

# **UTILIZING DOMAIN ADAPTATION FOR DEEP LEARNING BASED NECROSIS DETECTION IN MENINGIOMA AND GLIOBLASTOMA**

**Nagaraju Dasari**

**Master of Science by Research Thesis**  
June 2024



International Institute of Information Technology, Bangalore

**UTILIZING DOMAIN ADAPTATION FOR DEEP  
LEARNING BASED NECROSIS DETECTION IN  
MENINGIOMA AND GLIOBLASTOMA**

Submitted to International Institute of Information Technology,  
Bangalore  
in Partial Fulfillment of  
the Requirements for the Award of  
Master of Science by Research

by

**Nagaraju Dasari**  
**MS2021010**

International Institute of Information Technology, Bangalore  
June 2024

*Dedicated to  
the people who sacrificed their lives for research and innovation.*

## **Thesis Certificate**

This is to certify that the thesis titled **Utilizing Domain Adaptation for Deep Learning Based Necrosis Detection in Meningioma and Glioblastoma** submitted to the International Institute of Information Technology, Bangalore, for the award of the degree of **Master of Science by Research** is a bona fide record of the research work done by **Nagaraju Dasari, MS2021010**, under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma. The thesis conforms to plagiarism guidelines and compliance as per UGC recommendations.

---

Prof. TK Srikanth

Bengaluru,

The 20<sup>th</sup> of June, 2024.

## **UTILIZING DOMAIN ADAPTATION FOR DEEP LEARNING BASED NECROSIS DETECTION IN MENINGIOMA AND GLIOBLASTOMA**

### **Abstract**

Deep learning in healthcare has made diagnosing diseases faster and more accurate, especially in analyzing medical images. This technology has greatly improved patient care and hospital efficiency.

The digitalization of histopathology slides has transformed cancer diagnosis and treatment by creating digital images of tissue samples. This enables the use of computational methods, including machine learning and deep learning, for image classification. However, traditional methods may struggle to capture important biological features.

This thesis concentrates on utilizing deep learning methods to identify necrosis in brain tumor biopsies, particularly focusing on Meningioma. The proposed DeepNucleiNet model integrates pathology expertise to enhance performance. It combines features from nuclei-segmented images using XceptionNet and convolutional neural networks. Validation on a dataset confirms the model's high accuracy in detecting necrosis in Meningioma whole slide images.

It is then demonstrated that the DeepNucleiNet model, which was trained on meningioma images, also performs well in identifying necrosis in glioblastoma and low-grade gliomas. Further, it is shown how this technique can be used in the pathologist's workflow to efficiently and reliably detect regions of interest in whole slide images, thus improving productivity and reducing errors for pathologists.

This analysis sets a baseline for developing more robust models for analyzing histopathology images using domain knowledge.

## Acknowledgements

I sincerely thank and appreciate everyone who contributed to completing this thesis. First and foremost, I am immensely grateful to my supervisor, Dr. TK Srikanth, and my co-supervisor, Dr. Ramesh Kestur, for their unwavering guidance, expertise, and acceptance to fulfill my dream of MS by Research at IIIT Bangalore. I am sincerely thankful for their advice, continuous support, freedom, valuable comments, motivation, helpful recommendations, and guidance in finishing my thesis.

I would like to express my sincere gratitude to the Neuropathology lab at NIMHANS for their invaluable contribution to my thesis. Their willingness to share valuable data, including annotations, slides, and guidance, enabled me to conduct in-depth research on necrosis and its various subtypes of tumors.

I am deeply grateful to Dr. Shilpa Rao, Dr. Nandeesh, and Dr. Anitha Mahadevan for their invaluable contribution to my thesis. Their unwavering support and guidance have been instrumental in the successful completion of my research. Without their expertise and dedication, this achievement would not have been possible. I am truly grateful for their assistance and proud to have worked alongside such a talented team.

I sincerely thank TCGA for its contribution and dedication to advancing scientific knowledge in cancer research by sharing data.

I would like to thank EHRC and MINRO labs at IIITB for their funding, without which the theoretical ideas would not see the light of practicality.

I am also thankful to the faculty members of IIITB. Their profound knowledge and passion for teaching have greatly enriched my learning experience.

Last, I thank my parents and brother for their support and encouragement to do MS by Research. I want to thank my friends Vamshi, Vishnu, and Krishna for their support.

## Contents

<b>Abstract</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>List of Figures</b>	<b>xii</b>
<b>List of Tables</b>	<b>xv</b>
<b>List of Abbreviations</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Oncological Challenges: Cancer and Brain Tumors. . . . .	1
1.1.1 Primary Brain Tumors . . . . .	2
1.1.2 Secondary Brain Tumors . . . . .	3
1.2 Understanding Necrosis in Brain Tumors . . . . .	3
1.3 Preparation of Biopsy Slides . . . . .	4
1.4 Tissue Slide Digitalization . . . . .	6

1.5	Region in WSI . . . . .	9
1.6	Computational Pathology through Image Processing . . . . .	11
1.6.1	Early Detection and Scalability . . . . .	11
1.6.2	Traditional Methods in Image Processing for Feature Extraction	12
1.6.3	CNN: Deep Learning Method for Image Classification . . . . .	15
1.6.4	XceptionNet for Image Classification . . . . .	16
1.7	Support Vector Machines: . . . . .	18
1.8	Evaluation of prediction/detection performance . . . . .	20
1.9	Hypothesis . . . . .	23
1.10	Research Objectives . . . . .	23
1.11	Thesis Outline . . . . .	24
1.12	Summary . . . . .	25
<b>2</b>	<b>Literature Review</b>	<b>26</b>
2.1	AI for Medical Diagnosis . . . . .	26
2.2	Feature Extraction for Histopathology . . . . .	28
2.2.1	Conventional Methods for Histopathology Feature Extraction .	28
2.2.2	Limitations of Conventional Features for Histopathology . . .	30
2.2.3	Deep Learning Methods for Histopathology Feature Extraction .	30
2.2.4	Limitations of Deep Learning Features for Histopathology . .	32

2.2.5	Domain Feature Extraction . . . . .	33
2.2.6	Scope of using Domain Knowledge in Histopathology . . . . .	33
2.3	Image Classification . . . . .	34
2.3.1	Machine Learning for Image Classification . . . . .	34
2.3.2	Transfer Learning for Image Classification . . . . .	36
2.4	Summary of Limitations in the Discussed Papers . . . . .	39
2.5	Different Features in Histopathology of Brain Tumor . . . . .	41
2.6	Summary . . . . .	44
<b>3</b>	<b>Data Description</b>	<b>45</b>
3.1	NIMHANS Data Acquisition . . . . .	45
3.2	WSI Annotations Generation . . . . .	45
3.3	Generating tiles from whole slide images . . . . .	48
3.4	Generating labelled tiles from annotated whole slide images . . . . .	50
3.4.1	Processing tiles . . . . .	52
3.4.2	Removing Non-Tissue Areas . . . . .	52
3.4.3	Classifying tiles against annotations . . . . .	54
3.4.4	Border Tile Filtering Technique . . . . .	55
3.5	Data Preparation . . . . .	57
3.6	Brain Tumors - TCGA GBM and LGG . . . . .	60

3.7 Summary . . . . .	61
<b>4 Addressing the Challenges of Necrosis Detection in Histopathology</b>	<b>62</b>
4.1 Problem statement . . . . .	62
4.2 Baseline Model . . . . .	63
4.2.1 Performance Evaluation on inter-slide and intra-slide data . . . . .	65
4.2.2 Insights and Observations . . . . .	66
4.3 Fully Connected XceptionNet . . . . .	68
4.3.1 Model training and Performance Evaluation on inter-slide and inter-class . . . . .	69
4.3.2 GradCAM Analysis for Image Representations for XceptionNet	72
4.3.3 Insights and Observations . . . . .	73
4.4 XceptionNet Features with SVM . . . . .	75
4.4.1 Performance Evaluation on inter-slide and intra-slide . . . . .	78
4.4.2 Insights and Observations . . . . .	79
4.5 Domain Knowledge Extraction . . . . .	81
4.5.1 Nuclei Extraction . . . . .	81
4.5.2 Nuclei Detection and Segmentation . . . . .	81
4.5.3 Parameter Optimization for Nuclei Segmentation . . . . .	83
4.5.4 Analysis of Segmentation Results . . . . .	85

4.6 DeepNucleiNet . . . . .	86
4.6.1 Performance Evaluation on inter-slide and Inter- Class . . . . .	87
4.6.2 Insights and Observations . . . . .	88
4.7 Performance Analysis . . . . .	89
4.8 Partial Necrosis Performance Evaluation in Tiles . . . . .	93
4.9 Adapting Meningioma-trained Model for Necrosis Detection in GBM and LGG . . . . .	95
4.10 Summary . . . . .	96
<b>5 Enhancing Robustness in Necrotic Region Identification within WSI</b>	<b>98</b>
5.1 Introduction . . . . .	98
5.1.1 Region Extraction and Tile Generation . . . . .	99
5.1.2 Regional and Areal Analysis of Necrosis Detection . . . . .	100
5.1.3 Insights and Observations . . . . .	103
5.2 Necrosis Detection with Grid Positional Changes . . . . .	104
5.3 Summary . . . . .	105
<b>6 Conclusions and future work</b>	<b>107</b>
6.1 Findings: . . . . .	107
6.2 Conclusion: . . . . .	108
6.3 Future work: . . . . .	109

**Bibliography 110**

## List of Figures

FC1.1 Baseline Mode Architecture . . . . .	2
FC1.2 1)Tissue Processing 2) WSI Acquisition and Image Analysis . . . . .	7
FC1.3 High-resolution depiction of Meningioma tissue WSI at 0.4x magnification . . . . .	8
FC1.4 Image showing region of interest . . . . .	10
FC1.5 Image showing Annotated Necrosis Regions . . . . .	11
FC1.6 Visualization of key point descriptors and Grid Search . . . . .	14
FC1.7 Architecture of the XceptionNet model for ImageNet1k . . . . .	18
FC2.1 Steps for machine learning in digital pathological image analysis . . .	35
FC2.2 Scheme for transfer learning of histopathology image classification . .	37
FC2.3 Visualizing Necrosis and Other Annotations on ImageScope . . . . .	43
FC3.1 Visualizing Necrosis and Other Annotations on ImageScope . . . . .	46
FC3.2 Different levels of magnification in histopathlogy images . . . . .	51
FC3.3 Visualization of images with significant white space regions at border	56

FC4.1 Baseline Mode Architecture . . . . .	64
FC4.2 t-SNE Feature Visualization for Handcrafted Features . . . . .	68
FC4.3 XceptionNet Fully Connected Network for Necrosis Detection . . . . .	69
FC4.4 The Architecture of Fine-Tuned XceptionNet for Necrosis Detection . .	70
FC4.5 GradCAM Visualization of images with Feature Focus on Necrosis Images . . . . .	72
FC4.6 GradCAM Visualization of images with Feature Focus on Non-Necrosis Images . . . . .	73
FC4.7 Training and Validation Performance over Epochs for Batch 1 . . . . .	74
FC4.8 XceptionNet Features with SVM Network for Necrosis Detection . . .	76
FC4.9 t-SNE Feature Visualization for XceptionNet Features with SVM . . .	80
FC4.10Star Convex Polygon Identification with Convolutional Neural Net- works . . . . .	82
FC4.11Stardist Segmentation Methodology . . . . .	83
FC4.12Input and Segmented Images with Diameters 50 and 40, respectively, with a flow threshold of 0.4 . . . . .	84
FC4.13Input and Segmented Images with Diameters 30 and 25, respectively, with a flow threshold of 0.4 . . . . .	84
FC4.14Visualization Segmentation Methodology from Input to Binarized Im- age . . . . .	86
FC4.15Architecture of DeepNucleiNet . . . . .	87

FC4.16t-SNE Feature Visualization for Proposed Model . . . . .	90
FC4.17inter-slide, intra-slide and Overall Accuracy Comparison . . . . .	91
FC4.18inter-slide, intra-slide and Overall Recall Comparison . . . . .	91
FC4.19inter-slide, intra-slide and Overall Precision Comparison . . . . .	92
FC4.20inter-slide, intra-slide and Overall F1 Score Comparison . . . . .	93
FC4.21Performance of the DeepNucleiNet according to Necrosis area within the tile . . . . .	95
FC5.1 Distribution of Total Tiles, Interior, and Boundary tiles across regions .	100
FC5.2 Visualizing the distribution of Boundary Tiles, $<50\%$ , and $\geq 50\%$ Necrosis tiles across regions. . . . .	101
FC5.3 Recall Metrics for Interior and Boundary Necrosis Detection in Regions	102
FC5.4 Generating tiles using multiple initialization coordinates . . . . .	104

## List of Tables

TC1.1 Confusion Matrix . . . . .	20
TC2.1 Summary of Algorithms for Histopathology Feature Extraction . . . . .	39
TC3.1 Available annotations in Meningioma slides . . . . .	48
TC3.2 Distribution of Necrosis and Non-Necrosis Images across Samples . .	58
TC3.3 Batchwise distribution of slide samples (sample IDs) for intra-slide analysis . . . . .	59
TC3.4 Batchwise distribution of slide samples (sample IDs) for inter-slide analysis . . . . .	59
TC3.5 Necrosis and Non-Necrosis tile count for TCGA GBM and LGG Sam- ples . . . . .	60
TC4.1 Batch-wise Necrosis Detection Metrics for Baseline Model . . . . .	66
TC4.2 Weighted Average Metrics Evaluation for the Baseline Model . . . . .	67
TC4.3 Batch-wise Necrosis Detection Metrics for Fully Connected Xcep- tionNet . . . . .	71

TC4.4 Weighted Average Metrics for Fully Connected XceptionNet . . . . .	74
TC4.5 S metric values for different models for 512,1024,2018, and 4096 feature sizes . . . . .	77
TC4.6 Batch-wise Necrosis Detection Metrics for XceptionNet Features with SVM . . . . .	78
TC4.7 Weighted Average Metrics for XceptionNet Features with SVM . . . . .	79
TC4.8 Batch-wise Necrosis Detection Metrics for DeepNucleiNet Model . . . . .	88
TC4.9 Weighted Average Metrics for DeepNucleiNet . . . . .	89
TC4.10 Accuracy Metrics for intra-slide, inter-slide, and Overall for All Models . . . . .	90
TC4.11 Recall Metrics for intra-slide, inter-slide, and Overall for All Models . . . . .	90
TC4.12 Precision Metrics for intra-slide, inter-slide, and Overall for All Models . . . . .	92
TC4.13 F1 Score Metrics for intra-slide, inter-slide, and Overall for All Models . . . . .	92
TC4.14 Analysis of the DeepNucleiNet's performance in relation to the partial necrosis buckets . . . . .	94
TC4.15 Batch-wise Necrosis Detection Metrics on GGM and LGG using DeepNucleiNet . . . . .	96
TC5.1 Summary of Regions with Interior and Boundary Tile Statistics . . . . .	101
TC5.2 Distribution and Performance of Model with Multiple Tile Generation Steps . . . . .	105

## List of Abbreviations

- AI** ..... Artificial Intelligence
- BreakHis** ..... Breast Cancer Histopathological Image Classification
- CAD** ..... Computer-Aided Diagnosis
- CNN** ..... Convolutional Neural Network
- CT** ..... Computed Tomography Scan
- DenseNet** ..... Densely Connected Convolutional networks
- DNN** ..... Deep Neural Networks
- FNR** ..... False Negative Rate
- FPR** ..... False positive rate
- GBM** ..... Glioblastoma Multiforme
- GKS** ..... Gamma Knife Radiosurgery
- GLCM** ..... Gray Level Co-occurrence Matrix
- GradCAM** .... Gradient Weighted Class Activation Map
- H&E** ..... Hematoxylin and Eosin
- HPF** ..... High Power Field
- HSV** ..... Hue, Saturation and Value
- ISH** ..... In Situ Hybridization
- LBP** ..... Local Binary Patterns
- LGG** ..... Low Grade Glioma

<b>MRI</b>	Magnetic Resonance Imaging
<b>NIMHANS</b>	National Institute of Mental Health and Neurosciences
<b>PTE</b>	Peritumoral Edema
<b>ResNet</b>	Residual Network
<b>RBC</b>	Red Blood Cells
<b>SIFT</b>	Scale-Invariant Feature Transform
<b>STS</b>	Soft Tissue Sarcomas
<b>SVM</b>	Support Vector Machine
<b>TCGA</b>	The Cancer Genome Atlas Program
<b>t-SNE</b>	T-distributed Stochastic Neighbor embedding
<b>TIFF</b>	Tagged Image File Format
<b>UCSB</b>	University of California, Santa Barbara
<b>VGG</b>	Visual Geometry Group
<b>WHO</b>	World Health Organization
<b>WSI</b>	Whole Slide Imaging

## CHAPTER 1

### INTRODUCTION

#### **1.1 Oncological Challenges: Cancer and Brain Tumors.**

Cancer is a deadly disease that can effect any organ in the human body, and it is an essential barrier to increasing life expectancy. The American Association for Cancer (ASC) estimated there will be 27.5 million cases per year and 16.9 million deaths per year by 2040 [1]. Out of all cancers, brain cancer is the second largest death type in all cancers for ages below 15 years old, and it is the second largest growing cancer in patients with age above 65 years [1]. Brain cancers occur at low rates, but mortality is high compared to other cancerous tumors, creating a grave medical condition.

There are 120 tumor types in the brain based on the tissue that may be affected. Even some non-cancerous benign tumors are dangerous and lead to death because of their size and location in the brain. Blocking fluid flow between different brain regions can cause the tumor to spread [1].

The brain tumors have two types: primary and secondary tumors.

### 1.1.1 Primary Brain Tumors

Primary brain tumors originate within brain cells and can spread to other regions through the bloodstream or lymphatic system. These primary tumors are identified and classified based on the tissue they originate. The most common type of cancer starts in glial cells or supportive regions of glial cells. There are four types of glioma tumors [1] as shown in Figure FC1.1.

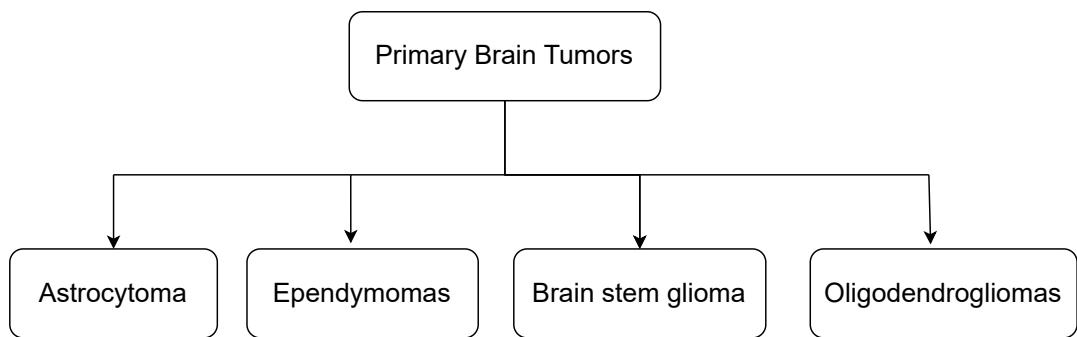


Figure FC1.1: Baseline Mode Architecture

**Astrocytoma:** Tumors can develop from specialized cells in the brain known as star cells. These cells can start or grow anywhere in the brain or the spinal cord. Among adults, such tumors are more likely to originate in the cerebrum. People commonly refer to Grade IV astrocytoma as Glioblastoma Multiforme (GBM) [1].

**Brain stem glioma:** Stem cells that originate from the stem part of the brain can develop into tumors in the brain, which can interfere with the vital functions performed by the brain. Doctors often identify these tumors as high-grade Astrocytomas [1].

**Ependymomas:** These tumors originate in the ventricles and spinal cord lining [1].

**Oligodendrogiomas:** Myelin-producing cells give rise to these tumors, a unique type of brain tumor. Unlike astrocytomas, they do not exhibit mild growth patterns or spread to other brain areas [1].

Other primary tumors that do not start from glioma cells are ;

**Meningiomas:** These tumors grow from the meningeal layers covering and protecting the brain and spinal cord. Most of these tumors are grade 1, meaning they are benign. Some meningiomas do not produce symptoms, and many are discovered incidentally during imaging studies [1].

### 1.1.2 Secondary Brain Tumors

The secondary brain tumors are called Metastasis. These tumors start from another part of the human body, spread toward the human brain, and can damage the brain tissues like primary tumors. This secondary tumor's name is decided based on the cancer cells coming from the organ. If the cancer cells come from the prostate, then that tumor is referred to as metastatic prostate cancer [1].

## 1.2 Understanding Necrosis in Brain Tumors

Necrosis is a term used to describe the death of cells in living tissue. This can occur in any type of cancer , including meningioma and glioblastoma, which are two types of brain tumors that have different structures [2].

The necrosis in meningioma is characterized by loss of cellular architecture, nuclear changes, and cellular debris [3]. In contrast, necrosis in glioma is identified with pleomorphic glial cells within a fibrillary background, accompanied by microvascular proliferation and necrotic regions [4].

In the case of brain tumors, including GBM and Low Grade Glioma (LGG), necrosis is particularly significant because it is associated with poor outcomes, such as increased tumor aggressiveness and resistance to therapy [5]. Necrosis is the death of cells in

a tissue or organ caused by injury, disease, or lack of blood supply. In brain tumors, necrosis can lead to a necrotic core, a dead tissue mass within the tumor that can promote tumor growth and spread [6]. Therefore, understanding the presence and extent of necrosis is essential in diagnosing and treating these tumors, as it can affect treatment options and patient prognosis [7].

Diagnosing necrosis in meningioma is primarily based on assessing nuclei density and tissue color [8]. In glioblastoma, additional biological characteristics such as microvascular proliferation, absence of tissue, high mitotic activity, and wavy patterns in the border regions of tissue slides may also indicate necrosis [9]. Necrosis in glioblastoma can take two forms: palisading, where the dead cells arrange themselves in a row-like pattern, or confluent, characterized by large zones of dead cells [10]. Despite these differences, it is essential to note that the fundamental nature of necrosis in meningioma and glioblastoma remains essentially the same [10].

### 1.3 Preparation of Biopsy Slides

During biopsies or surgeries, doctors extract tissue samples from the affected region to analyze the changes in the cells' shape, size, and structure. Tissue biopsy helps doctors identify specific proteins or biomolecules, making a more accurate diagnosis [11]. However, tissue biopsies only provide a small, three-dimensional representation of a larger tissue or organ, making it challenging to analyze the entire specimen in its original form for better diagnosis and treatment planning. Further, studying cell structures requires examining cross-section of tissues under a microscope. Preparing biopsy slides involves slicing thin sections of tissue and mounting them onto glass slides, typically 4-5 micrometers thick, for microscopic examination [11]. This process helps to address the challenges of examining tissue samples.

The slide preparation process involves several steps, including fixing the tissue with

a solution, stabilizing its cellular components, removing water by dehydration, and embedding it in a paraffin wax block. The tissue sections are then immersed in xylene and gradually exposed to lower alcohol concentrations to deparaffinize and rehydrate them. Then, a microtome cuts the block into thin sections and places it onto a glass slide. After the slide preparation, the staining process begins, using various techniques to highlight different tissue components. Hematoxylin and Eosin (H&E) staining, for example, can underline cellular structures, while immunohistochemistry can identify specific proteins. After staining, the pathologist covers the slides with a transparent slip and examines them under a microscope to identify and treat various diseases, including cancer and inflammatory disorders [11].

Staining techniques play a crucial role in histological analysis and help researchers better understand the structure and function of tissues in health and disease. For instance, fluorescent staining can detect particular molecules, while silver staining highlights nerve fibers. Other advanced methods are available, such as In Situ Hybridization (ISH) , which can detect specific nucleic acid sequences within a tissue, and electron microscopy , which can provide high-resolution images of tissue structures at the cellular and subcellular level. Part 1 in Figure FC1.1 shows slide preparation and staining

Hematoxylin and Eosin (H&E) staining is one of histology's most commonly used techniques. Hematoxylin is a basic dye that stains the nuclei of cells blue-purple, while eosin is an acidic dye that stains cytoplasm and extracellular structures pink. Combined, these two dyes allow for the visualization of cell nuclei, cytoplasm, and other structures within a tissue sample. In addition to highlighting typical tissue structures, H&E staining can reveal abnormalities such as inflammation, necrosis, and tumor growth. It is an essential tool in diagnosing and classifying diseases and is routinely used in clinical pathology to help guide treatment decisions [12].

Stained slides are examined under a microscope to identify the characteristic fea-

tures of the tissue sample. Pathologists look for changes in the shape, size, and structure of cells and the presence of specific proteins or other biomolecules [12].

Examining stained slides involves careful observation of cells and tissues at different levels of magnification. Pathologists use various techniques to identify the features of interest, including varying the microscope's focus, lighting, and contrast [12].

Pathologists also rely on their experience and knowledge of specific diseases to interpret the observed features. For example, in the case of cancer, pathologists look for abnormal cell growth patterns and the presence of cancerous cells with characteristics such as enlarged nuclei or irregular shapes and necrosis [12].

## 1.4 Tissue Slide Digitalization

Whole Slide Imaging (WSI) technology scans entire tissue slides to create high-resolution digital images. It enables the comprehensive and remote examination of pathology samples in fields such as medicine and research. WSI technology has revolutionized pathology by allowing pathologist's to view tissue samples digitally instead of using microscopes to view physical slides [13].

WSI technology has several advantages, including easier collaboration between pathologist's and the ability to analyze tissue samples remotely [14]. In addition, WSI technology has opened up new possibilities for research and education. With digital slides, researchers can analyze large amounts of tissue data quickly and efficiently [15]. At the same time, educators can use digital slides to teach pathology to students more interactively and engagingly [16]. With WSI, pathologists, researchers, and educators have a powerful tool to advance the pathology field and improve patient care [17]. The process steps from biopsy to slide visualization [18] are visually represented by Figure FC1.2. In Figure FC1.2, part 1 depicts the flow from biopsy to tissue block generation.

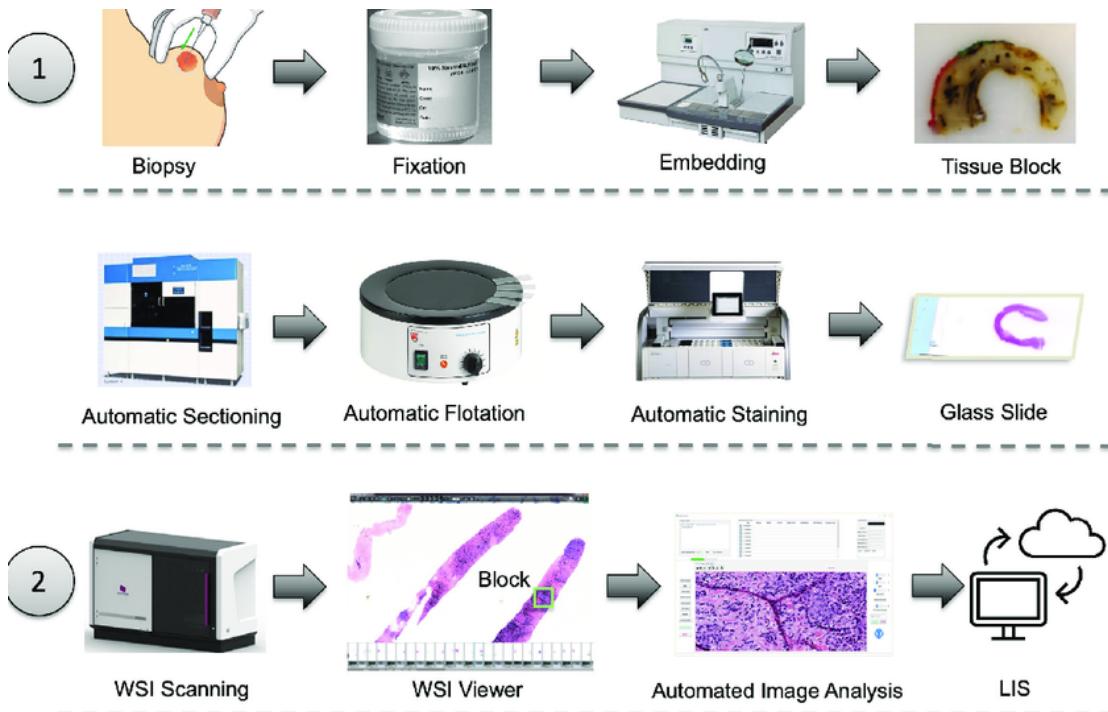


Figure FC1.2: 1)Tissue Processing 2) WSI Acquisition and Image Analysis

The middle flow represents operations on the tissue block, such as sectioning, flotation, and staining. The part 2 in Figure FC1.2 shows scanning and WSI slide generation.

**Slide Scanning:** The system automatically captures images of the entire slide at high magnification and resolution once it loads the slide onto the scanner. The high-resolution camera captures microscopic images of the slide, which can take several hours to complete, depending on the size and complexity of the sample.

Scanning typically operates at 20x or 40x magnification to capture detailed sample images. The high magnification and resolution of the images provide a wealth of information about the sample, including its cellular makeup, structure, and any abnormalities present [19]. Quality control measures are in place throughout the scanning and processing stages to ensure the image's accuracy and consistency. These measures include regular calibration of the equipment, visual inspection of the slides before and after scanning, and review of the final images by trained professionals [19] [20]. Figure

FC1.3 shows an example of a Whole Slide Image.

WSI enables statistical modeling on these slides to gain insights and automate diagnostic processes. Technology and machine learning integration have revolutionized pathology, improving diagnostic accuracy, identifying subtle changes in tissue morphology and crucial biomarkers, and a more nuanced understanding of complex tissue structures. This has led to better patient outcomes through improved diagnostic precision. This combination of WSI and machine learning highlights the transformative potential of cutting-edge technologies in medical diagnostics [21].

