012; Vink et al., 2013). Although these models show efficiency in nuclei detection, it is still very challenging to find proper seed points or deciding initial contours. Furthermore, false or miss detected nuclei regions will be propagated during segmentation steps which causes inaccurate segmentation results (Jia-Mei Chen, 2017). Approaches to separate clumps of nuclei including splitting them along points of high concavity (Fatakdawala et al., 2010; Wienert et al., 2012) or by unsupervised Bayesian classification (Jung, Kim, Chae, & Oh, 2010). Usually the choice of segmentation methods depends on the intended application and the feasible time and computational resources. For example, graph cuts may yield superior nuclear segmentation, however, it is really time consuming and power consuming, which limit their use for high throughput applications, especially for the whole slide images' segmentation which contains too much nucleus. Therefore, if such high throughput processing is needed, a simple method like thresholding followed by morphological operations might be used with cost of less accurate segmentation results. Meanwhile, different segmentation methods are used for different cell types, such as Bayesian classifier was used for meningioma subtype discrimination (O. S. Al-Kadi, 2010). Region based approach, which is also named as seeded volume growing is used by M.

MuthuRama Krishnan (Muthu Rama Krishnan et al., 2009) for segmenting and classifying subepithelial connective tissue (SECT) cells in oral mucosa, and also used for application correlated with breast, lung cancer, prostate (A. N. Basavanhally et al., 2010; Y. Y. Wang, Chang, Wu, Tsai, & Sun, 2007). Segmentation methods for shape based algorithm such as expectation maximization (EM) algorithm can be used for detecting centers of lymphocytes in breast cancer images (Fatakawala et al., 2010); color and texture-based segmentation method is used for colon cancer detection (Gunduz-Demir, Kandemir, Tosun, & Sokmensuer, 2010). Since the tissue architecture and cell shape, textures are different for different organs, therefore the methods applied for one cell type may not work on another.

2.4.4 Mitosis Detection

Mitotic cells are hyperchromatic objects due to the nature of DNA duplication, and it loses nuclear membrane with “hairy” protrusions around the edges and basophilia instead of eosinophilia in the surrounding cytoplasm. However, all these features are more instructive than definitive, and it takes a lot of training and practicing for pathologists to become an expert, due to the fact that many other objects such as apoptotic and necrotic nuclei may appear similarly, or prophase may be detected as nonmitotic cells, while later stages like anaphase and telophase of the mitotic cells may appear to split into two dark-blue spots which will count as one, all of these lead to variability in measurements and high inter-observer error. By far, the most distinctive feature of the mitotic figures is their hyperchromicity, which means that the intensity of the staining of the mitotic nucleus is noticeably darker than normal nuclei (illustrated in figure 6a).



(a) Mitotic figures (arrow).

(b) Apoptotic cells (arrow).

Figure 6. Mitotic figures and apoptotic figures.

Therefore, these features are often used to perform thresholding, pixel-wise classification or local intensity minima detection, followed by morphological operations and/or active contours segmentation. However, many other objects such as apoptotic and necrotic nuclei may appear similar features, which cause the false detection (illustrated in figure 6b). Moreover, performance of mitosis detection largely depends on image quality. However, digital images are acquired at single focal plane, which may cause some of the nucleus out of focus and affect the pathologist's judgment (illustrated in Figure 7). Therefore, we anticipate that in the near future, multiple focal planes can be acquired and stored in a new image compression technique.



Figure 7. Example digital images with different focus offset values(Caglar Senaras, 2018)

2.4.5 Feature Extraction and Dimension Reduction

After segmented the single cells, feature extraction is subsequently performed to encode morphological parameters of the nucleus into a set of features, which is important for later classification. To mimic the approach of reading a histology images by pathologists, different levels of features will be extracted. These features includes pixel level, object level, and spatial level features (Gurcan et al., 2009; He et al., 2012). Pixel-level features include color features and texture features, which are the least interpretable in terms of pathological knowledge. And for mitotic cell detection, color features are usually widely used since one of the remarkable features of mitotic nucleus is hyperchromicity (Irshad et al., 2013). Texture features including features like Haralick entropy, Gabor filter, power spectrum, co-occurrence matrices, and wavelets. Object-level features consists size and shape features, which can be used to describe morphological characteristics of individual microstructure. Spatial-level features are used to describe the relationship of cells with neighbor cells, which are statistical domain-specific information (e.g. Voronoi diagrams, Delaunay triangulation, and minimum spanning trees). To derive information from images as more as possible, a huge number of features need to be extracted. However, it is not true that more features make the classification more accurate, because first of all, high-dimensional feature are computationally more consuming, second it may contain irrelevant and redundant features which may hinder in achieving high classification accuracy. Therefore, feature dimensionality reduction methods such as direct feature selection based on domain-specific knowledge or ranking methods such as maximum Relevance Minimum Redundancy (mRMR) (Peng, Long, & Ding, 2005) are usually used to identify the most discriminative features (Bouzas, Arvanitopoulos, & Tefas, 2015; Mignotte, 2011). Moreover, feature dimensionality reduction tools consist linear and nonlinear techniques. Linear techniques such as principal component analysis (PCA) (Richard O. Duda, 2000), linear discriminant analysis (LDA) (Richard O. Duda,

2000), and multidimensional scaling (MDS) (Venna & Kaski, 2006), which use the Euclidean distance between the feature points. As to nonlinear dimensionality reduction methods, it consists of spectral clustering, isometric mapping (Isomap) (Tenenbaum, de Silva, & Langford, 2000), locally linear embedding (LLE) (Roweis & Saul, 2000), and Laplacian eigenmaps (LEM) (Belkin, 2003), among which Euclidean relationship among the feature points is not assumed.

2.4.6 Classification

For histopathology images, after segmentation and feature extraction, selection of different classifiers are important for diagnosing abnormality in image. Usually, specialized features are used to train a classification model, which includes supervised and unsupervised methods. With labelled training sets, supervised methods are typically use ‘ground truth’ reference images for training, and use trained models to predict the classes of unlabeled data. In contrast, without a set of labeled data, unsupervised classification methods are used which do not require pre-labelled training sets for their learning but instead rely on certain similarity measures to group data into separate homogeneous clusters. Unsupervised classification methods includes K-means, fuzzy c-means, ISODATA clustering, self-organizing map (Haykin, 1999), and adaptive resonance theory (Carpenter & Grossberg, 2003), such as neural network, k-nearest neighborhood algorithm, logistic regression method, fuzzy systems etc (Amoli D. Belsare, 2012). A support vector machine (SVM) classifier is most commonly used classifier for cancer diagnosis including differentiating between cancerous and non-cancerous images (Raza, Parry, Moffitt, Young, & Wang, 2011) and distinguishing between different stages of cancer (A. Basavanhally et al., 2013; Raza et al., 2011). Better accuracy has been achieved by using supervised Deep Neural Networks(DNN) in breast cancer histology images (Ciresan, Giusti, Gambardella, & Schmidhuber, 2013). After training the DNN automatically with a set of visual features

from the training data, it directly operates on raw RGB data sampled from a square to differentiate patches centered on the nucleus between mitotic from non-mitotic. The approach that was proposed in (Ciresan et al., 2013) is unique since it uses deep convolutional neural networks instead of candidate detection such as segmentation as an initial stage. Instead, it performs classification at every pixel location, which achieves excellent results. However, this method needs a lot of ground truth data for training, hence it is usually time and power consuming.

As in segmentation by clustering methods, supervised algorithms can be applied to determine and grade disease. SVM, for example, which is one of the widely used machine learning tools, finds the hyperplane with maximum distance to the nearest training samples. The original linear SVM can also be extended to nonlinear feature space with the Kernel Trick, i.e., kernel functions based on inner products of two vectors. In feature similarity computation, a number of metrics (Richard O. Duda, 2000) can be applied besides the commonly used Euclidean distance, such as Mahalanobis distance, Manhattan distance, and Chebyshev distance.

3. METHODOLOGY

3.1 Overall Approach

This thesis proposes robust automated methods to identify and quantify mitotic cells in meningioma whole slides. The approach for dealing with mitotic cells is explained in the Figure 8. As shown in the figure, biopsy images are acquired, and pathologists provide annotations on mitotic cells on few biopsy images manually. We used the whole slide biopsy images and get image patches that mimics the pathologist's procedure. And perform segmentation for identifying each single cell in an image patch. Attributes are extracted for each cell region. Since the labeled images for meningioma are so little, which is also the major problem of getting annotated images for training in real world. we used unsupervised method, and also supervised methods. For supervised method, we used labeled breast tumor mitotic cells from public source and used attributes from breast tumor for model training since mitotic cells have same features among all cell type. And then predicted on meningioma images to calculate the precision, accuracy and F1 score.

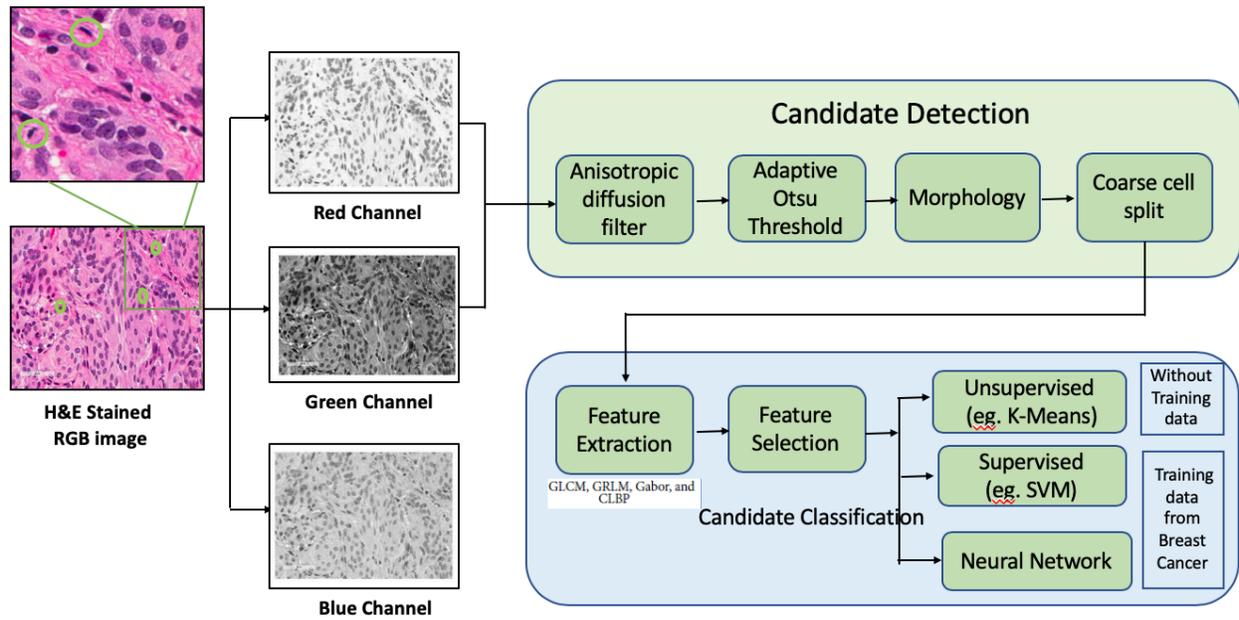


Figure 8. Overall Approach

3.2 Data

3.2.1 Images

The data set used in this research are provided by Dr. Dibson D. Gondim in Indiana University Health, which consists H&E stained biopsy section images from meningiomas patients. After H&E slides are acquired, they have been scanned by Aperio Scanscope by Kyle Christopher McElyea at the IU School of Medicine. Aperio scanners generate a semi-proprietary file format called SVS, which is a multi-page tiff file that stored a pyramid of smaller tiff files of the original image. To read an SVS image, Aperio ImageScope can be used to open the image in Windows operating system, while QuPath can be used in Mac operation system. If we open one image in Aperio ImageScope, the figure below will be displayed (Figure 9.)

