Multi-Scale Deep Learning Models for Pathology Image Analysis

ISLAM ALZOUBI

Doctor of Philosophy (Computer Science)

Supervisor: Xiuying Wang
Associate Supervisor: Yuk Chung

A thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy

School of Computer
Science
The University of Sydney
Australia

15 November 2024

Declaration

This is to certify that to the best of my knowledge; the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes. I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Abstract

Tumors exhibit significant morphological heterogeneity at the macroscopic, histological, and cellular levels, which complicates their analysis. Pathological examination remains the gold standard for cancer diagnosis, but manual microscopic evaluation of whole slide images (WSIs) is time-consuming, labor-intensive, and prone to error. Recent advancements in digital imaging have led to the emergence of computational histopathology, which aims to enhance clinical workflows by leveraging advanced image processing and analysis techniques.

Early computational approaches relied on manually extracted, handcrafted features from histopathological images, which were limited in their generalizability and robustness. The advent of artificial intelligence (AI), particularly deep learning (DL), has transformed this process by automating feature extraction, improving model accuracy, and enhancing the efficiency of cellular and tissue-level analysis. These techniques have demonstrated significant potential in supporting cancer diagnosis and prognosis.

However, significant challenges remain in capturing and interpreting the morphological heterogeneity across different scales, from individual cells to larger tissue structures.

- **Cellular-Level Analysis:** Variations in staining techniques and tissue preparation can create inconsistencies across datasets, complicating the development of generalizable cell detection models. Overlapping cells, blurred boundaries, and artifacts also hinder accurate segmentation. Furthermore, patch-based convolutional neural network (CNN) approaches often result in a loss of contextual information, limiting their effectiveness in capturing the broader tissue architecture. On the other hand, incorporating information from different imaging modalities (e.g., immunohistochemical and histological data) presents another challenge due to differences in resolution, contrast, and staining

techniques. These variations make it difficult to align and analyze data across modalities consistently, hindering a comprehensive understanding of the relationship between cellular markers and tissue structure.

- **Tissue-Level Analysis and Survival Prediction:** Despite the promise of CNNs in predicting patient outcomes, they face significant limitations in capturing global contextual information and modeling complex tissue compositions. This inefficiency arises when dealing with relation-aware representations, leading to challenges in processing gigapixel-sized WSIs. However, processing and analyzing WSIs, presents a major challenge due to their large data size and the complex spatial relationships between cells. Efficiently encoding these massive images while preserving the essential spatial and contextual information is crucial but computationally intensive. Additionally, tissue sections exhibit significant spatial and morphological heterogeneity, which is further complicated by the organization of cells (e.g., cancer clones). Accurately modeling and interpreting these relationships, especially in the context of cancer prognosis and survival prediction, is a complex task.
- **Multi-Graph Modeling and Cancer Grading**: Traditional single-graph approaches are typically limited to small structures and often struggle to capture the intricate relationships within tissue microenvironments effectively. Accurately modeling these complex interactions is crucial, as both local and global structures influence tumor grading. Furthermore, integrating multiple graphs, each representing different aspects of tissue microenvironments, such as cellular morphology and spatial organization, poses additional challenges in maintaining key information and identifying relevant features for cancer grading.

Addressing these challenges is essential for unlocking the full potential of deep learning in providing reliable, fast, and accurate decision-making in cancer diagnosis. This thesis presents a multi-stage framework for comprehensive WSI analysis at both cellular and tissue levels to enhance tumor detection, classification, and survival prediction. Specifically, our major contributions include:

- **At the Cellular-Level,** we proposed a multi-scale CNN-based model enhances cellular-level detection and segmentation by capturing both fine-grained cellular details and broader tissue context, improving accuracy in identifying and counting cells in WSIs. Additionally, a fusion method based on Discrete Wavelet Transform (DWT) was developed for cross-modality analysis, enabling the exploration of relationships between immunohistochemical signals and histological features.
- **At the Tissue-Level and for Survival Prediction,** we developed an innovative method to encodes gigapixel WSIs into a memory-efficient representation that maintains critical spatial relationships while significantly reducing data size. This facilitates the creation of graph-based structures to model the intricate organization of cells within tissue sections. Furthermore, we implemented a graph-based attention mechanism that adaptively adjusts to feature similarities and spatial proximities, enhancing the model's capacity to capture tissue heterogeneity and predict patient survival. The insights gained from attention distribution provide a deeper understanding of key regions within the tissue, offering valuable guidance for prognosis and advancing pathology research. To further improve performance, we introduced a novel loss function that encourages balanced attention distribution, resulting in enhanced image-level classification accuracy.
- **Multi-Graph Modeling and Cancer Grading**, we proposed a multi-graph framework which reduces reliance on the quality of a single graph and enables the integration of multiple graphs to capture complex tissue relationships and combining multilevel structural information. The model employs advanced fusion techniques learns to balance the contributions of different graphs, addressing the uncertainty of their relative importance, ensuring that the most relevant features from each graph are prioritized, and enhancing the overall classification performance. Furthermore, a dual-level loss optimization strategy is proposed, incorporating intra-graph and inter-graph similarity measures, ensuring consistency and robustness in the learned embeddings and significantly enhancing the model's performance.