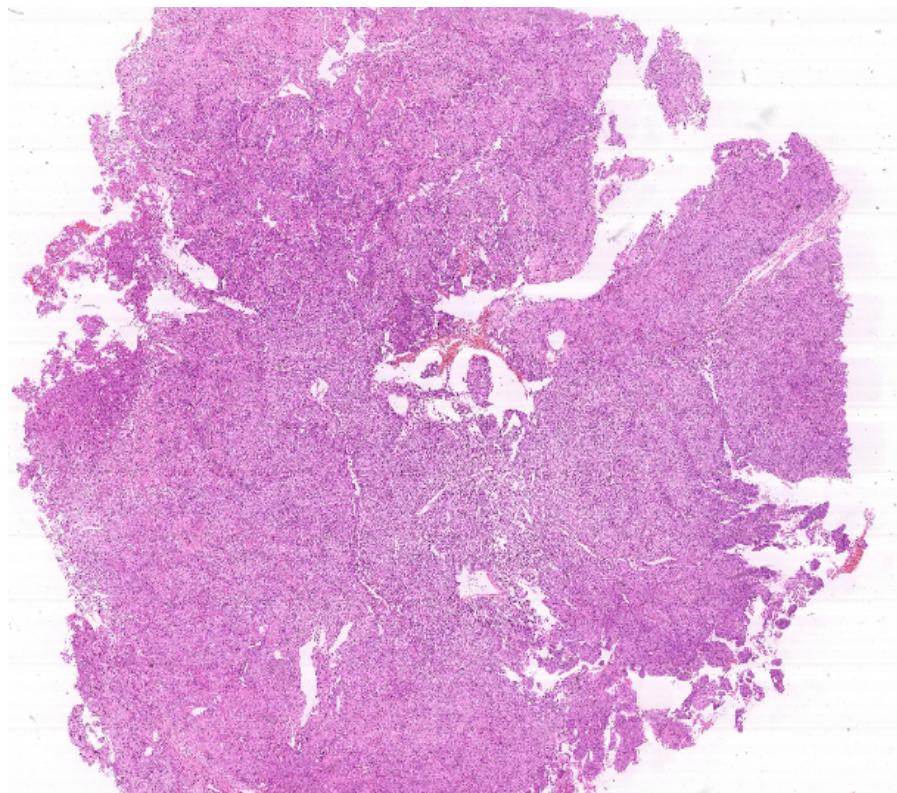


Figure FC1.3: High-resolution depiction of Meningioma tissue WSI at 0.4x magnification

## 1.5 Region in WSI

In WSI, a region is a specific area of interest on a digital slide selected for further analysis or examination. These regions are crucial in extracting meaningful information from large-scale digital pathology datasets as they allow researchers and pathologists to focus on specific regions of diagnostic or research significance.

**Region of Interest (ROI):** It refers to a specific part or section within the slide image chosen for closer scrutiny or analysis. In medical imaging, especially in pathology, not all parts of a slide may be relevant for diagnosis or study. As a result, pathologists or researchers mark particular regions as ROIs to concentrate their attention on areas that contain critical pathological features or structures. This targeted method simplifies the analysis process and enables a more effective and accurate assessment of the critical elements within the slide. ROIs can differ depending on the specific objectives of the analysis, such as identifying tumors, lesions, or other significant patterns.

A WSI can be divided into regions in different ways.

**Region by ROI:** One method is to extract a group of pixels or areas at a larger scale, focusing on a particular anatomical location or region of interest, without considering its specific characteristics. For instance, it may involve selecting a rectangular area that covers a tumor mass or a blood vessel network. In Figure FC1.4, the marked region includes both necrotic and non-necrotic tissue.

**Region by disease:** An alternative method is to create a region defined by disease-related attributes. This targets pixels or areas with one-of-a-kind characteristics associated with a particular disease type. This approach could involve manual region identification or image processing techniques to recognize areas with specific staining patterns, texture, or morphological features that suggest a particular disease. The green-colored polygons as shown in Figure FC1.5 indicate different disease-specific regions.

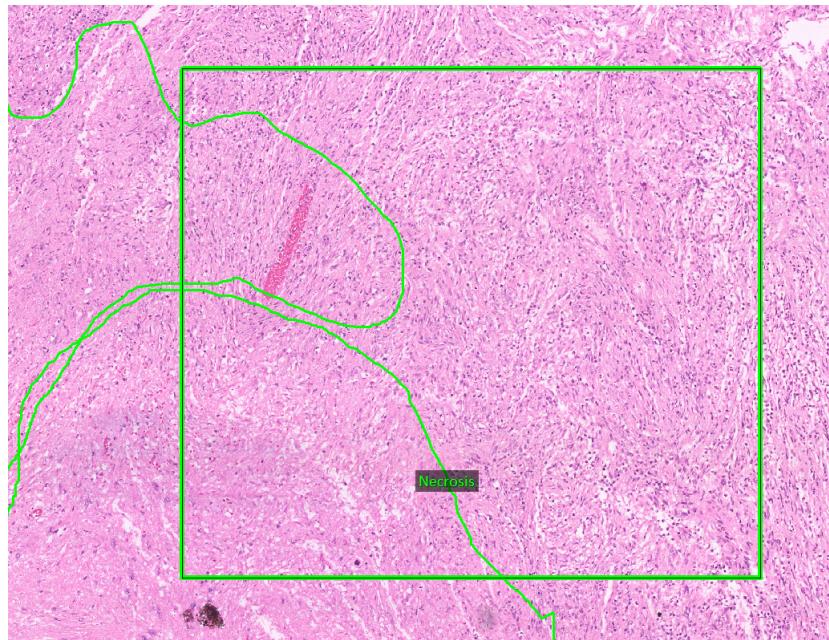


Figure FC1.4: Image showing region of interest

The computation of the area of a region in WSI is a critical step in quantitative analysis. The area measurement provides valuable quantitative data that can be utilized for various purposes, such as assessing the extent of tissue damage, determining the percentage of tissue involvement in a disease, or quantifying specific cellular features.

Computing the area of a region in WSI involves advanced image analysis techniques. These techniques may include segmentation algorithms, which delineate the boundaries of the region of interest, and pixel-wise calculations to determine the area occupied by the selected region. By employing these computational methods, researchers can obtain accurate and reproducible measurements, enhancing the reliability of the digital pathology analysis.

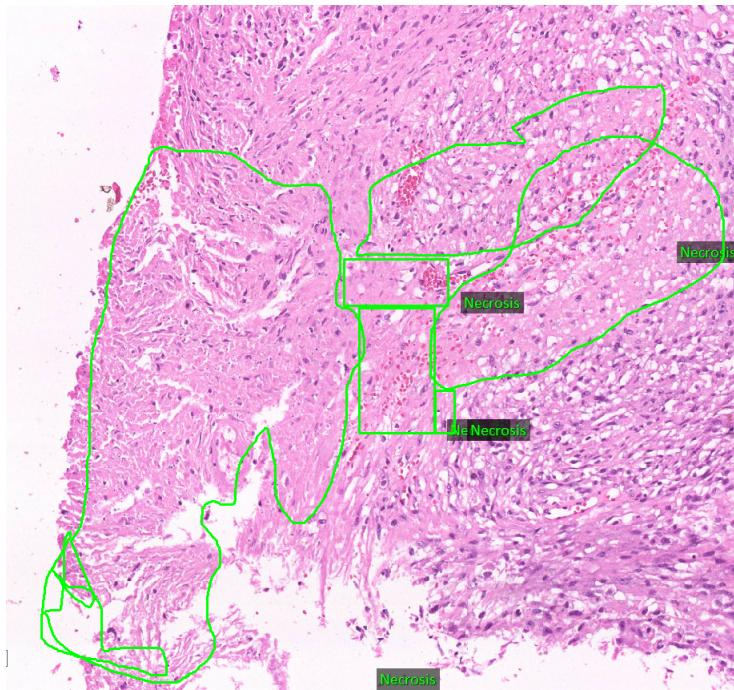


Figure FC1.5: Image showing Annotated Necrosis Regions

## 1.6 Computational Pathology through Image Processing

### 1.6.1 Early Detection and Scalability

Improving brain cancer survival rates is a significant focus area for medical professionals, and early advances in treatment approaches, such as targeted therapy and personalized medicine, can play a crucial role in this [22]. Early detection enables doctors to proceed with minimally invasive surgical techniques that allow for more precise removal of tumors while minimizing the damage to healthy brain tissue, leading to faster patient recovery times and reduced risks associated with surgery [22]. However, several factors, such as accessibility of clinics, availability of trained experts, and financial considerations, can impact the early detection of brain tumors [23]. Identifying brain tumors requires well-trained neuropathologists, but there is a shortage of such experts, which poses a significant challenge to early detection or identification [23]. Addition-

ally, traditional necrosis detection methods rely on time-consuming and labor-intensive histopathology examination, which can be subject to inter-observer variability [24].

The proliferation of digital histopathology images has increased the need for their analysis, specifically computer-aided diagnosis through machine learning techniques [25]. Digitizing biopsy glass slides to generate WSI presents significant potential for applying digital image processing in the large-scale diagnosis of diseases [26].

### 1.6.2 Traditional Methods in Image Processing for Feature Extraction

Researchers have reported that the earlier developments in medical image processing, which focused on traditional features such as color histograms, Gray Level Co-occurrence Matrix (GLCM) features, Texture and binary patterns, Scale-Invariant Feature Transform (SIFT), and statistical features, effectively diagnose a disease. However, these features have limitations such as complexity, difficulty handling variability, learning complex patterns, dependence on expertise, manual feature engineering, sensitivity to noise, and limited generalization [27].

**Color Histogram:** A color histogram is a visual representation that shows how frequently different color intensities appear in an image. It does this by dividing the intensity values of each color channel (red, green, and blue) into separate bins and then counting how many pixels fall into each bin. Understanding an image's color composition can provide valuable insights and benefit various applications, including image recognition, classification, and segmentation. The histogram can be considered a feature vector that can provide valuable insights into the image's color content [28].

When studying histopathology images, it is essential to remember that tissue staining can vary. This can impact the features observed from the histogram, as pathology slides are often stained over multiple days, affecting the amount of information in each pixel [29]. Furthermore, pathology images may contain several artifacts, includ-

ing air bubbles, tissue folds, and staining imperfections. These artifacts can introduce noise into the histogram, making distinguishing between fundamental features and artifacts [29] difficult.

**GLCM Features:** GLCM is a method used to analyze the Texture and variability of an image. This technique works by examining the occurrence of pixel values within an image grid. GLCM features are calculated by dividing an image into a grid of cells and computing a GLCM for each cell to represent the co-occurrence of pixel values within that cell. The image analysis uses these GLCMs as features to represent the image [30].

The GLCM technique helps extract information from histopathology images of small tissue samples, but its features are limited. GLCM relies on the co-occurrence of pixel values, which can sometimes result in the extraction of artifacts such as air bubbles that resemble tumor cells. Therefore, when analyzing small tissue samples, it is essential to use other image analysis techniques in conjunction with GLCM to obtain a more accurate assessment of the sample [30].

**SIFT Features:** SIFT is widely used in computer vision to extract features from images. It analyzes the pixel value distribution in a ROI to represent its visual content. SIFT is a method for extracting features from an image. The algorithm divides the picture into a grid of cells and computes a histogram of pixel values for each cell. Then, it generates a SIFT feature vector using the histogram. The SIFT feature extraction process includes generating the scale space, detecting key points, assigning key-point orientation, and generating key-point descriptors [31]. The Figure FC1.6 shows grid search and descriptor selection for SIFT feature extraction method.

In the field of histopathology, the interpretation of the same image can vary among multiple pathologists, leading to significant differences. This variability can make it challenging to develop robust SIFT features that account for inter-observer differences. Additionally, the lack of a standard staining protocol can result in substantial variations

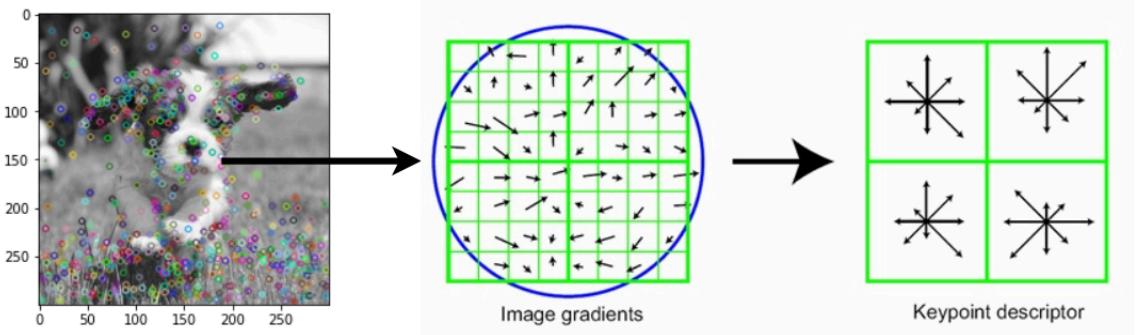


Figure FC1.6: Visualization of key point descriptors and Grid Search

in the appearance of images from different hospitals and laboratories, making it difficult to create transferable SIFT features for different datasets [32].

**Local Binary Patterns:** The LBP algorithm extracts image texture features effectively by using a straightforward approach. Texture, a crucial property of images, can serve various purposes, such as image classification, object detection, and image segmentation. The algorithm relies on the concept that the local distribution of pixel values in a specific area can express an image's texture. Texture features provide information about an image's contrast, homogeneity, directionality, roughness, and regularity [33].

The extraction of texture features in histopathology, particularly for machine learning applications, encounters challenges. Identifying these features is intricate, as the texture characteristics of tumors may closely resemble those present in the surrounding healthy tissue. The similarity between pathological and normal textures introduces complexity, making it challenging to accurately differentiate between the two. Additionally, the presence of histopathological image artifacts contributes to noise, further complicating the extraction of meaningful texture features for subsequent machine learning analyses [34].

The methods used in this section are used in the baseline model architecture shown in section 4.2. The features from these methods are combined and then used to train

machine learning model.

### 1.6.3 CNN: Deep Learning Method for Image Classification

Convolutional neural networks excel at image classification tasks by extracting image features via convolutional and pooling layers. The convolutional layers identify patterns, while the pooling layers reduce image dimensionality without losing critical information. CNN's are versatile and valuable for several image-related tasks, including image classification, object detection, and image segmentation. Many smartphones use CNN's for facial recognition, and medical imaging applications use them to detect tumors.

**Convolution** Convolution and pooling are essential operations that form the backbone of a CNN. Convolution helps to extract intricate features from images by applying a filter matrix to every pixel in the image [35]. The filter can identify specific patterns such as edges, corners, or textures, and the convolution operation produces a feature map that highlights where these patterns are present in the original image. This process helps to create a more abstract representation of the image, allowing the neural network to understand better and classify it.

**Pooling** After performing the convolution operation, the next step is pooling, which involves selecting essential features from the image. This process helps reduce the image's size while retaining critical information. It combines the data from all the pixels in a filter block into a single value. The pooling layer reduces the computational cost of the network and enhances its ability to generalize [35].

**Flatten Layer** The flattening layer is an essential component in neural networks, particularly for tasks such as image classification, object detection, and natural language processing. This layer takes the output from the previous layer and converts it into a one-dimensional vector [36]. The flattening layer is necessary because the final layer

of a neural network is typically fully connected and can only accept one-dimensional inputs.

**Transfer Learning** Transfer learning is a popular approach to solving complex machine-learning problems, including medical image classification. In transfer learning, a pre-trained neural network is used as a starting point for a new task, allowing the new model to leverage the knowledge and insights gained from the previous task.

When it comes to medical image classification, transfer learning has proven to be a powerful tool for improving accuracy and reducing the need for large amounts of labeled data. By using pre-trained models on large datasets such as ImageNet have shown state-of-the-art performance on a range of medical image classification tasks.

One of the key benefits of transfer learning for medical image classification is that it allows models to generalize more effectively to new datasets. This is especially important in medical imaging, where data can be scarce and expensive to collect. By leveraging pre-existing models one can reduce labeled data needed to achieve high accuracy and reduce the time and cost of model development.

XceptionNet is a neural network architecture that uses depth-wise separable convolutions, making it popular for medical imaging tasks [37]. It leverages transfer learning to use pre-trained knowledge from datasets like ImageNet, making it effective in identifying tumors in MRI scans and anomalies in X-rays. Its streamlined design boosts computational efficiency and interpretability, providing healthcare professionals with valuable insights for improved diagnostics and patient care.

#### 1.6.4 XceptionNet for Image Classification

XceptionNet [37] is a highly efficient and accurate convolutional neural network architecture for image classification and recognition tasks. It was developed by François

Chollet in 2017 as an extension of the Inception architecture, among the most popular CNN models. The name "Xception" stands for "Extreme Inception" and reflects the network's focus on enhancing the power of deep networks through depth-wise separable convolutions.

One of the critical features of XceptionNet is its use of depth-wise separable convolutions, which separate the spatial and channel-wise dimensions of the input data. This convolution type makes feature learning more efficient. It reduces the required parameters, resulting in a more streamlined and efficient architecture. The network consists of multiple depth-wise separable convolutional blocks followed by fully connected layers for classification. The architecture [38] of XceptionNet for the ImageNet1k [39] dataset is illustrated in Figure FC1.7.

The depth and width of XceptionNet enable it to capture intricate hierarchical features in images with high accuracy, making it well-suited for image classification tasks. Its efficient design has also made it highly influential in various computer vision applications, including object recognition, image segmentation, and fine-grained classification.

One of the critical advantages of XceptionNet is its ability to facilitate transfer learning scenarios. By leveraging pre-trained models on large datasets, practitioners can improve performance on domain-specific tasks with smaller datasets. Additionally, the architecture has demonstrated competitive performance on benchmark datasets such as ImageNet1k, where it achieved top-tier accuracy.

In summary, XceptionNet is a powerful and efficient CNN architecture that significantly contributes to computer vision. Its depth-wise separable convolutions have improved accuracy and efficiency, making it a popular choice for image classification tasks and transfer learning scenarios.

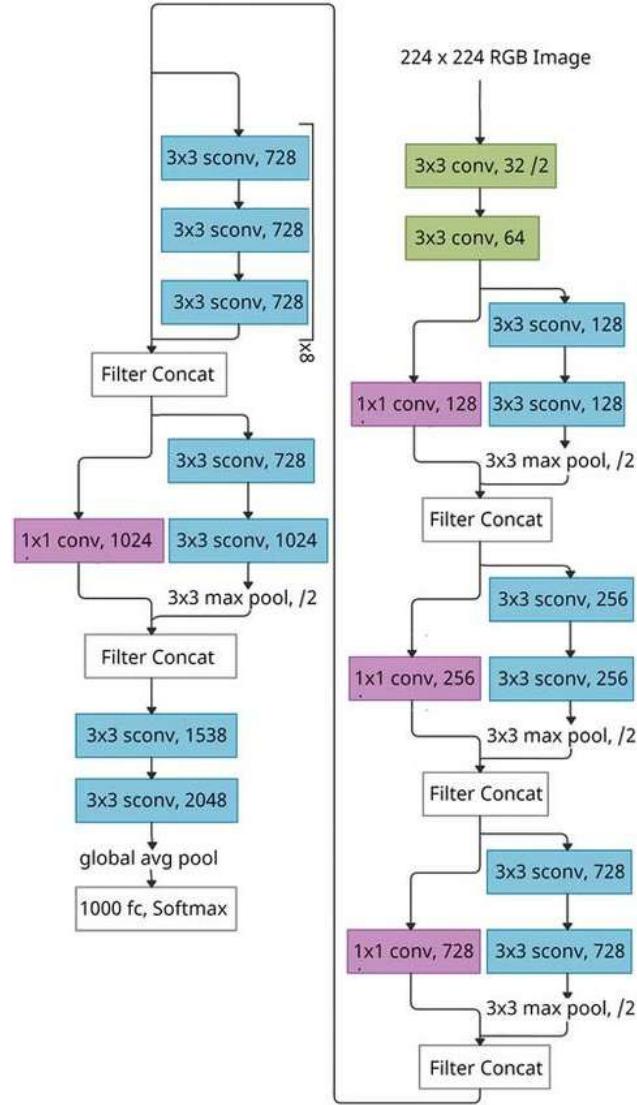


Figure FC1.7: Architecture of the XceptionNet model for ImageNet1k

## 1.7 Support Vector Machines:

Support Vector Machines (SVMs) [40] are a powerful subset of machine learning models designed specifically for classification tasks. They excel at predicting the category or class of a given input by establishing robust decision boundaries. These decision boundaries are effective in separating different classes in the input space.

SVMs [40] are a type of machine learning algorithm that can effectively handle

complex and non-linear relationships between input features and class labels. To classify data, SVMs use a technique called the kernel trick [40], which enables them to capture non-linear relationships without explicitly mapping input features into higher-dimensional spaces. This is beneficial because performing the explicit mapping can be computationally expensive. Instead of mapping the input features, the kernel trick implicitly computes the dot product between the input features in the higher-dimensional space. By doing so, SVMs can operate effectively in a higher-dimensional feature space without explicitly computing the transformation. This saves computational resources and enables SVMs to handle non-linear relationships in data efficiently. The kernel trick is a powerful tool that enhances the flexibility of SVMs, making them well-suited for a wide range of applications.

SVMs work by identifying a hyperplane that best separates different classes while maximizing the margin, which is the distance between the hyperplane and the nearest data points from each class. This margin maximization contributes to the model's robustness and generalization capability.

The decision function in a non-linear SVM:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \quad (\text{Eqn 1.1})$$

Here,  $x$  represents the input data,  $K(x, x_i)$  is the kernel function,  $\alpha_i$  are the Lagrange multipliers,  $y_i$  is the class label of data point  $x_i$ , and  $b$  is a bias term. The kernel function computes the similarity between pairs of data points in the transformed space, allowing SVM to capture complex decision boundaries.

Commonly used kernel functions include the polynomial kernel ( $K(x, x_i) = (x \cdot x_i + c)^d$ ) and the Radial Basis Function (RBF) kernel ( $K(x, x_i) = \exp(-\frac{\|x - x_i\|^2}{2\sigma^2})$ ). These kernel functions introduce non-linearity, enabling SVM to more effectively model intricate

relationships within the data.

In summary, SVM extends its capability to handle non-linear decision boundaries through the kernel trick by transforming input features into a higher-dimensional space, where it can efficiently find hyperplanes for adequate classification.

## 1.8 Evaluation of prediction/detection performance

The prediction performance of the models are evaluated using the performance measures obtained from the confusion or contingency matrix and weighted scores.

**Confusion Matrix:** A confusion matrix is a table that is used to evaluate the performance of a classification model. It provides a detailed breakdown of the model's predictions and the actual outcomes for different classes as shown in Table TC1.1. The matrix is particularly useful when dealing with binary classification problems, where there are two possible outcomes (positive and negative).

Table TC1.1: Confusion Matrix

		Predicted	
		Positive	Negative
Actual	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

- **TP (True Positive):** The model correctly predicts a positive case.

Example: A patient actually has a tumor, and the model accurately predicts it as positive.

- **FN (False Negative):** The model incorrectly predicts a case as negative, when it is actually positive.

Example: A patient has a tumor, but the model incorrectly classifies it as negative. This can have serious consequences as it might delay necessary treatment.

- **FP (False Positive):** The model incorrectly predicts a case as positive, when it is actually negative.

Example: A patient does not have a tumor, but the model incorrectly classifies it as positive. This can lead to unnecessary stress, procedures, and costs.

- **TN (True Negative):** The model correctly predicts a negative case.

Example: A patient does not have a tumor, and the model correctly predicts it as negative.

**Accuracy:** This is a measure of the overall correctness of the model. It represents the ratio of correctly predicted instances to the total instances.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (\text{Eqn 1.2})$$

**Precision:** This is the ratio of correctly predicted positive observations to the total positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (\text{Eqn 1.3})$$

**Recall:** This is the ratio of correctly predicted positive observations to all observations in the actual class. It measures the ability of the model to capture all the relevant instances. This metric is also known as sensitivity or True Positive Rate (TPR), emphasizing the model's effectiveness in identifying positive cases

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (\text{Eqn 1.4})$$

**F1 Score:** This is the harmonic mean of precision and recall. It provides a balance between precision and recall, especially when there is an uneven class distribution.

$$F1\text{-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (\text{Eqn 1.5})$$

**Specificity:** This is the ratio of correctly predicted negative observations to all observations in the actual negative class. It measures the ability of the model to accurately exclude instances not belonging to the positive class, emphasizing its ability to avoid false positive predictions.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (\text{Eqn 1.6})$$

**Weighted Metric Evaluation:** It is crucial to adopt a weighted scoring approach to ensure an unbiased evaluation of model performance, especially when dealing with imbalanced datasets. This approach reduces the impact of varying dataset sizes on performance metrics for positive and negative instances. By using weighted scores, each metric (such as accuracy, recall, precision, or f1 score) is multiplied by the respective number of images in both positive and negative classes. Afterward, the weighted metrics are averaged to ensure a fair evaluation process, regardless of potential imbalances in the dataset distribution.

Consider the three datasets P, Q, and R. Each contains a distinct number of images denoted as  $P_i$ ,  $Q_i$ , and  $R_i$ . A performance metric is reported on these datasets, denoted as X, Y, and Z. The objective is to calculate weighted scores that account for the dataset sizes.

The Weighted Metric (WM) is computed as follows:

$$WM = \frac{X \cdot P_i + Y \cdot Q_i + Z \cdot R_i}{P_i + Q_i + R_i}$$

In this formula, each metric (X, Y, Z) is multiplied by the corresponding dataset size

( $P_i$ ,  $Q_i$ ,  $R_i$ ), and the results are summed. The total is then divided by the total number of images when three datasets are combined, yielding the weighted metric. This approach ensures that the contribution of each dataset to the overall metric is proportionate to its size.

## 1.9 Hypothesis

1. **Hypothesis 1:** Our hypothesis is that combining domain-based features with pre-trained deep learning model-based features will significantly improve the accuracy and efficiency of necrosis detection in Meningioma, compared to using either feature type alone.
2. **Hypothesis 2:** We hypothesize that the model trained for necrosis detection in Meningioma will exhibit a transferable capability to detect necrosis in Glioblastoma Multiforme (GBM) and Low-Grade Glioma (LGG) effectively.

## 1.10 Research Objectives

The following are the stated objectives of this research.

1. We aim to improve the accuracy and efficiency of detecting necrosis in Meningioma by combining domain-based features with pre-trained deep-learning model-based features to develop a novel feature extraction method.
2. We aim to assess the capability of the model, which has been trained to detect necrosis in Meningioma, to detect necrosis in GBM and LGG.
3. We aim to aid pathologists by identifying regions of interest that contains necrosis in a WSI.

## 1.11 Thesis Outline

The thesis comprises five chapters, each addressing distinct aspects as outlined below:

Chapter 1: This chapter provides an overview of brain tumors and their pathologies, especially necrosis. It also explores image processing techniques and advanced methods in computational pathology. These foundational concepts will be essential for understanding the subsequent chapters.

Chapter 2: This chapter provides a comprehensive survey of literature on computational pathology methods, ranging from conventional to deep learning approaches. We explore image classification methods for histopathology images.

Chapter 3: This chapter describes acquiring data and generating tiles from WSI slides obtained from NIMHANS and TCGA. We will discuss the preparation of inter-slide and intra-slide batches from meningioma slides. Additionally, we will provide information on labeled data for other brain tumors, including LGG and GBM.

Chapter 4: This chapter explores challenges in classifying histopathological images using deep learning and proposes incorporating domain knowledge to enhance model performance. We will assess the model's effectiveness on inter-slide and inter-class categories within the meningioma dataset. Additionally, we will evaluate the meningioma-trained model's ability to identify necrosis in other brain tumors, such as glioblastoma and low-grade glioma.

Chapter 5: This chapter examines how well the model performs when handling different grid positions and evaluating its performance on specific regions within a slide.

Chapter 6: This chapter summarizes the contributions made throughout the thesis and presents the conclusions derived from our work. Additionally, we discuss potential future directions for further research in this field.

## 1.12 Summary

This chapter discusses brain cancers, including primary and secondary tumors, focusing on necrosis in meningioma and glioblastoma. We have explored the current issues in pathology and how technology is helping to address them, including artificial intelligence and machine learning that enable solutions for digitized medical data. We have also examined feature extraction for machine learning and deep learning algorithms, looking at both handcrafted and automated methods applied to medical images and the approach to evaluate prediction performance. We also discussed various metrics that we would use to analyze the model's performance.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 AI for Medical Diagnosis

Deep Learning has proven effective in recent years across various fields, including computer vision, speech recognition, and Natural Language Processing. This success has translated into rapid advancements in applying Artificial Intelligence (AI) to medical diagnosis. By analyzing patterns in medical records, AI algorithms can now help detect and diagnose diseases with remarkable efficiency [41] [42]. Early applications of DL in medical record analysis focused on three main areas: Dermatology, Radiology, and Ophthalmology. These fields primarily utilize images and scans, making them well-suited for AI-based detection systems [43]. DL algorithms have proven highly successful in identifying and classifying skin lesions (Dermatology) [44], detecting abnormalities in X-rays, CT scans, and MRIs (Radiology) [45], and identifying diabetic retinopathy from retinal scans (Ophthalmology) [46].

The growing demand for medical services has made manual diagnosis increasingly challenging, requiring high resources and skilled personnel [43]. DL algorithms present a robust solution, automating parts of the diagnostic process [47]. Deep Learning improves efficiency and opens up new possibilities for providing timely and accurate diagnoses to a broader population. While the initial focus was on image-based data, Deep

Learning is now expanding its reach to analyze other forms of medical data, such as EHR and genomics data [48]. By combining these diverse data sources, researchers are developing even more efficient algorithms that predict patient outcomes, personalize treatment plans, and identify individuals at high risk for developing specific diseases.

There are several advantages of incorporating AI in healthcare, such as:

1. **Increased accuracy:** Early and accurate diagnosis allows for prompt treatment, improving patient outcomes [47].
2. **Reduced costs:** Identifying at-risk individuals enables preventive measures, reducing the need for expensive treatment interventions [49].
3. **Improved access to healthcare:** AI-powered solutions can bridge geographical and economic barriers, providing diagnostic services to underserved communities [50].

The authors in [51] explores how digital pathology and whole slide imaging have enabled the development and application of AI algorithms. Researchers have utilized hand-crafted feature-based and deep-learning approaches to detect, grade, predict prognosis, and treat cancer. Several studies have used morphological features extracted from histology images to predict survival outcomes, gene mutations, and immunotherapy responses.

While AI shows potential to increase efficiency, accuracy, and reproducibility in pathology, challenges remain regarding data quality, algorithm validation and regulation, interpretability, and clinical adoption. A recent study tested the effectiveness of Paige Prostate [52], an AI-powered detection system approved by the Food and Drug Administration (FDA) for identifying prostate cancer in whole slide images of prostate biopsies from 100 patients. The system was compared to pathologists in its ability to differentiate between cancerous and benign slides and to improve diagnostic accuracy

and efficiency. The study found that Paige Prostate had higher sensitivity and negative predictive value for cancer detection compared to human pathologists. Moreover, it was able to detect cancer in some cases that were initially missed by pathologists. Using Paige Prostate also led to a significant reduction in diagnostic time by approximately 65%.

## 2.2 Feature Extraction for Histopathology

### 2.2.1 Conventional Methods for Histopathology Feature Extraction

Automated analysis of histopathology images is a rapidly growing and impactful area of research [53] [54]. However, these image's rich visual details and diverse structures require specialized analysis methods depending on the target organ and specific task. Authors in [55] used Gabor filter-based texture analysis to identify relevant regions. This method leverages pathologists' reliance on magnified images of cervix biopsies to grade cervical cancers based on the spread of abnormal cells within the epithelium layers.