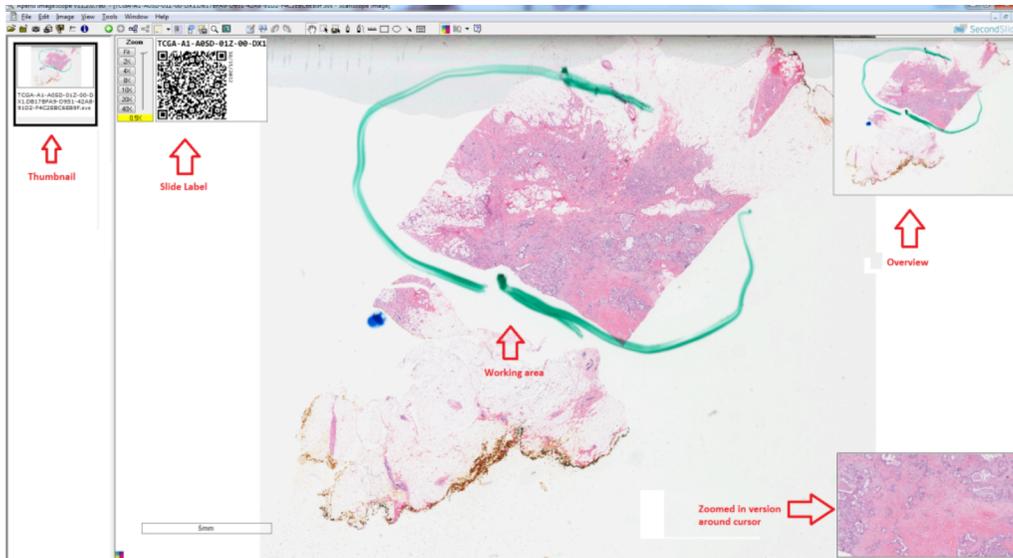


Figure 9. SVS image view at Aperi ImageScope.¹

For each view, there are five different views (in arrows) can be seen, from left to right are: a thumbnail view, a slide label view, a working area view which can be scrolled in and out of regions of interest (ROI), a high-level overview which can be used to choose ROI fast, and a low level view which shows the zoomed in version around cursor. Actually, all different levels of the images are compressed in one SVS file, and those levels are illustrated in figure 9.

Pyramid	Width	Height	Ratio
Thumbnail	899	768	104:1
Level 3	2,939	2,508	32:1
Level 2	5,879	5,017	16:1
Level 1	23,518	20,071	4:1
Base	94,075	80,287	

Figure 10. Various layers in the pyramid.

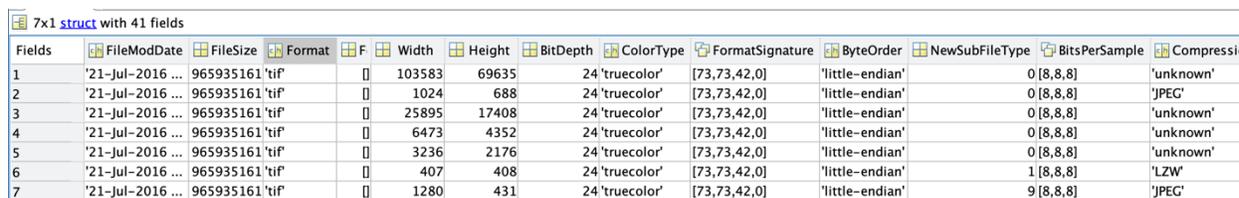
¹ <http://www.andrewjanowczyk.com/working-with-aperio-svs-files-in-matlab-introduction/>

Matlab natively supports multi-page tiff reading by simply providing index like:

```
01. io=imread('TCGA-A1-A0SD-01Z-00-DX1.DB17BFA9-D951-42A8-91D2-
F4C2EBC6EB9F.svs','Index',2);
```

Figure 11. Matlab commands for open svx file

However, our images cannot be opened by typing this simple command since several levels of images are not compressed correctly and corresponding information are lack for those levels (illustrated in last column at Figure 11.) Fields 1 stands for original image, fields 2 stands for thumbnail images, fields 3 stands for level 1 image which is 1:4 ratio to original image, fields 4 stands for level 2 image which is 1:16 to original image, fields 5 stands for level 5 image which is 1:32 to original image, fields 6 stands for slide label which is the ID information for the slide, fields 7 stands for whole image.



Fields	FileModDate	FileSize	Format	F	Width	Height	BitDepth	ColorType	FormatSignature	ByteOrder	NewSubFileType	BitsPerSample	Compressio
1	'21-Jul-2016 ...	965935161	'tif'	[]	103583	69635	24	'truecolor'	[73,73,42,0]	'little-endian'	0	[8,8,8]	'unknown'
2	'21-Jul-2016 ...	965935161	'tif'	[]	1024	688	24	'truecolor'	[73,73,42,0]	'little-endian'	0	[8,8,8]	'JPEG'
3	'21-Jul-2016 ...	965935161	'tif'	[]	25895	17408	24	'truecolor'	[73,73,42,0]	'little-endian'	0	[8,8,8]	'unknown'
4	'21-Jul-2016 ...	965935161	'tif'	[]	6473	4352	24	'truecolor'	[73,73,42,0]	'little-endian'	0	[8,8,8]	'unknown'
5	'21-Jul-2016 ...	965935161	'tif'	[]	3236	2176	24	'truecolor'	[73,73,42,0]	'little-endian'	0	[8,8,8]	'unknown'
6	'21-Jul-2016 ...	965935161	'tif'	[]	407	408	24	'truecolor'	[73,73,42,0]	'little-endian'	1	[8,8,8]	'LZW'
7	'21-Jul-2016 ...	965935161	'tif'	[]	1280	431	24	'truecolor'	[73,73,42,0]	'little-endian'	9	[8,8,8]	'JPEG'

Figure 12. Matlab 'imfinfo' command for image information.

To solve this problem, we imported OpenSlide Python module into matlab, and set the environment to run python command in matlab. The original resolution images consume a large amount of memory, i.e. up-to to gigabytes for each image. To optimize the disk and memory usage during image processing, high level of images will be used for acquiring 10 HPF like pathologists' routine and then used for following image process.

3.2.2 Pathologist Annotation

For the purpose of this research, Dr. Dibson D. Gondim who is a practicing Anatomic Pathologist in Indiana University Health provided us with some cases of annotation of mitotic cells from meningiomas patient slides. And the annotations are divided into two categories, one is definitely mitosis, and the other is definitely not mitosis. For definitely mitosis, there are total 10 cases, while for definitely not mitosis, there are 5 cases (Figure 12).

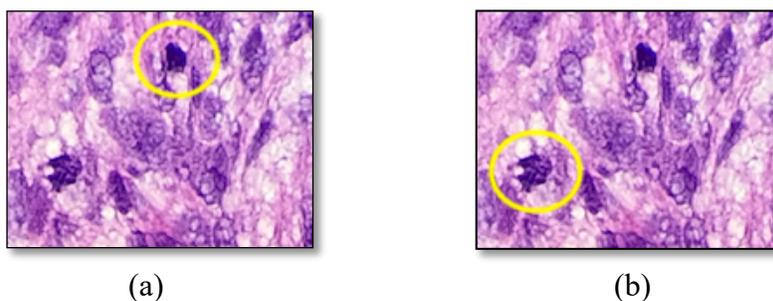


Figure 13. (a) Definitely mitosis cell (b) Definitely not mitosis cell

Since we lack availability of large volumes of labelled data (Ground Truth) which requires Dr. Dibson a lot of time and effort for the annotations. We also refer the MITOS datasets provided for the MITOS-ATYPIA contest 2012 and 2014. And the beauty of MITOS datasets are all the annotations come from the opinions of two pathologists (or three in case of disagreement), providing to contestants some additional clues about objects that are clearly mitosis (agreement between pathologists) and those that are not so easy to identify (disagreement between pathologists). In each slide, the pathologists selected 10 high power fields (HPF) at 40X magnification. An HPF has a size of $512 \times 512 \text{ m}^2$ (that is an area of 0.262 mm^2), which is the equivalent of a microscope field diameter of 0.58 mm , and equivalent of 2084×2084 pixels. These 50 HPFs contain more than 500 mitosis in total, with 686659 non-mitotic cells in total. As there are several possible shapes for mitosis, it is necessary to have a large dataset to be able to cover all the cases.

3.3 Image Preprocessing

This section describes the technical process used to preprocess the images, which prepare for the segmentation.

1. Obtain Image – Patients tumor samples are obtained and fixed onto slides, and then stained with a hematoxylineosin (H&E) stain. All slides are scanned and stored as RGB images with 8 bits per color channel.
2. Read highest level of raw image – Used OpenSlides module to open images in matlab for level 5 image which is 1:32 to original image. And exported saved images into matlab.
3. For each raw image, convert RGB image to gray scale by using the following formula to create : $0.2989 * R + 0.5870 * G + 0.1140 * B$ (Pratt, 2001). Get the mask image which excludes the white region for later process, which is to distinguish between tissue sample and white background. The filtering process is to get the foreground with pixel bigger than 200.
4. Create blocks for 10 high power fields selection – Based on the mask image that created from last process, small image patches with 2084 x 2084 pixels in raw images will be created if more than 90% of the content are not background (white region), which equals to 65 x 65 pixel in level 5 images (1:32 to original image). Meanwhile, the location of the image patches are created for later allocation.
5. Get rid of the image label marks – Due to the whole image features, the labels created by blue markers are also included in the image patches for the high pixel value. Moreover, those markers usually appear at the corner. Therefore to avoid being selected as real tissue images for later process, four corner image patches plus one more column and one more row and their connected images which are recognized as real images will be excluded for later random selection.

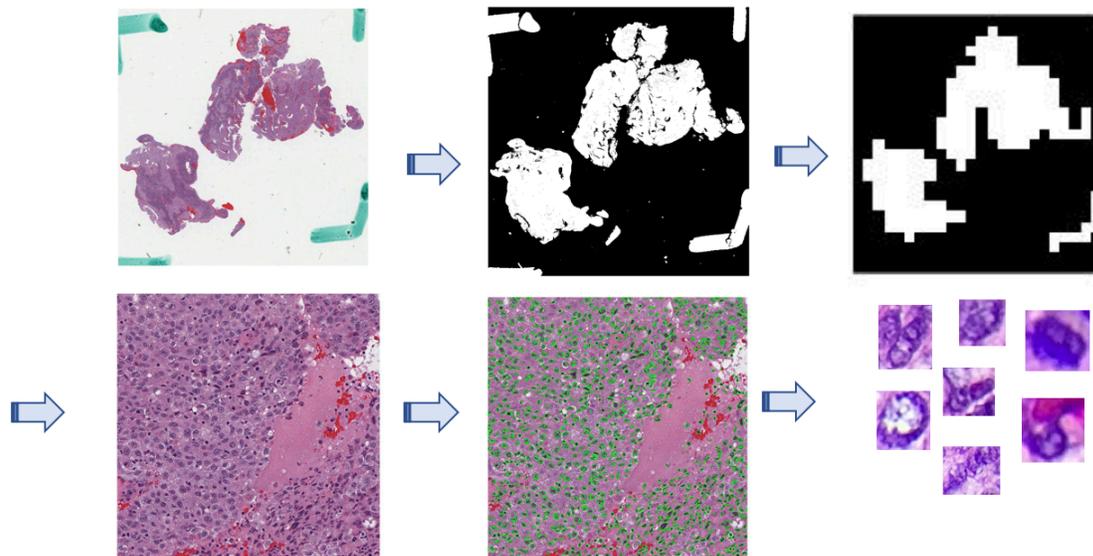


Figure 14. Image preprocess steps for whole slide images:

Step1: Whole slide image in lowest resolution(1:32); Step2: Get foreground images by filtering out the background part; Step3: Generate image patches in low resolution for randomly pick 10 areas.

6. Randomly pick 10 rectangle blocks that generated from last step and projected back to the original images to get 2084 x 2084 image patches for later cell segmentation since it is the size for 1 HPF under 40x magnification.
7. For each image, first normalize the image with the ref_image with function imhistmatch, so that each image will be normalized to the similar histogram, which has a better distinguish between foreground and background. See figure below.

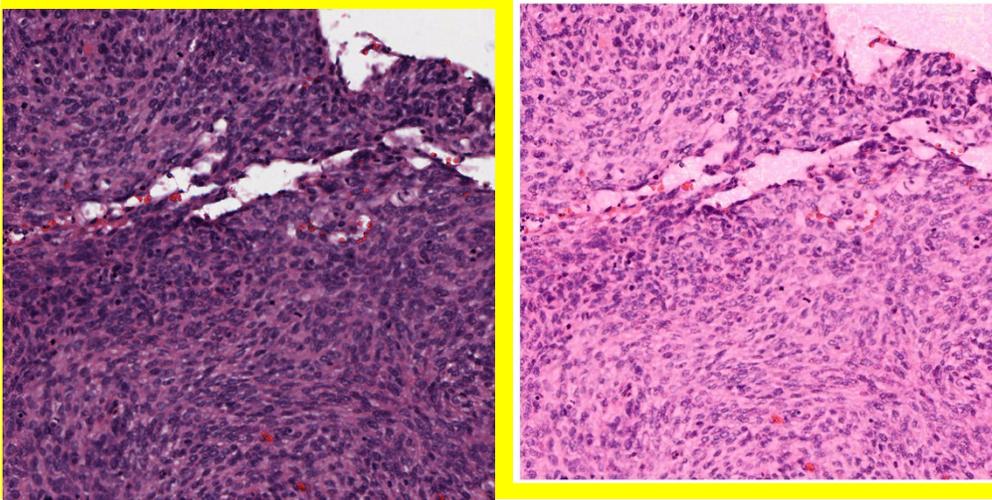


Figure 15. Before normalization and after normalization

3.4 Cell Segmentation

This step is to segment the cell nuclei from H&E stained backgrounds for feature extractions and deep learning.