Publications

Journal Papers

1. **Islam Alzoubi**, Guoqing Bao, Rong Zhang, Christina Loh, Yuqi Zheng, Svetlana Cherepanoff, Gary Gracie et al. " An Open-Source AI Framework for the Analysis of Single Cells in Whole-Slide Images with a Note on CD276 in Glioblastoma." *Cancers (Basel),* vol. 14, no. 14, Jul 15, 2022, doi: 10.3390/cancers14143441.

2. **Islam Alzoubi**, Guoqing Bao, Yuqi Zheng, Xiuying Wang, and Manuel B. Graeber. " Artificial Intelligence Techniques for Neuropathological Diagnostics and Research." *Neuropathology,* vol. 43, Nov 28, 2022, doi: 10.1111/neup.12880.

3. **Islam Alzoubi**, Lin Zhang, Yuqi Zheng, Christina Loh, Xiuying Wang, and Manuel B. Graeber. " PathoGraph: An Attention-Based Graph Neural Network Capable of Prognostication Based on CD276 Labeling of Malignant Glioma Cells." *Cancers,* vol. 16, no. 4, Feb 11, 2024, p.750.

4. Enqi Liu, Lin Zhang, **Islam Alzoubi**, Haneya Fuse, Manuel B. Graeber, and Xiuying Wang. " A Self-Adaptive Framework for Efficient Cell Detection and Segmentation in Histopathological Images with Minimal Expert Input." In *Australasian Joint Conference on Artificial Intelligence*, Nov 20, 2024, pp. 323-334. Singapore: Springer Nature Singapore.

5. **Islam Alzoubi**, Bowen Xin, Rolf Bjerkvig, Jian Wang, and Xiuying Wang, 2024. "An

Attentive Multi-Graph Fusion for Tumor Grading in Pathology Images" (under review).

Poster presentation:

1. Submitted a poster presentation to the Brain and Mind Centre (BMC) Symposium, held on 4–5 November 2021.

2. Submitted abstract to Digital Health Week 2022, held in February 2022.

   Islam Alzoubi, Rong Zhang, Guoqing Bao, Christina Loh, Svetlana Cherepanoff, Michael Buckland, Xiuying Wang, and Manuel B. Graeber. "Deep Learning for Automated Detection of Cellular Profiles in Whole-slide Immunohistochemical Images." *Digital Health Week*, vol. 1, no. 1, Jul 27, 2022.

3. Submitted a poster presentation to the Brain and Mind Centre (BMC) Symposium, held on 31 November 2022.

4. Presented an oral presentation in 2023 Brain Tumour Research Symposium, held on 20 September 2023.

Authorship Attribution Statement

Chapter 1&2 of this thesis is published as **Publication 2**. I designed the review framework with the assistance of my co-author, collaborated to identify key areas of focus, conducted a thorough literature review, analyzed the existing research and wrote the manuscript.

Chapter 3 of this thesis is published as **Publication 1**. I designed the study, developed the methodology, analyzed the data, conducted the experiments and wrote the manuscript.

Chapter 4 of this thesis is published as **Publication 3**. I designed the study, developed the methodology, analyzed the data, conducted the experiments and wrote the manuscript.

Chapter 5 of this thesis is 'under review' as **Publication 5**. I designed the study, conducted the experiments, analyzed the data, conducted the experiments and wrote the manuscript.

Acknowledgements

I would like to express my deepest gratitude to the University of Sydney for awarding me the **University of Sydney International Stipend Scholarship** and the **University of Sydney Tuition Fee Scholarship**. These scholarships have provided invaluable financial support throughout my studies, allowing me to fully focus on my research and academic pursuits. I am sincerely appreciative of the opportunities and resources these scholarships have afforded me.

Acknowledgements

I am deeply thankful to everyone who provided invaluable support and guidance throughout my Ph.D. journey and in the completion of this thesis.