The research paper [56] presents a novel approach to analyzing colon cancer by utilizing a homogeneity measure based on the texture of distinct tissue regions. Identifying regions through the spatial arrangement of their cellular and connective tissue components is crucial in accurately detecting and diagnosing diseases. Additionally, these organizational patterns vary across organs. Addressing prostate cancer diagnosis [57], utilized gland area as a feature, recognizing its discriminatory power between benign and malignant glands. The authors in [58] developed a hierarchical algorithm to detect cancerous regions at lower resolutions and refine gleason grades at higher resolutions. The authors proposed an automated breast cancer grading model using texture and topological features, which demonstrated discriminate capability and yielded an

interpretable model based on graph theory.

Researchers have conducted quantitative analyses of the prognostic value of tumor necrosis in histological specimens for lung cancer, colorectal cancer, upper urinary tract carcinoma, breast carcinoma, gastrointestinal stromal cancer, and renal carcinoma concerning biological variables in studies [59], [60], [61], [62], [63], [64]. Tumor necrosis is a crucial prognostic factor at the histological level. These studies indicate that tumor necrosis characteristics can be correlated with the type and extent of respective tumors and show a significant impact on tumor prognosis. Furthermore, there are several studies on texture-based analysis in histology, for example, in [65] for classifying the grade of neuroblastic differentiation in hematoxylin and eosin-stained neuroblastoma images and in [66] for classifying cervical nuclei in Papanicolaou stained cervical tissue images. In [67], researchers have explored different orders of textures for pattern recognition to contribute towards pathoanatomic diagnosis in histopathology, and they have described their findings in detail. However, these studies chiefly address the problem of diagnosis-oriented classification for the respective tumors.

Research has been conducted on classifying histological image signatures of Glioblastoma images into necrosis, apoptosis, and viable regions [68]. However, the method used in this research was based on a small set of training images consisting of only 50 patches of size 80x80. The authors used a classification approach based on reconstruction from subspace analysis, which is known to be a slow process and takes several hours to learn small images. Therefore, this method is not suitable to be applied to datasets containing numerous gastric cancer image patches of various sizes. Representing histopathology images poses significant challenges due to high visual variability from diverse acquisition processes, anatomical differences, staining procedures, image magnification, and tissue section type [69]. Despite these challenges, automated image representation and analysis methods have achieved remarkable success in medical imaging.

While handcrafted features have been the cornerstone of histopathology image analysis for decades, their limitations are becoming increasingly apparent in the face of complex diagnostic tasks.

Here are some fundamental limitations to consider while using handcrafted features

### **2.2.2 Limitations of Conventional Features for Histopathology**

1. **Limited Feature Representation:** Handcrafted features often rely on pre-defined parameters and assumptions, potentially overlooking subtle or unexpected patterns [70]. Using conventional features can lead to missing important information and limited generalizability to other tasks or datasets [71].
2. **High Development Cost and Expertise:** Creating robust handcrafted features requires significant expertise in image processing and domain knowledge [71]. Handcrafted feature extraction can be time-consuming and expensive, hindering the development and adoption of new diagnostic tools.
3. **Feature Engineering Challenges:** Selecting the optimal set of handcrafted features can be challenging and involve trial and error [70]. The optimal set can vary depending on the disease, image quality, and staining protocol
4. **Difficulty in Handling High-Dimensional Data:** Handcrafted features often need manual adaptation to different imaging conditions and tissue types [71]. Handcrafted feature extraction can be time-consuming and limit the scalability of diagnostic tools.

### **2.2.3 Deep Learning Methods for Histopathology Feature Extraction**

The deep learning-based methods to extract features involve understanding the essential features and extracting them without any manual intervention. This article [72]

describes a deep-learning approach for detecting necrosis in whole slide images of Canine Perivascular Wall Tumors (cPWTs). The researchers extracted patches from annotated regions of necrosis and non-necrosis in cPWT slides. They trained several deep-learning models to classify patches, including a DenseNet-161 convolutional neural network and an ensemble model.

This paper [73] compares three deep learning algorithms (SSD, Mask R-CNN, and DeepLabV3+) for detecting hepatic necrosis in histopathology images from non-clinical toxicity studies. Researchers treated rats to induce hepatic necrosis and digitized their liver tissue slides. They trained the algorithms on 5,750 image tiles containing necrosis annotations and evaluated their performance on large test images based on precision, recall, accuracy, and mask IoU. The two segmentation models, Mask R-CNN and DeepLabV3+, achieved over 90% accuracy. DeepLabV3+ distinguished necrosis from other features best. This study [74] investigated the use of machine learning to assist pathologists in detecting and quantifying necrosis in digitized WSIs of canine STSs. Pathologists manually annotated necrotic regions in 21 WSIs. The authors used annotations to train a pre-trained DenseNet161 model to classify whether image patches contain necrosis. The model achieved 92.7% accuracy in detecting necrosis in validation data from three-fold cross-validation experiments.

The authors in [75] propose a new machine learning model called Hierarchical Deep Learning based Brain Tumor Classifier to accurately detect and classify brain tumors into four types: glioma, meningioma, pituitary, and no-tumor using biopsy images. The proposed model uses a convolutional neural network and achieves 92.13% accuracy in detecting brain tumors, outperforming previous methods. The authors in [76] aimed to develop a deep neural network –based predictive model to forecast Peritumoral Edema (PTE) after Gamma Knife Radiosurgery (GKS) for meningioma. The researchers trained and evaluated a 50-layer residual neural network model at a single center using clinical and radiological data from 202 patients who underwent GKS for

meningioma. The authors in [77] discuss applications of deep learning techniques in tumor pathology. It reviews studies that have used algorithms to perform tasks like tumor detection, classification, grading, diagnosis, prognosis prediction, analysis of pathological features, and identification of genetic mutations. While deep learning shows promise for improving accuracy and efficiency, challenges remain around validating models on larger multi-institution datasets and explaining results.

Here are some limitations to consider while using Deep Learning based features

#### 2.2.4 Limitations of Deep Learning Features for Histopathology

1. **Overfitting and Generalizability:** Deep learning models are susceptible to overfitting, where they memorize the training data rather than learning generalizable patterns [78]. Deep Learning can lead to poor performance on unseen data, limiting their clinical utility and real-world impact.
2. **Data Requirements and Bias::** Training deep learning models requires massive amounts of high-quality, labeled data, which can be expensive and time-consuming to acquire [79]. Limited or biased training data can lead to models performing poorly on unseen data that differs from the training set, raising concerns about generalizability [80].
3. **Domain Adaptation and Transfer Learning Challenges:** Adapting deep learning models to different imaging protocols, tissue types, and staining variations can be challenging [80]. While transfer learning can help bridge some gaps, it may only sometimes overcome these differences effectively, requiring further research into robust adaptation techniques.
4. **Regulatory and Ethical Considerations:** Implementing deep learning models in clinical practice raises ethical and regulatory concerns around data privacy, the potential for bias, and misdiagnosis risks [80].

**5. Explainability and Trust:** Deep learning's lack of explainability makes it hard to trust and interpret. This is particularly problematic in healthcare and finance, where transparency and accountability are vital. Improving the interpretability of deep learning models is a significant challenge [80].

### 2.2.5 Domain Feature Extraction

This authors in [81] presents a study on domain shift in convolutional neural networks trained for tumor classification on histopathology whole-slide images. They analyze how data augmentation and normalization strategies affect model performance and representations. Based on a model's learned representation, they introduce a novel measure to evaluate the distance between domains.

This authors in [82] proposes a domain-adversarial training framework for convolutional neural networks to learn invariant representations of histological images and improve generalization to unseen domains. The framework models each slide image as a distinct domain and trains feature extractors to make representations independent of domains while optimizing for classification. This article [83] surveys research works that integrate medical domain knowledge into deep learning models for various medical image analysis tasks such as disease diagnosis, lesion detection, and segmentation. It categorizes existing works based on the type of task, domain knowledge introduced, and how the knowledge is incorporated.

### 2.2.6 Scope of using Domain Knowledge in Histopathology

**1. Improved Feature Extraction::** Pathologists know extensively about tissue morphology, cell types, and disease patterns. This knowledge can guide the design of deep learning models, leading to more accurate and relevant feature extraction [84].

2. **Enhanced Model interpretability:** Integrating domain knowledge into deep learning models can make pathologists' decision-making process more transparent and understandable, facilitating trust and acceptance in clinical settings [85].
3. **Improved Generalizability :** By incorporating domain knowledge into model design, we can address challenges like data bias and overfitting, leading to models that perform well on unseen data and generalize across different imaging protocols and tissue types [86].

## 2.3 Image Classification

### 2.3.1 Machine Learning for Image Classification

A typical architecture of using Machine Learning for histopathological image classification is shown in Figure FC2.1. In [87], the authors discuss applying machine learning techniques to analyze histopathological images. The authors of [88] explain in their paper how they utilized algorithms such as deep learning to precisely segment neurons in histology images of macaque brains. They presented a method for selecting a set of stable and concise features that accurately segment the neurons in the digital pathological images. They compared multiple machine learning algorithms such as SVM, Random Forest (RF), and CNN for classifying breast cancer histopathology images based on extracted features from various domains, including color and texture. The methodology involved an exhaustive search of all possible feature combinations using brute force to assess each combination's efficacy in segmenting neurons.

The authors in [89] proposes an automatic breast cancer classification system that uses a SVM classifier based on integrated features such as texture, geometrical, and color features extracted from histopathology images. The authors tested the system on two publicly available datasets - UCSB and BreakHis.

In a recent study, the authors [90] explored machine-learning techniques to classify cervical histopathology images. They extracted texture and morphological features from the images and used them as input for an SVM classifier. In a study [91], the performance of different machine learning algorithms, including SVM, Random Forest, and CNN, was compared for the classification of breast cancer histopathology images. The study sheds light on the potential of machine learning in medical image analysis. The authors in [92] discuss using machine learning approaches like SVM and LR with pre-trained Convolutional Neural Networks for the automatic classification of breast cancer histopathology images. The researchers proposed a two-phase model to classify images based on magnification, distinguishing between 40x, 100x, 200x, and 400x and then classifying the samples as benign or malignant.

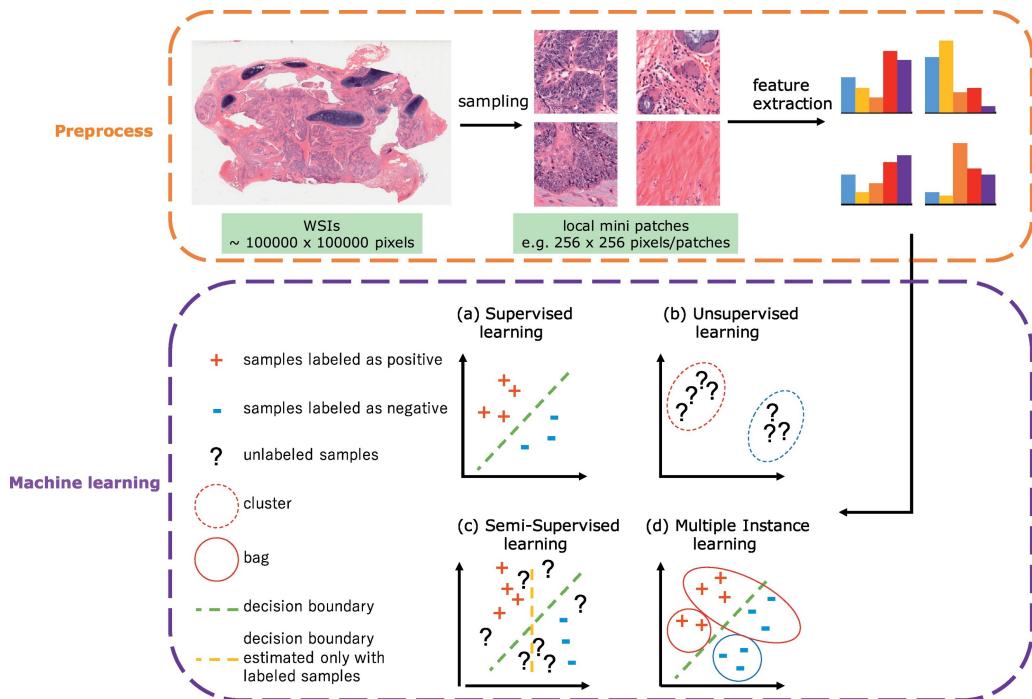


Figure FC2.1: Steps for machine learning in digital pathological image analysis

The paper referenced as [93] describes a new technique that employs image processing to aid pathologists in diagnosing breast cancer with greater accuracy and efficiency. The approach consists of two modules: an anomaly detection method utilizing a sup-

port vector machines, which enhances the second module's training, and a RANet. In another study, the authors [94] developed a machine learning algorithm that employs a fruitfly optimization algorithm to classify breast cancer using histopathological images. The method employs a patch-classification model that first categorizes image patches into normal, benign, in situ, and invasive categories. It then employs this patch information to classify all images into cancerous and non-cancerous groups. Furthermore, [95] explores the performance of different SVM and KNN algorithms for sub-size gastric pathology images to improve classification accuracy using ensemble learning.

### 2.3.2 Transfer Learning for Image Classification

A typical architecture of using transfer learning for histopathological image classification is shown in Figure FC2.2. In their study [96], the authors explored the use of various pre-trained ResNet models for histopathology image classification tasks, focusing particularly on metastatic cancer in lymph node sections. The authors compared different transfer learning strategies. The authors discovered that deeper ResNet models, such as ResNet-152, did not necessarily outperform shallower ones like ResNet-34. The best results were obtained by fine-tuning all layers of the pre-trained ResNet model. This study highlights that the depth of the pre-trained histopathological models is an important factor to consider when working with them. The authors of [97] uses pre-trained convolutional neural network models like DenseNet-161 and ResNet-50 for automated detection and classification of diseases from digital histopathology images via transfer learning. The DenseNet-161 model achieved 97.89% accuracy on grayscale images, while the ResNet-50 model obtained 98.87% accuracy on color images, outperforming other state-of-the-art methods.

This study [98] introduces a novel CNN for automatic breast cancer classification in histopathological images. The model achieves impressive accuracy between 98.87%

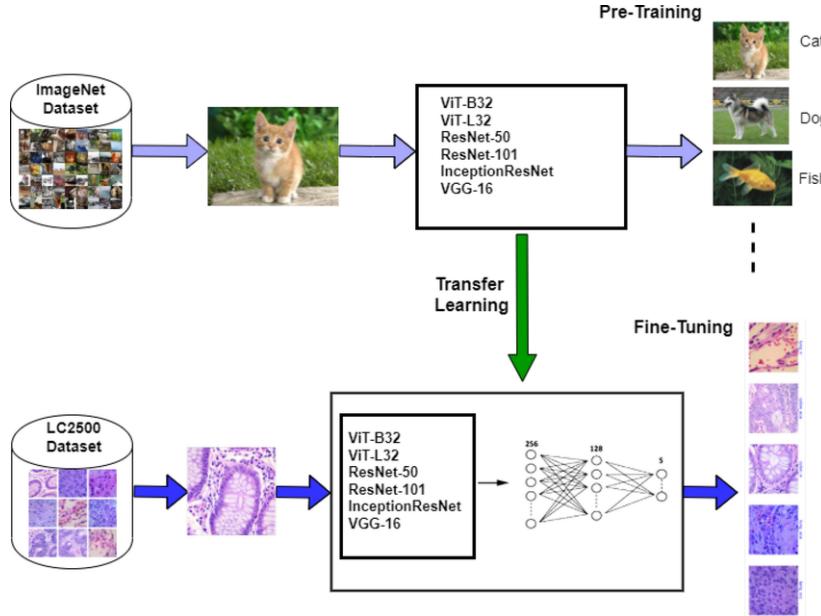


Figure FC2.2: Scheme for transfer learning of histopathology image classification

and 99.34% for binary classification and 90.66% to 93.81% for multi-class classification using the BreakHis dataset. The proposed CNN incorporates a small SE-ResNet module and a new learning rate scheduler, showcasing its potential for enhancing breast cancer diagnosis accuracy and efficiency.

The authors in [99] used a ResNet-50-based method for early cholangiocarcinoma detection using hyperspectral microscopy images. The dataset, annotated by pathologists, consists of 6,800 training and 210 testing images. After preprocessing, ResNet-50 achieves 82.4% accuracy in automatically classifying cancerous and non-cancerous choledoch tissue regions, demonstrating the potential for computer-aided diagnosis. This study [100] introduces a deep learning transfer learning approach for automated histopathology image classification. Using pre-trained models (DenseNet-161 and ResNet-50) on color and grayscale images, DenseNet-161 achieves 97.89% accuracy on grayscale, and ResNet-50 attains 98.87% on color images. These pre-trained models show promise for swift and accurate histopathology image classification, aiding pathologists in clinical tasks.

This study [101] investigates histopathological image classification using deep ResNet models, assessing the impact of network depth and transfer learning strategies. Findings show that deeper layers do not consistently lead to better performance, and transfer learning with deep ResNet outperforms traditional methods. Fine-tuning all layers is more effective than freezing most layers. The proposed model in [102], based on an enhanced XceptionNet with a swish activation function and depthwise separable convolutions, exhibits improved classification accuracy compared to the original Xception and other architectures. Simulation results demonstrate the method’s superior accuracy compared to state-of-the-art skin cancer diagnosis solutions. This study [103] explores computer-assisted pathology analysis, emphasizing the efficacy of the pre-trained Xception model in magnification-dependent breast cancer histopathological image classification. The Xception model and SVM classifier achieve consistently high performance, with accuracies of 96.25%, 96.25%, 95.74%, and 94.11% at magnifications of 40X, 100X, 200X, and 400X, respectively.

In their paper [104], the authors introduce DIRXNet, a highly effective CAD system for classifying breast cancer histopathology images. By utilizing preprocessing techniques and combining three pre-trained neural networks, DIRXNet achieves exceptional accuracy rates of 98.12%, 96.15%, and 94.8% on the BreakHis, PatchCamelyon, and BACH datasets. The authors in [105] explore how H&E stain normalization affects deep learning (DL) models in cancer image classification. Assessing VGG19, VGG16, ResNet50, MobileNet, Xception, and InceptionV3 on an H&E-stained cancer image dataset, VGG16 performs strongly, while VGG19 and ResNet50 show limitations. Stain normalization significantly improves less complex models like MobileNet and Xception, making them efficient alternatives. In a study presented in [106], the authors propose a novel fine-tuning approach for analyzing histopathological images, which differs from traditional methods. They use the Inception-v3 architecture and intra-domain fine-tuning, which involves training the model on six histopathological

source datasets and four target sets. The study highlights that pre-trained histopathological models perform better than those trained on ImageNet.

## 2.4 Summary of Limitations in the Discussed Papers

Histopathology analysis requires using particular datasets, which often need more diversity in staining techniques, tissue types, and disease presentations [107]. As a result, models trained on such homogeneous datasets may need help in generalizing effectively to a broader range of real-world scenarios. This limitation could hinder their practical applicability in various clinical settings.

Extracting meaningful features from histopathological images is a significant challenge, whether using handcrafted methods or convolutional neural networks [108]. Using inefficient feature extraction methods may compromise the accurate representation of intricate tissue structures, thereby leading to suboptimal diagnostic performance of the models. Hence, developing efficient and effective feature extraction techniques for histopathological image analysis is crucial. Several studies have been conducted in this area, exploring techniques such as transfer learning, multi-scale analysis, and attention-based mechanisms [108]. The summary of studies along with advantages and disadvantages is presented in Table TC2.1.

Table TC2.1: Summary of Algorithms for Histopathology Feature Extraction

Algorithm	Framework	Advantages	Limitations	References
Gabor Filter	Conventional	Identifying relevant regions	Limited feature representation, High cost	[55]
Homogeneity Measure	Conventional	Analyzing tissue texture	Limited feature representation, High cost	[56]
Hierarchical Algorithm	Conventional	Detecting cancerous regions	Limited feature representation, High cost	[58]
Paige Prostate	AI-based	High sensitivity	Overfitting, Data bias, Regulatory concerns	[52]
DenseNet-161	Deep Learning	Accurate necrosis detection	Overfitting, Data bias, Regulatory concerns	[72]
Mask R-CNN	Deep Learning	Hepatic necrosis detection	Overfitting, Data bias, Regulatory concerns	[73]
Brain Tumor Classifier	Deep Learning	Accurate tumor classification	Overfitting, Data bias, Regulatory concerns	[75]
Peritumoral Edema Prediction	Deep Learning	Forecasting edema after GKS	Overfitting, Data bias, Regulatory concerns	[76]
Tumor Pathology Applications	Deep Learning	Various pathology tasks	Overfitting, Data bias, Regulatory concerns	[77]
Domain-Adversarial Training	Domain Feature	Invariant image representations	Overfitting, Data bias, Regulatory concerns	[82]
Domain Shift Analysis	Domain Feature	Effects of domain shift	Overfitting, Data bias, Regulatory concerns	[81]
Medical Knowledge Integration	Domain Feature	Incorporating domain knowledge	Overfitting, Data bias, Regulatory concerns	[83]

Histopathological specimens inherently exhibit variability in cellular morphology

and tissue structure, which could impact the accuracy of deep-learning models that fail to account for this variability across different patients and diverse pathological conditions [109]. Furthermore, histopathological slides undergo diverse staining procedures, leading to variations in color and contrast, which can limit the applicability of deep learning models trained on specific staining protocols to unstained or differently stained slides, as well as across various laboratories and settings [109].

Deep Learning models are gaining popularity for predicting outcomes in clinical settings. However, these models often lack interpretability, which makes it difficult for healthcare professionals to trust their predictions. As transparency is essential for gaining the confidence of healthcare professionals in clinical settings, the limited interpretability and explainability of deep learning models can hinder their acceptance. Therefore, it is crucial to keep developing new methods to enhance the interpretability and explainability of deep learning models in healthcare [110] [111].

Histopathological analysis has been used to detect necrosis, but it has faced technical challenges due to the reliance on handcrafted and Convolutional Neural Network features. Handcrafted features depend on predetermined parameters and may struggle to detect subtle patterns that indicate necrotic regions, especially concerning intricate local features such as shapes and sizes. Moreover, developing robust features requires specialized expertise and incurs high development costs, which limits their adaptability to the nuanced nature of necrosis detection. Although CNN features are powerful, they have limitations in comprehending complex global features of tissues, making them prone to overfitting. This susceptibility necessitates extensive labeled data, which may be challenging to obtain in the context of necrosis with its varied manifestations. The complexity of high-dimensional histopathological data and the diverse conditions in which necrosis appears further complicate the manual adaptation of features, hindering the scalability of these approaches.

## 2.5 Different Features in Histopathology of Brain Tumor

Diagnosing disease or its grading in histopathology images involves identifying histological structures like cancer cell nuclei, glands, etc. [112]. The histological criteria indicating a more aggressive behavior and, thus, an increase in the malignancy grade include frequent mitoses, regions of hypercellularity, sheet-like growth, high nuclear–cytoplasmic ratio, prominent nucleoli, and spontaneous necrosis [112]. The other morphological appearance of these structures, like size, shape, and color intensity, are also essential factors for the presence of meningioma [113].

Some of the important histological structures that pathologists look for include:

- **Mitosis** - It is an individual cell characteristic. Atypical mitosis is present in a person with a malignant tumor. It indicates a region with very active cell division and multiplication. A High Power Field (HPF) is the portion of a slide visible under the high magnification of the microscope. If the number of mitosis is less than 4 per High Power Field, then it is Grade I Meningioma; if it is greater than 4 per HPF, then it is Grade II Meningioma [114]. Atypical (WHO grade II) meningiomas exhibit 4 or more but fewer than 20 mitoses per HPF, and anaplastic (WHO grade III) meningiomas show high mitotic activity with 20 or more mitoses per 10 HPF [115].

Perry et al. described the mitotic index (MI) as the sum of MFs per 10 consecutive HPFs in the area of the highest mitotic activity and reported that it is the most reliable histological predictor of the likelihood of meningioma recurrence [115]. Figure FC3.2 provides an example of mitosis. Meningiomas are usually graded based on the number of mitotic counts per unit area on H&E sections. However, this approach may be complicated by tumor cellularity. The median mitotic indices for WHO grade I, II, and III malignancies are 1, 4, and 12, respectively [114].

- **Necrosis** – This region lacks nuclei and is considered dead. It appears as a small dust-like structure with broken-down cells and fewer connective tissues. There is also a larger presence of cell debris. Necrosis is a dynamic process that results in gradual texture changes. Different phases of necrosis have varying colors and textures. When necrosis occurs in tissues affected by malignancies, it is referred to as tumor necrosis. Tumor necrosis is a common feature of solid tumors due to chronic ischemic injury caused by rapid tumor growth. Depending on the extent of its presence in the tumor area, tumor necrosis can be classified as focal, moderate, or extensive [116]. An example of necrosis is shown in Figure FC2.3.
- **Whorls** – Whorls are lobules of cells separated by a variable amount of connective tissue. It is a typical characteristic of grade I tumors. With increasing grades, it loses benign character. Whorls (arrows) are well-developed, and tumor cells are more elongated. The image of Whorls is shown in Figure FC2.3.
- **Small Cell Change** – Small cell changes are challenging to interpret in whorled and hypercellular areas. Small cell change occurs when tumor cells show increased nuclear to cytoplasmic ratio (lymphocyte-like morphology) [117]. The image of Small Cell Change is shown in Figure FC2.3.
- **Sheeting Pattern** - Sheeting is defined as uninterrupted pattern-less or sheet-like growth (loss of whorled or fascicular growth pattern). In the sheeting pattern, we primarily observe a loss in architecture and disorganization of character. Meningothelial meningioma grows in syncytial sheets and may simulate a sheeting pattern [118]. Detection of sheeting is challenging since, on occasion, it coexists with other processes, such as inflammatory infiltrates, blood vessels, or regions with syncytial architecture mimicking sheet-like growth [118]. The image of Sheeting Pattern is shown in Figure FC2.3.
- **Crowding of Cells** – This is generally observed in tumors of a higher grade. A

higher rate of growth is manifested as crowding of cells. In humans, this atypical feature may imply a more aggressive clinical course [119].

- **Nucleoli** – In Nucleoli, the cytoplasm ratio will be high in Grade II Meningioma. Macronucleoli is considered to be present when it is observable under 10x objective [117]. Grade 2 lesions are atypical lesions that include three or more of the following characteristics: high cellularity; conspicuous nucleoli (i.e., easily visible with a 10x objective); necrosis; patternless sheets of tumor cells; and tumor cells with little cytoplasm compared to nuclear size (small cell changes) [120]. The image of Nucleoli is shown in Figure FC2.3.
- **Fascicles** The loss of standard internal architecture seen in atypical meningiomas, such as the absence of whorls and fascicles, results in less directionality in atypical meningiomas than benign meningiomas [121]. The image of Fascicles is shown in Figure FC2.3.

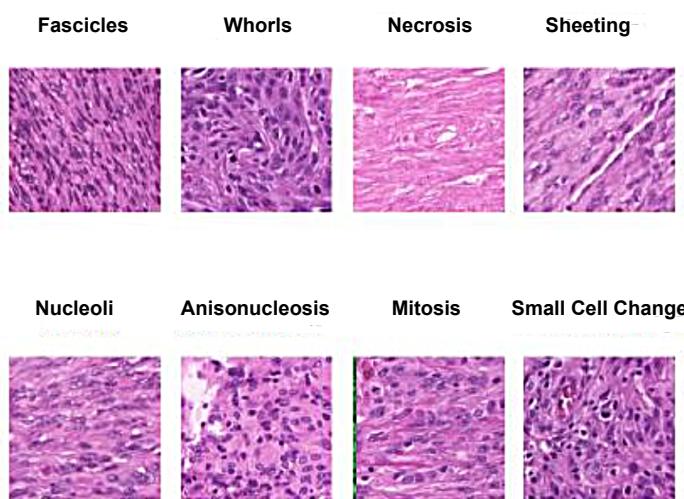


Figure FC2.3: Visualizing Necrosis and Other Annotations on ImageScope

## 2.6 Summary

Pathologist's expertise is crucial in addressing the inefficiencies in medical imaging analysis. Pathologist's in-depth knowledge of tissue morphology, cellular structures, and disease patterns can help improve the design of deep-learning models, resulting in more accurate feature extraction. By incorporating domain knowledge, the models become more transparent and comprehensible to pathologists, improving their interpretation of the results. Moreover, domain knowledge can help improve the models' generalizability across different imaging conditions and tissue types, dealing with data bias and overfitting issues. Thus, combining deep learning with domain knowledge could produce models that are robust, clinically relevant, and applicable in real-world healthcare scenarios.

## CHAPTER 3

### DATA DESCRIPTION

#### 3.1 NIMHANS Data Acquisition

An initial set of slides were analyzed and additional slides were provided by NIMHANS to improve the variety of data. A total of 11 slides were collected, which were seen as sufficient to train and classify meningioma. These slides are valuable resources for studying and understanding various brain disorders and conditions, such as necrosis. Ten of the slides are of meningioma, and the other is of GBM. Nine of the ten meningioma slides have been annotated and verified by expert pathologist's at NIMHANS. The slides provided by NIMHANS were shared in accordance with NIMHANS' policy on the use of patient data for research purposes.

#### 3.2 WSI Annotations Generation

WSI transforms how we study tissue samples by capturing high-definition digital images of the entire slide, which we can view on a computer. In contrast to traditional microscopy, WSI represents a technological advancement in histopathology. Rather than examining tissue samples on individual glass slides, WSI utilizes digital scanning technology to capture high-resolution images of entire slides. Digital slides are usually

larger in file size and pixel dimensions when compared to traditional microscopy slides. A typical WSI can have a size of 50000 x 50000 or even higher. The larger size of WSI images allows for a more detailed examination and analysis of the specimen at different magnifications.

A range of tools are available to create annotations for WSI. A tool called ImageScope, developed by Aperio, is a software program used in pathology to view, analyze, and annotate digital images of tissue samples. Pathologists and researchers can utilize ImageScope's many features, such as zooming in on high-resolution images, adjusting brightness and contrast, and performing measurements and annotations, to make it a powerful tool. The Figure FC3.1 shows an example of annotations created using Image Scope.