1. First red channel is extracted from R,G,B channel (Figure 16), and x 1.3 is used for adjustment to make the contrast better for segmentation (Figure 17)

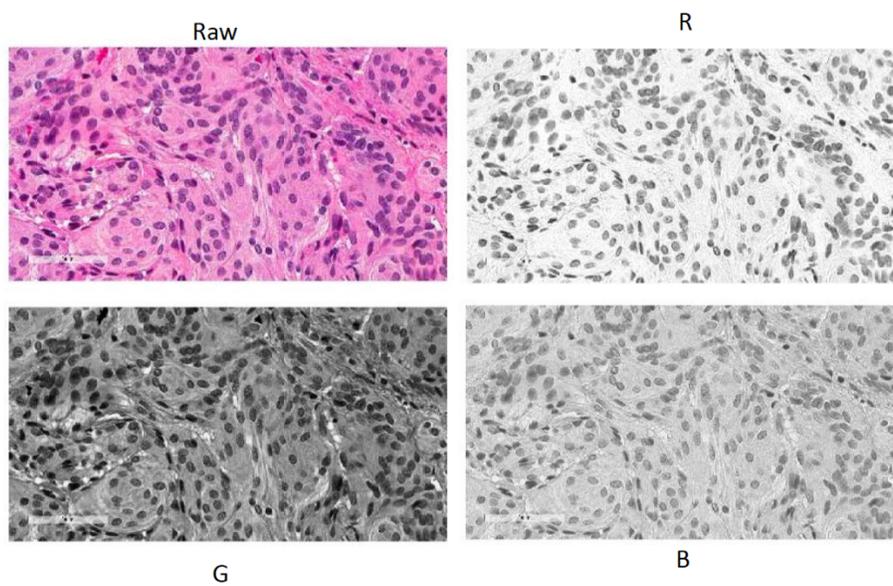


Figure 16. RGB channel of raw images

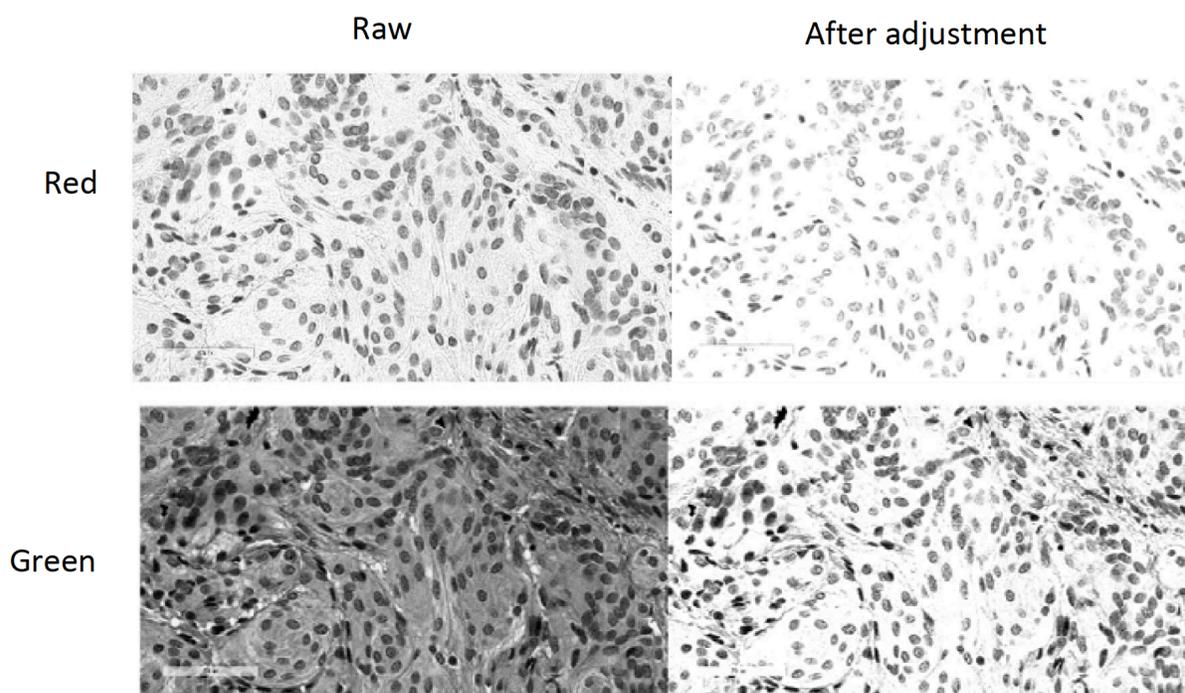


Figure 17. Green channel before and after adjustment

2. Anisotropic diffusion filter. For this filter method (Malik, 1990) it removes noise from images without blurring edges which can greatly preserve the edge information of nuclei. A function called `anisodiff2D` from mathworks is used, in this function:

A 2D network structure of 8 neighboring nodes is considered for diffusion conduction.

(function: `diff_im = anisodiff2D(IM, NUM_ITER, DELTA_T, KAPPA, OPTION)`)

ARGUMENT DESCRIPTION:

IM - gray scale image (MxN).

NUM_ITER - number of iterations.

DELTA_T - integration constant ($0 \leq \text{delta}_t \leq 1/7$).

KAPPA - gradient modulus threshold that controls the conduction.

OPTION - conduction coefficient functions proposed by Perona & Malik:

Since Kappa controls the strength of edges, and with Kappa increases, the diffusion will become linear. Therefore, after getting the average gradient magnitude of the image, and experimenting different K that is lower than the threshold, I found 30 is a good number for my image. Also with diffusion time increases, the background become clearer. After experimenting different iteration numbers, 20 was chosen (Figure below)

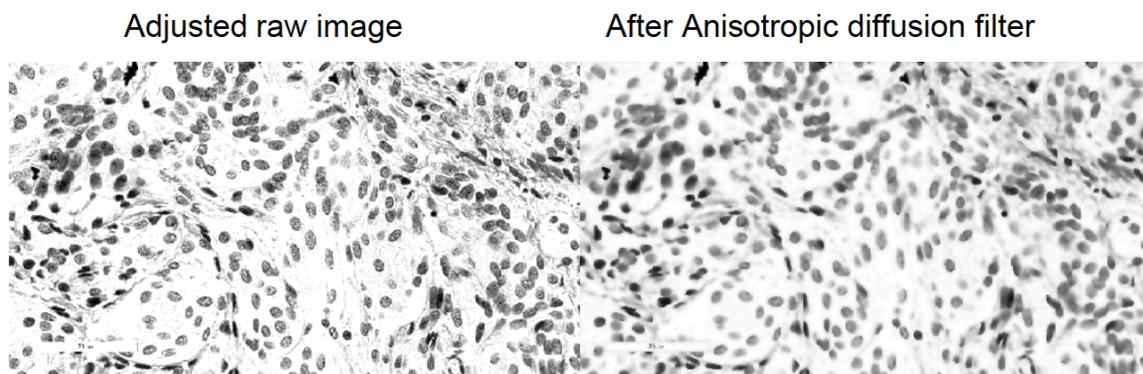


Figure 18. Before and after anisotropic diffusion

3. Use adaptive otsu threshold method to segment cells. The block size of the window depends on the row and column of the image, therefore for different images, the size will be different. Meanwhile, overlapping block ('border') also tried to see whether it will give a better result for covering the overlaps, however, it did not work as good as blockproc function without overlapping. Time is recorded for each different block size to monitor the efficiency. Binary images of the nuclei were obtained after this step, shown in yellow.

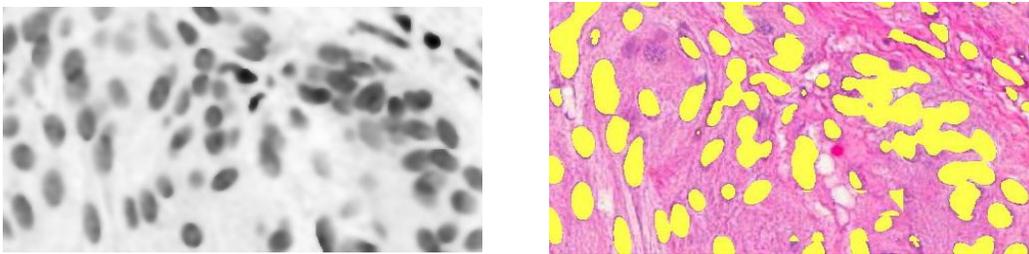


Figure 19. Before and after cell segmentation

4. Remove small region areas: After obtaining the binary images, small region areas are measured by using image tool, and 80 pixels are used as a threshold for getting rid of these small regions, since they are usually come from the background (Image below)



Figure 20. Before and after small region removal, and the display of removed region on raw image

5. Get the boundary of each cell by using “bwboundaries” function obtained at previous step. (See figure below). The red channel will not include the red part as arrow shows, which is not the cell.

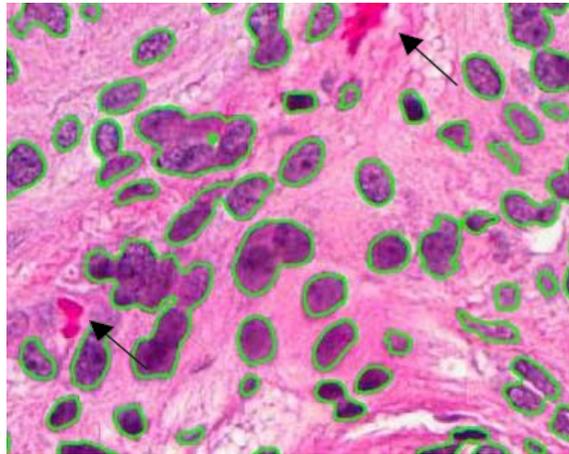


Figure 21. Cell boundaries by using red channel

6. Cell cluster splitting.

To split the cell cluster, the distance transform of the complement of the binary image is first computed using “bwdist”, and then compute the watershed transform and display the resulting label matrix as an RGB image (Fig 22). Most of the cells are splitted (right square box) while there are still some cells are not split correctly (left square box), however those sub-cluster cells will be analyzed differently in the later procedure.

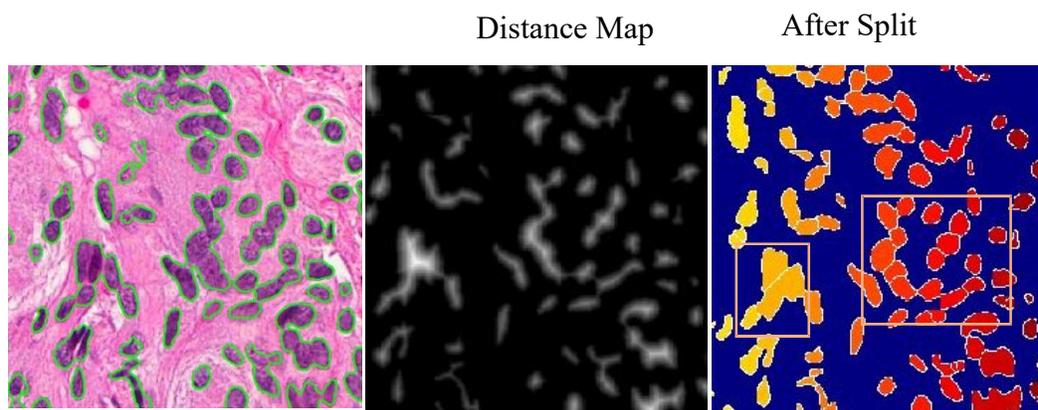


Figure 22. Cells before and after splitting

7. Feature extraction

7.1 After single nuclei region obtained, their morphological features can be extracted for differentiating between different types of nucleus. Those morphological features including:

- 1) Area – total number of pixels in the region
- 2) MajorAxisLength -- The length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region
- 3) MinorAxisLength -- The length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region
- 4) Perimeter – The distance (in pixels) around the boundary of the region
- 5) Eccentricity --The eccentricity of the ellipse that has the same second-moments as the region, the eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1.

6) Extent -- The ratio of pixels in the region to pixels in the total bounding box

7) EquivDiameter -- The diameter of a circle with the same area as the region

Since texture is an import character of differentiating cells between different types, texture statistics are also computed for the gray scale image of a given region.

8) Solidity -- Proportion of the pixels in the convex hull that are also in the region, returned as a scalar. Computed as $\text{Area}/\text{ConvexArea}$.