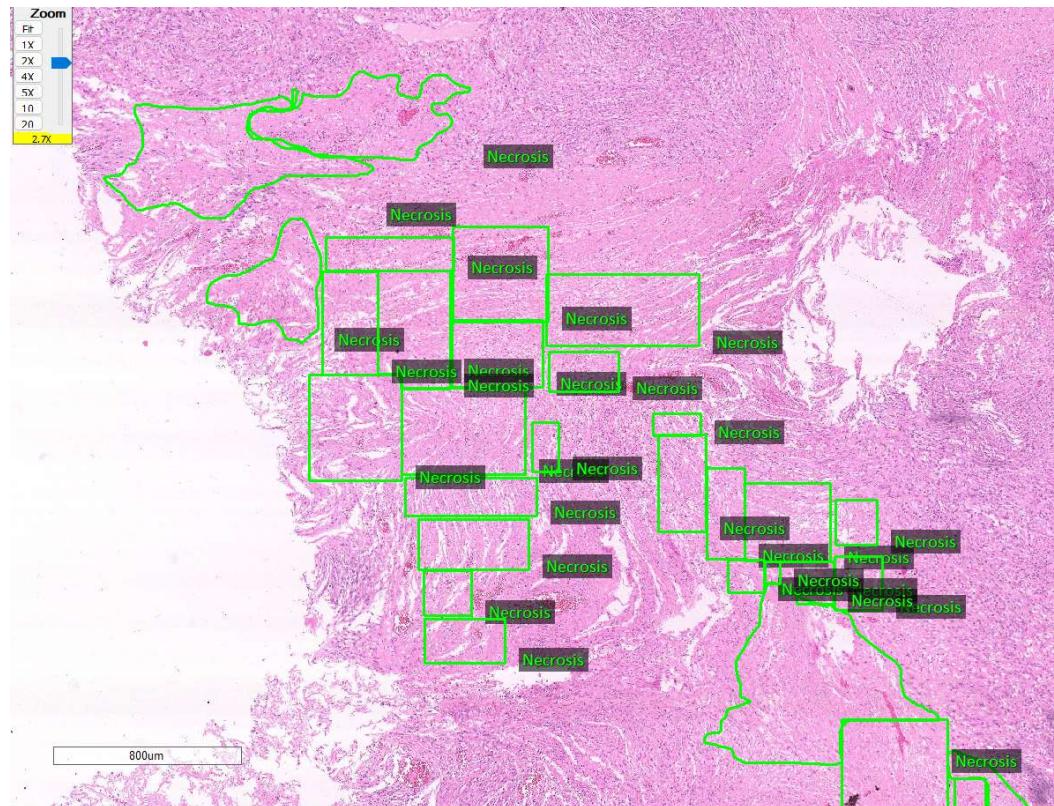


Figure FC3.1: Visualizing Necrosis and Other Annotations on ImageScope

We used Image Scope to view slides in our project at different levels of magnifi-

cations. This tool allowed us to create detailed annotations and accurately calculate the area of specific regions within the slides. Additionally, we were able to use Image Scope to view predictions at the WSI level, enabling us to make informed decisions and improve the accuracy of our analysis.

The steps involved in creating annotations on WSI using ImageScope are as follows:

1. Open the digital image of the tissue sample in ImageScope.
2. Use the zoom tool to adjust the magnification of the image to the desired level.
3. Select the annotation tool from the toolbar to draw shapes like rectangles, circles, and polygons.
4. Draw the shape around the region of interest in the image to annotate.
5. Add a label to the annotation, providing information about the tissue sample, such as the cell type or disease state.
6. Save the annotation to the image file or export it separately in xml format.

We have collected nine WSI's of meningioma: Sample\_001, Sample\_002, Sample\_003, Sample\_004, Sample\_005, Sample\_006, Sample\_007, Sample\_008, and Sample\_009. One or more expert pathologists from the Neuropathology Lab at NIMHANS have carefully examined and marked various pathological features.

Table TC3.1: Available annotations in Meningioma slides

Name of the Slide	Labels_Covered
Sample_001	Mitosis,MicroNecrosis
Sample_002	Fascicles,Whorls,Sheeting,Nucleoli,Mitosis,Necrosis,PartialNecrosis,RBC
Sample_003	Necrosis,NonNecrosis,Stroma,PartialNecrosis,EarlyNecrosis
Sample_004	Fascicles,Whorls,Sheeting,Nucleoli,Anisonucleosis,Mitosis,Necrosis
Sample_005	Fascicles,Whorls,Sheeting,Nucleoli,Anisonucleosis,Mitosis,Smallcellchange
Sample_006	Whorls,Sheeting,Nucleoli,Anisonucleosis,Mitosis,Necrosis,MicroNecrosis
Sample_007	Necrosis,RBC
Sample_008	Fascicles,Nucleoli,SuspiciousMitosis,cellcrowding
Sample_009	Nucleoli,Anisonucleosis,Mitosis,SuspiciousMitosis

The set of sample slides we have used, Sample\_001 to Sample\_009, contain different combinations of these pathology features, some of which have been identified by the pathologists. The distribution of these features across the samples is shown in Table TC3.1. After careful observation by pathologists, Sample\_003 has been excluded from the analysis because it was deemed unsuitable by the pathologists. It should be noted that sample IDs are arbitrary and solely used to provide a unique identifier for each slide, so the numbers may not follow a sequential pattern.

### 3.3 Generating tiles from whole slide images

A WSI is usually a large image that can reach several gigapixels, with an average resolution of 50,000 x 50,000 pixels or even higher. This size can put a strain on computational resources, but by dividing the WSI into smaller tiles, it becomes more manageable.

Generating tiles is crucial in developing machine learning algorithms designed explicitly for WSI analysis. Dividing the image into smaller tiles makes it easier to load the data into memory and train and test machine learning models more efficiently. This approach significantly reduces the demand for computational resources, leading to more precise and efficient analysis of WSI data. Further, since most pathology analysis is per-

formed on small crucial regions of the slide (such as the visible field of a microscope and its neighboring regions), tile-based computational analysis can be potentially effective.

Several studies have investigated the optimal tile size for WSI processing. A survey by Magee et al. (2009) found that a tile size of 256x256 pixels provided the best results for image analysis tasks such as tissue segmentation and cell counting [122]. Similarly, a study by Janowczyk and Madabhushi (2016) found that a tile size of 256x256 pixels provided the best trade-off between computational efficiency and image quality for deep learning-based analysis of WSI [123]. Using tiles of size 256x256 pixels to process WSI is efficient [123]. This size provides a good balance between computational efficiency and image quality. A smaller tile size may result in a large number of tiles, leading to longer processing times and increased storage requirements. On the other hand, a larger tile size may result in loss of image quality due to down-sampling. Hence, for our studies, we have generated tiles of 256x256 pixels.

In our study, two approaches were used for generating labelled tiles:

- As described above, pathologists used slide visualisation software (in our case, ImageScope) to annotate features of the WSI, and save these labelled polygonal regions in XML files. A grid of uniformly sized tiles of the required size is extracted from these images, and their labels are determined by comparing them to the annotation polygons. The following slides have slide-level annotations: Sample 1, Sample 6, Sample 7, Sample 8, and Sample 9. This is described in detail in the below sections.
- Tiles of appropriate size were first generated from the WSI, and subsets of these were presented to the pathologist using slide visualisation software (in our case, ImageScope) that allowed them to view the tile and its neighbors. The tools allowed them to assign one or more pre-defined labels to each tile. These labels

were saved to a csv file, along with the tile name and assigned label(s). Such labelled tiles can be directly added to the dataset to be used for machine learning. The following slides have tile-level annotations: Sample 2, Sample 4, and Sample 5.

### 3.4 Generating labelled tiles from annotated whole slide images

Our study uses non-overlapping tiles of  $256 \times 256$  pixels extracted from the WSI. This is done by first superimposing a uniform grid on the WSI, and using functionality of the Python OpenSlide library to extract tiles corresponding to the grid cells.

**OpenSlide:** The OpenSlide [124] tool provides a comprehensive set of functions and methods within its software library, offering a standardized interface that enables users to efficiently access, navigate, and manipulate high-resolution, multi-resolution WSIs in diverse digital pathology formats, encompassing Aperio, Hamamatsu, Leica, Sakura, and other commonly utilized file formats.

The OpenSlide API provides systematic access to multi-resolution visual data using pyramid levels [125] as shown in Figure FC3.2. An organized list of pyramid levels describes a digital slide. The highest-resolution level is level 0, and each successive level is a down-sampled category of the preceding level. The Openslide library also executes other functionalities like image scaling. However, the only levels accessible to developers via the API are those accumulated in the slide file.

The internal constructors and function pointers associate the software program with the vendor's code. When a whole slide image opens initially, the vendor-specific driver program interprets and translates the slide's metadata and locates the compact image data for each region in the whole slide image.

ImageScope and OpenSlide use different coordinate systems to identify pixel loca-

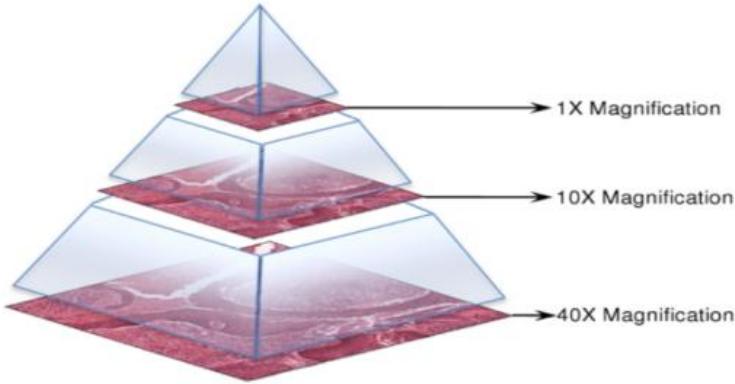


Figure FC3.2: Different levels of magnification in histopathology images

tions on the image. Pixel coordinates in one system must be transformed through a rotation and offset transform to map to the coordinates in the other system. The equations Eqn 3.1 and Eqn 3.2 shows transformation.

$$\text{Actual}_x = \text{WSI}_{ht} - | \text{Tile}_y - \text{Bounds}_y | \quad (\text{Eqn 3.1})$$

$$\text{Actual}_y = | \text{Tile}_x - \text{Bounds}_x | \quad (\text{Eqn 3.2})$$

Where  $\text{Actual}_x$  and  $\text{Actual}_y$  represent ImageScope pixel coordinates,  $\text{Tile}_x$  and  $\text{Tile}_y$  represent pixel coordinates in OpenSlide space, and  $\text{Bounds}_x$ , and  $\text{Bounds}_y$  are image parameters extracted by OpenSlide from the SCN files.  $\text{WSI}_{ht}$ (height) is the height of the Whole Slide Image.

These equations allow us to map tiles between the two platforms, ensuring they are correctly placed back to their original location in ImageScope. This process guarantees the integrity and correctness of the transferred data.

### 3.4.1 Processing tiles

The patches extracted from the Openslide Library are in the RGB channel [126]. We index into the loaded Openslide instance and use the grid to retrieve the patches (or tiles in Openslide). The retrieved tile is then turned into a  $256 \times 256 \times 3$  Numpy array [126]. These tiles are then processed as follows

1. Tiles with mostly white space and minor artifacts are removed during pre-processing, with human intervention and verification by pathologists at NIMHANS.
2. The remaining tiles are classified based on their biological features, such as necrosis, mitosis, etc.
3. Labeled tiles obtained from both annotation-based and tile-based processing are merged to create a comprehensive input dataset.
4. The generated tiles are manually viewed to identify those where labelling could be ambiguous, for example tiles on the boundary of annotation regions or where there are certain tissue artefacts.

### 3.4.2 Removing Non-Tissue Areas

To identify relevant tissue, we used a threshold-based segmentation approach that automatically segregates the tissue from the background white spaces in all slide images while determining background regions [127]. Subsampling of patches was carried out to ensure that the pipeline only received patches primarily composed of tissue. These preprocessing steps allowed us to save computation time and ensure that we only focused on sections of the slide that were most likely to have necrotic regions while using good-quality patches as input for the training pipeline.

The process of filtering out non-tissue tiles involves a few necessary steps, which are described below:

1. Convert the image from RGB to HSV color space. The OpenCV COLOR\_RGB2HSV function is used to convert . OpenCV utilizes mathematical formulas to perform the conversion from RGB to HSV color space.
2. Define a mask and determine the optimal lower and upper thresholds through experimentation. After analyzing the data, the lower threshold value is set to 200 based on saturation and hue.
3. Calculate the number of pixels with values below 200, referred to as non-zero pixels. This is carried out using coutning all pixels with its value if the value is below 200 then it is counted and referred as non-zero pixel.
4. The tile will be discarded if the percentage of non-zero pixels is below 5%. This is calculated by dividing the number of non-zero pixels from point 3 to the total number of pixels.

### **RGB to HSV Conversion Process:**

The following steps outline the mathematical process for converting RGB values to HSV as uses in point 1 of section 3.5.2.

First, normalize the RGB values:

$$R' = \frac{R}{255}, \quad G' = \frac{G}{255}, \quad B' = \frac{B}{255}$$

Calculate the Value (V):

$$V = \max(R', G', B')$$

Calculate the Saturation (S):

$$S = \begin{cases} 0 & \text{if } V = 0 \\ \frac{V - \min(R', G', B')}{V} & \text{otherwise} \end{cases}$$

Calculate the Hue (H):

$$H = \begin{cases} 60^\circ \times \left(0 + \frac{G' - B'}{V - \min(R', G', B')} \right) & \text{if } V = R' \\ 60^\circ \times \left(2 + \frac{B' - R'}{V - \min(R', G', B')} \right) & \text{if } V = G' \\ 60^\circ \times \left(4 + \frac{R' - G'}{V - \min(R', G', B')} \right) & \text{if } V = B' \end{cases}$$

If  $H$  is negative, add 360 degrees:

$$H = H + 360^\circ$$

Normalize the Hue:

$$H = \frac{H}{360^\circ}$$

### 3.4.3 Classifying tiles against annotations

For the slides that were annotated at the WSI level, the next step is the classification of each WSI tile against the WSI's polygonal annotated regions. Since these polygons do not necessarily follow grid cell or tile boundaries, additional computation is needed to check if a tile is inside, outside, or crosses the boundary of an annotated region.

To efficiently compute the spatial relationship between a tile and an annotation poly-

gon, we use Shapely [128], the Python package for set-theoretic analysis and manipulation of planar features, employing methods from the well-known and extensively used GEOS library [129] (through Python’s types module). This package provides efficient implementations of operations on polygons, such as computing if a point is contained in a polygon or computing the intersection area of two polygons. For the label generation step, each tile (a rectangle) was classified relative to each annotation region (polygon) and labeled based on the extent of overlap with the polygon:

- If the tile is entirely contained within the polygon, we classify it as ”in” the polygon with an area of overlap of 100%.
- If the tile does not overlap or intersect with the polygon area, we classify it as ”out” of the polygon, and it has 0% overlap with the polygon.
- When a tile intersects the boundary of the polygon, it also needs to intersect with the interior of the polygon. We consider such tiles to be ”on” the polygon’s boundary and calculate the overlap of their interiors to determine the area of intersection. Since the annotation has some uncertainty at the boundaries of necrotic regions, the tiles on the boundary were grouped into 10 ”buckets” based on percentage area of overlap, i.e. [0-10], [10-20].... [90-100] among them.

#### **3.4.4 Border Tile Filtering Technique**

In section 3.5.2, we explained a technique for eliminating non-tissue regions. The process involves removing all tiles that might be considered white space by manipulating tiles at the pixel level. In this section, we concentrated on eliminating border tiles by analyzing groups of pixels at once, and we are referring to that group of pixels as a block. While extracting patches from WSIs, some generated tiles contain sections of the slide’s borders, referred to as border tiles. These regions of the border tiles may

contain non-representative information such as empty spaces near border of WSI, causing noise and bias in the machine-learning models trained on such data. Including these border tiles in the training dataset is problematic for digital pathology models. It results in noisy data that may make it hard for the models to generalize correctly. This leads to identifying patterns associated with the border regions rather than the tissue structures of interest. This can negatively impact the models' accuracy, robustness, and efficiency in performing subsequent image analysis tasks. The Figure FC3.3 shows an example case of an image with large area of white regions

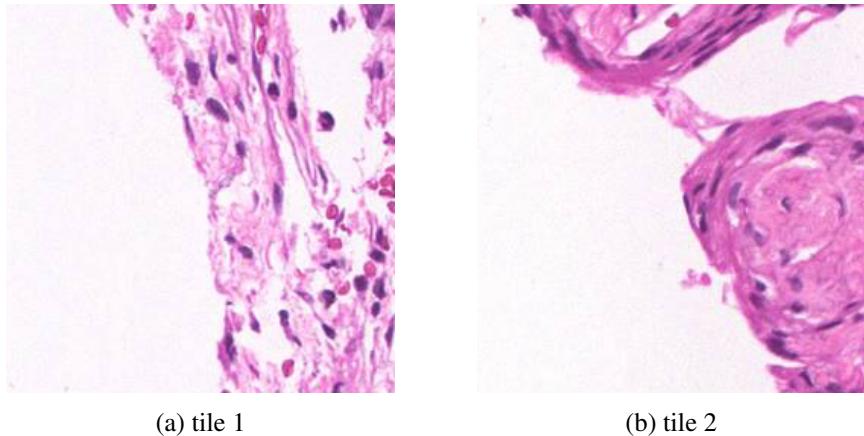


Figure FC3.3: Visualization of images with significant white space regions at border

Excluding border tiles during pre-processing optimizes machine learning models by reducing noise and biases, improving generalization capabilities and efficiency, and focusing on learning meaningful tissue features. This is critical for optimizing performance and robustness in digital pathology image analysis. A straightforward approach for eliminating border tiles involves tallying white pixels across the entire image and applying a threshold for removal. However, this method poses a risk of discarding tiles with dispersed white spaces, potentially undermining our focus on the complete tissue region.

In order to mitigate any adverse effects caused by border-tiles, we have implemented a pixel-level manipulation method to remove them from our pool of images. The pro-

cedure comprises a few essential steps:

1. convert it into grayscale
2. For pixel analysis, choose a window block size of 16 for rolling on the image.  
Selecting a window block size of 16 allows for a fine-grained analysis of the image, capturing local variations in pixel intensity
3. Iterate the same process through the entire tile with a window size of 16. In each iteration, calculate the number of non-zero pixels with a pixel value greater than 240 to identify WSI border tiles.
4. The non-white pixel percentage is calculated by dividing the count of non-zero pixels by 65536 (256 multiplied by 256), allowing for standardized comparisons across images and datasets.
5. The tile is discarded if non-white pixels exceed a certain percentage threshold. We have determined that a threshold of 60% non-white pixels provides a good balance between retaining informative tissue regions and filtering out less relevant border tiles.

### 3.5 Data Preparation

We used Sample 1, Sample 2, Sample 4, Sample 5, Sample 6, Sample 7, Sample 8, and Sample 9 to generate labeled tiles during the tile generation process. Our primary focus was on the necrosis label. In contrast, all other labels, such as Fascicles, Whorls, Sheeting, Nucleoli, Anisonucleosis, RBC, Mitosis, and Small cell change, were grouped as single-label non-necrosis. Some slides were labeled as necrosis only, others were labeled non-necrosis only, and some had both labels.

Table TC3.2 provides detailed tile-level information for each slide in meningioma derived after removing border tiles as described in section 3.1.4.

Table TC3.2: Distribution of Necrosis and Non-Necrosis Images across Samples

Sample ID	Total Tiles	Necrosis	Non-Necrosis
Sample_001	17	0	17
Sample_002	4089	338	3751
Sample_004	674	72	602
Sample_005	918	0	918
Sample_006	224	76	148
Sample_007	2983	2983	0
Sample_008	111	0	111
Sample_009	23	0	23
Total	9039	3469	5570

**Sample ID:** Up until this point, we have been referring to the slide name as Sample\_xxx, where xxx represents the slide number. Moving forward, we will only use the last xx in xxx to indicate the slide number. For instance, when we use slide 01, we refer Sample\_001.

**Batch:** A batch is a subset of the dataset created from a pool of tiles from all slides based on certain conditions. Each batch has training and testing tiles with necrosis and non-necrosis classes.

To improve the accuracy and reliability of our machine learning model for neuropathological image analysis, we create datasets that mimic real-world situations and challenges commonly faced in clinical practice. Our main objective is to thoroughly evaluate the model's effectiveness by training it on diverse datasets. We are also preparing our data for inter-slide and intra-slide analysis to more comprehensively assess the model's performance.

In intra-slide preparations, data from the same slide is used for training and testing. The goal is to capture the variation that may occur in the classes within a slide, such as the presence of necrotic and non-necrotic regions in different batches. This way,

the model is exposed to diverse scenarios, enhancing its ability to accurately generalize and classify necrotic and non-necrotic areas. Table TC3.3 displays the slide IDs for the distribution of Necrosis and Non-Necrosis across both training and testing for Inter-class.

Table TC3.3: Batchwise distribution of slide samples (sample IDs) for intra-slide analysis

<b>Batch</b>	<b>Train Necrosis</b>	<b>Train Non-Necrosis</b>	<b>Test Necrosis</b>	<b>Test Non-Necrosis</b>
6	02(50%),07,06,04	02,04(50%),05,06,08,09	02(50%)	04(50%)
7	04(50%),07,06,02	02,01,04(50%),06,08,09	04(50%)	04(50%)
8	06(50%),07,04,02	02,04,05,06,08(50%),01	06(50%)	08(50%)
9	02(50%),07,04,06	02,04,05(50%),06,09,01,08	02(50%)	05(50%)

Similarly, in inter-slide preparations, the model encounters previously unseen testing samples, and the goal is to evaluate its adaptability to different slides with varying colors and structural orientations. This approach tests the model's robustness and explores potential inter-slide dependencies, which is crucial for reliable predictions in real-world applications. Table TC3.4 shows the distribution of Necrosis and Non-Necrosis across both training and testing for inter-slide, with the corresponding slide IDs.

Table TC3.4: Batchwise distribution of slide samples (sample IDs) for inter-slide analysis

<b>Batch</b>	<b>Train Necrosis</b>	<b>Train Non-Necrosis</b>	<b>Test Necrosis</b>	<b>Test Non-Necrosis</b>
1	02,04,07	01,02,04,05,08,09	06	06
2	06,04,07	01,04,05,08,09,06	02	02
3	06,02,07	01,02,05,08,09,06	04	04
4	06,02,04	04,02,05,08,09,06	07	01
5	06,02,07	01,02,05,09,06	04	08

The Tables TC3.3 and TC3.4 illustrate the variation in class distribution across different batches for intra-slide and inter-slide distributions in training and testing data. These preparation strategies ensure that the model is well-equipped to handle the intricacies of neuropathological images. This way, the model's reliability and effectiveness in clinical applications increase.

### 3.6 Brain Tumors - TCGA GBM and LGG

For experimentation purposes, the meningioma-trained model was tested to validate its ability to detect necrosis in glioma slides. A large and popular whole slide image (WSI) dataset, namely TCGA, was chosen for this purpose. The TCGA Program [130], consisting of 132 projects, has a diverse collection of Whole Slide Images from different human organs of both males and females. This program contains various data types, including diagnoses and tissue slides with slide-level labeling. However, there were no region-specific annotations or labeling provided. For the analysis of brain tumors, diagnostic slides from two types of brain tumors - GBM and LGG - were carefully selected. The slide numbers are presented sequentially, but they have been assigned arbitrarily for organizational purposes.

Four slides from GBM (Sample 21, Sample 22, Sample 23, and Sample 24) and three from LGG (Sample 25, Sample 26, and Sample 27) were chosen for analysis. A subset of tiles from each slide of GBM and LGG from TCGA has been labeled by pathologists in the Neuropathology lab at NIMHANS. Table TC3.5 shows the necrosis and non-necrosis distributions for each sample ID.

Table TC3.5: Necrosis and Non-Necrosis tile count for TCGA GBM and LGG Samples

Sample ID	Sample	Necrosis	Non-Necrosis
Sample_021	GBM	27	64
Sample_022	GBM	125	34
Sample_023	GBM	52	45
Sample_024	GBM	65	12
Sample_025	LGG	0	156
Sample_026	LGG	0	285
Sample_027	LGG	0	35

### 3.7 Summary

This chapter provides a detailed account of the data used in the study. The dataset comprises Whole Slide Images of brain tissue slides from two sources: the NIMHANS and TCGA datasets.

The NIMHANS dataset encompasses 11 brain tissue slides, out of which ten slides correspond to meningioma. The remaining slide corresponds to GBM, a more aggressive variant of brain tumor. Expert pathologists have carefully annotated nine meningioma slides with information about the tissue.

On the other hand, the TCGA dataset comprises seven diagnostic slides extracted from glioma cases, four of which are GBM, while the remaining three are LGG.

For comprehensive evaluation and robustness across different datasets, the planned models will undergo validation against all batches in the upcoming chapter.

## CHAPTER 4

### ADDRESSING THE CHALLENGES OF NECROSIS

### DETECTION IN HISTOPATHOLOGY

#### **4.1 Problem statement**

Detecting necrosis plays a crucial role in diagnosing medical conditions such as cancer, heart disease, and stroke and contributes significantly to fundamental research on cell death and disease. However, identifying necrosis computationally within viable tumor cells poses a substantial challenge, particularly when employing different feature extraction methods.

Creating a robust necrosis detection model becomes complicated due to limited data and varying distributions, especially in non-necrotic regions where diverse cell types and limited information may be present. Misclassifying non-necrotic regions as necrotic can compromise accuracy. Therefore, a nuanced understanding of cell shape and size beyond conventional features is necessary.

Detecting necrosis is a significant hurdle due to the limited generalizability of detection models across different types of necrosis and slides. Innovative approaches are required to ensure adaptability and effectiveness in diverse pathological contexts. A crucial challenge that needs to be addressed is grasping the structural organization of necrotic regions. Traditional methods and even features derived from Convolutional

Neural Networks have limitations, emphasizing global or local features. Overcoming these challenges requires an efficient feature extraction approach integrating domain-specific and non-domain features to guide accurate necrosis pattern identification.

To address these challenges, we aim to develop a robust necrosis detection model that performs well in challenging scenarios. The approach involves incorporating domain features to enhance accuracy. We have curated a comprehensive dataset focused on Meningioma brain tumors to facilitate algorithm development and testing. The dataset encompasses inter-slide and intra-slide samples, providing a rich variety for training and evaluation. This initiative aims to contribute to the advancement of accurate and adaptable necrosis detection in medical imaging.

In the evaluation phase, we use a notation of values between 0 and 1 to represent accuracy, precision, recall, and F1 score. For example, 0.256 represents 25.6%.

## 4.2 Baseline Model

In a past study [131], authors created a model to identify necrosis using handcrafted features. This model, proposed in [131], was tested on a specific set of samples from nine meningioma cases. Our study aims to check how well a similar model, using handcrafted features, performs on different data for inter-slide and intra-slide batches beyond the samples used in the prior study.

The SVM technique was chosen for this study based on the findings of [131]. The study builds on the work of [131] by focusing on feature selection while using the same models. Various models, such as Logistic Regression, Naive Bayes, and Decision Trees, were tested, but SVM was ultimately selected due to its superior performance. SVM outperformed the other models on the testing set and also achieved a higher recall score in k-fold cross-validation.

The previous study [131] uses a dataset with training, testing, and validation samples stratified from all nine meningioma slides for necrosis and non-necrosis. However, the current study uses the datasets as discussed in 3.5.1 and 3.5.2.

The features used in [131] consist of traditional color distribution, textures, and morphology, extracted using techniques like GLCM, histograms, and region-based measures. The support vector classification model is trained and validated using these features to classify new test samples.

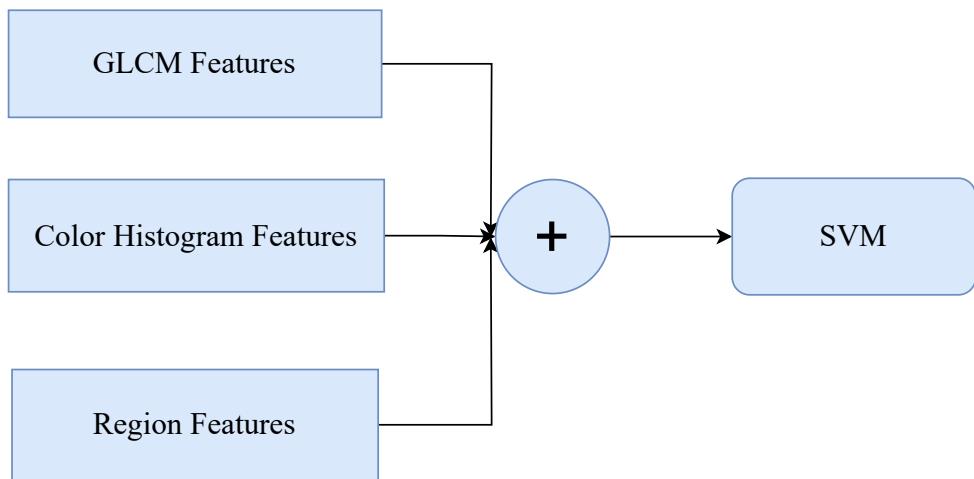


Figure FC4.1: Baseline Mode Architecture

This model for identifying necrotic and non-necrotic tiles, which we refer to as the baseline model, comprises two distinct parts: a feature extraction module and a machine learning model, as depicted in Figure FC4.1. In Figure FC4.4, the left side, depicted within a circle, corresponds to the feature extraction process, while the right side, also within a circle, signifies the machine learning phase. The '+' symbol acts as the separator between these two components.

**Feature Extraction Module:** This module takes an input image and extracts various features, including color histogram, texture, GLCM, and region-based features as described in Section 1.6.2.

**Machine Learning Model:** This module utilizes the extracted features to identify

necrotic and non-necrotic tiles by discerning patterns within these features. The supervised support vector classification model was trained with inter-slide and intra-slide, as mentioned in sections 3.5.1 and 3.5.2.

#### 4.2.1 Performance Evaluation on inter-slide and intra-slide data

For our baseline model, we implemented a supervised machine learning model that utilizes traditional image features and validated its performance against inter-slide and intra-slide datasets. We optimized hyperparameters using grid search to improve the performance of the conventional necrosis detection model. Our objective was to determine the optimal values for parameters related to three fundamental feature extraction techniques: histogram features, LBP, and texture. We adopted this robust optimization approach to find the parameter combination that maximized the model's performance and enhanced the discriminative power of the feature extraction techniques, leading to an optimal SVM model for accurate image classification.

We examined the impact of varying bin sizes and ranges for histogram features. The bin size determines the width of intervals in the histogram, while the range of bins defines the extent of pixel intensity values considered. For LBP, we focused on the number of points considered and the neighborhood size. These parameters influence the granularity and scale of texture patterns captured by the LBP operator.

We optimized the hyperparameters of the SVM classifier, which is essential in the final classification process. We conducted a grid search and explored parameters such as kernel choice, regularization strength ( $C$ ), and kernel coefficient ( $\gamma$ ). This systematic variation helped us identify an optimal combination of hyperparameters.

The necrosis detection metrics for the baseline model, presented in Table TC4.1, show that the model performs differently across different batches and sample types. In the inter-slide analysis (Batches 1 to 5), the model shows varying levels of accuracy,

Table TC4.1: Batch-wise Necrosis Detection Metrics for Baseline Model

<b>Batch</b>	<b>Sample Type</b>	<b>Necrosis</b>	<b>Non-Necrosis</b>	<b>Accuracy</b>	<b>Recall</b>	<b>Precision</b>	<b>F1 Score</b>
Batch 1	inter-slide	71	848	0.489	0.873	0.209	0.334
Batch 2	inter-slide	57	23	0.575	0.807	0.730	0.767
Batch 3	inter-slide	152	596	0.737	0.934	0.590	0.720
Batch 4	inter-slide	152	148	0.503	0.934	0.656	0.771
Batch 5	inter-slide	57	148	0.268	0.807	0.380	0.517
Batch 6	intra-slide	71	596	0.706	0.873	0.388	0.537
Batch 7	intra-slide	71	109	0.344	0.873	0.512	0.647
Batch 8	intra-slide	71	23	0.659	0.873	0.795	0.832
Batch 9	intra-slide	71	15	0.849	0.873	0.905	0.889

ranging from 26.8% to 73.7%. Batch 3 performs the best, with an accuracy of 73.7% and a recall of 93.4%. However, Batch 5 performs the worst with an accuracy of 26.8%, indicating difficulties in identifying necrotic regions for this specific batch.

On the other hand, in the intra-slide analysis (Batches 6 to 9), the model consistently achieves higher accuracy, ranging from 65.9% to 84.9%. Batch 9 performs the best with an significant accuracy of 84.9%, demonstrating the model's ability to detect necrosis within individual slides. Additionally, across all intra-slide batches, the model shows good precision, exceeding 79.5%, highlighting its capability to identify necrotic regions precisely. The lower accuracy of Batch 5 may be due to improper feature representation for necrosis and non-necrosis. A few tiles from the testing set of Batch 5 have high partial necrosis tiles.

#### 4.2.2 Insights and Observations

The method described in Section 1.6.4 was used to calculate weighted metrics for both inter-slide and intra-slide cases. The Table TC4.2 shows the weighted average metrics for each of inter-slide , inter-class and overall. The analysis of the classification results demonstrates that the baseline model has unique characteristics in distinguishing between necrotic and non-necrotic samples. The model can identify necrotic areas within the same slide with a high recall rate of 87.3%. However, the model has a lower

precision rate of 49%, indicating that it can mistakenly classify some non-necrotic regions as necrotic. The inter-slide evaluation further highlights this imbalance, indicating that the model recalls with high accuracy at 89.3% but exhibits a precision rate of only 42.9%. This implies a potential for misclassification across different slides.

Table TC4.2: Weighted Average Metrics Evaluation for the Baseline Model

<b>Sample Type</b>	<b>Accuracy</b>	<b>Recall</b>	<b>Precision</b>	<b>F1 Score</b>
intra-slide	0.650	0.873	0.490	0.626
inter-slide	0.557	0.893	0.429	0.579
overall	0.586	0.886	0.448	0.597

The baseline model shows robust sensitivity in detecting necrotic regions but suffers from imprecision in recognizing non-necrotic areas, leading to many false positives. As a result, the model's overall accuracy is only 58.6%, limiting its clinical applicability. Therefore, to achieve more reliable detection of tissue death in various pathological settings, further iterations of the model or the incorporation of additional features may be necessary to address these nuances and achieve a more balanced performance.

We analyzed the quality of feature representation by using t-SNE distributions in addition to quantitative results. We combined the features derived from color histograms, texture, and local binary patterns. The t-SNE model efficiently compressed the data into two components, enabling a detailed and insightful visualization. The scatter plot in Figure FC4.2 shows the t-SNE distribution of Necrosis and Non-Necrosis for a subset of the dataset.

The scatter plot shows that the extracted features cannot effectively distinguish between Necrosis and Non-Necrosis. There is an overlap in the t-SNE distribution for both classes, indicating the need for further exploration and refinement in feature engineering or the consideration of alternative feature sets to improve the discriminative power of the model. Addressing these aspects will be essential to enhance the overall efficacy of the feature representation in capturing the nuances of necrosis detection.

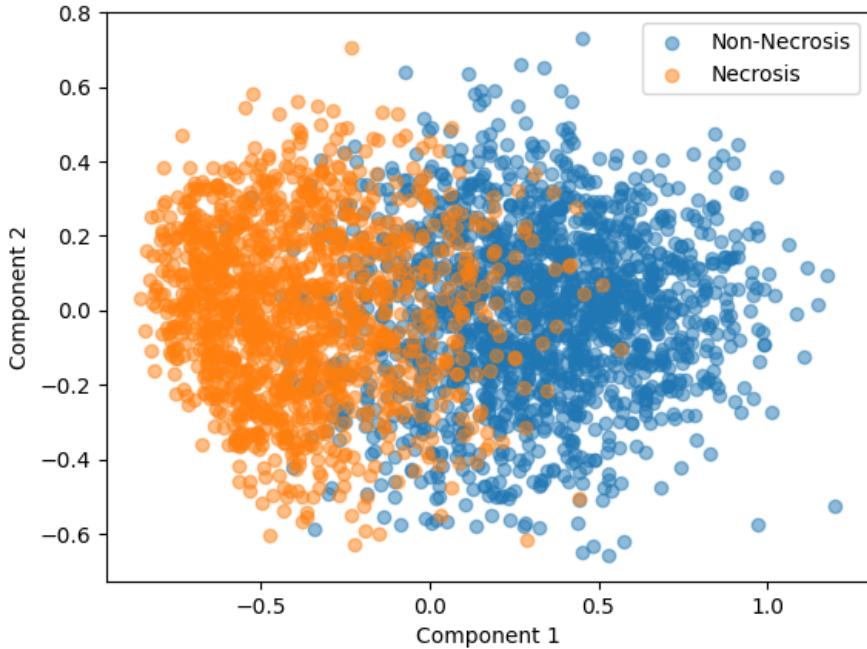


Figure FC4.2: t-SNE Feature Visualization for Handcrafted Features

### 4.3 Fully Connected XceptionNet

We explored the use of Convolutional Neural Networks to efficiently extract crucial features for our necrosis detection model. The XceptionNet model is a Convolutional Neural Network pre-trained on the ImageNet1K dataset. The ImageNet1k dataset contains over one million images that belong to 1,000 different categories. This pre-training process has enabled the XceptionNet model to efficiently identify important image features and categorize them into various classes. We implemented a fine-tuning process to allow the use of this model for necrosis detection. The architectural flow of the Fully Connected XceptionNet for necrosis identification is shown in Figure FC4.3.

During the fine-tuning phase, we adjust the XceptionNet model to process images that are 256x256 pixels in size and have three color channels (RGB). We initialize the model with weights from its original ImageNet training to improve its ability to recognize general features.

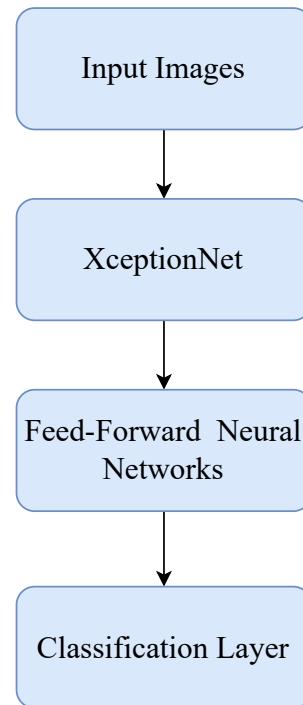


Figure FC4.3: XceptionNet Fully Connected Network for Necrosis Detection

The XceptionNet creates a group of 2048 unique features after the global average pooling layer, as shown in Figure FC4.4. The classification or output layer uses the softmax activation function to estimate the probability of classifying the image as either necrotic or non-necrotic. The fine-tuned architecture, including the input image size and the number of nodes in the output layer, is illustrated in Figure FC4.4.

#### **4.3.1 Model training and Performance Evaluation on inter-slide and inter-class**

During the optimization process of the XceptionNet model, a comprehensive grid search is conducted to identify the optimal values for the learning rate, optimizer, and batch size. The grid search involves testing three different learning rates (0.1, 0.05, and 0.001) and three optimizers (Adam, SGD, and Adagrad), each with three different batch sizes (16, 32, and 64). After performing the grid search, the best values are selected to develop the model. The model is trained for 20 epochs, with a batch size of 32, using

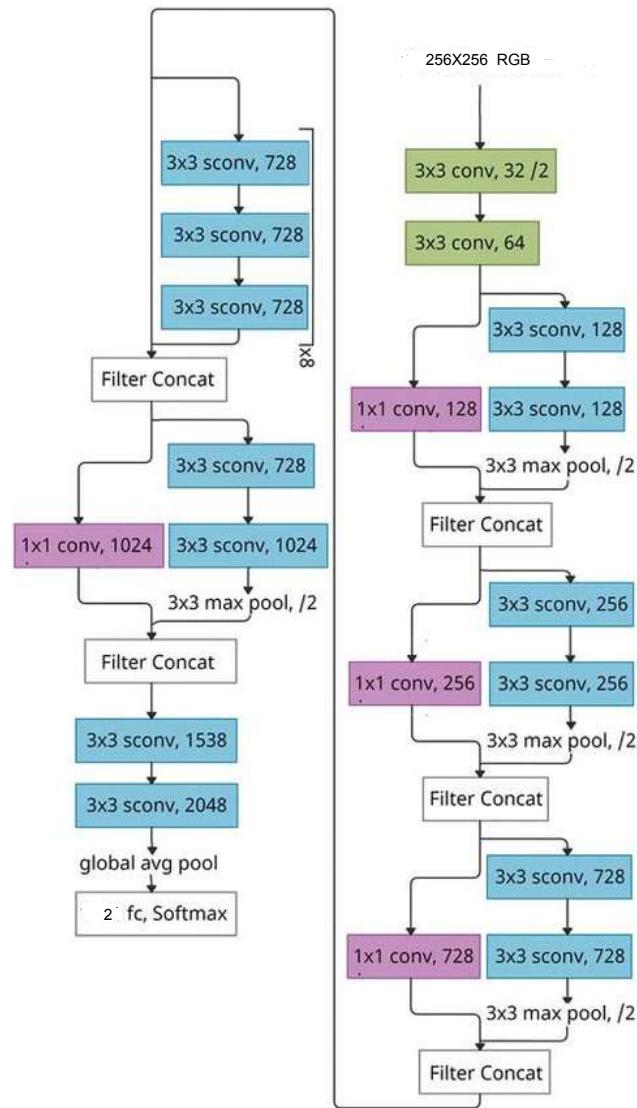


Figure FC4.4: The Architecture of Fine-Tuned XceptionNet for Necrosis Detection

the Adam optimizer with a learning rate of 0.05. Early stopping is used to prevent overfitting and reduce unnecessary training time. Validation loss is monitored during each training batch, and early stopping is triggered when necessary.

The Table TC4.3 shows the necrosis detection metrics and their performance across different batches and sample types. In the analysis of different batches (Batches 1 to 5), the model's accuracy varies from 17.4% to 99.5%. Batch 5 has the highest accuracy at 99.5%, with high recall (1.000) and high precision (0.991). However, Batch 8 has the

Table TC4.3: Batch-wise Necrosis Detection Metrics for Fully Connected XceptionNet

<b>Batch</b>	<b>Sample Type</b>	<b>Necrosis</b>	<b>Non-Necrosis</b>	<b>Accuracy</b>	<b>Recall</b>	<b>Precision</b>	<b>F1 Score</b>
Batch 1	inter-slide	71	848	0.934	0.397	0.265	0.338
Batch 2	inter-slide	57	23	0.900	0.860	0.925	0.892
Batch 3	inter-slide	152	596	0.797	0.246	0.335	0.321
Batch 4	inter-slide	152	148	0.493	0.283	0.290	0.286
Batch 5	inter-slide	57	148	0.995	1.000	0.991	0.996
Batch 6	intra-slide	71	596	0.894	0.256	0.318	0.382
Batch 7	intra-slide	71	109	0.994	0.986	0.993	0.990
Batch 8	intra-slide	71	23	0.245	0.372	0.390	0.267
Batch 9	intra-slide	71	15	0.174	0.362	0.294	0.324

lowest accuracy at 24.5%, reflecting challenges in distinguishing necrotic regions for this particular batch.

Moving on to the analysis of individual slides (Batches 6 to 9), the model consistently achieves higher accuracy, ranging from 17.4% to a significant 99.4%. Batch 7 has the highest accuracy at 99.4%, showing the model's capability in detecting necrosis within individual slides. Additionally, across all intra-slide batches, the model has a robust precision rate, consistently exceeding 97.5%, indicating its ability to identify necrotic regions precisely.

To analyze the model's behavior about overfitting, we divided the training data into five folds and trained the model. Although we achieved high accuracy results, we validated this experiment by conducting cross-validation and calculating an accuracy metric. No overfitting was observed, as the cross-validation results consistently showed high accuracy across all folds. This consistency indicates that the model performed well on both the training and validation sets without significant discrepancies, leading us to conclude that the model did not overfit the data. However, trade-offs are evident in precision and F1 score, highlighting the model's challenges in simultaneously achieving high precision and recall. These results emphasize the need to measure false positives and false negatives in necrosis detection. Further refinement and optimization may be necessary to strike a better balance between precision and recall and enhance the overall

effectiveness of the model.

#### 4.3.2 GradCAM Analysis for Image Representations for XceptionNet

To better understand the behaviour of the XceptionNet model, we utilized visualization techniques on the output of intermediate layers. Gradient maps visualize the model's focus in differentiating between necrotic and non-necrotic regions. Feature heat maps highlight high and low importance areas for classification decisions, using color coding for clarity. The analysis uses the layer before flattening and draws maps regarding the last convolutional block in the pre-trained XceptionNet model.

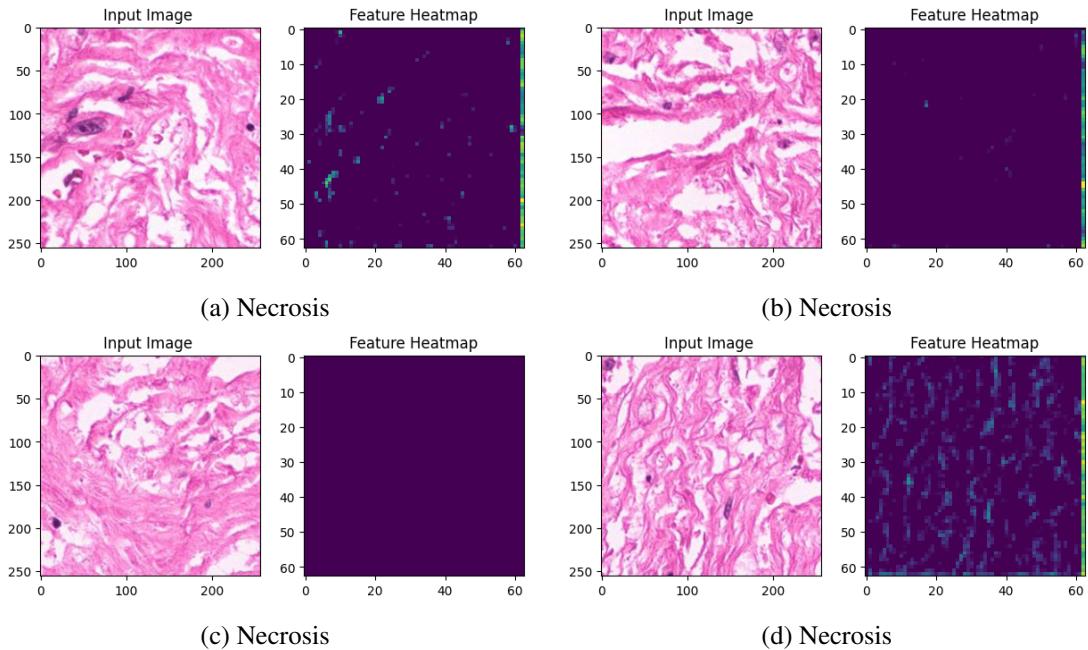


Figure FC4.5: GradCAM Visualization of images with Feature Focus on Necrosis Images

As we can observe from the gradient maps shown in Figure FC4.5, the model focuses on black regions in some images. For example, in Image A and Image B, the model focuses only on black regions. However, the model focuses on white lines and the black region in Image D. This shows that the model focuses on the global context, while necrosis is of local interest. Therefore, CNNs are limited where the local patterns

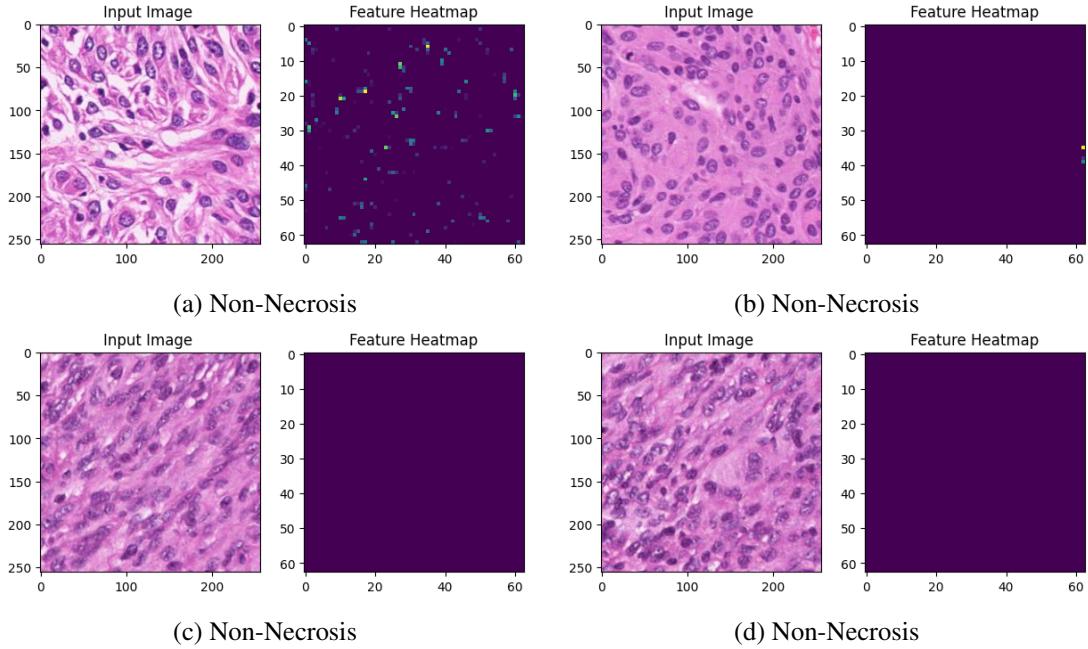


Figure FC4.6: GradCAM Visualization of images with Feature Focus on Non-Necrosis Images

have more importance than global information. This shows that there may be more efficient models for identifying necrosis than XceptionNet.

As we can observe from the images shown in Figure FC4.6, it appears that the XceptionNet model focuses only on the intensity of black pixels when deciding on a non-necrotic region. In images A, B, C, and D, nuclei exist for the entire image, but the model chooses higher-intensity nuclei. This shows a bias towards high-intensity black pixels when identifying non-necrotic regions. However, in the remaining images, the model completely ignores nuclei, focusing only on the background, similar to its understanding of necrotic regions. This shows that the features captured by XceptionNet are either improper or insufficient to identify necrosis.

### 4.3.3 Insights and Observations

During training, we observed that the model exhibited satisfactory performance on both training and validation data in the first epoch. The Figure FC4.7 displays the

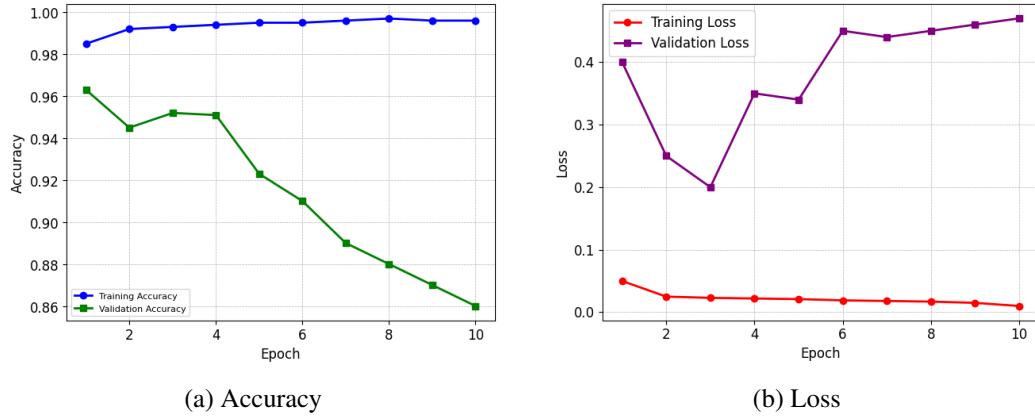


Figure FC4.7: Training and Validation Performance over Epochs for Batch 1

accuracy and loss for both training and validation throughout training. However, the performance and loss declined as the training progressed through the following epochs. As a result, the model was trained for only three epochs in most of the batches, as determined by the early stopping method.

Table TC4.4: Weighted Average Metrics for Fully Connected XceptionNet

Sample Type	Accuracy	Recall	Precision	F1 Score
inter-slide	0.834	0.304	0.381	0.338
intra-slide	0.791	0.403	0.440	0.421
Overall	0.820	0.335	0.399	0.363

The Fully Connected XceptionNet model's performance scores are presented in Table TC4.4. The model can differentiate between necrotic and non-necrotic samples. However, it has a low recall rate (40.3%) and average accuracy (79.1%) in identifying necrotic regions in the same slide. Further, the precision rate of 44% suggests a tendency to misclassify non-necrotic regions as necrotic in intra-slide assessments. In inter-slide evaluations, the model maintains a high accuracy of 83.4% but demonstrates a precision rate of 38.1%, indicating potential misclassification across different slides.

Although the Fully Connected XceptionNet model is sensitive to necrotic regions, its lower precision poses a challenge in minimizing false positives, which is reflected in its overall accuracy of 82%. In medical diagnostics, it is crucial to balance sensitivity

and precision accurately to identify pathological conditions. Therefore, the clinical utility of the model may benefit from addressing this precision-recall trade-off. The model may require further improvements or feature enhancements to deliver a more nuanced and dependable performance in different pathological scenarios, distinguishing between necrosis and non-necrosis.

The model performs better when trained on batches of tissue samples from the same slide than those from different slides. Compared to a traditional hand-crafted features-based SVM model, the fully connected XceptionNet model shows improved accuracy. However, the model's recall has significantly decreased. Deep learning models have shown the potential to recognize local patterns in medical imaging. However, they may face challenges when differentiating between necrotic and non-necrotic regions. This limitation arises because these models focus on overall tissue patterns rather than specific changes in individual cells or local patterns. As a result, deep learning models may make incorrect predictions, which can have profound implications for medical diagnosis and treatment. One can use extracted features from images to enhance recognition of necrotic and non-necrotic regions, incorporate contextual information, or consider using ensembles of models. These techniques have the potential to significantly increase the accuracy of deep learning models, making them more effective in medical diagnosis and treatment.

#### 4.4 XceptionNet Features with SVM

XceptionNet can also be helpful in providing vector representations of features in addition to fully connected classification models. XceptionNet's deep layers learn complex and hierarchical features for image recognition by being trained on large datasets. These features can be extracted and used in various machine-learning tasks. Using pre-trained features, we can adopt a transfer learning approach, allowing simpler algorithms

to be applied on top of the learned representations. This strategic utilization leads to a more efficient machine learning model, making it a streamlined yet effectual approach for various image-related applications. The architecture of XceptionNet features with SVM is illustrated in Figure FC4.8.

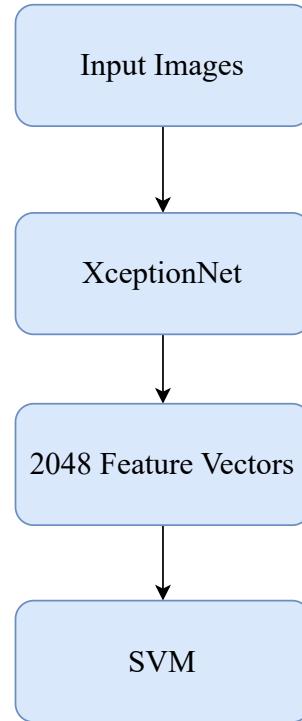


Figure FC4.8: XceptionNet Features with SVM Network for Necrosis Detection

The original XceptionNet design subjected the input images to a series of convolutional and pooling layers, extracting 2048 features. The XceptionNet architecture allows us to choose a variable-length output feature vector size based on our specific requirements, enabling the extraction of enhanced visual features at various dimensions, including dimensionality reduction. Also, selecting feature size is crucial to deep learning, especially in image processing applications. It significantly influences the model's ability to identify meaningful patterns from the data. In our experiment, we chose an output vector size of 2048. We based this decision on the model's performance regarding the designed metric on sizes 256, 512, 1024, and 2048. To compare the performance of the model for different vector sizes, we defined a metric named S score as  $FPR+FNR$ ,

the sum of FPR and FNR. A lower S\_metric value indicates better performance.

$$S\_metric = FPR + FNR$$

Table TC4.5: S metric values for different models for 512,1024,2018, and 4096 feature sizes

inter-slide	model_Name	S_metric	intra-slide	model_Name	S_metric
Batch_1	model_XceptionSVM_512	0.361	Batch_8	model_XceptionSVM_512	0.421
Batch_1	model_XceptionSVM_1024	0.317	Batch_8	model_XceptionSVM_1024	0.359
Batch_1	model_XceptionSVM_2048	<b>0.268</b>	Batch_8	model_XceptionSVM_2048	<b>0.215</b>
Batch_1	model_XceptionSVM_4096	0.482	Batch_8	model_XceptionSVM_4096	0.547
Batch_2	model_XceptionSVM_512	0.537	Batch_9	model_XceptionSVM_512	0.291
Batch_2	model_XceptionSVM_1024	0.444	Batch_9	model_XceptionSVM_1024	0.317
Batch_2	model_XceptionSVM_2048	<b>0.249</b>	Batch_9	model_XceptionSVM_2048	<b>0.188</b>
Batch_2	model_XceptionSVM_4096	0.573	Batch_9	model_XceptionSVM_4096	0.499

To evaluate a model's overall performance, the S\_metric is used, which integrates the False Positive Rate (FPR) and the False Negative Rate (FNR) equally. This approach provides a holistic measure that assigns equal importance to FPR and FNR. It ensures a fair assessment of the model's effectiveness in detecting necrosis, helping fine-tune its performance to meet specific requirements and priorities in histopathology.

In the case of the above Table TC4.5, which presents S\_metric values for different feature vector sizes across inter-slide and intra-slide variations, it is evident that the 2048 feature size consistently outperforms the 512, 1024, and 4096 feature sizes in terms of the S\_metric. The S\_metric values indicate that the 2048 feature size achieves a better balance between FPR and FNR, emphasizing its effectiveness in the necrosis detection task.

The S\_metric was chosen for selecting top features from necrotic and non-necrotic images due to its effective differentiation between these two categories. Choosing different metrics could change the feature extraction process and might require changes to pre-trained models. The selection of the XceptionNet architecture was based on its performance on the dataset, rather than its popularity in previous research. This decision was made to ensure optimal results by aligning the chosen model with the dataset's

characteristics.

#### 4.4.1 Performance Evaluation on inter-slide and intra-slide

We use 2048-dimensional feature vectors obtained from the Fully Connected layer of XceptionNet for training and validating batches. To achieve this, we modify the XceptionNet model by adding a dense layer after the global average pooling layer. This modification ensures that we get 2048-dimensional feature vectors.

We systematically searched the parameter space using a grid search technique, with carefully selected values of regularization parameter ( $C$ ) = 0.1, 0.01, 1 and kernel = 'gaussian,' 'linear', to determine the optimal SVM parameters. This approach allowed us to identify the most suitable configuration for our model. Subsequently, we utilized a SVM with a Gaussian kernel for the classification task. During the training phase, we set the regularization parameter to 1 to enable the SVM to generalize patterns effectively.

Table TC4.6: Batch-wise Necrosis Detection Metrics for XceptionNet Features with SVM

<b>Batch</b>	<b>Sample Type</b>	<b>Necrosis</b>	<b>Non-Necrosis</b>	<b>Accuracy</b>	<b>Recall</b>	<b>Precision</b>	<b>F1 Score</b>
Batch 1	inter-slide	71	848	1.000	1.000	1.000	1.000
Batch 2	inter-slide	57	23	0.988	0.982	0.991	0.987
Batch 3	inter-slide	152	596	0.934	0.678	0.808	0.738
Batch 4	inter-slide	152	148	0.837	0.678	0.808	0.738
Batch 5	inter-slide	57	148	0.995	0.982	0.991	0.987
Batch 6	intra-slide	71	596	1.000	1.000	1.000	1.000
Batch 7	intra-slide	71	109	1.000	1.000	1.000	1.000
Batch 8	intra-slide	71	23	1.000	1.000	1.000	1.000
Batch 9	intra-slide	71	15	1.000	1.000	1.000	1.000

Table TC4.6 showcases the necrosis detection metrics for XceptionNet features with SVM. The model's performance across different batches and sample types is reflected in these metrics.

The model demonstrates varied outcomes in the inter-slide analysis (Batches 1 to 5). Batch 1 performs well with 100% across metrics like accuracy, recall, precision,

and F1 score. Batch 2 maintains a high accuracy of 98.8% and excels in recall (98.2%), precision (99.1%), and F1 score (98.7%). However, Batch 3 shows a trade-off with good precision (80.8%) but lower recall (67.8%). Batch 4 continues the trend of significant results. In contrast, Batch 5 performs well, underscoring the model's consistency and effectiveness in inter-slide analysis.

Moving on to the intra-slide analysis (Batches 6 to 9), the model consistently achieves high accuracy, ranging from 100% to 84.9%. Batch 6 demonstrates perfect performance, emphasizing the model's proficiency in discerning necrosis within individual slides. Across all intra-slide batches, the model maintains robust precision, exceeding 79.5%, showcasing its ability to identify necrotic regions precisely.

Despite these achievements, there are noticeable trade-offs, particularly in precision and F1 score, revealing the model's challenges in simultaneously achieving high precision and recall. The results underscore the importance of considering false positives and negatives in necrosis detection. Further refinement and optimization may be required to strike a better balance between precision and recall, enhancing the overall effectiveness of the model in diverse settings.

#### 4.4.2 Insights and Observations

The method described in section 1.6.4 was used to calculate the weighted average metrics for XceptionNet2048+SVM. Table TC4.7 presents the weighted average metrics for XceptionNet2048+SVM.