7.2 Basic texture statistics including:

- 1) Mean Intensity – The average pixel intensity
- 2) Maximum intensity—The maximum pixel intensiy
- 3) Minimum intensity – The minimum pixel intensiy

4) Standard Deviation – The standard deviation of the intensity in the region, computed as distance Map After Split the square root of the variance, σ^2

5) Smoothness (R)– It equals 0 for images of constant intensity and 1 for large value of

$\sigma^2(z)$, computed as $R = 1 - \frac{1}{1+s}$, where $s = \frac{\sigma^2(z)}{(L-1)^2}$, L is the number of image intensity level.

6) Third moment – Central sample moment of the region is calculated for $n=3$, it is computed as $M_k = E[(x-\mu)^k]$, where $E[x]$ is expected value of x.

7) Uniformity - Measure the relative smoothness or coarseness of a region, computed as $U = \sum p^2$, where p is the histogram of the image.

8) Entropy -- A statistical measure of randomness that can be used to characterize the texture of the input image. Computed as $e = -\sum p * \log_2 p$, where p is histogram of the image

9) Skewness -- Skewness is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data spreads out more to the left of the mean than to the right. If skewness is positive, the data spreads out more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero. The skewness of a distribution is defined as

$$s = \frac{E(x-\mu)^3}{\sigma^3}, \text{ where } \mu \text{ is the mean of } x, \sigma \text{ is the standard deviation of } x, \text{ and } E(t)$$

represents the expected value of the quantity t. The skewness function computes a sample version of this population value.

10) Kurtosis -- Kurtosis is a measure of how outlier-prone a distribution is. The kurtosis of the normal distribution is 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3; distributions that are less outlier-prone have kurtosis less than 3. Some definitions of kurtosis subtract 3 from the computed value, so

that the normal distribution has kurtosis of 0. The kurtosis function does not use this convention.

The kurtosis of a distribution is defined as :

$$K = \frac{E(x-\mu)^4}{\sigma^4}, \text{ where } \mu \text{ is the mean of } x, \sigma \text{ is the standard deviation of } x, \text{ and } E(t)$$

represents the expected value of the quantity t . The kurtosis function computes a sample version of this population value.

8. Feature Standardization

All extracted features are normalized by standardization method to zero mean and unit variance, so that they'll have the properties of a standard normal distribution with $\mu=0$ and $\sigma=1$. It defined as:

$$Z = \frac{x - \mu}{\sigma}$$

Without feature standardization, features can be on different scales, certain weights may update faster than others during since the feature values x_j play a role in the weight updates

$$\Delta w_j = -\eta \frac{\partial J}{\partial w_j} = \eta \sum_i (t^{(i)} - o^{(i)}) x_j^{(i)},$$

so that $W_j := w_j + \Delta w_j$, where η is the learning rate, t the target class label, and o the actual output. All the clustering algorithms that use Euclidean distance measures matters with or without feature scaling – in fact, tree-based classifier are probably the only classifiers where feature scaling doesn't make a difference.

9. Feature Selection/Ranking

Use correlation method to check any of the texture features is correlated or not, if true, then only one of them will be used. And Attribute Selection- InfoGainAttributeEval from Weka

is used for ranking the features. It evaluates the worth of an attribute by measuring the information gain with respect to the class.

$$\text{InfoGain}(\text{Class}, \text{Attribute}) = H(\text{Class}) - H(\text{Class} | \text{Attribute}).$$

3.5 K-Means Clustering

Since we have very few labeled data, probably 10 samples for positive, and 1700 samples for negative, it will be very hard to do supervised methods. Therefore, the first method I tried is the unsupervised method k-means. Typically, unsupervised algorithms make inferences from datasets using only input vectors without referring to known, or labelled, outcomes. Typically, unsupervised algorithms make inferences from datasets using only input vectors without referring to known, or labelled, outcomes. A cluster refers to a collection of data points aggregated together because of certain similarities. And we know that my dataset needs to be separated into two clusters, one is mitotic cell cluster, another is non-mitotic cell cluster, but we don't know whether the feature we extracted will be enough to separate the cells into two clusters. Therefore, we used the elbow method to determine the optimal number of clusters for k-means first (Demidenko, 2018).

The idea of the elbow method is to run k-means clustering on the dataset for a range of values of k (say, k from 1 to 10 in the examples above), and for each value of k calculate the sum of squared errors (SSE). Like this:

```
var sse = {};  
for (var k = 1; k <= maxK; ++k) {  
  sse[k] = 0;  
  clusters = kmeans(dataset, k);  
  clusters.forEach(function(cluster) {  
    mean = clusterMean(cluster);  
    cluster.forEach(function(datapoint) {  
      sse[k] += Math.pow(datapoint - mean, 2);  
    });  
  });  
}
```

Then, plot a line chart of the SSE for each value of k . If the line chart looks like an arm, then the "elbow" on the arm is the value of k that is the best.

The idea is that we want a small SSE, but that the SSE tends to decrease toward 0 as we increase k (the SSE is 0 when k is equal to the number of data points in the dataset, because then each data point is its own cluster, and there is no error between it and the center of its cluster). So our goal is to choose a small value of k that still has a low SSE, and the elbow usually represents where we start to have diminishing returns by increasing k .

So the elbow method for my dataset looks like following:

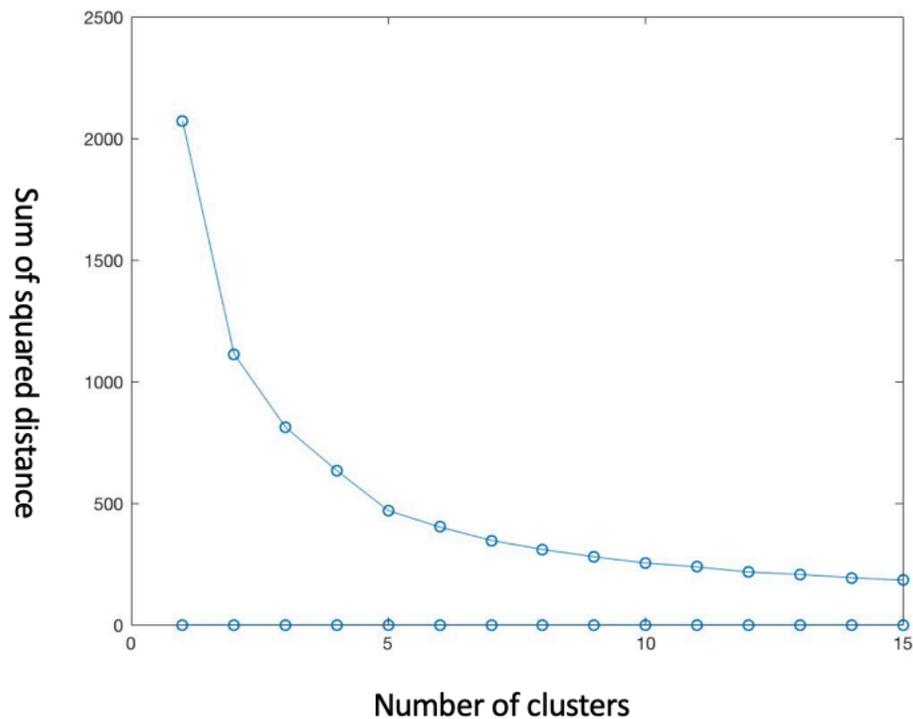


Figure 23. Elbow method for k-means

So based on the elbow method, $k = 5$ is chosen for k-means methods.

3.6 Classification

Mitoses are characteristic of malignant cells. The number of mitoses increases, atypical mitosis forms with defects in the mitotic spindle appear, which results in triple or quadruple asters and dissymmetrical structures and atypical forms of chromosomes. Morphologically, the cancerous cell is characterized by a large nucleus, having an irregular size and shape, the nucleoli are prominent, the cytoplasm is scarce and intensely colored or, on the contrary, is pale (Baba AI, 2007). Therefore, this morphological features are the same among all kinds of cells. Although we don't have enough labeled meningioma data for feature extraction to do supervised learning, we found labeled mitotic cells in breast cancer histological images online. So we will extract positive

and negative features from labeled breast tumor cells, and use different classification methods to predict on brain tumor cells.

3.6.1 C4.5 Pruned Decision Tree

C4.5, developed by Ross Quinlan (Quinlan, 1993), is a classification algorithm to generate a decision tree. C4.5 is an extension of ID3 algorithm which also developed by Quinlan earlier. It became very popular after ranking #1 in the Top 10 Algorithms in Data Mining pre-eminent paper published by Springer LNCS in 2008 (Wu, 2008). Furthermore, in 2011, authors of the Weka machine learning software commented the C4.5 algorithm as "a landmark decision tree program that is probably the machine learning workhorse most widely used in practice to date" (Ian H. Witten, 2011). In Weka data mining tool, J48 is an open source Java implementation of the C4.5 algorithm.

In the similar way as ID3, C4.5 uses the concept of information entropy to build decision trees with a set of training data. The training data contains a set of already classified samples. Each sample consists of a n-dimensional vector, which represents features of the sample and class label. At each node of the decision tree, C4.5 chooses the attribute with the highest normalized information gain to split.

This algorithm includes several base cases.

[1], If all the samples belong to the same class. The algorithm simply creates one leaf node with class label in the decision.

[2], If none of the features provide any information gain. C4.5 creates a node higher up the tree with the expected value of the class.

[3], In case of previously-unseen class, C4.5 creates a decision node higher up the tree using the expected value.

The pseudocode for building decision trees is: (Kotsiantis, 2007)

- 1, Check all the base cases.
- 2, For each attribute a, find the normalized information gain ratio from splitting on a.
- 3, Let a_best be the attribute with the highest normalized information gain.
- 4, Create a decision node that splits on a_best.
- 5, Recur on the sublists obtained by splitting on a_best, and add those nodes as children of node.

All the equations of C4.5 algorithm are as follow: calculate Entropy Info(S) to identify the class in the training set S,

$$Info(S) = - \sum_{i=1}^k \left\{ \left[\frac{freq(C_i, S)}{|S|} \right] \log_2 \left[\frac{freq(C_i, S)}{|S|} \right] \right\}$$

Where |s| is the number of cases in the training set, Ci is a class, I = 1,2,...k, k is the number of classes, freq(Ci, S) is the number of cases in Ci.

3.6.2 SMO: Sequential Minimal Optimization

Sequential minimal optimization (SMO) was published by John Platt at Microsoft Research in 1998 (Platt, 1998). SMO is an algorithm for solving the quadratic programming (QP) problem that exists in the training step of support-vector machines (SVM). Before the publication of the SMO algorithm in 1998, previously available algorithm for SVM were very complicated and required

expensive third-party QP solvers (Rifkin, 2002). SMO is implemented by the famous LIBSVM tool (Chang & Lin, 2011; Zanni, 2006).

Algorithm SMO is an iterative algorithm for solving the binary optimization problem. SMO breaks the whole problem into series of smallest possible sub-problems. The sub-problems are then solved analytically.

Since the linear equality constraint involving the Lagrange multipliers a_i , the smallest possible problem contains two such multipliers. Therefore, for any two multipliers a_1 and a_2 , the constraints are reduced to: $0 \leq a_1, a_2 \leq C, y_1 a_1 + y_2 a_2 = k$,

the reduced problem can be solved analytically: simply needs to find the minimum of a 1D quadratic function. k is the negative of the sum over the rest of terms in the equality constraint.

The algorithm proceeds as follows:

Repeat steps 1 and 2 until convergence.

Step 1: For the optimization problem, find a Lagrange multiplier which violates the Karush–Kuhn–Tucker (KKT) conditions.

Step 2: Pick a second multiplier a_2 and optimize the pair (a_1, a_2) .

The problem will be solved after all the Lagrange multipliers satisfy the KKT conditions. Although convergence is guaranteed for this algorithm, heuristics are used to pick up the pair of multipliers to accelerate the speed of convergence. This is extremely useful for large data sets because there are $n(n-1)/2$ possible ways for a_i and a_j .

3.6.3 JRip: Rules Classifieres

JRip was developed by William W. Cohen in 1995 as an optimized version of IREP. JRip implements a propositional rule learner Repeated Incremental Pruning to Produce Error Reduction (RIPPER).

It is based in association rules with reduced error pruning (REP), which is a common and effective technique among decision tree algorithms. In REP, the training data is split into growing set and pruning set. At first, an initial rule set is formed for the growing set, by using certain heuristic method. This overlarge rule set is then repeatedly simplified by applying one of a set of pruning operators. Typical pruning operators would be used to delete any single condition or rule. At each stage of simplification, the pruning operator that yields the greatest reduction of error on the pruning set is chosen. Simplification will end when applying any pruning operator would increase error on the pruning set.

3.6.4 Libsvm Polynomial

Support-vector machines (Chang & Lin, 2011) are supervised learning models for classification and regression analysis. Given a set of training examples, which are belong to one of two categories, an SVM training algorithm builds a model that assigns new examples to one of the two categories, making it a binary linear classifier.

An SVM model represents the examples as points in the space, which is mapped so that the examples of the two categories are divided by a gap as wide as possible. Then new examples are mapped into that same space and its category will be predicted by the side of the gap it falls to.