Table TC4.7: Weighted Average Metrics for XceptionNet Features with SVM

Sample Type	Accuracy	Recall	Precision	F1 Score
inter-slide	0.955	0.847	0.905	0.875
intra-slide	1.000	1.000	1.000	1.000
Overall	0.969	0.8955	0.937	0.915

In inter-slide evaluations, the model maintains a high accuracy of 95.5%, supported

by an 84.7% recall and a 90.5% precision. It also maintains a good balance between precision and recall.

However, it is essential to acknowledge potential areas for improvement. The model may encounter challenges in some scenarios, as reflected in slightly lower scores in some metrics, such as the 89.5% recall in the overall evaluation.

The XceptionNet has extracted 2048 features, which are then transformed into two components using t-SNE distribution. Figure FC4.9 shows the scatter plot for the t-SNE distribution of Necrosis and Non-Necrosis.

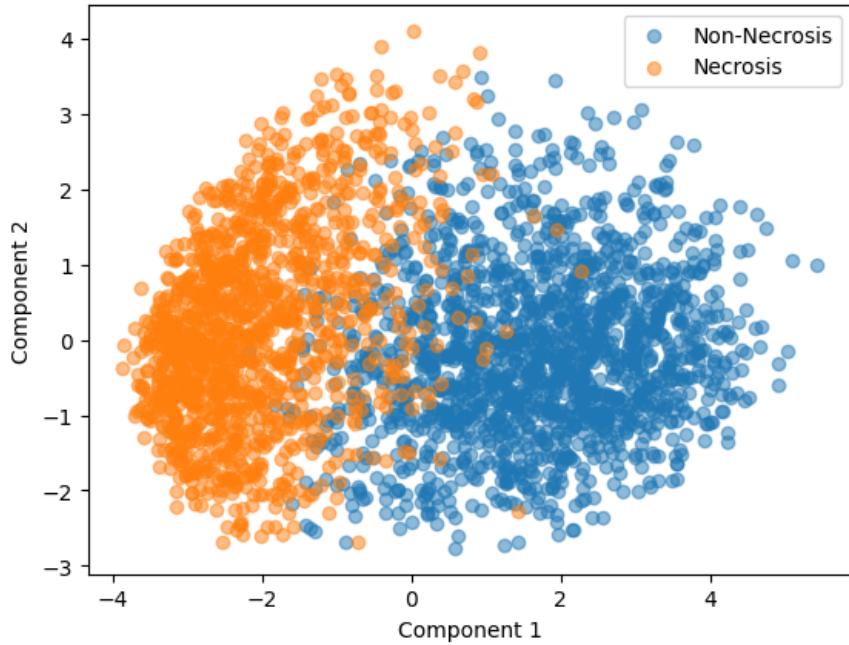


Figure FC4.9: t-SNE Feature Visualization for XceptionNet Features with SVM

While the t-SNE plot shows improved discrimination between necrosis and non-necrosis tiles, significant overlap regions can also be observed. Continuous refinement and optimization of the model can further enhance its sensitivity and precision, ensuring robust performance in diverse pathological contexts.

## 4.5 Domain Knowledge Extraction

### 4.5.1 Nuclei Extraction

The Identification of necrosis in WSI slides involves verifying the presence of healthy cells, tissue appearance, and arrangement patterns. Out of all these methods, the nuclei information is a crucial criterion for distinguishing necrotic and non-necrotic regions in tiles. In addition to unhealthy or diseased cells like mitotic cells, there are also other types of cells, such as viable and tumor cells. Therefore, extracting cell-related information can significantly enhance the model’s performance, as conventional and deep learning-based feature extraction methods may not be effective in comparing with domain-specific features to identify particular regions of interest.

Our methodology for extracting features from histopathological image tiles utilizes nucleus information. We focus on isolating and highlighting nuclei within the image tiles. To achieve this, we follow a unique approach of selectively retaining nuclei while uniformly coloring all other elements within the image black. This includes all background colors and insignificant white spaces, which are also converted to black. The purpose of this strategy is to emphasize the presence and distribution of nuclei in a tile. By marking all non-nuclear elements as inactive features represented by black, the focus is on the crucial nuclear structures, allowing for a more targeted analysis of their distribution and patterns within the images.

### 4.5.2 Nuclei Detection and Segmentation

In our feature extraction pipeline, we utilize StarDist [132], an advanced framework for precise object detection and segmentation, specifically for nuclei segmentation.

**StarDist** is a deep learning-based image analysis method designed for the detec-

tion of star-convex objects, particularly well-suited for applications in biological image analysis, such as the detection of cell nuclei. It utilizes a deep neural network to perform instance segmentation, identifying individual objects in images based on their star-convex shape.

StarDist provides a variety of pre-trained models that are specifically designed for accurate 2D image segmentation. Two specialized variants among these models are dedicated to nucleus segmentation. The first model, 2D\_versatile\_fluo [132], is designed to excel in segmenting fluorescently stained nuclei. On the other hand, the second model, 2D\_versatile\_he [132], is proficient at segmenting nuclei, particularly in Hematoxylin and Eosin stained images. Both models use multiple convolutional neural networks to efficiently identify nuclei of various shapes, resulting in robust and precise segmentation outcomes. In the nuclei detection and segmentation process, we have used the 2D\_versatile\_he model.

The nuclei segmentation of a given tile involves three steps:

**Nuclei Identification:** Multiple CNNs in 2D\_versatile\_he model detect cell types to identify cell polygons. These individual detections are combined to generate a final cell prediction. Each CNN identifies pixel classes and estimates whether a pixel belongs to a cell by calculating radial distances to the object's boundary at that location (i.e., a star-convex polygon representation).

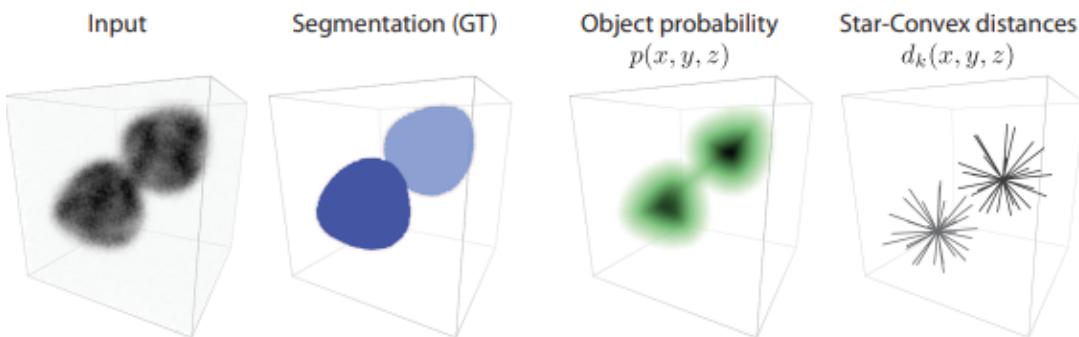


Figure FC4.10: Star Convex Polygon Identification with Convolutional Neural Networks

Multiple CNNs and star-convex polygon representation allow for more accurate and robust nuclei identification, even in complex tissue structures. The process of Nuclei identification is illustrated in Figure FC4.10.

**Nuclei segmentation:** To separate the nuclei from the surrounding tissue, we group the pixels belonging to the same nucleus based on their color, intensity, or texture similarity. We use a back end that relies on CNN instance segmentation to group cells of similar size by color. The process of Nuclei segmentation is illustrated in Figure FC4.11.

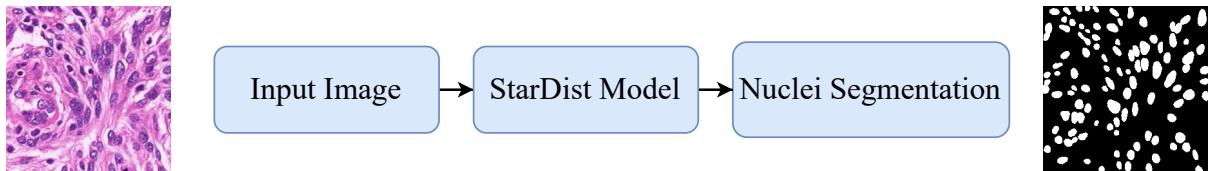


Figure FC4.11: Stardist Segmentation Methodology

**Binarization:** In order to convert the segmented nuclei image into a binary image, the nuclei are represented with white pixels, while the background are represented with black pixels. After completing this step, XceptionNet extracts features from the binary image.

#### 4.5.3 Parameter Optimization for Nuclei Segmentation

The Stardist segmentation module enables us to set the minimum cell diameter and the flow threshold for consideration. The utilization of the minimum diameter parameter serves a crucial purpose in the context of nuclei segmentation. By discarding nuclei below a specific size threshold, typically set to exclude smaller structures, we aim to improve the accuracy and relevance of the segmentation results. In medical imaging, particularly in histopathology, smaller nuclei might be considered less diagnostically significant or represent artifacts and noise in the image rather than actual cellular structures. Therefore, filtering out these smaller nuclei helps prioritize identifying and ana-

lyzing more substantial and potentially more clinically relevant nuclei. This approach contributes to a more accurate representation of the underlying tissue morphology. It aids in extracting meaningful information for subsequent analyses or diagnostic purposes. The flow threshold parameter controls the merging of predicted nuclei into a single segmentation mask.

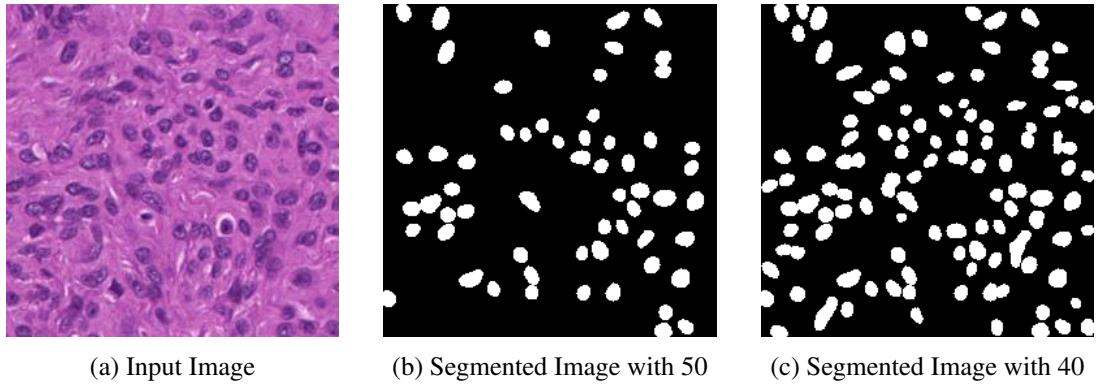


Figure FC4.12: Input and Segmented Images with Diameters 50 and 40, respectively, with a flow threshold of 0.4

Selecting inappropriate parameter values can alter the segmentation outcome and lead to improper results, potentially causing misclassification by the chosen model. We conducted experiments where we determined the optimal diameter and flow threshold values through a trial-and-error approach. The experimental values and input and segmented data images are presented in Figures FC4.12 and FC4.13.

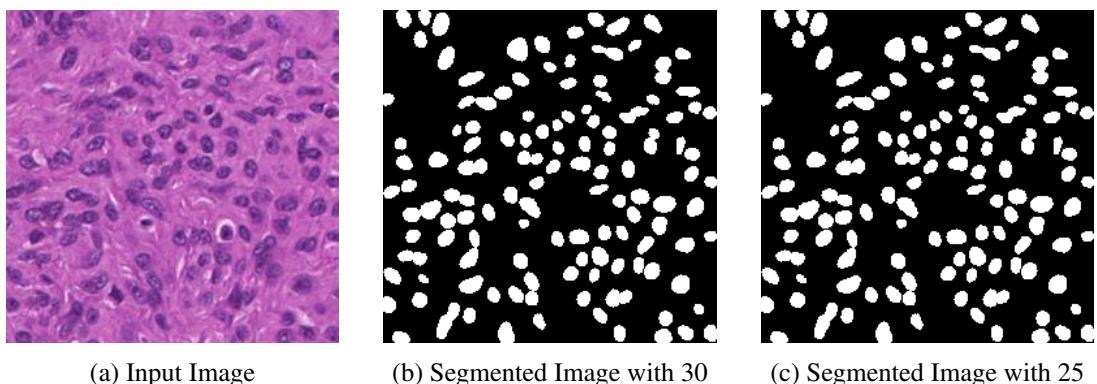


Figure FC4.13: Input and Segmented Images with Diameters 30 and 25, respectively, with a flow threshold of 0.4

The Figures FC4.12 and FC4.13 indicated that the segmentations were missing several nuclei when the diameter was 50. However, there was a significant improvement when the diameter decreased to 40, though some nuclei remained unnoticed. The segmentation was significantly better with diameters of 30 and 25, where almost all nuclei were covered. To maintain consistency in diameter segmentation across all tiles and prevent mitotic cells from being misidentified as regular cells during the segmentation process, we chose 30 as the control parameter for diameter segmentation. The default threshold value of 0.4 was chosen for nuclei segmentation. It was experimented with to strike a balance between capturing object details and minimizing over-segmentation. The current parameters are suitable for a tile size of 256x256, and the size of the nuclei depends on parameters that may need to be adjusted accordingly. The diameter of the nuclei will increase as the magnification increases.

#### **4.5.4 Analysis of Segmentation Results**

The 2D\_versatile\_he module, integral to the Stardist framework, goes beyond conventional segmentation by independently normalizing image channels. This normalization step optimizes the input, setting the stage for subsequent segmentation tasks. Once applied, the module exhibits its prowess by producing instance segmentation results. This segmentation assigns distinct colors to nuclei of varying sizes within a tile. Each unique color represents nuclei of the same size within that particular tile. The tiles are then binarized using threshold techniques that set the color based on certain thresholds and divided into black and white pixels. This transformation demonstrates the significance of the 2D\_versatile\_he module in seamlessly translating raw input images into finely segmented representations, showcasing the nuanced details of nuclei in different size categories. The image presented in Figure FC4.14 illustrates an input image, its corresponding instance segmentation, and nuclei segmentation images.

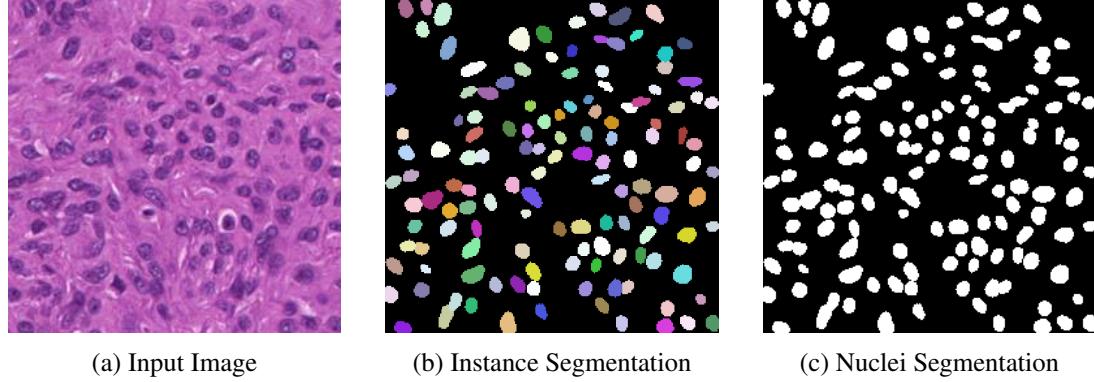


Figure FC4.14: Visualization Segmentation Methodology from Input to Binarized Image

## 4.6 DeepNucleiNet

The models discussed in sections 4.2, 4.3, and 4.4, which use handcrafted features, fully connected convolutional neural networks, and features from pre-trained models, have demonstrated promising performance. However, they need more robustness in handling variations between slides. Additionally, they have exhibited lower performance on specific data batches, such as Batch 3.

Including domain-specific features, particularly those extracted from nuclei, will enhance the model’s ability to make accurate detection decisions. We propose DeepNucleiNet, a two-stage feature extraction process that uses features extracted from pre-trained models and segmented nuclei images. For the pre-trained model, XceptionNet is used in both phases to extract 2,048 features each from the original and segmented images. The features from both phases are then combined to form the desired features. The process increases the feature size to 4,096. Then, it uses it to build the model further for analyzing inter-slide and intra-slide variations. The architecture of DeepNucliNet is presented in Figure FC4.15.

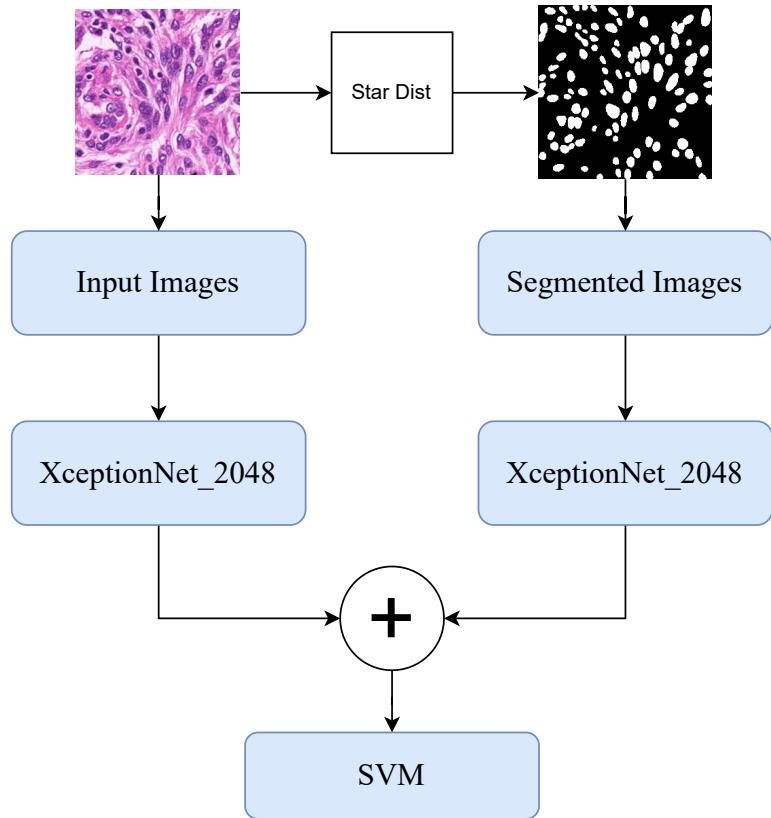


Figure FC4.15: Architecture of DeepNucleiNet

#### 4.6.1 Performance Evaluation on inter-slide and Inter- Class

The features were extracted for each batch and combined to train a support vector classification model. We used the same parameter settings we had prepared in the previous method for XceptionNet and SVM. The model's performance was analyzed using accuracy, precision, recall, and F1 score. The Table TC4.8 shows the necrosis detection metrics and their performance across different batches and sample types.

The DeepNucleiNet model performs well in necrosis identification across intra-slide and inter-slide batches. The performance metrics for DeepNucleiNet are shown in Table 4.8. In the inter-slide analysis (Batches 1 to 5), the model demonstrates high accuracy, ranging from 98.8% to 100%.

Table TC4.8: Batch-wise Necrosis Detection Metrics for DeepNucleiNet Model

<b>Batch</b>	<b>Sample Type</b>	<b>Necrosis</b>	<b>Non-Necrosis</b>	<b>Accuracy</b>	<b>Recall</b>	<b>Precision</b>	<b>F1 Score</b>
Batch 1	inter-slide	71	848	1.000	1.000	1.000	1.000
Batch 2	inter-slide	57	23	0.988	0.982	0.991	0.987
Batch 3	inter-slide	152	596	1.000	1.000	1.000	1.000
Batch 4	inter-slide	152	148	1.000	1.000	1.000	1.000
Batch 5	inter-slide	57	148	0.995	0.982	0.991	0.987
Batch 6	intra-slide	71	596	1.000	1.000	1.000	1.000
Batch 7	intra-slide	71	109	1.000	1.000	1.000	1.000
Batch 8	intra-slide	71	23	1.000	1.000	1.000	1.000
Batch 9	intra-slide	71	15	1.000	1.000	1.000	1.000

Similarly, in the intra-slide analysis (Batches 6 to 9), the model maintains a perfect score across all metrics, emphasizing its robust performance in discerning necrosis within individual slides. Batch 9 stands out with exceptional accuracy, showcasing the model's consistency in accurate necrosis detection. The precision and F1 score metrics further underscore the model's ability to precisely identify necrotic regions while maintaining a balance between false positives and false negatives.

The overall performance of the proposed model is highly encouraging, with perfect or near-perfect scores in all evaluated metrics. The results indicate the model's effectiveness in necrosis detection across diverse settings, demonstrating its potential for applications in medical diagnostics and histopathological analysis. Further refinement and optimization may contribute to even more nuanced performance in specific scenarios.

#### 4.6.2 Insights and Observations

The performance metrics in Table TC4.9 demonstrate that the DeepNucleiNet model differentiates between necrotic and non-necrotic samples more accurately compared to all experimented models described in sections 4.2, 4.3, and 4.6. The model consistently demonstrates strong performance in intra-slide assessments, showcasing accuracy, recall, precision, and F1 scores across the board. This indicates the model's ability to

identify necrotic regions within the same slide. In inter-slide evaluations, the model demonstrates high accuracy at 99.7%, supported by a high recall of 99.7% and precision of 97.1%.

Table TC4.9: Weighted Average Metrics for DeepNucleiNet

<b>Sample Type</b>	<b>Accuracy</b>	<b>Recall</b>	<b>Precision</b>	<b>F1 Score</b>
inter-slide	0.997	0.997	0.971	0.984
intra-slide	1.000	1.000	1.000	1.000
Overall	0.996	0.999	0.981	0.990

The overall evaluation further validates the robustness of the DeepNucleiNet model, achieving an accuracy of 99.6%, a recall of 99.9%, and a precision of 98.1%, resulting in an F1 score of 99.0%. These metrics demonstrate the model's precision and balance in identifying both necrotic and non-necrotic regions. The DeepNucleiNet model's higher performance positions it as a promising tool for necrosis detection in various neuropathological settings, underscoring its potential for dependable clinical applications. The t-SNE distribution, shown in Figure FC4.16, is used to analyze the quality of feature representation with DeepNucleiNet.

## 4.7 Performance Analysis

The feature analysis is conducted for three models including baseline , Xceptionnet with SVM and DeepNucleiNet and their t-SNE scatter plots for two components are shown in Figures FC4.2, FC4.9 and FC4.16. The t-SNE plot of DeepNucleiNet has a much better representation of features than Baseline, and XceptionNet features with SVM.

In terms of overall accuracy, the proposed model achieved a 74.7% improvement over the baseline model and a 40.28% improvement over the fully connected Xception-Net. It also showed a 2.68% improvement over the XceptionNet features trained on

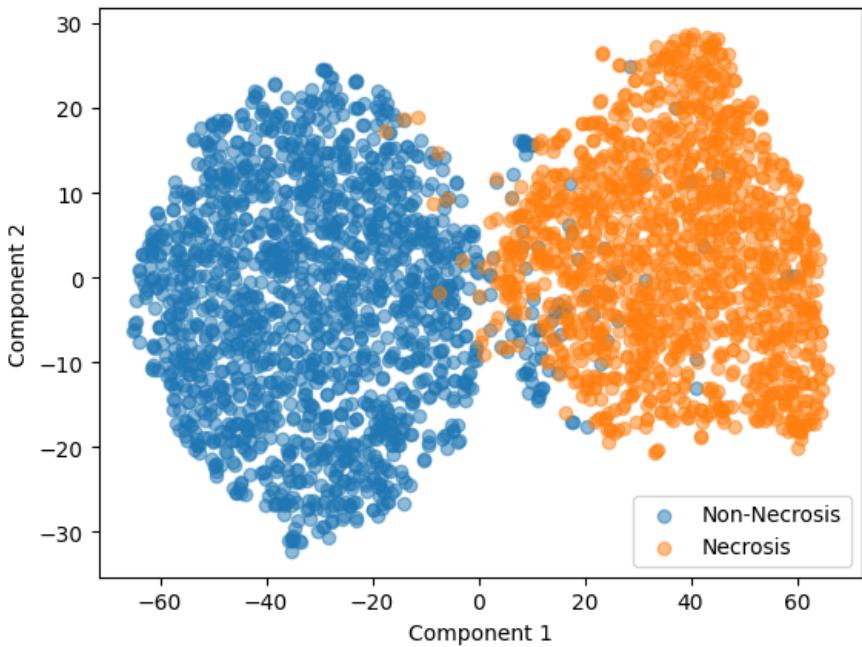


Figure FC4.16: t-SNE Feature Visualization for Proposed Model

Table TC4.10: Accuracy Metrics for intra-slide, inter-slide, and Overall for All Models

Model	intra-slide	inter-slide	Overall
Baseline	0.65	0.557	0.586
XceptionNet	0.834	0.791	0.82
XceptionNet_2048+SVM	0.955	1	0.969
<b>DeepNucleiNet</b>	<b>0.997</b>	<b>1</b>	<b>0.996</b>

an SVM model. The accuracy values for all models are shown in Table TC4.10 and visualized in Figure FC4.17.

Table TC4.11: Recall Metrics for intra-slide, inter-slide, and Overall for All Models

Model	intra-slide	inter-slide	Overall
Baseline	0.873	0.893	0.886
XceptionNet	0.304	0.403	0.335
XceptionNet_2048+SVM	0.847	1	0.895
<b>DeepNucleiNet</b>	<b>0.997</b>	<b>1</b>	<b>0.999</b>

Regarding overall recall, the proposed model achieved a 13.79% improvement over the baseline model and a 98% improvement over the fully connected XceptionNet. It

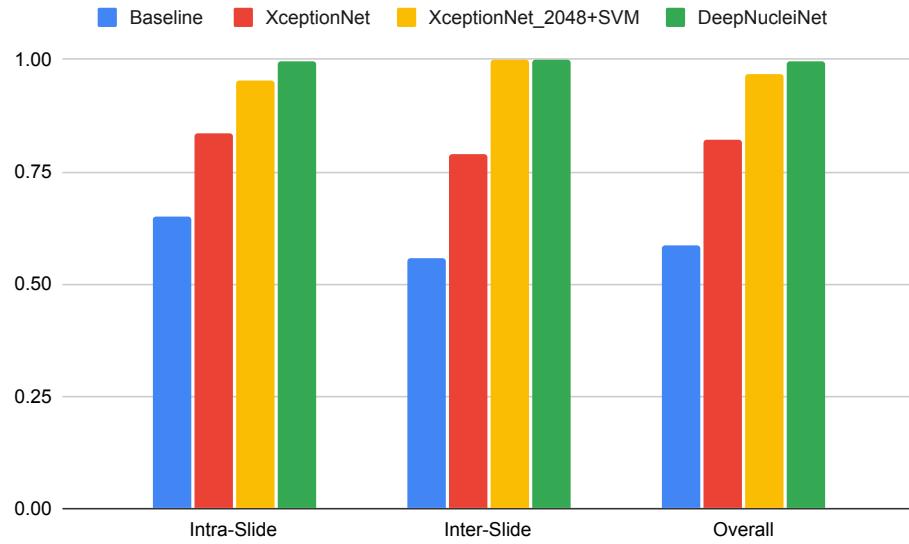


Figure FC4.17: inter-slide, intra-slide and Overall Accuracy Comparison

also showed a 7.61% improvement over the XceptionNet features trained on an SVM model. The recall values for all models are shown in Table TC4.11 and visualized in Figure FC4.18.

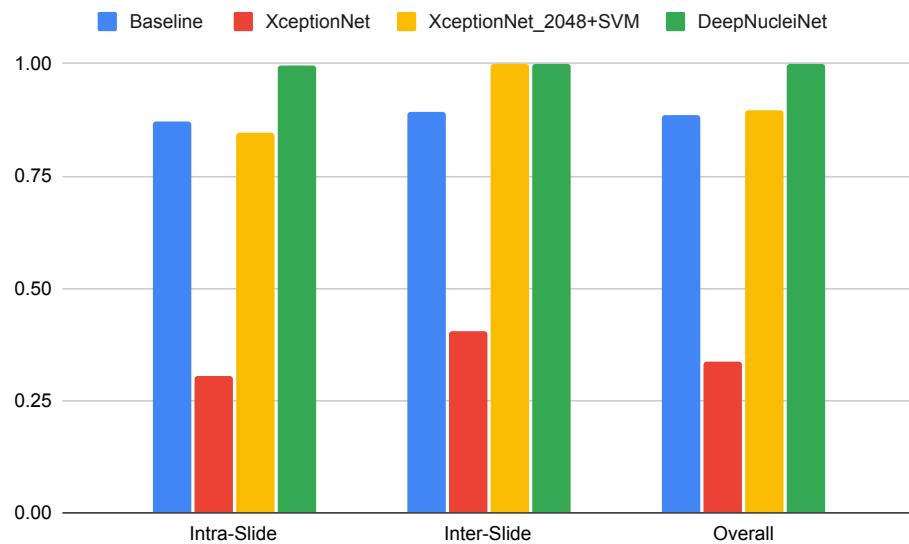


Figure FC4.18: inter-slide, intra-slide and Overall Recall Comparison

In terms of overall precision, the proposed model achieved a 73.68% improvement

Table TC4.12: Precision Metrics for intra-slide, inter-slide, and Overall for All Models

Model	intra-slide	inter-slide	Overall
Baseline	0.490	0.429	0.448
XceptionNet	0.381	0.440	0.399
XceptionNet_2048+SVM	0.905	1	0.937
<b>DeepNucleiNet</b>	<b>0.971</b>	<b>1</b>	<b>0.981</b>

over the baseline model and a 40.28% improvement over the fully connected Xception-Net. It also showed a 2.68% improvement over the XceptionNet features trained on an SVM model. The precision values for all models are shown in Table TC4.12 and visualized in Figure FC4.19.

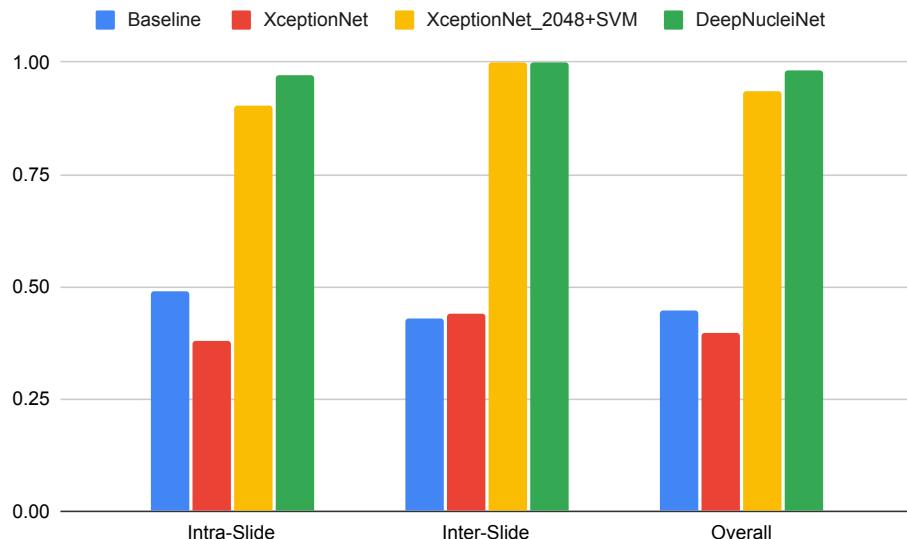


Figure FC4.19: inter-slide, intra-slide and Overall Precision Comparison

Table TC4.13: F1 Score Metrics for intra-slide, inter-slide, and Overall for All Models

Model	intra-slide	inter-slide	Overall
Baseline	0.626	0.579	0.597
XceptionNet	0.421	0.338	0.363
XceptionNet_2048+SVM	0.875	1	0.915
<b>DeepNucleiNet</b>	<b>0.984</b>	<b>1</b>	<b>0.990</b>

Regarding the overall F1 score, the proposed model achieved a 66.08% improve-

ment over the baseline model and a 172.40% improvement over the fully connected XceptionNet. It shows performance approximately similar to the XceptionNet features trained on an SVM model. The recall values for all models are shown in Table TC4.13 and visualized in Figure FC4.20

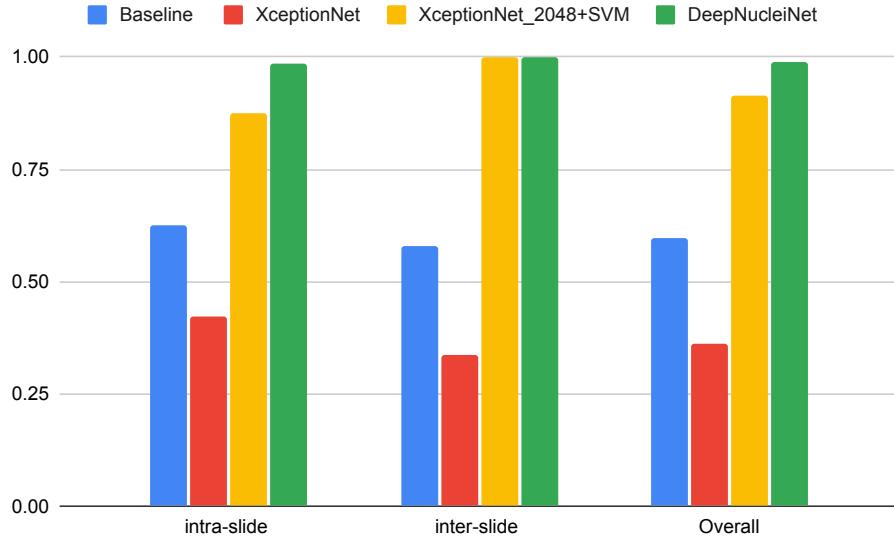


Figure FC4.20: inter-slide, intra-slide and Overall F1 Score Comparison

Our proposed model consistently improved accuracy for both inter-slide and intra-slide classification across multiple batches. Although the baseline and experimental models also achieved good accuracy on some batches, they required assistance in understanding and classifying entirely new samples within a slide. This led to lower accuracy, particularly for batches 3 and 4.

## 4.8 Partial Necrosis Performance Evaluation in Tiles

When generating and labeling tiles for meningioma images, it is crucial to ensure that the predictions are as accurate as possible. In some instances, tiles may drop onto the border region of annotated areas of necrosis that can have partial necrosis ranging from very small areas to close to 100%. We analyze the ability of DeepNucleiNet to

identify necrotic tiles with varying area of necrosis within the tile, using the bucketing into ranges of 10% as described earlier.

Table TC4.14: Analysis of the DeepNucleiNet’s performance in relation to the partial necrosis buckets

<b>Threshold (%)</b>	<b>Tiles</b>	<b>Recall</b>
0-10	277	0.45
10-20	161	0.60
20-30	137	0.70
30-40	145	0.83
40-50	156	0.86
50-60	199	0.97
60-70	207	1
70-80	213	1
80-90	274	1
90-100	2465	1

The analysis of the model’s performance in relation to the partial necrosis buckets, as presented in Table TC4.14, provides insights into its ability to detect varying degrees of necrosis. The model has a higher recall rate as the threshold increases, which indicates a higher likelihood of necrosis. At the lowest threshold (0-10), the recall rate is 0.45, steadily increasing to 1.00 at the highest threshold range (90-100). This indicates that the model effectively identifies and captures necrosis instances, especially in scenarios where necrotic regions are more prominent.

The Table TC4.14 also shows the number of tiles processed for each threshold, which provides additional context to the model’s performance. The gradual increase in recall rate, as shown in Figure FC4.21, along with the corresponding threshold levels and the number of tiles processed, highlights the model’s ability to adapt to varying levels of necrotic involvement. This information can help optimize the model’s deployment in scenarios where accurate identification of necrotic regions is essential, such as in neuropathological analyses.

The DeepNucleiNet performed well when the threshold of necrosis reached 40% to 50%. This suggests that the model’s accuracy is significantly affected by the necrosis

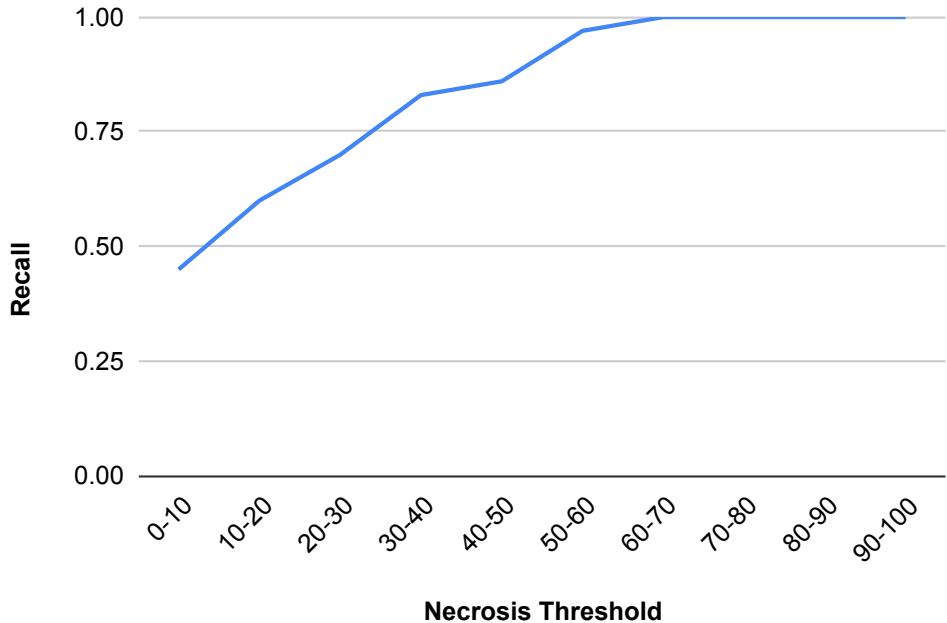


Figure FC4.21: Performance of the DeepNucleiNet according to Necrosis area within the tile

threshold percentage in tiles.

Analyzing the impact of necrosis threshold percentages on the model's accuracy can benefit pathologists. It helps them make informed decisions on the optimal threshold percentage for analysis, leading to better patient outcomes and identifying areas for improvement in the current model.

## 4.9 Adapting Meningioma-trained Model for Necrosis Detection in GBM and LGG

The DeepNucleiNet model and the performance metrics discussed here were obtained by training on meningioma image tiles. We now evaluate the effectiveness of this model to detect necrosis in other brain regions using GBM and LGG data extracted from TCGA platform.

Tiles from glioma diagnostic slides are used in this prediction process to identify necrosis in each of the LGG and GBM manifolds. Seven slides are analyzed, including four from GBM and three from LGG. When tested on these slides, the model's performance is validated using an accuracy score.

Table TC4.15: Batch-wise Necrosis Detection Metrics on GGM and LGG using DeepNucleiNet

Sample	Sample Type	Necrosis	Non-Necrosis	Accuracy
Sample_021	GBM	27	64	0.964
Sample_022	GBM	125	34	0.955
Sample_023	GBM	52	45	0.982
Sample_024	GBM	65	12	1
Sample_025	LGG	0	156	1
Sample_026	LGG	0	285	1
Sample_027	LGG	0	35	1

An expert pathologist at NIMHANS, India, reviewed 424 GBM and 476 LGG tiles to validate the predictions made by the DeepNucleiNet model. Table TC4.15 summarizes the validation process results, presenting the model's performance.

## 4.10 Summary

In this chapter, we have discussed the design and execution of the baseline, experimental, and DeepNucleiNet, including feature extraction and model development. We focused mainly on the segmentation process for tiles generated from WSI images using the Stardist nuclei segmentation model for the proposed model.

Next, we evaluated the performance of all models baseline, proposed, and experimental in terms of accuracy, recall, precision, and F1 score. We analyzed their performance and compared them to demonstrate the efficiency of the DeepNucleiNet model over the baseline model.

We conclude that the DeepNucleiNet model outperformed the experimental and baseline models in all inter-slide and intra-slide batches.

In addition, we found that the DeepNucleiNet model trained on meningioma exhibits efficient results when used to detect necrosis in glioma slides, such as GBM and LGG. The DeepNucleiNet is particularly effective in indicating necrosis, which is not a prominent feature in LGG [132].

# CHAPTER 5

## ENHANCING ROBUSTNESS IN NECROTIC REGION IDENTIFICATION WITHIN WSI

### 5.1 Introduction

In chapters 3 and 4, we presented a domain adaptation pipeline based on deep learning, specifically designed to analyze WSI images. The focus of our approach was to analyze tiles of size 256x256 pixels. Our analysis results demonstrate our approach's practical effectiveness, especially in automating the detection of necrosis regions within WSI images.

Our technology can help pathologists examine slide images more efficiently. We propose evaluating the effectiveness of DeepNuceliNet as a pre-processor within a pathology workflow. We wish to evaluate if this model can identify potential necrosis regions, either completely or partially. This can then allow pathologists to zoom in for more nuanced analysis. This can significantly reduce the need for pathologists to manually search for potential necrotic areas across the entire slide. This reliable pre-processing step can not only enhance the efficiency of pathologists but also mitigate errors in slide labeling.

From the viewpoint of the pathologist, it is crucial to address the following critical questions:

1. Can the tool consistently detect at least some part of all necrosis regions in a slide?
2. What percentage of each necrosis region can the tool accurately detect?
3. What is the minimum necrotic area the tool can reliably identify?
4. How does the tool's performance vary based on grid-related parameters such as position, orientation, and size since it relies on generating tiles from a grid?

In this Chapter, we attempt to analyze DeepNucleiNet, described in section 4.7, from the standpoint of these questions.

### **5.1.1 Region Extraction and Tile Generation**

In this analysis, we have used Sample\_007, which has 248 regions of necrosis marked in different shapes, such as polygons and rectangles. The steps involved in the process from region selection to tile extraction are given below.

1. Load annotations from the XML file into ImageScope for Slide Sample 7. Choose 'Detailed Annotations View' and export all annotations as a CSV file.
2. Identify a region by clicking on the Region column, copying the region ID and area.
3. Locate the corresponding region ID in the Sample\_007.xml file, note the X and Y coordinates, and map them to the prepared tiles. Select the tiles for each region accordingly.

For this experiment, we extracted 15 specific regions from Sample 7. Then, we extracted tiles of those regions from the pool of tiles belonging to Sample 7, guided by the annotation file. We used ImageScope to extract and measure the areas of these regions.

### 5.1.2 Regional and Areal Analysis of Necrosis Detection

The tile generation process will create two types of tiles: interior and boundary. Based on our earlier experiments, we observed that the model performed well on those border tiles with approximately 50% or higher necrosis area. Further, the expert pathologists indicated that labeling would be ambiguous in the boundary regions, and tiles with low necrosis regions may have ambiguity in labels. Hence, for analysis purposes, we decided to choose the 50% threshold. The boundary tiles are further categorized based on a 50% necrosis threshold. Table TC5.1 displays detailed information on the regions, area, and tile count for each type of interior and boundary and the recall metrics for their tiles.

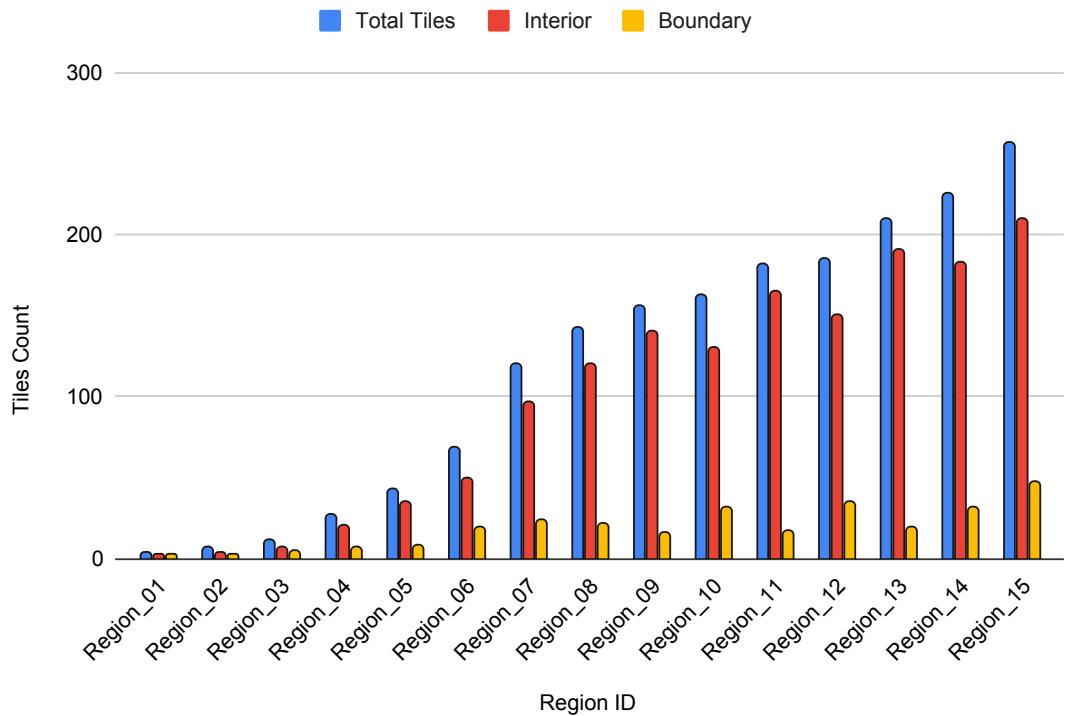


Figure FC5.1: Distribution of Total Tiles, Interior, and Boundary tiles across regions

The detailed information regarding interior and boundary tiles is illustrated in Figure FC5.1. The key metrics include the total number of tiles, interior tiles, and boundary

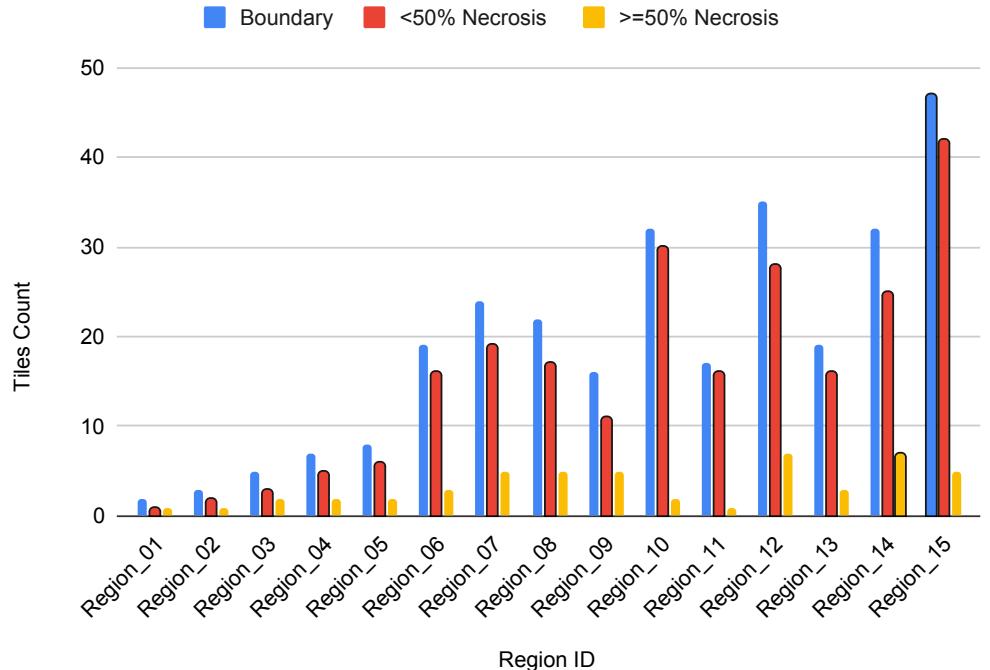


Figure FC5.2: Visualizing the distribution of Boundary Tiles,  $<50\%$ , and  $\geq 50\%$  Necrosis tiles across regions.

Table TC5.1: Summary of Regions with Interior and Boundary Tile Statistics

Region ID	Area (um2)	Total Tiles	Interior		Boundary			
			Tiles	Recall	<50%	Recall	≥50%	Recall
Region_01	997	4	2	1	1	0	1	1
Region_02	1375	7	4	1	2	0	1	1
Region_03	2857	12	7	1	3	0.333	2	1
Region_04	4772	27	20	1	5	0.2	2	1
Region_05	7921	43	35	1	6	0.6	2	1
Region_06	15615	68	49	1	16	0.5	3	1
Region_07	30805	120	96	1	19	0.375	5	1
Region_08	57923	142	120	1	17	0.421	5	1
Region_09	63495	156	140	1	11	0.588	5	1
Region_10	74525	162	130	1	30	0.533	2	1
Region_11	79936	182	165	1	16	0.6	1	1
Region_12	85045	185	150	1	28	0.562	7	1
Region_13	89054	209	190	1	16	0.642	3	1
Region_14	91578	225	183	1	25	0.625	7	1
Region_15	97850	257	210	1	42	0.404	5	1

tiles for each region. Among all regions, Region\_01 has the smallest area (997 um<sup>2</sup>). It consists of four tiles divided into two interior and two boundary tiles. On the other hand, Region\_15 has the most significant area (97850 um<sup>2</sup>). It features 257 tiles, of which 210 are interior tiles and 47 are boundary tiles. The boundary tiles and tiles with percentages of necrosis are shown in Figure FC5.2.

The DeepNucleiNet described in section 4.6, is utilized to predict the results on tiles prepared according to section 5.1.1. The recall performance metric is then calculated for both interior and boundary tiles.

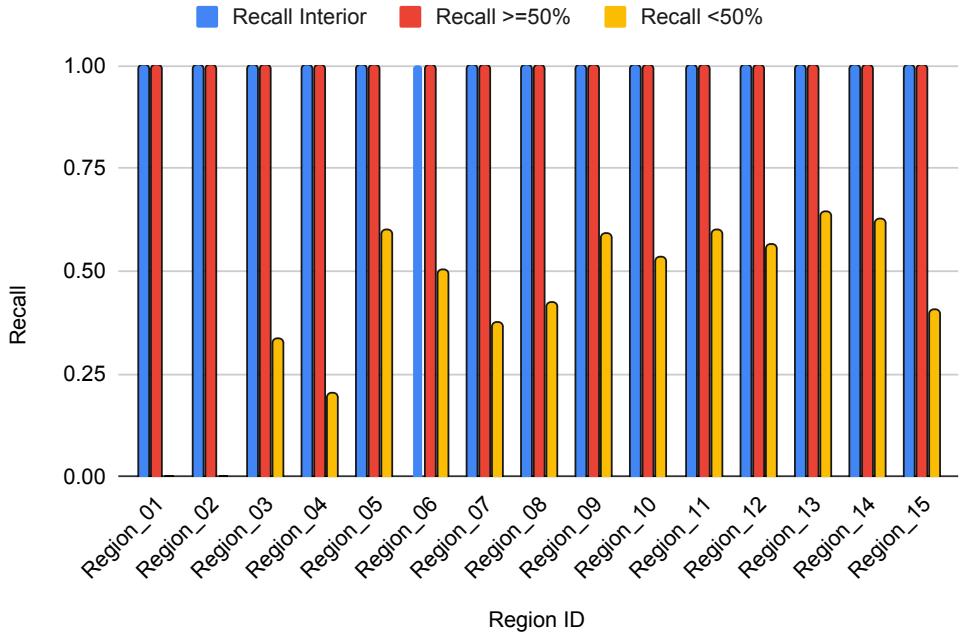


Figure FC5.3: Recall Metrics for Interior and Boundary Necrosis Detection in Regions

We can see from Figure FC 5.3 that the model consistently identifies interior tiles across all regions, indicating a reliable method for detecting the interior of necrosis regions. When the necrosis threshold for a tile is less than 50%, the model's accuracy decreases compared to the interior tiles. Regions such as \_04 and \_15 have a lower percentage of necrosis in the interior, which suggests a potential difficulty in accurately detecting parts of the necrotic regions, significantly when they extend to the boundaries.

The model faces complexities in detecting necrotic boundaries accurately, which can affect accuracy, especially when necrosis levels are relatively lower. Despite these challenges, the model's consistent performance in detecting interior tiles highlights its overall reliability.

### 5.1.3 Insights and Observations

After analyzing the results from Table TC5.1, it is apparent that DeepNucleiNet performed well, achieving a high recall of approximately 1 for interior and boundary tiles with more than 50% necrosis. These results indicate the model's ability to accurately identify regions containing necrosis, whether in the interior or at parts of the boundary of the tissue samples. The high recall suggests that the model successfully detects all regions of necrosis with at least one interior tile.

This observation is crucial as it showcases the model's ability to detect all necrotic regions within WSI. Identifying necrotic areas is critical in pathology and medical diagnostics, and the high recall values indicate the potential effectiveness of DeepNucleiNet in this context.

Table TC5.1 shows that DeepNucleiNet can identify necrosis with high probability if the tile contains 50% necrosis or more. Therefore, to ensure reliable detection by the model, at least one of the tiles must contain more than 50% necrosis. This discovery is significant and provides a valuable threshold for understanding the model's sensitivity to necrotic regions. This understanding will contribute to refining the application of DeepNucleiNet, ensuring robust detection even when necrotic regions span multiple tiles. It is important to note, however, that depending on the geometry of the region and the position of tiles, some necrosis regions may span multiple tiles, with all of these tiles containing less than 50% necrosis.

## 5.2 Necrosis Detection with Grid Positional Changes

Since the choice of grid origin is arbitrary relative to the geometry of the tissue, we would also try to analyse the sensitivity of the approach to the choice of grid position. For the samples we have analyzed, we try to evaluate if shifting the origin of the grid impacts the performance of the technique.

In Figure FC5.4, we can see that we produced tiles five times on the same slide by modifying the initial coordinate. The X represents the starting point for the first tile generation step and processing. After generating the tiles, we labeled them using annotations provided for Sample\_007. Sample\_007 includes multiple labels such as necrosis, RBC, and fascicles. We segmented the tiles using the method proposed in Section 4.6. XceptionNet extracted the features from both the segmented and original tiles. The next iteration of tile generation was performed by offsetting the grid origin by 100 pixels to avoid repeating tiles in the subsequent steps. We carried out this process for X+100, X+200, X+300, and X+400. At each stage of tile generation, we calculated the model's accuracy.

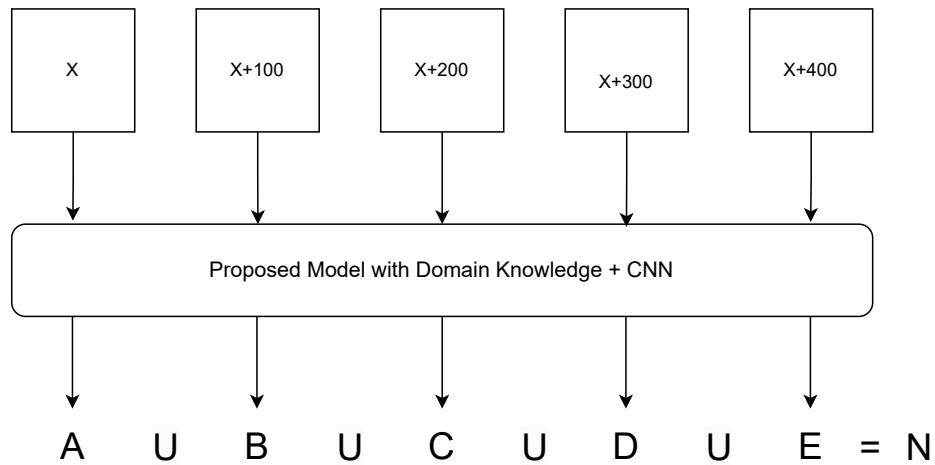


Figure FC5.4: Generating tiles using multiple initialization coordinates

We use the following process to identify all the necrosis regions in a WSI. Firstly,

we make predictions for each image tile, including X, X+100, X+200, and X+300. We then separate the tiles that we predict to be necrosis and compute the union of all pixels in these tiles to determine the final necrosis regions in the WSI. In Figure FC5.3, the necrosis-predicted tiles in each phase are referred to as A, B, C, D, and E. The calculation determined the union of these inputs and produced an output of N. This method reduces the likelihood of missing any necrosis region caused by the position of the grid.

The analysis of results in Table TC5.2 provides a detailed evaluation of the model's performance based on different tile generation points

The DeepNucleiNet has consistently shown high accuracy in identifying necrosis. The model's performance metrics indicate its robustness in effectively recognizing smaller necrotic regions and detecting all necrosis regions with at least one interior tile. These findings further support the reliability of DeepNucleiNet in identifying necrotic areas regardless of the specific tile generation point.

Table TC5.2: Distribution and Performance of Model with Multiple Tile Generation Steps

<b>Threshold</b>	<b>Tiles</b>	<b>Necrosis</b>	<b>Non-Necrosis</b>	<b>Performance</b>
X	3162	3014	148	0.9539
X+100	3160	3012	148	0.96
X+200	3164	3015	149	0.975
X+300	3165	3015	151	0.95
X+400	3167	3017	150	0.952

### 5.3 Summary

In this chapter, we have explored the ability of DeepNucleiNet to detect areas and regions with necrosis by selecting 15 regions, i.e., tiles that had been labeled as overlapping the selected region from Sample\_007. The DeepNucleiNet demonstrated good performance over the area and could identify even small areas with necrosis.

Furthermore, we evaluated the robustness of DeepNucleiNet by generating tiles from grids with multiple initial coordinates, which showed good performance and promising stability toward tile generation.

## CHAPTER 6

### CONCLUSIONS AND FUTURE WORK

#### 6.1 Findings:

- The fully connected convolutional neural networks will not be efficient in understanding images where local importance is more significant than global importance. This conclusion was based on the results obtained in Chapter 3.3.
- The inclusion of domain features in this case, including segmented nuclei images in addition to the original images, showed good classification abilities in terms of feature distribution. This can be concluded from Figures FC4.2, FC4.9, and FC4.16.
- The DeepNucliNet model supports hypothesis 1 as it shows significant improvement when domain features are used in addition to neural network-based features. This conclusion can be drawn from Tables TC4.10, TC4.11, TC4.12, and TC4.13.
- The DeepNucleiNet also showed good ability in identifying necrosis in GBM and LGG, which supports hypothesis 2.

## 6.2 Conclusion:

The healthcare field has greatly benefited from the recent advancements in computer vision, particularly in image classification and segmentation. This thesis delves into the effectiveness of incorporating domain knowledge to enhance the accuracy of necrosis detection in deep-learning models. The key contributions of this research can be summarized as follows:

In Section 3.3, we demonstrated effective data preparation techniques to validate the models across unseen slides and classes. This involved preparing data in batches for inter-slide and inter-class (or intra-slide) scenarios, providing a robust and suitable dataset for training and validation.

In Sections 4.2, 4.3, and 4.4, we presented an analysis of the performance of necrosis identification algorithms using handcrafted features, Fully Connected XceptionNet, and XceptionNet features with SVM. The performance of these three algorithms was evaluated on both inter-slide and intra-slide data batches.

In Section 4.5, we demonstrated the process of nuclei segmentation using the pre-trained StarDist nuclei segmentation backend by a CNN. This segmented data served as one of our domain features.

In Section 4.6, we presented the DeepNucliNet design with a two-stage feature extraction network using domain, and CNN features for necrosis identification using the SVM model.

In Section 4.7, we demonstrated the performance of the DeepNucliNet model over the baseline and experimented with models for necrosis detection in meningioma. Our DeepNucliNet model achieved an accuracy improvement of 74.7%, a recall improvement of 13.79%, a precision improvement of 73.69%, and an F1 score improvement

of 110.64% compared to the baseline model. This indicates significant performance improvement by incorporating nuclei-based domain knowledge into the deep learning model.

In Section 4.9, we validated the meningioma-trained model on other brain regions, particularly glioma (GBM and LLG). The DeepNucliNet model successfully detected glioma necrosis, even though it was only trained on meningioma data. This suggests using a meningioma-trained model to detect necrosis in other brain tumors, such as LGG and GBM.

In Section 5.1, we have measured the performance of the DeepNucliNet model in identifying necrosis at both regional and areal levels. The DeepNucliNet reported good performance in both analyses.

In Section 5.2, we analyzed the model’s robustness by generating tiles with multiple initialization coordinates. The model performed well in all cases, demonstrating robustness towards tile generation.

### **6.3 Future work:**

The proposed model focuses on necrosis detection. Subsequent research will apply this model to detect fascicles, whorls, and sheeting patterns. Further extensions could extract more meaningful and robust domain information from viable, tumor, and healthy cells.

The CNN-based model was designed with a fixed image size of 256x256x3. Future research aims to develop methods for processing regions of varying sizes simultaneously and accurately classifying them as necrotic or non-necrotic.

## Bibliography

- [1] Mayfield Clinic Oxford. Brain tumours: An introduction. Oxford, UK, 2018.  
<https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084>.
- [2] Quinn T Ostrom, Haley Gittleman, Gabrielle Truitt, Alexander Boscia, Carol Kruchko, and Jill S Barnholtz-Sloan. Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2011–2015. *Neuro-Oncology*, 20(suppl\_4):iv1–iv86, October 2019.  
<http://doi.org/10.1093/neuonc/noy131>.
- [3] Quinn T. Ostrom, Haley Gittleman, Peter M. de Blank, Jonathan L. Finlay, James G. Gurney, Roberta McKean-Cowdin, Duncan S. Stearns, Johannes E. Wolff, Max Liu, Yingli Wolinsky, Carol Kruchko, and Jill S. Barnholtz-Sloan. American brain tumor association adolescent and young adult primary brain and central nervous system tumors diagnosed in the united states in 2008-2012. *Neuro-Oncology*, 18(suppl\_1):i1–i50, January 2016.  
<https://doi.org/10.1093/neuonc/nov297>.
- [4] D. N. Louis, A. Perry, and G. et al. Reifenberger. The 2016 world health organization classification of tumors of the central nervous system: A summary. *Acta Neuropathol*, 131(6):803–820, 2016.  
<https://doi.org/10.1007/s00401-016-1545-1>.