Besides performing linear classification, SVMs can perform non-linear classifications using kernel trick, implicitly mapping the features to higher dimensional spaces.

The polynomial kernel is commonly used in support vector machines (SVMs). Intuitively, the polynomial kernel looks not only at the given features of the samples for determining their similarity, but also combinations of the features. In the regression analysis, the combinations are known as interaction features. The implicit feature space of a polynomial kernel is equivalent to that of polynomial regression.

3.7 Deep Learning

3.7.1 Data Pre-processing and Data Augmentation

Since we do not have enough labeling data, we will use augmentation methods to augment the data numbers through random transformations such as rotate and flip, so that we can feed more training data for better model accuracy, meanwhile it can help to prevent overfitting the model therefore to make the model more generic. `keras.preprocessing.image.ImageDataGenerator` class is used to a) instantiate generators of augmented image batches via `.flow(data, labels)` or `.flow_from_directory(directory)`. b) configure random transformations and normalization operations on image data during training.

For augmentation methods are as follows:

- a. Rotate images in degrees (0-180).
- b. `Width_shift` and `height_shift` to translate pictures vertically or horizontally
- c. Rescale the image with 1/255 factor, so all images will be rescaled to range 0-1 instead of 0-255, which is too high for models to process.

- d. `shear_range` to randomly apply shear transformations
- e. `zoom_range` to randomly zoom the pictures
- f. `horizontal_flip` to randomly flip half of the images horizontally
- g. `fill_mode` to fill in newly created pixels which appears after a rotation or a width/height shift.

All cells are segmented according to previously method, and each boundingbox around cell is cropped with normalized histogram, and then scaled by normalized with 255 for later steps. And there are 4576 cell images are for training, among them there are 220 are positive; While for validation, there are 3670 cell images are for validation, and 216 are positives.

3.7.2 Transfer Learning with Breast Cancer Training Data

I am using the VGG16 architecture, which is pre-trained on the ImageNet dataset. Because the ImageNet dataset contains so many classes, so the pretrained model may already have learned features that are relevant to my classification problem. Here only convolutional part of the model up to the fully connected layers will be initiated. And then the weights will be used to train the new model by my training and validation data, and recording the output in two NumPy arrays (class 0, class 1), which is the the last activation maps before the fully-connected layers in VGG16 model. Following image is the VGG16 architecture

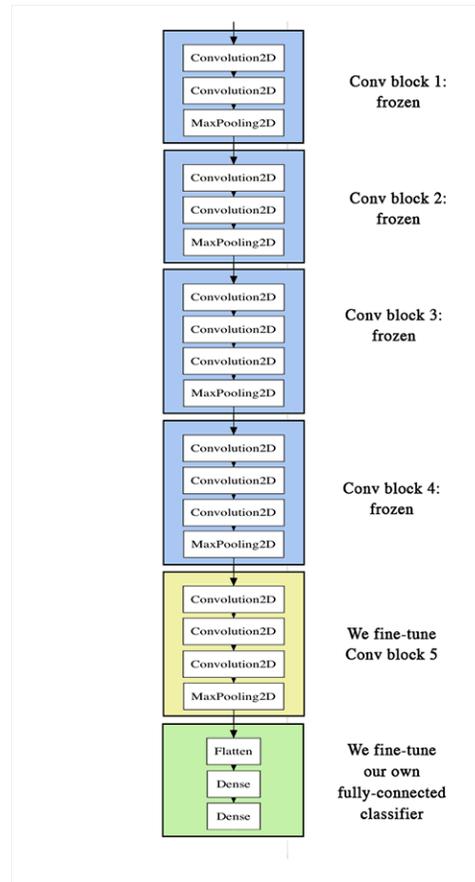


Figure 24. VGG16 architecture

For computational efficacy, first I stored the features offline rather than adding the fully-connected model directly on top of a frozen convolutional base instead of running the whole thing. Then I load saved data and trained the small fully-connected model with 50 epoch. For each epoch, I checkpoint the best validation accuracy, and saved the weight for later prediction. Before it trained on fully-connected model, the weights looks like following:

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 150, 150, 3)	0
block1_conv1 (Conv2D)	(None, 150, 150, 64)	1792
block1_conv2 (Conv2D)	(None, 150, 150, 64)	36928
block1_pool (MaxPooling2D)	(None, 75, 75, 64)	0
block2_conv1 (Conv2D)	(None, 75, 75, 128)	73856
block2_conv2 (Conv2D)	(None, 75, 75, 128)	147584
block2_pool (MaxPooling2D)	(None, 37, 37, 128)	0
block3_conv1 (Conv2D)	(None, 37, 37, 256)	295168
block3_conv2 (Conv2D)	(None, 37, 37, 256)	590080
block3_conv3 (Conv2D)	(None, 37, 37, 256)	590080
block3_pool (MaxPooling2D)	(None, 18, 18, 256)	0
block4_conv1 (Conv2D)	(None, 18, 18, 512)	1180160
block4_conv2 (Conv2D)	(None, 18, 18, 512)	2359808
block4_conv3 (Conv2D)	(None, 18, 18, 512)	2359808
block4_pool (MaxPooling2D)	(None, 9, 9, 512)	0
block5_conv1 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv2 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv3 (Conv2D)	(None, 9, 9, 512)	2359808
block5_pool (MaxPooling2D)	(None, 4, 4, 512)	0
sequential_1 (Sequential)	(None, 1)	2097665
=====		
Total params: 16,812,353		
Trainable params: 9,177,089		
Non-trainable params: 7,635,264		

Figure 25. Model overview before trained on fully-connected layer

In order to prevent overfitting, I chose to fine-tune the last convolutional block instead of the entire, since it will have a strong tendency to overfit with the entire network which have a very large

entropic capacity. The features learned by low-level convolutional blocks are more general, less abstract than those found higher-up, so it is sensible to keep the first few blocks fixed (more general features) and only fine-tune the last one (more specialized features). I used slow learning rate to fine tune with the SGD optimizer, instead of an adaptive learning rate optimizer such as RMSProp, which is to make sure the magnitude of the updates stays very small. The image below shows the fine tuning result after each epoch. And the epoch 45 gets to validation accuracy of 0.98 as shown below.

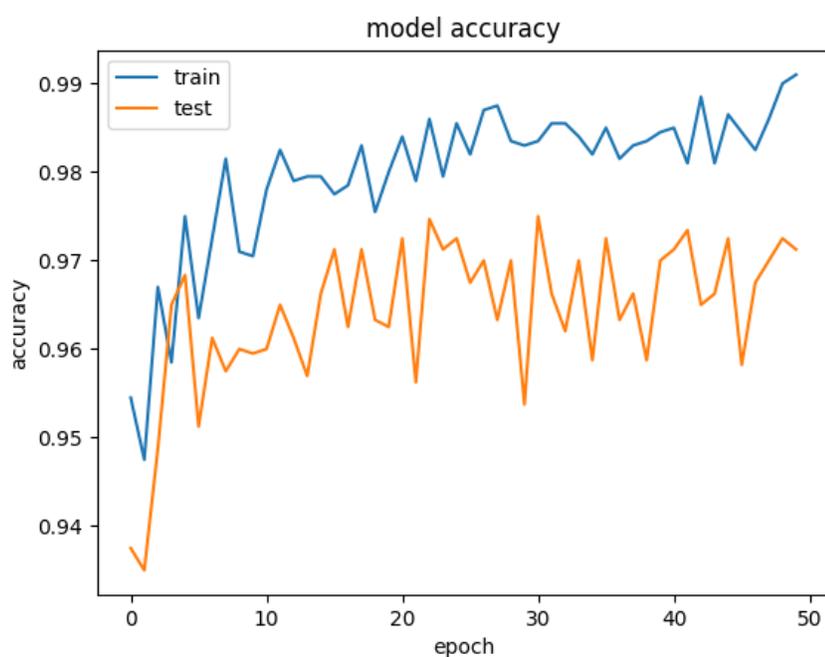


Figure 26. Fine tune results after each step

```
Epoch 45/50  
- 11s - loss: 0.0334 - accuracy: 0.9880 - val_loss: 0.0115 - val_accuracy: 0.9812  
Epoch 00045: val_accuracy improved from 0.98000 to 0.98125, saving model to weights.best.hdf5
```

Figure 27. Epoch 45 - validation accuracy: 0.9812

4. EXPERIMENTS AND RESULTS

All predictions will be made on the two images that have doctor's labels. And the precision, recall and F1 score will be calculated.

4.1 K-Means Results

To calculate how much number for cluster k is a good number, we used elbow method: which is to calculate the sum of the squared distance within each group and use $k = 5$ for clustering.

4.1.1 Negative Cluster

The following images show the four classes that are not including the positive cells. Green boundaries are the predicted clusters

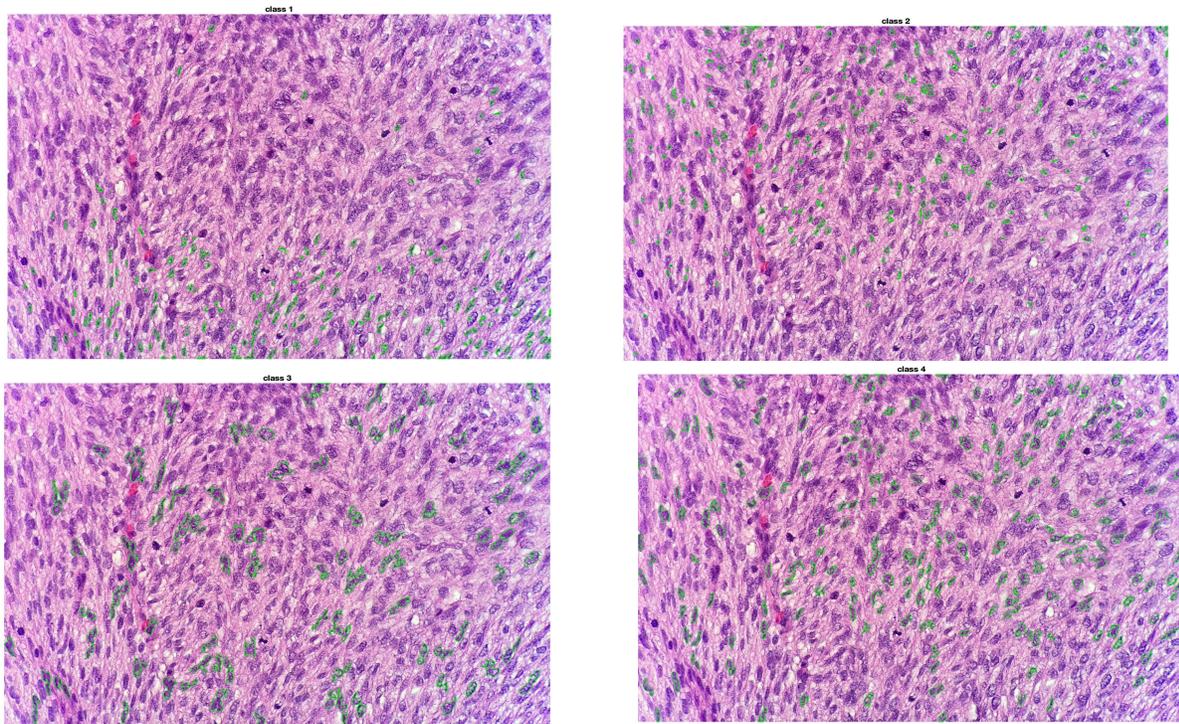


Figure 28. Negative clusters from k-means method.

4.1.2 Positive Clusters

Following is the positive cluster that includes the mitotic cells. And for positive clusters the precision is 0.055 since a lot of false positive. And the Recall is 1 since it includes all true positive. And the $F1 = 0.104$. Yellow bounding box is ground truth.