- [5] Alf M.D. Giese and Manfred M.D. Westphal. Glioma invasion in the central nervous system. *Neurosurgery*, 39(2):235–252, August 1996. <https://doi.org/10.1097/00006123-199608000-00001>.
- [6] Li-Na Wang, Li Wang, Guanghui Cheng, Mingkai Dai, Yunhai Yu, Guoxin Teng, Jingjie Zhao, and Dawei Xu. The association of telomere maintenance and tert expression with susceptibility to human papillomavirus infection in cervical epithelium. *Cellular and Molecular Life Sciences*, 79(2), 2022. <https://doi.org/10.1007/s00018-021-04113-0>.
- [7] R. Stupp, M. E. Hegi, W. P. Mason, M. J. van den Bent, M. J. Taphoorn, R. C. Janzer, S. K. Ludwin, A. Allgeier, B. Fisher, K. Belanger, P. Hau, A. A. Brandes, J. Gijtenbeek, C. Marosi, C. J. Vecht, K. Mokhtari, P. Wesseling, S. Villa, E. Eisenhauer, and R.-O. Mirimanoff. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase iii study: 5-year analysis of the eortc-ncic trial. *The Lancet Oncology*, 10(5):459–466, 2009. [https://doi.org/10.1016/S1470-2045\(09\)70025-7](https://doi.org/10.1016/S1470-2045(09)70025-7).
- [8] A. Perry, D. N. Louis, and A. von Deimling. *WHO Classification of Tumours of the Central Nervous System (Revised 4th Edition)*. International Agency for Research on Cancer, 2016. <https://doi.org/10.1007/s00401-016-1545-1>.
- [9] D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, and P. C. Burger, editors. *WHO Classification of Tumours of the Central Nervous System. Revised 4th Edition*. World Health Organization, 2016. <https://doi.org/10.1007/s00401-016-1545-1>.
- [10] R. A. Towner, N. Smith, S. Doblas, Y. Tesiram, P. Garteiser, D. Saunders, and D. Robertson. In vivo detection of experimental cerebral metastases using magnetic resonance imaging and a 2-conjugated gadolinium

- metallofullerene mri contrast agent. *Neuro-oncology*, 14(1):6–18, 2012. <https://doi.org/10.1093/neuonc/nor186>.
- [11] SM Dintzis. Digital pathology: Advantages, limitations and emerging perspectives. *Journal of Laboratory and Precision Medicine*, 3:22, 2018. <https://doi.org/10.21037/jlpm.2018.05.01>.
- [12] P. Bankhead, MB Loughrey, JA Fernández, and et al. Qupath: Open source software for digital pathology image analysis. *Scientific Reports*, 7(1):16878, 2017. <https://doi.org/10.1038/s41598-017-17204-5>.
- [13] L. Pantanowitz, J. H. Sinard, and WH Henricks. *Medical Laboratory Informatics*. American Society of Clinical Pathologists Press, 2014. <https://doi.org/10.1016/j.cll.2007.07.011>.
- [14] J. Ho. *Digital Pathology: A Comprehensive Guide to the Future of Precision Medicine*. Springer International Publishing, 2018. <https://doi.org/10.4103/2153-3539.255259>.
- [15] A. Madabhushi and G. Lee. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis*, 33:170–175, 2016. <https://doi.org/10.1016/j.media.2016.06.037>.
- [16] E. A. Krupinski and R. S. Weinstein. Digital pathology: Image analysis and beyond. *The Journal of Pathology Informatics*, 8:27, 2017. <https://doi.org/10.4103/2153-3539.68332>.
- [17] A. Momeni-Boroujeni and F. Saedi. Digital pathology: Current status and future perspectives. *Journal of Medical Systems*, 42(4):70, 2018. <https://doi.org/10.1111/j.1365-2559.2011.03814.x>.
- [18] Md Hossain, Galib Muhammad Shahriar Himel, Mahbubul Syeed, Faisal Uddin, Mahady Hasan, and Rubina Bari. Tissue artifact segmentation and sever-

- ity assessment for automatic analysis using wsi. 11:21977–21991, 2023. <https://doi.org/10.48550/arXiv.2401.01386>.
- [19] L. Pantanowitz, A. Sharma, AB Carter, and et al. Twenty years of digital pathology: An overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. *Journal of Pathology Informatics*, 9:40, 2018. [://doi.org/10.1038/s41746-024-01106-8](https://doi.org/10.1038/s41746-024-01106-8).
- [20] N. Farahani, AV Parwani, and L. Pantanowitz. Whole slide imaging in pathology: Advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International*, 7:23–33, 2015. <https://doi.org/10.2147/PLMI.S59826>.
- [21] Ostrom QT, Gittleman H, Truitt G, and et al. Cbtrus statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2011-2015. *Neuro Oncol*, 20(suppl\_4):iv1–iv86, 2018. <https://doi.org/10.1093/neuonc/noy131>.
- [22] Argyriou AA, Karnezis S, Makridou A, and et al. Brain cancer diagnosis: The role of the neuropathologist. *Oncol Lett*, 15(1):3–5, 2018. <https://doi.org/10.3892/ol.2017.7411>.
- [23] Arvind V, Chennupati SK, Pillai V, and et al. Computer-aided diagnosis of brain tumor necrosis using random subspace method on mri images. *PLoS One*, 13(7):e0199617, 2018. <https://doi.org/10.1371/journal.pone.0199617>.
- [24] Razzak MI, Naz S, and Zaib A. Deep learning for medical image processing: Overview, challenges and future. *IEEE Access*, 6:92460–92490, 2018. <https://doi.org/10.48550/arXiv.1704.06825>.
- [25] Amanatiadis A, Kandaris D, Giatili S, and et al. Whole slide imaging for digital pathology: Advantages and limitations. *Arch Comput Methods Eng*, 25(2):399–418, 2018. <https://doi.org/10.3390/jcm9113697>.

- [26] S. Khan et al. A review of feature extraction techniques in medical image processing. *Journal of Healthcare Engineering*, 2018:2715395, 2018. [https://doi.org/10.1007/978-3-030-46943-6<sub>3</sub>](https://doi.org/10.1007/978-3-030-46943-6_3).
- [27] M. H. Fathy and M. H. Korashy. A comparative study of image feature extraction techniques for detection of colon polyps. In *2019 3rd International Conference on Communications, Signal Processing, and Their Applications (ICCS-PA)*, pages 1–6, 2019. <https://doi.org/10.5815/ijigsp.2015.03.03>.
- [28] J. Chen et al. A simple and effective pathology image retrieval method based on multi-feature fusion. *IEEE Access*, 8:4224–4234, 2020. <https://doi.org/10.1108/24699322.2018.1560087>.
- [29] G. Singh and H. B. Kekre. Image retrieval using glcm based texture features. In *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, pages 1006–1010, 2019. <https://doi.org/10.1109/ICACCS.2019.8728411>.
- [30] K. C. Santosh and B. K. Tripathy. Exploring glcm and lbp features for histopathological image analysis. In *2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, pages 1–4, 2019. <https://doi.org/10.1109/ANTS47819.2019.9118117>.
- [31] X. Li and et al. A multi-modality medical image fusion method based on multi-scale feature learning. *IEEE Access*, 8:19660–19670, 2020. <https://doi.org/10.1155/2021/1544955>.
- [32] S. A. N. Waheed and et al. Deep learning based automated diagnosis of breast cancer using histopathological images. *Journal of Healthcare Engineering*, 2020:Article ID 8896867, 12 pages, 2020. <https://doi.org/10.3390/healthcare10010010>.

- [33] T. Ahonen and et al. Face recognition with local binary patterns. In *Computer Vision - ECCV 2004*, pages 469–481, 2004. <https://doi.org/10.1109/IIHMSP.2010.87>.
- [34] B. C. Russell and A. Torralba. Building a database of texture in natural scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(4):747–761, 2010. <https://doi.org/10.1109/TPAMI.2009.118>.
- [35] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. <https://doi.org/10.1109/5.726791>.
- [36] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [37] F. Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017. <https://doi.org/10.1109/CVPR.2017.195>.
- [38] K. Srinivasan and et al. Performance comparison of deep cnn models for detecting driver’s distraction. *CMC - Tech Science Press*, 68:4109–4124, 2021. <https://doi.org/10.32604/cmc.2021.016736>.
- [39] J. Deng and et al. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. <https://doi.org/10.1109/CVPR.2009.5206848>.
- [40] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. <https://doi.org/10.1007/BF00994018>.
- [41] G. Wang, Y. Huang, L. Yao, W. Jiang, L. Wang, and X. Zeng. Using machine learning to detect diseases from medical data. *Journal of Healthcare Engineering*, 2019:1–11, 2019. <https://doi.org/10.2174>

- [42] A. Rajkomar, J. Dean, and I. Kohane. Machine learning in medicine. *New England Journal of Medicine*, 380(14):1347–1358, 2019. <https://doi.org/10.1056/NEJMra1814259>.
- [43] Yogesh Kumar and et al. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of Ambient Intelligence and Humanized Computing*, 14(7):8459–8486, 2023. [://doi.org/10.1007/s12652-021-03612-z](https://doi.org/10.1007/s12652-021-03612-z).
- [44] S. G. Honavar. Artificial intelligence in ophthalmology - machines think! *Indian J. Ophthalmol.*, 70(4):1075–1079, 2022. [https://doi.org/10.4103/ijo.ijo\\_644\\_2](https://doi.org/10.4103/ijo.ijo_644_2).
- [45] T. Davenport and R. Kalakota. The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2):94–98, 2019. <https://doi.org/10.7861/futurehosp.6-2-94>.
- [46] A. Hosny and et al. Artificial intelligence in radiology. *Nature Reviews Cancer*, 18(8):500–510, 2018. <https://doi.org/10.1038/s41568-018-0016-5>.
- [47] D. K. Sharma and et al. Deep learning applications for disease diagnosis. In *Deep Learning for Medical Applications with Unique Data*, page 31–51. Elsevier, 2022. <https://doi.org/10.1016/b978-0-12-824145-5.00005-8>.
- [48] A. Samaras and et al. Artificial intelligence-based mining of electronic health record data to accelerate the digital transformation of the national cardiovascular ecosystem: design protocol of the cardiomining study. *BMJ Open*, 13(4):E068698, 2023. <https://doi.org/10.1136/bmjopen-2022-068698>.
- [49] J. G. Richens, C. M. Lee, and S. Johri. Author correction: Improving the accuracy of medical diagnosis with causal machine learning. *Nature Communications*, 12(1):2108, 2021. <https://doi.org/10.1038/s41467-020-17419-7>.

- [50] N. N. Khanna and et al. Economics of artificial intelligence in healthcare: Diagnosis vs. treatment. *Healthcare (Basel)*, 10(12):2493, 2022. <https://doi.org/10.3390/healthcare10122493>.
- [51] K. Liopyris and et al. Artificial intelligence in dermatology: Challenges and perspectives. *Dermatology and Therapy*, 12(12):2637–2651, 2022. <https://doi.org/10.1186/s13000-023-01375-z>.
- [52] S. Shafi and A. V. Parwani. Artificial intelligence in diagnostic pathology. *Diagnostic Pathology*, 18(1):109, 2023. <https://doi.org/10.1186/s13000-023-01375-z>.
- [53] S. Doyle and et al. Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2008. <https://doi.org/10.1109/ISBI.2008.4541077>.
- [54] L. He and et al. Histology image analysis for carcinoma detection and grading. *Computer Methods and Programs in Biomedicine*, 107(3):538–556, 2012. <https://doi.org/10.1016/j.cmpb.2011.12.005>.
- [55] L. M. da Silva and et al. Independent real-world application of a clinical-grade automated prostate cancer detection system. *The Journal of Pathology*, 254(2):147–158, 2021. <https://doi.org/10.1002/path.5665>.
- [56] Rahmadwati and et al. Cervical cancer classification using gabor filters. In *2011 IEEE First International Conference on Healthcare Informatics, Imaging and Systems Biology*, 2011. <https://doi.org/10.1109/HISB.2011.11>.
- [57] A. B. Tosun and et al. Object-oriented texture analysis for the unsupervised segmentation of biopsy images for cancer detection. *Pattern Recognition*, 42(6):1104–1112, 2009. <https://doi.org/10.1016/j.patcog.2008.08.016>.

- [58] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener. Histopathological image analysis: a review. *Medical Image Analysis*, 2:147–171, October 2009. <https://doi.org/10.1016/j.media.2009.05.004>.
- [59] J. Xu and et al. Deep learning for histopathological image analysis: Towards computerized diagnosis on cancers. In *Deep Learning and Convolutional Neural Networks for Medical Image Computing, Advances in computer vision and pattern recognition*, page 73–95. Springer International Publishing, 2017. [https://doi.org/10.1007/978-3-319-42999-1\\_6](https://doi.org/10.1007/978-3-319-42999-1_6).
- [60] C. Gkogkou and et al. Necrosis and apoptotic index as prognostic factors in non-small cell lung carcinoma: a review. *Springerplus*, 3(1):120, 2014. <https://doi.org/10.1186/2193-1801-3-120>.
- [61] K. Chen and et al. Pathological features and prognostication in colorectal cancer. *Current Oncology*, 28(6):5356–5383, 2021. <https://doi.org/10.3390/curroncol28060452>.
- [62] M. Culpan and et al. The prognostic impact of tumor necrosis in non-muscle invasive bladder cancer. *Rev. Assoc. Med. Bras.*, 68(11):1587–1592, 2022. <https://doi.org/10.1590>
- [63] G. Bredholt and et al. Tumor necrosis is an important hallmark of aggressive endometrial cancer and associates with hypoxia, angiogenesis, and inflammation responses. *Oncotarget*, 6(37):39676–39691, 2015. <https://doi.org/10.18632/oncotarget.5667>.
- [64] P. P. Yee and W. Li. Tumor necrosis: A synergistic consequence of metabolic stress and inflammation. *Bioessays*, 43(7):e2100029, 2021. <https://doi.org/10.1002/bies.202100029>.
- [65] S. Shafi and A. V. Parwani. Artificial intelligence in diagnostic pathology. *Diagnostic Pathology*, 18(1):109, 2023. <https://doi.org/10.1186/s13000-023-01375-z>.

- [66] S. Sengupta and et al. Histologic coagulative tumor necrosis as a prognostic indicator of renal cell carcinoma aggressiveness. *Cancer*, 104(3):511–520, 2005. <https://doi.org/10.1002/cncr.21204>.
- [67] R. F. Walker and et al. Classification of cervical cell nuclei using morphological segmentation and textural feature extraction. In *Proceedings of ANZIIS '94 - Australian New Zealand Intelligent Information Systems Conference*, 2002. <https://doi.org/10.1109/ANZIIS.1994.396981>.
- [68] T. Aitazaz and et al. Transfer learning for histopathology images: an empirical study. *Neural Computing and Applications*, 35(11):7963–7974, 2023. <https://doi.org/10.1007/s00521-022-07516-7>.
- [69] M. N. Gurcan and et al. Histopathological image analysis: a review. *IEEE Reviews in Biomedical Engineering*, 2:147–171, 2009. <https://doi.org/10.1109/RBME.2009.2034865>.
- [70] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener. Histopathological image analysis: a review. *Medical Image Analysis*, 2:147–171, October 2009. <https://doi.org/10.1016/j.media.2009.05.004>.
- [71] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener. Histopathological image analysis: a review. *Medical Image Analysis*, 2:147–171, Oct 2009. <https://doi.org/10.1109/RBME.2009.2034865>.
- [72] K. Kayser and H. Hoffgen. Pattern recognition in histopathology by orders of textures. *Medical Informatics*, 9(1):55–59, 1984. <https://doi.org/10.3109/14639238409010938>.
- [73] Taranpreet Rai, Ambra Morisi, Barbara Bacci, Nicholas J Bacon, Michael J Dark, Tawfik Aboellail, Spencer Angus Thomas, Miroslaw Bober, Roberto La Ragione, and Kevin Wells. Deep learning for necrosis detection using canine perivascu-

- lar wall tumour whole slide images. *Scientific Reports*, 12(1):10634, Jun 2022. <https://doi.org/10.1038/s41598-022-13928-1>.
- [74] Ji-Hee Hwang, Minyoung Lim, Gyeongjin Han, Heejin Park, Yong-Bum Kim, Jinseok Park, Sang-Yeop Jun, Jaeku Lee, and Jae-Woo Cho. A comparative study on the implementation of deep learning algorithms for detection of hepatic necrosis in toxicity studies. *Toxicological Research*, 39(3):399–408, Jul 2023. <://doi.org/10.1007/s43188-023-00173-5>.
- [75] Ambra Morisi, Taran Rai, Nicholas J Bacon, Spencer A Thomas, Miroslaw Bober, Kevin Wells, Michael J Dark, Tawfik Aboellail, Barbara Bacci, and Roberto M La Ragione. Detection of necrosis in digitised whole-slide images for better grading of canine soft-tissue sarcomas using machine-learning. *Veterinary Sciences*, 10(1):45, Jan 2023. <https://doi.org/10.3390/vetsci10010045>.
- [76] Abdul Hannan Khan, Sagheer Abbas, Muhammad Adnan Khan, Umer Farooq, Wasim Ahmad Khan, Shahan Yamin Siddiqui, and Aiesha Ahmad. Intelligent model for brain tumor identification using deep learning. *Applied Computational Intelligence and Soft Computing*, 2022:1–10, Jan. <https://doi.org/10.1155/2022/8104054>.
- [77] Kyung Hwan Kim, Sangkeun Jung, Han-Joo Lee, Hyon-Jo Kwon, Seung-Won Choi, Hyeon-Song Koh, Jin-Young Youm, and Seon-Hwan Kim. A deep neural network-based model predicting peritumoral edema after radiosurgery for meningioma. *World Neurosurgery*, 164:e280–e289, Aug 2022. <https://doi.org/10.1016/j.wneu.2022.04.125>.
- [78] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.

- [79] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, and et al. A guide to deep learning in healthcare. *Nature Medicine*, 25(1):24–29, 2019. <https://doi.org/10.3390/mi13122197>.
- [80] Y. Wang, C. Huang, Y. Peng, Y. Chen, and Y. Li. Deep learning applications in histopathological image analysis: Challenges and opportunities. *Frontiers in Bioengineering and Biotechnology*, 9:672907, 2021. <https://doi.org/10.3389/fbioe.2021.672907>.
- [81] Alhassan Ali Ahmed, Mohamed Abouzid, and Elzbieta Kaczmarek. Deep learning approaches in histopathology. *Cancers (Basel)*, 14(21):5264, Oct 2022. <https://doi.org/10.3390/cancers14215264>.
- [82] Karin Stacke, Gabriel Eilertsen, Jonas Unger, and Claes Lundstrom. A closer look at domain shift for deep learning in histopathology. *arXiv*, 2019. <https://doi.org/10.48550/arXiv.1909.11575>.
- [83] Maxime W Lafarge, Josien P W Pluim, Koen A J Eppenhof, and Mitko Veta. Learning domain-invariant representations of histological images. *Front. Med. (Lausanne)*, 6:162, Jul 2019. <https://doi.org/10.3389/fmed.2019.00162>.
- [84] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, and et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017. <https://doi.org/10.1016/j.media.2017.07.005>.
- [85] A. Janowczyk and A. Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7:29, 2016. <https://doi.org/10.4103/2153-3539.186902>.
- [86] D. Wang, A. Khosla, R. Gargya, H. Irshad, and A. H. Beck. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv:1606.05718*, 2016. <https://doi.org/10.48550/arXiv.1606.05718>.

- [87] Xiaozheng Xie, Jianwei Niu, Xuefeng Liu, Zhengsu Chen, Shaojie Tang, and Shui Yu. A survey on incorporating domain knowledge into deep learning for medical image analysis. *Med. Image Anal.*, 69:101985, Apr 2021. <https://doi.org/10.1016/j.media.2021.101985>.
- [88] Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Comput. Struct. Biotechnol. J.*, 16:34–42, Feb 2018. <https://doi.org/10.1016/j.csbj.2018.01.001>.
- [89] C. Bouvier, N. Souedet, J. Levy, C. Jan, Z. You, A-S. Herard, G. Mer-goil, B H. Rodriguez, C. Clouchoux, and T. Delzescaux. Reduced and stable feature sets selection with random forest for neurons segmentation in histological images of macaque brain. *Sci. Rep.*, 11(1):22973, Nov 2021. <https://doi.org/10.1038/s41598-021-02344-6> .
- [90] M. A. Aswathy and M. Jagannath. An svm approach towards breast cancer classification from h&e-stained histopathology images based on integrated features. *Med. Biol. Eng. Comput.*, 59(9):1773–1783, 2021. <https://doi.org/10.1007/s11517-021-02403-0> .
- [91] Rachana Yogitha, N. G. Milina, Sabhahit, and M. Ainain. Breast histopathology image classification: A comparative evaluation of svm, random forest and cnn. In *2023 3rd International Conference on Intelligent Technologies (CONIT)*. IEEE, 2023. <https://doi.org/10.1109/CONIT59222.2023.10205567>.
- [92] K. Gupta and N. Chawla. Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained cnn. *Procedia Computer Science*, 167:878–889, 2020. <https://doi.org/10.1016/j.procs.2020.03.270>.
- [93] Y. Zhou, C. Zhang, and S. Gao. Breast cancer classification from histopathological images using resolution adaptive network. *IEEE Access*, 10:35977–35991, 2022. <https://doi.org/10.1155/2022/8904768>.

- [94] A. Bagchi, P. Pramanik, and R. Sarkar. A multi-stage approach to breast cancer classification using histopathology images. *Diagnostics (Basel)*, 13(1):126, 2022. <https://doi.org/10.3390/diagnostics13010126>.
- [95] W. Hu, H. Chen, W. Liu, X. Li, H. Sun, X. Huang, M. Grzegorzek, and C. Li. A comparative study of gastric histopathology sub-size image classification: From linear regression to visual transformer. *Frontiers in Medicine*, 9:1072109, 2022. <https://doi.org/10.3389/fmed.2022.1072109>.
- [96] Z. Wang, J. Gao, H. Kan, Y. Huang, F. Tang, and W. Li. Resnet for histopathologic cancer detection, the deeper, the better? *Journal of Data Science and Intelligent Systems*, 2023. <https://doi.org/10.47852/bonviewJDSIS3202744>.
- [97] M. Talo. Convolutional neural networks for multi-class histopathology image classification. *ArXiv*, 2019. <https://doi.org/10.48550/arXiv.1903.10035>.
- [98] Y. Jiang, L. Chen, H. Zhang, and X. Xiao. Breast cancer histopathological image classification using convolutional neural networks with small se-resnet module. *PLoS One*, 14(3):e0214587, 2019. <https://doi.org/10.1371/journal.pone.0214587>.
- [99] Y. Deng, J. Yin, Y. Wang, J. Chen, L. Sun, and Q. Li. Resnet-50 based method for cholangiocarcinoma identification from microscopic hyperspectral pathology images. *Journal of Physics: Conference Series*, 1880(1):012019, 2021. <https://doi.org/10.1088/1742-6596/1880/1/012019>.
- [100] M. Talo. Automated classification of histopathology images using transfer learning. *Artificial Intelligence in Medicine*, 101:101743, 2019. <https://doi.org/10.1016/j.artmed.2019.101743>.
- [101] X. Lu and Y. A. Firoozeh Abolhasani Zadeh. Deep learning-based classification for melanoma detection using xceptionnet. *Journal of Health Engineering*, 2022:2196096, 2022. <https://doi.org/10.1155/2022/2196096>.

- [102] S. Sharma and S. Kumar. The xception model: A potential feature extractor in breast cancer histology images classification. *ICT Express*, November 2021. <https://doi.org/10.1016/j.icte.2021.11.004>.
- [103] J. A. A. Jothi and K. Damania. Dirxnet: A hybrid deep network for classification of breast histopathology images. *SN Computer Science*, 5(1), December 2023. <https://doi.org/10.1007/s42979-023-02413-9>.
- [104] N. Madusanka, P. Jayalath, D. Fernando, L. Yasakethu, and B.-I. Lee. Impact of he stain normalization on deep learning models in cancer image classification: Performance, complexity, and trade-offs. *Cancers (Basel)*, 15(16), August 2023. <https://doi.org/10.3390/cancers15164144>.
- [105] N. Dif and Z. Elberrichi. A new intra fine-tuning method between histopathological datasets in deep learning. *International Journal of Service Science, Management, Engineering, and Technology*, 11(2):16–40, April 2020. <https://ideas.repec.org/a/igg/jssmet/v11y2020i2p16-40.html>.
- [106] T. Aitazaz, A. Tubaishat, F. Al-Obeidat, B. Shah, T. Zia, and A. Tariq. Transfer learning for histopathology images: an empirical study. *Neural Computing and Applications*, 35(11):7963–7974, 2023. <https://doi.org/10.1007/s00521-022-07516-7>.
- [107] J. Saltz, R. Gupta, L. Hou, T. Kurc, P. Singh, V. Nguyen, and et al. Spatial organisation and molecular correlation of tumour-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep.*, 23(1):181–193.e7, 2018. <https://doi.org/10.1016/j.celrep.2018.03.086>.
- [108] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the*

- IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017. <https://doi.org/10.1109/CVPR.2017.369>.
- [109] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, 2017. <https://doi.org/10.1038/nature21056>.
- [110] J. Kim and Y. C. Kim. Interpretable deep neural network for classification of alzheimer’s disease using hippocampal magnetic resonance imaging. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 2154–2162, 2018. <https://doi.org/10.1109/CVPRW.2018.00275>.
- [111] J. Z. Cheng, D. Ni, Y. H. Chou, J. Qin, C. M. Tiu, Y. C. Chang, and et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in ct scans. *Sci Rep.*, 6(1):24454, 2016. <https://doi.org/10.1038/srep24454>.
- [112] M. W. Hart. Histological grading of tumours of the central nervous system. *Acta Neuropathol.*, 1(3):316–323, 1962. <https://doi.org/10.1007/BF00689597>.
- [113] R. R. N. Pitts and S. H. Cooperberg. Modern pathology: anatomic pathology. *Modern Pathology*, 15(4):432–443, 2002. <https://doi.org/10.1038/modpathol.3880564>.
- [114] A. Perry, D. N. Louis, A. Scheithauer, B. W. Scheithauer, and D. W. R. B. W. Budka. Meningiomas. *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of the Nervous System*, 5:176–184, 2000. <https://doi.org/10.1093/jnen/61.3.215>.
- [115] R. A. Green and D. A. Kroemer. Necrosis: What is it, how is it caused, and how can it be prevented? *Nat. Rev. Mol. Cell Biol.*, 5(7):647–655, 2004. <https://doi.org/10.1038/nrm1432>.

- [116] J. P. Higgins and P. A. R. Anderson. Morphology and biology of meningiomas. In *Meningiomas*, pages 1–17. Springer New York, 2009. <https://doi.org/10.1007/BF00165650>.
- [117] D. J. Brat, J. E. Parisi, B. K. Kleinschmidt-DeMasters, A. T. Yachnis, T. J. Monterey, P. J. Boyer, and et al. Surgical neuropathology update. a textbook of cns pathology. *Sci. Rep.*, 11(1):22973, Nov 2021. <https://doi.org/10.5858/2008-132-993-SNUARO>.
- [118] H. S. et al. Levin. High-grade meningioma: the importance of histologic features on survival. *J. Neurooncol.*, 10(2):107–118, 1991. <https://doi.org/10.1007/BF00177405>.
- [119] J. A. Lee, S. J. Kim, J. H. Kim, S. H. Park, C. H. Kim, and H. K. Kim. Atypical meningioma: Mri, ct, and angiographic findings in 23 patients. *J. Comput. Assist—Tomogr.*, 22(2):215–221, 1998. <https://doi.org/10.1097/00004728-199803000-00008>.
- [120] M. P. Hedley-Whyte. Meningioma. In G. Paxinos, editor, *The Human Nervous System: A Handbook for Students and Clinicians*, pages 1729–1732. Academic Press, 1990. <https://doi.org/10.1016/B978-0-12-547620-8.50066-1>.
- [121] D. Magee, D. Treanor, D. Crellin, M. Shires, and J. Smith. Reliable automated measurement of er/pr and her2 receptor expression in breast cancer. *Human Pathology*, 40(4):609–617, 2009. <https://doi.org/10.1016/j.humpath.2008.09.014>.
- [122] A. Janowczyk and A. Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7:29, 2016. <https://doi.org/10.4103/2153-3539.183390>.
- [123] A. Goode, B. Gilbert, J. Harkes, D. Jukic, M. Satyanarayanan, and A. Sharma. Openslide: A vendor-neutral software foundation for digital pathology. *Jour-*

- nal of Pathology Informatics*, 4:27, 2013. <https://doi.org/10.4103/2153-3539.119005>.
- [124] Yinhai Wang, Kathleen Williamson, Paul Kelly, Jacqueline James, and Peter Hamilton. Surfaceslide: A multitouch digital pathology platform. *PloS one*, 7:e30783, 2012. <https://doi.org/10.1371/journal.pone.0030783>.
- [125] Daisuke Komura and Shigeto Seno. Deep learning for microscopic image analysis: A comprehensive tutorial with selected use cases. *Pattern Recognition*, 105:107316, 2020. <https://doi.org/10.1016/j.patcog.2020.107316>.
- [126] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, and et al. A survey on deep learning in medical image analysis. *Med. Image Anal.*, 42:60–88, 2017. <https://doi.org/10.1016/j.media.2017.07.005>.
- [127] Shapely: Python package for set-theoretic analysis and manipulation of planar features. Available at: <https://pypi.org/project/Shapely/>.
- [128] Geos: Geometry engine, open source - a c++ port of the java topology suite (jts). Available at: <https://trac.osgeo.org/geos/>.
- [129] The glioblastoma multiforme (GBM) and low-grade glioma (LGG) data utilized in this study were obtained from the official TCGA website: <https://www.cancer.gov/tcga>.
- [130] Hitesha Mukherjee, T.K Srikanth, Ramesh Kestur, Shilpa Rao, Nandeesh, and Anitha Mahadevan. Necrosis detection in histopathological images of meningioma. Unpublished masters dissertation, International Institute of Information Technology, Bangalore, 2022.
- [131] U. Schmidt, M. Weigert, C. Broaddus, and G. Myers. Cell detection with star-convex polygons. In *Medical Image Computing and*

- Computer Assisted Intervention – MICCAI 2018*, pages 265–273, 2018.  
<https://doi.org/10.48550/arXiv.1806.03535>.
- [132] D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, Cavenee, and H. Ohgaki. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*, 131(6):803–820, 2016. <https://doi.org/10.1007/s00401-016-1545-1>.