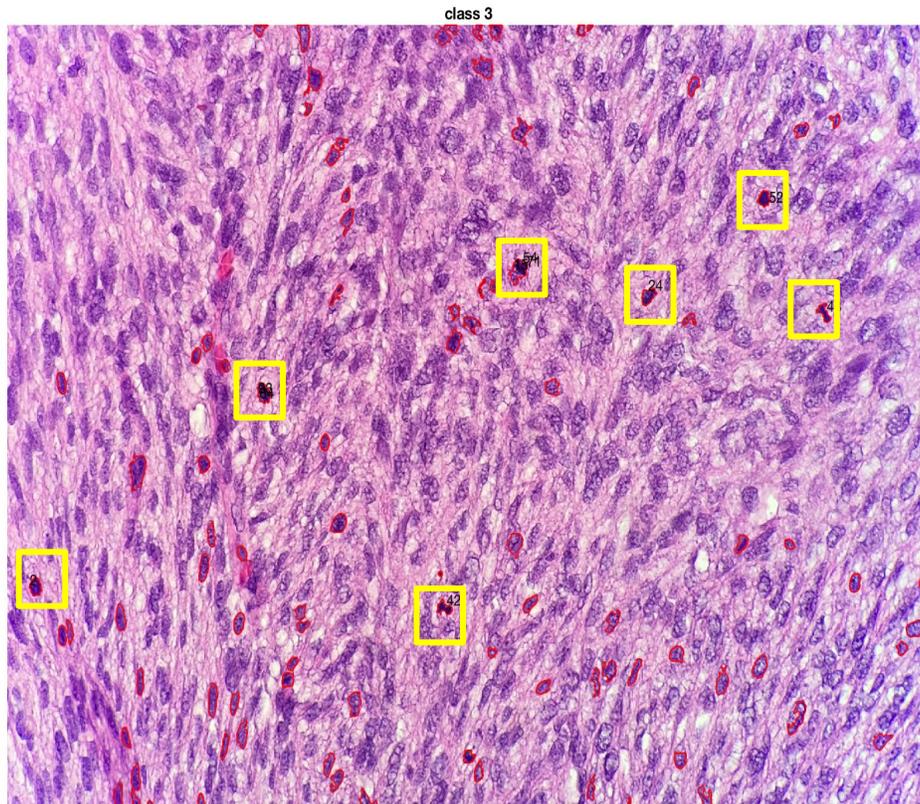


Figure 29. Cluster for positive candidates (red boundaries).

For positive clusters on img2, the precision is 0.018, and the recall is 1, while the F1 score is 0.035. Red boundaries are the positive candidates, yellow bounding box is ground truth.

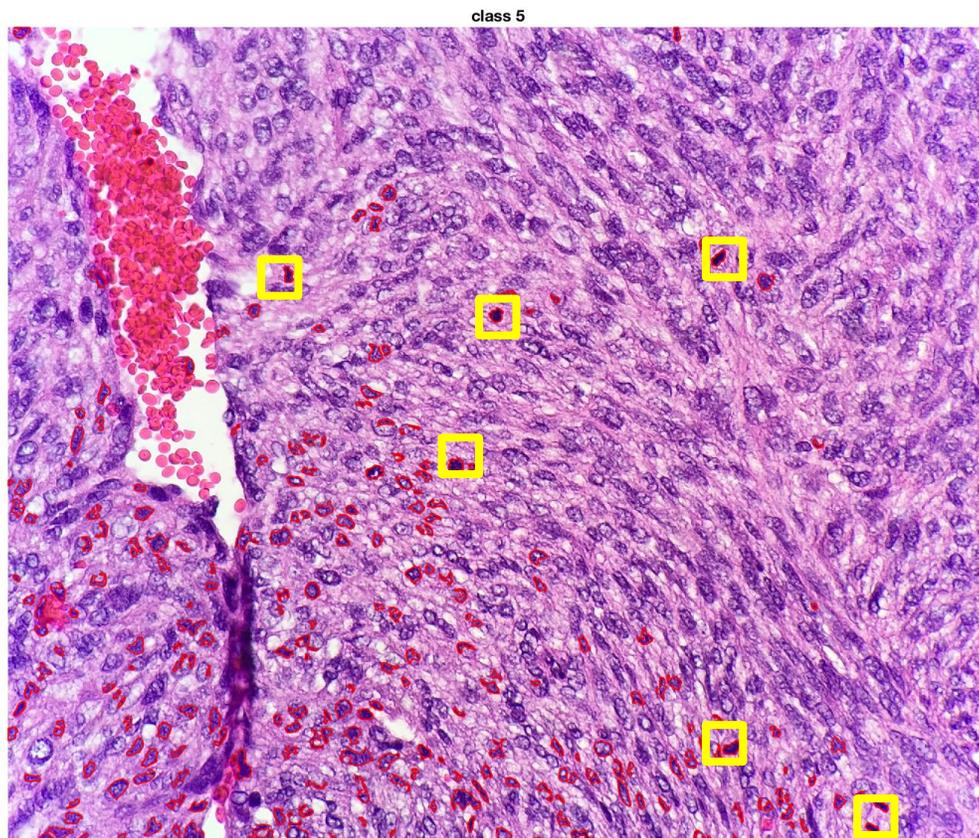


Figure 30. Candidate Clusters on Img2.

4.1.3 K-means on Breast Cancer Slides

The following image also shows the k-means cluster on breast cancer. Yellow bounding box is ground truth.

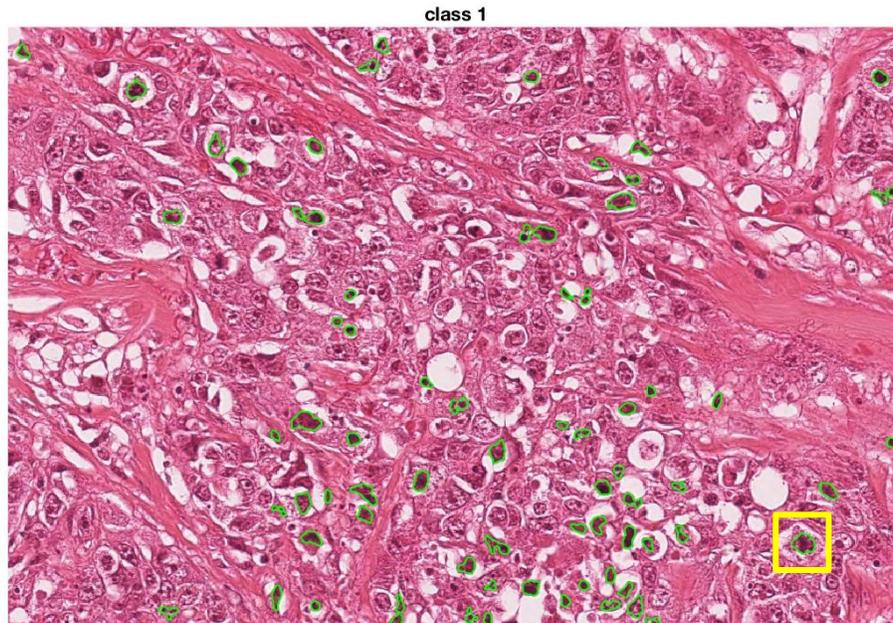


Figure 31. K-means method on breast cancer slides. Classification Results

4.1.4 J48 Results

The following image is the C4.5 prune decision tree result, the yellow circles are ground truth. For image1, positive precision = 0.06, recall = 1, F1 = 0.11. Red box are positive class, yellow circles are ground truth.

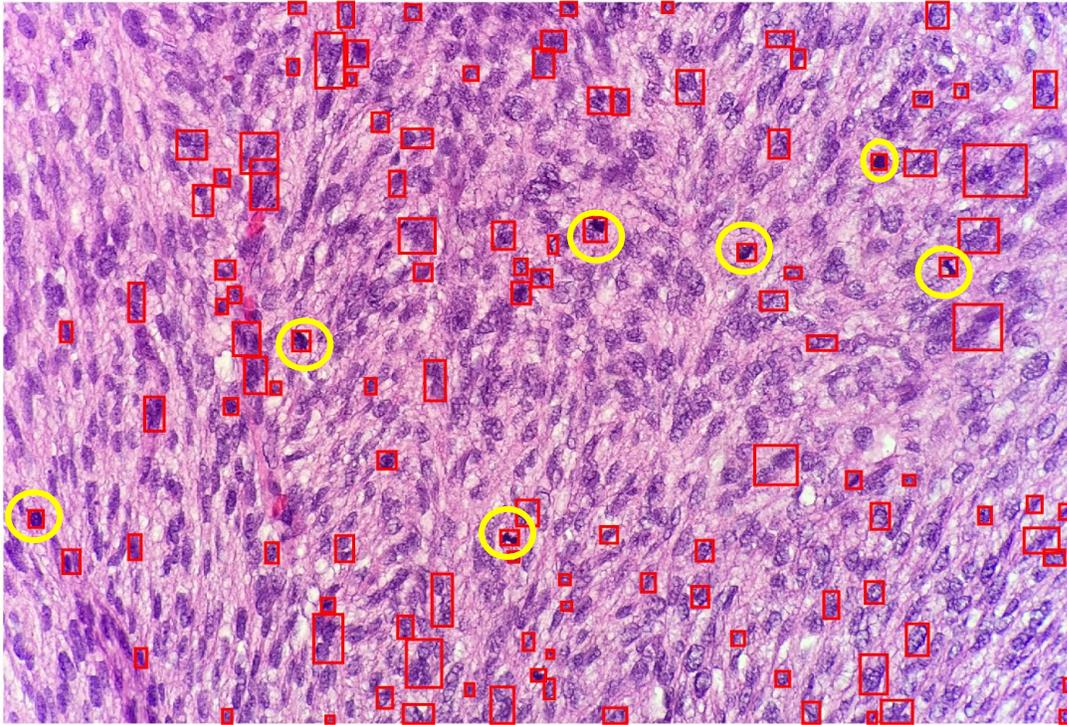


Figure 32. Classification results with J48 method on img1.

4.1.5 SMO Result

The following image is SMO result on image1, for positive cluster, the precision = 0.125, recall = 1, F1 = 0.22. Yellow bounding box of positive class, red circles are ground truth.

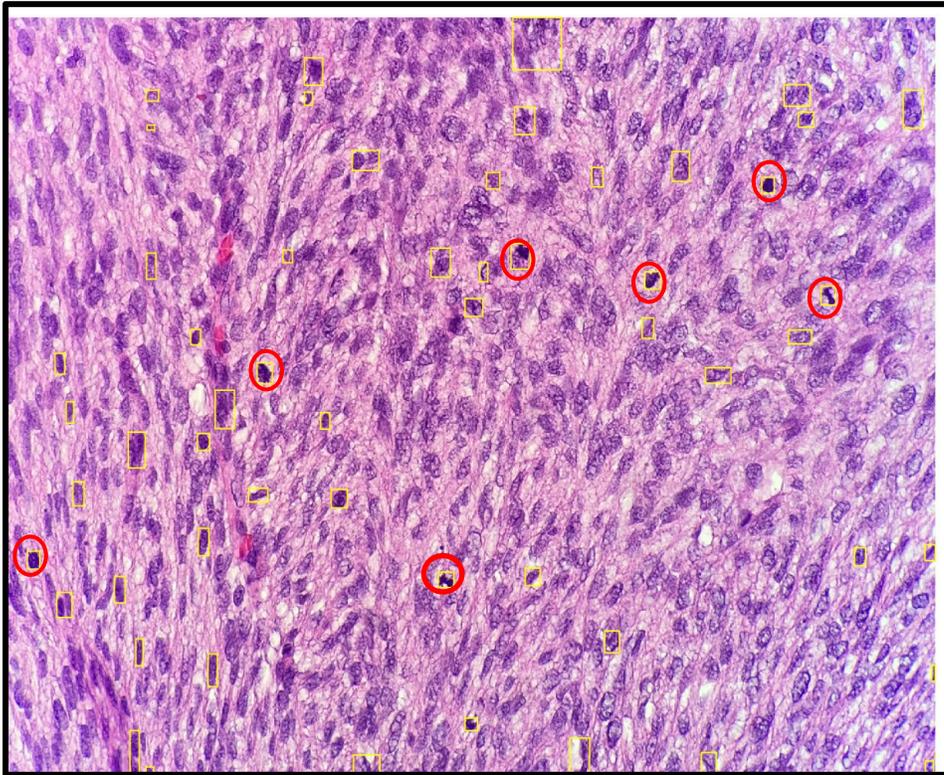


Figure 33. SMO result on image1.

4.1.6 JRip Result

The following image is the result for image1, for positive class, the precision = 0.08, recall = 1, F1 = 0.144.

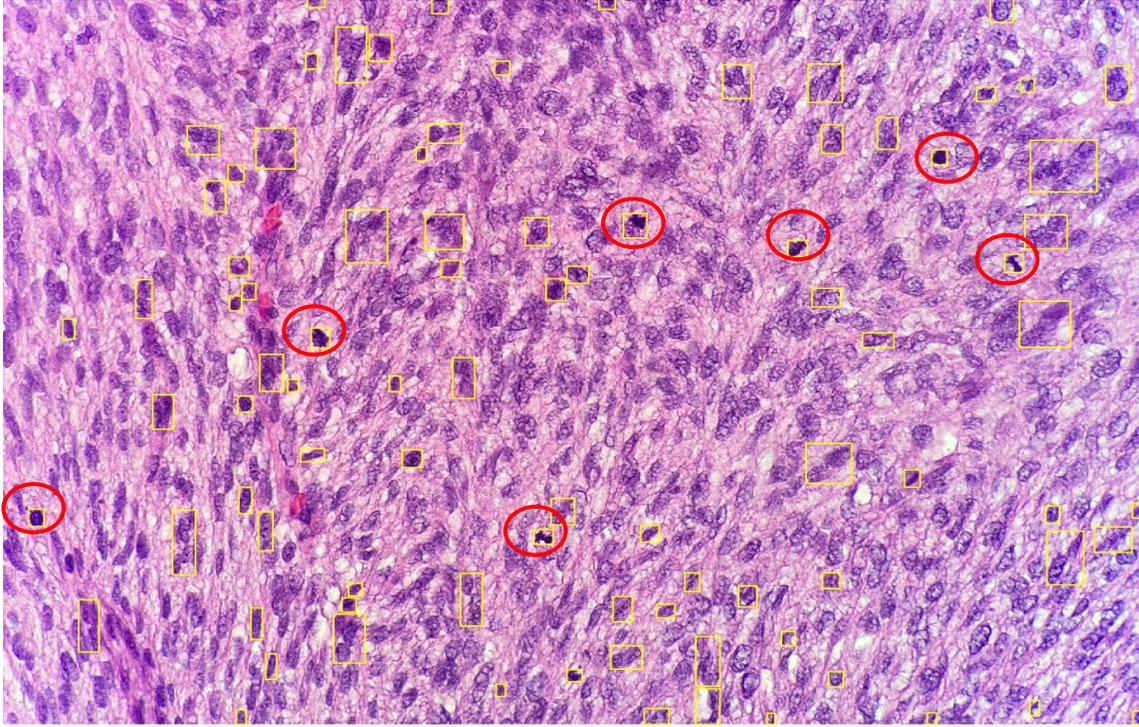


Figure 34. JRip result on image1, red circles are ground truth.

4.1.7 Libsvm Results

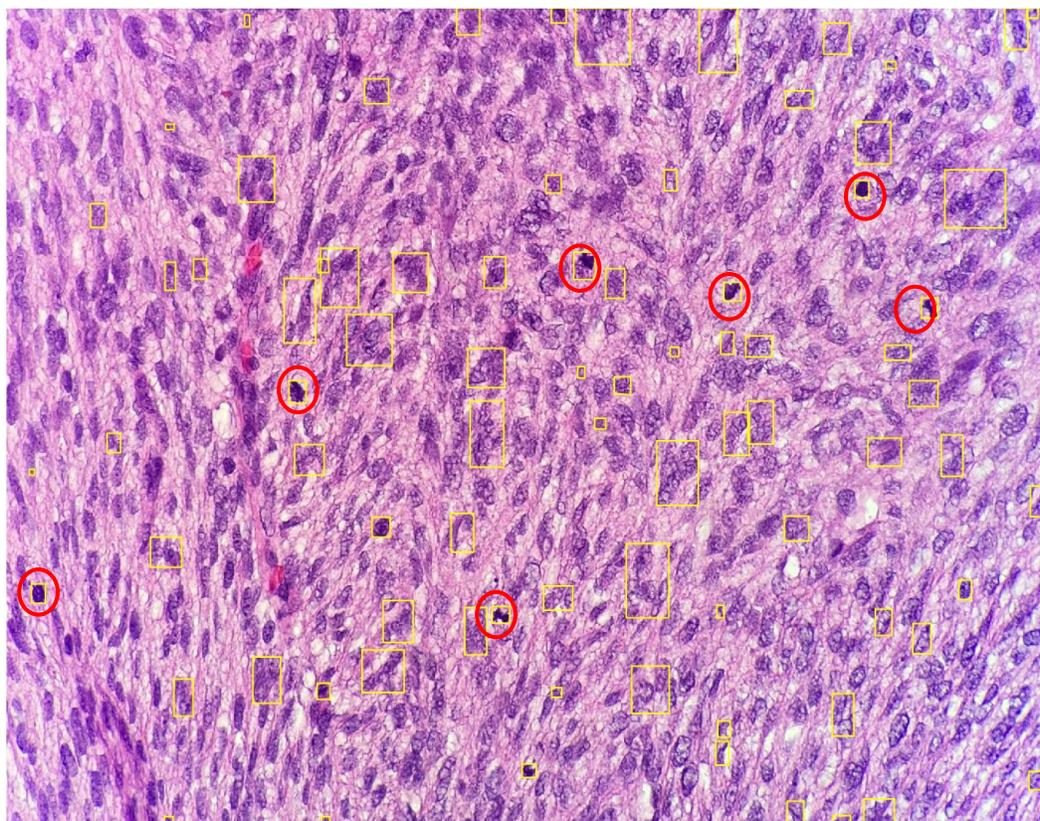


Figure 35. Libsvm Result

4.2 Deep Learning Results

With best validation accuracy weights from brain cancer, the model is used on predicting brain tumor image1 and image2. Green box is the result with possibility between $[0.3,0.5]$, while yellow box stands for possibility bigger than 0.5. The positive precision is 0.83, while recall is 0.71, and F1 score is 0.77.

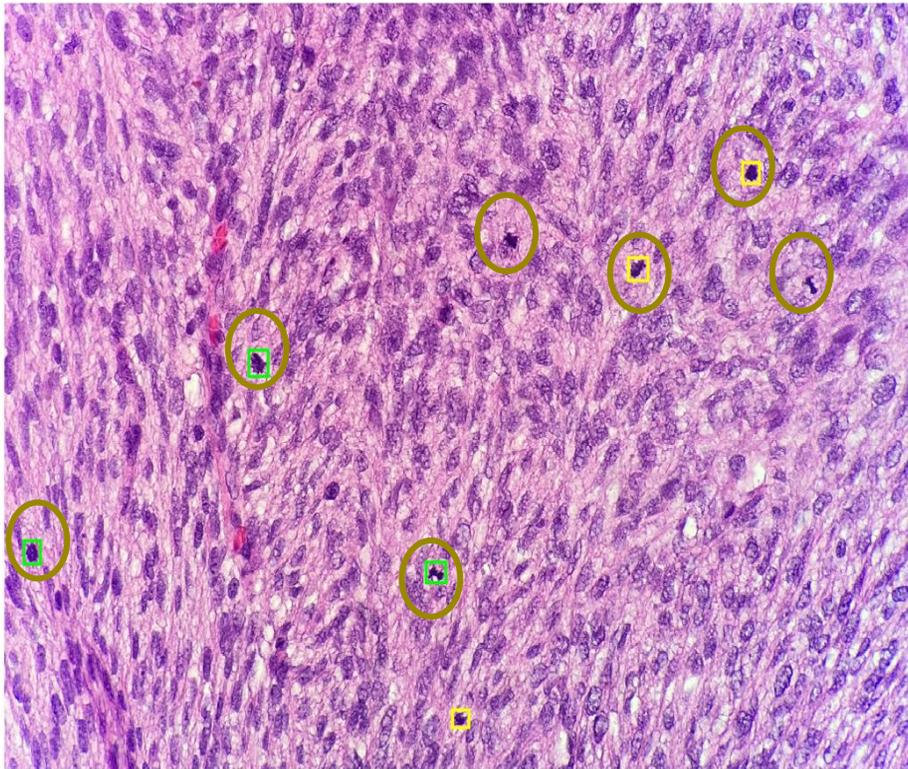


Figure 36. Prediction results on image1

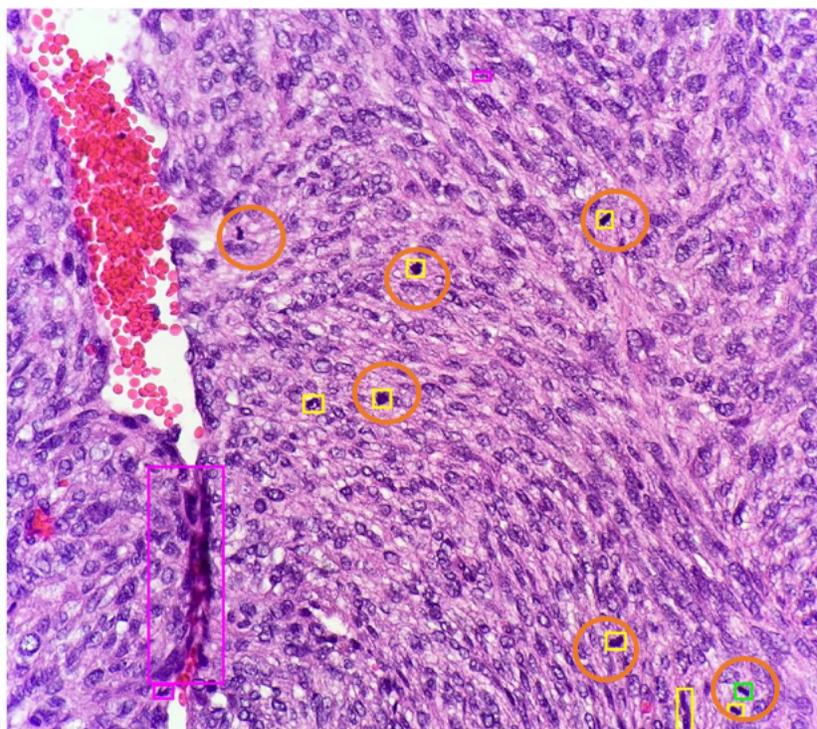


Figure 37. Prediction results on image2

5. CONCLUSION

This thesis set out to determine the feasibility of developing automated computer models for classify specific cell types without enough labeled data. We want to determine if it is feasible to take the whole slide image of H&E stained image as an input, and mimic the steps of pathologist procedure to aid the cancer staging. We also want to determine if it is feasible to use public data to transfer learn the features, and use the model to help the diagnosis the cancer cells

5.1 Whole Slide Image Process

And for mitotic detection, the pathologists selected 10 high power fields (HPF) at 40X magnification and count the cancerous cells to determine the stage of the cancer. Although the whole slide images are used, not all the cells are analyzed. Therefore, by using this domain knowledge, we invented the automated high power fields detection methods to get 10 cell patches that are the similar size of the high power filed under 40X. And by using the lowest resolution image of the whole slide, it is super-fast to crop the cell patches under highest resolution, which highly increase the computation speed and reduce the computation power by mimicking the real procedure of pathologist techniques.

5.2 Transfer Learning

The main problem of supervised method in biomedical data is the labels, which require a lot of time and efforts from doctors. Meanwhile it is very hard to find public training data with specific cell types. Therefore, it is urgent to find a way to use public data which has labels, and use those data to aid the diagnosis of specific cell types. Our result shows that by using the public available breast cancer cells, we retrained the model and predict on brain tumor slides, which gives a very

promising result. So, by using this model, we are able to predict the meningioma whole slide images with pretty good F1 score. Therefore, we can use this model to help pathologist to detect cancer cell candidate, which greatly reduce their efforts in finding mitotic cells in lower stage tumors. Also, with the whole process, we will be able to collect more labeled data, and reinforce the model to make it more accurate during the whole process.

REFERENCES

- Al-Kadi, O. S. (2010). Assessment of texture measures susceptibility to noise in conventional and contrast enhanced computed tomography lung tumour images. *Comput Med Imaging Graph*, 34(6), 494-503. doi:10.1016/j.compmedimag.2009.12.011
- Al-Kadi, O. S. (2017). A Gabor Filter Texture Analysis Approach for Histopathological Brain Tumor Subtype Discrimination. *ISESCO Journal of Science and Technology*.
- Al-Kofahi, Y., Lassoued, W., Lee, W., & Roysam, B. (2010). Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Trans Biomed Eng*, 57(4), 841-852. doi:10.1109/TBME.2009.2035102
- Ali, S., & Madabhushi, A. (2012). An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. *IEEE Trans Med Imaging*, 31(7), 1448-1460. doi:10.1109/TMI.2012.2190089
- Amoli D. Belsare, M. M. M. (2012). HISTOPATHOLOGICAL IMAGE ANALYSIS USING IMAGE PROCESSING TECHNIQUES : AN OVERVIEW. *Signal & Image Processing : An International Journal (SIPIJ)*, 3, 23 - 36.
- Arteta, C., Lempitsky, V., Noble, J. A., & Zisserman, A. (2012). Learning to detect cells using non-overlapping extremal regions. *Med Image Comput Comput Assist Interv*, 15(Pt 1), 348-356. doi:10.1007/978-3-642-33415-3_43
- Baba AI, C. C. (2007). "Chapter3 : Tumor Cell Morphology". *Comparative Oncology*.
- Basavanhally, A., Ganesan, S., Feldman, M., Shih, N., Mies, C., Tomaszewski, J., & Madabhushi, A. (2013). Multi-field-of-view framework for distinguishing tumor grade in ER+ breast cancer from entire histopathology slides. *IEEE Trans Biomed Eng*, 60(8), 2089-2099. doi:10.1109/TBME.2013.2245129
- Basavanhally, A. N., Ganesan, S., Agner, S., Monaco, J. P., Feldman, M. D., Tomaszewski, J. E., . . . Madabhushi, A. (2010). Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. *IEEE Trans Biomed Eng*, 57(3), 642-653. doi:10.1109/TBME.2009.2035305
- Belkacem-Boussaid, K., Pennell, M., Lozanski, G., Shana'ah, A., & Gurcan, M. (2010). Computer-aided classification of centroblast cells in follicular lymphoma. *Anal Quant Cytol Histol*, 32(5), 254-260. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/21509147>

- Belkin, M. N., P. . (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, *15*, 1373–1396.
- Bhattacharyya, S. (2011). A Brief Survey of Color Image Preprocessing and Segmentation Techniques. *J Pattern Recognit Res*, *6*, 120-129.
- Black, M., Sapiro, G., Marimont, D., & Heeger, D. (1998). Robust anisotropic diffusion. . *IEEE Trans. Image Processing*, *7*, 421-432.
- Bouzas, D., Arvanitopoulos, N., & Tefas, A. (2015). Graph embedded nonparametric mutual information for supervised dimensionality reduction. *IEEE Trans Neural Netw Learn Syst*, *26*(5), 951-963. doi:10.1109/TNNLS.2014.2329240
- Buerki, R. A., Horbinski, C. M., Kruser, T., Horowitz, P. M., James, C. D., & Lukas, R. V. (2018). An overview of meningiomas. *Future Oncol*, *14*(21), 2161-2177. doi:10.2217/fo-2018-0006
- Caglar Senaras, D. c., Formal analysis and et al. (2018). DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. *Plos One*, *10*.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*.
- Cheng, J., & Rajapakse, J. C. (2009). Segmentation of clustered nuclei with shape markers and marking function. *IEEE Trans Biomed Eng*, *56*(3), 741-748. doi:10.1109/TBME.2008.2008635
- Ciresan, D. C., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2013). Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Comput Assist Interv*, *16*(Pt 2), 411-418. doi:10.1007/978-3-642-40763-5_51
- Demidenko, E. (2018). The next-generation K-means algorithm. *Stat Anal Data Min*, *11*(4), 153-166. doi:10.1002/sam.11379
- Donoho, D. L. (1995). De-noising by soft thresholding. *IEEE Trans. Information Theory*, *41*, 613-627.
- Esmailsabzali, H., Sakaki, K., Dechev, N., Burke, R. D., & Park, E. J. (2012). Machine vision-based localization of nucleic and cytoplasmic injection sites on low-contrast adherent cells. *Med Biol Eng Comput*, *50*(1), 11-21. doi:10.1007/s11517-011-0831-2

- Fatakdwala, H., Xu, J., Basavanhally, A., Bhanot, G., Ganesan, S., Feldman, M., . . . Madabhushi, A. (2010). Expectation-maximization-driven geodesic active contour with overlap resolution (EMaGACOR): application to lymphocyte segmentation on breast cancer histopathology. *IEEE Trans Biomed Eng*, 57(7), 1676-1689. doi:10.1109/TBME.2010.2041232
- Fauzi, M. F., Pennell, M., Sahiner, B., Chen, W., Shana'ah, A., Hemminger, J., . . . Gurcan, M. N. (2015). Classification of follicular lymphoma: the effect of computer aid on pathologists grading. *BMC Med Inform Decis Mak*, 15, 115. doi:10.1186/s12911-015-0235-6
- Gilboa, G., Sochen, N., & Zeevi, Y. Y. (2002). Forward-and-backward diffusion processes for adaptive image enhancement and denoising. *IEEE Trans Image Process*, 11(7), 689-703. doi:10.1109/TIP.2002.800883
- Gonzalez, R. C. W., R. E. . (2008). Digital Image Processing. *Pearson Prentice Hall, 3rd Ed.*
- Gunduz-Demir, C., Kandemir, M., Tosun, A. B., & Sokmensuer, C. (2010). Automatic segmentation of colon glands using object-graphs. *Med Image Anal*, 14(1), 1-12. doi:10.1016/j.media.2009.09.001
- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image analysis: a review. *IEEE Rev Biomed Eng*, 2, 147-171. doi:10.1109/RBME.2009.2034865
- Gurcan, M. N., Kong, J., Sertel, O., Cambazoglu, B. B., Saltz, J., & Catalyurek, U. (2007). Computerized pathological image analysis for neuroblastoma prognosis. *AMIA Annu Symp Proc*, 304-308. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/18693847>
- Haykin, S. (1999). "9. Self-organizing maps". . *Neural networks - A comprehensive foundation (2nd ed.)*.
- He, L., Long, L. R., Antani, S., & Thoma, G. R. (2012). Histology image analysis for carcinoma detection and grading. *Comput Methods Programs Biomed*, 107(3), 538-556. doi:10.1016/j.cmpb.2011.12.007
- Hoque, A., Lippman, S. M., Boiko, I. V., Atkinson, E. N., Sneige, N., Sahin, A., . . . Hittelman, W. N. (2001). Quantitative nuclear morphometry by image analysis for prediction of recurrence of ductal carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev*, 10(3), 249-259. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/11303595>

- Ian H. Witten, E. F. a. M. A. H. (2011). Data Mining: Practical Machine Learning Tools and Techniques. *ScienceDirect, 3rd Edition*.
- Irshad, H., Jalali, S., Roux, L., Racoceanu, D., Hwee, L. J., Naour, G. L., & Capron, F. (2013). Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach. *J Pathol Inform, 4(Suppl)*, S12. doi:10.4103/2153-3539.109870
- Jia-Mei Chen, Y. L., Jun Xu and et al. (2017). Computer-aided prognosis on breast cancer with hematoxylin and eosin histopathology images: A review. *Tumor Biology*, 1-12.
- Jung, C., & Kim, C. (2010). Segmenting clustered nuclei using H-minima transform-based marker extraction and contour parameterization. *IEEE Trans Biomed Eng, 57(10)*, 2600-2604. doi:10.1109/TBME.2010.2060336
- Jung, C., Kim, C., Chae, S. W., & Oh, S. (2010). Unsupervised segmentation of overlapped nuclei using Bayesian classification. *IEEE Trans Biomed Eng, 57(12)*, 2825-2832. doi:10.1109/TBME.2010.2060486
- Karacali, B., & Tozeren, A. (2007). Automated detection of regions of interest for tissue microarray experiments: an image texture analysis. *BMC Med Imaging, 7, 2*. doi:10.1186/1471-2342-7-2
- Khan, A. M., Rajpoot, N., Treanor, D., & Magee, D. (2014). A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. *IEEE Trans Biomed Eng, 61(6)*, 1729-1738. doi:10.1109/TBME.2014.2303294
- Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. *Proceedings of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, 3 - 24.
- Lei He, L. R. L., Sameer Antani, George R. Thoma. (2010). *Computer Assisted Diagnosis in Histopathology*: National Institutes of Health, USA.
- Li, F., Zhou, X., Ma, J., & Wong, S. T. (2010). Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. *IEEE Trans Med Imaging, 29(1)*, 96-105. doi:10.1109/TMI.2009.2027813

- Liu, L., Tian, Z., Zhang, Z., & Fei, B. (2016). Computer-aided Detection of Prostate Cancer with MRI: Technology and Applications. *Acad Radiol*, 23(8), 1024-1046. doi:10.1016/j.acra.2016.03.010
- Malik, P. P. J. (1990). Scale-Space and Edge Detection Using Anisotropic Diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12, 629-639.
- Mignotte, M. (2011). MDS-based multiresolution nonlinear dimensionality reduction model for color image segmentation. *IEEE Trans Neural Netw*, 22(3), 447-460. doi:10.1109/TNN.2010.2101614
- Muthu Rama Krishnan, M., Pal, M., Bomminayuni, S. K., Chakraborty, C., Paul, R. R., Chatterjee, J., & Ray, A. K. (2009). Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis-an SVM based approach. *Comput Biol Med*, 39(12), 1096-1104. doi:10.1016/j.combiomed.2009.09.004
- Ostrom QT, G. H., Liao P, et al. (2017). CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2010-2014.
- Parvin, B., Yang, Q., Han, J., Chang, H., Rydberg, B., & Barcellos-Hoff, M. H. (2007). Iterative voting for inference of structural saliency and characterization of subcellular events. *IEEE Trans Image Process*, 16(3), 615-623. doi:10.1109/tip.2007.891154
- Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell*, 27(8), 1226-1238. doi:10.1109/TPAMI.2005.159
- Platt, J. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines.
- Pratt, W. K. (2001). Digital Image Processing, 3rd. *Wiley, New York*.
- Qi, X., Xing, F., Foran, D. J., & Yang, L. (2012). Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. *IEEE Trans Biomed Eng*, 59(3), 754-765. doi:10.1109/TBME.2011.2179298
- Quinlan, J. R. (1993). C4.5: programs for machine learning. *Morgan Kaufmann Publishers Inc. San Francisco, CA, USA*.

- Raza, S. H., Parry, R. M., Moffitt, R. A., Young, A. N., & Wang, M. D. (2011). An analysis of scale and rotation invariance in the bag-of-features method for histopathological image classification. *Med Image Comput Comput Assist Interv*, 14(Pt 3), 66-74. doi:10.1007/978-3-642-23626-6_9
- Richard O. Duda, P. E. H., David G. Stork. (2000). Pattern Classification, 2nd Edition. *Wiley*.
- Rifkin, R. M. (2002). Everything old is new again : a fresh look at historical approaches in machine learning. *Ph.D. Thesis, Massachusetts Institute of Technology*.
- Rockhill, J., Mrugala, M., & Chamberlain, M. C. (2007). Intracranial meningiomas: an overview of diagnosis and treatment. *Neurosurg Focus*, 23(4), E1. doi:10.3171/FOC-07/10/E1
- Roux, L. (2014). Mitosis atypia 14 grand challenge. <https://mitos-atypia-14.grand-challenge.org/>.
- Roux, L., Racoceanu, D., Lomenie, N., Kulikova, M., Irshad, H., Klossa, J., . . . Gurcan, M. N. (2013). Mitosis detection in breast cancer histological images An ICPR 2012 contest. *J Pathol Inform*, 4, 8. doi:10.4103/2153-3539.112693
- Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2323-2326. doi:10.1126/science.290.5500.2323
- Rudin, L., Osher, L., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60, 259-268.
- Schnorrenberg, F., Pattichis, C. S., Kyriacou, K. C., & Schizas, C. N. (1997). Computer-aided detection of breast cancer nuclei. *IEEE Trans Inf Technol Biomed*, 1(2), 128-140. doi:10.1109/4233.640655
- Sertel, O., Kong, J., Shimada, H., Catalyurek, U. V., Saltz, J. H., & Gurcan, M. N. (2009). Computer-aided Prognosis of Neuroblastoma on Whole-slide Images: Classification of Stromal Development. *Pattern Recognit*, 42(6), 1093-1103. doi:10.1016/j.patcog.2008.08.027
- Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500), 2319-2323. doi:10.1126/science.290.5500.2319
- Venna, J., & Kaski, S. (2006). Local multidimensional scaling. *Neural Netw*, 19(6-7), 889-899. doi:10.1016/j.neunet.2006.05.014

- Veta, M., van Diest, P. J., Willems, S. M., Wang, H., Madabhushi, A., Cruz-Roa, A., . . . Pluim, J. P. (2015). Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal*, 20(1), 237-248. doi:10.1016/j.media.2014.11.010
- Vink, J. P., Van Leeuwen, M. B., Van Deurzen, C. H., & De Haan, G. (2013). Efficient nucleus detector in histopathology images. *J Microsc*, 249(2), 124-135. doi:10.1111/jmi.12001
- Wang, N., & Osswald, M. (2018). Meningiomas: Overview and New Directions in Therapy. *Semin Neurol*, 38(1), 112-120. doi:10.1055/s-0038-1636502
- Wang, Y. Y., Chang, S. C., Wu, L. W., Tsai, S. T., & Sun, Y. N. (2007). A color-based approach for automated segmentation in tumor tissue classification. *Conf Proc IEEE Eng Med Biol Soc*, 2007, 6577-6580. doi:10.1109/IEMBS.2007.4353866
- Wienert, S., Heim, D., Saeger, K., Stenzinger, A., Beil, M., Hufnagl, P., . . . Klauschen, F. (2012). Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. *Sci Rep*, 2, 503. doi:10.1038/srep00503
- Wu, X. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14, 1 - 37.
- Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., & Madabhushi, A. (2016). Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. *IEEE Trans Med Imaging*, 35(1), 119-130. doi:10.1109/TMI.2015.2458702
- Yan, C., Li, A., Zhang, B., Ding, W., Luo, Q., & Gong, H. (2013). Automated and accurate detection of soma location and surface morphology in large-scale 3D neuron images. *PLoS One*, 8(4), e62579. doi:10.1371/journal.pone.0062579
- Zanni, L. (2006). A Parallel Software for Training Large-Scale Support Vector Machines on Multiprocessor Systems. *Journal of Machine Learning Research*, 7, 1467-1492.