First and foremost, I would like to express my heartfelt thanks to my supervisor, **Assoc. Prof. Xiuying Wang**, for her unwavering support, insightful advice, and patient guidance. Her immense knowledge and professional expertise have been a source of inspiration throughout my research. Without her constant encouragement, meticulous feedback, and invaluable mentorship, this thesis would not have been possible.

Furthermore, I extend my deepest gratitude to all fellow students in our research groups for their treasured support, encouragement, and companionship along the way. Their collaboration and shared enthusiasm for research have been invaluable.

I wish to express my deepest thanks to my family, especially **my parents**, and my husband, whose endless love and faith in me have sustained me through every step of this journey. To my husband **Musab**, who has stood by my side throughout this entire process, your unwavering support has been my greatest strength, and I am incredibly grateful to have shared this incredible journey with you. Last but not least, I would like to dedicate this thesis to my little kids, **Mariam and Kareem**.

Abbreviations

ACC: Accuracy
AMGF-GNN: Adaptive Multi-Graph Fusion based Graph Neural Network
AI: Artificial Intelligence
ANNs: Artificial Neural Networks
AUC: Area Under the Curve

BCNN: Bifocal Convolutional Neural Network

CAD: Computer Aided Detection
CGNNs: Cell Graph Neural Networks
CNN: Convolutional Neural Network
CT: Tomography

DL: Deep Learning
DT: Decision Tree
DWT: Discrete Wavelet Transform

ED: Euclidean Distance

F1: F1 Score
FC: Fully Connected
FCN: Fully Convolutional Network

GAN: Generative Adversarial Network
GBM: Glioblastoma
GCN: Graph Convolutional Network
GIN: Graph Isomorphic Network
GLCM: Gray Level Co-Occurrence Matrix
GNN: Graph Neural Network
GONNs: Genetically Optimized Neural Networks
GRUs: Gated Recurrent Units
GSCs: Glioblastoma Stem Cells

H&E: Hematoxylin and Eosin

HSIs: Hyperspectral Images

IDC: Invasive Ductal Carcinoma
IHC: Immunohistochemistry

KM: Kaplan-Meier
KNN: K-Nearest Neighbor

LGG: Lower Grade Glioma
LLSMD: Linear Least Squares Minimum Distance
LSTM: Long Short-Term Memory

MIL: Multi Instance Learning
MIS: Multiple Instance Selection
ML: Machine Learning
MLP: Multi-Layer Perceptron
MRI: Magnetic Resonance Imaging

NA: Nelson-Aalen Estimator

RF: Random Forest
RNN: Recurrent Neural Network
ROC: Receiver Operating Characteristic
ROI: Region of Interest

SGD: Stochastic Gradient Descent
SVM: Support Vector Machine

TCGA: The Cancer Genome Atlas
TME: Tumor Microenvironment

WHO: World Health Organization
WSIs: Whole Slide Images

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Cancer and Glioblastoma Overview

Cancer ranks among the top causes of mortality globally. It is a disease that unfortunately involves the uncontrolled spread of cells and their growth within the body. The World Health Organization (WHO) reports that there were approximately 9.7 million cancer-related deaths worldwide in 2022, and the global cancer burden is projected to increase significantly, with over 35 million new cancer cases predicted by 2050 [2]. This rise in cases can be attributed to reasons such, as a growing population, aging and shifts in the occurrence of specific cancer types influenced by social and economic progress. While brain cancer is less prevalent than other types, it presents significant challenges due to its complexity and the vital functions of the brain. Annually it impacts around 308,000 people worldwide and ranks among the top contributors to cancer related fatalities [3]. In the United States it is estimated that over 24,000 new instances of brain and spinal cord tumors will be diagnosed in 2023 alone [4].

Glioblastoma (GBM) stands out as the most prevalent and aggressive form of brain tumor (Figure 1.1). On average patients have an estimated survival period of around months post treatment with the median age for diagnosis being 64 years. In rare instances some individuals may live up to five years after their diagnosis. A report from the Central Brain Tumor Registry of the United States (CBTRUS) indicates that GBM occurs at a rate of 3.19 cases per 100,000 people annually [5,6]. There are two classifications of GBM: primary or de novo and secondary. Primary GBMs arise without any known precursor while secondary GBMs develop from tumors

that progress into GBM. The diagnostic process usually starts with a computed tomography (CT) or magnetic resonance imaging (MRI) scan followed by a biopsy or surgical resection to confirm the diagnosis. The reasons behind GBM are still unclear and there is no evidence linking lifestyle habits like smoking, drinking, drug use or mobile phone usage to GBM. Despite the use of treatments like surgery in GBM cases most tumors tend to come back [7].

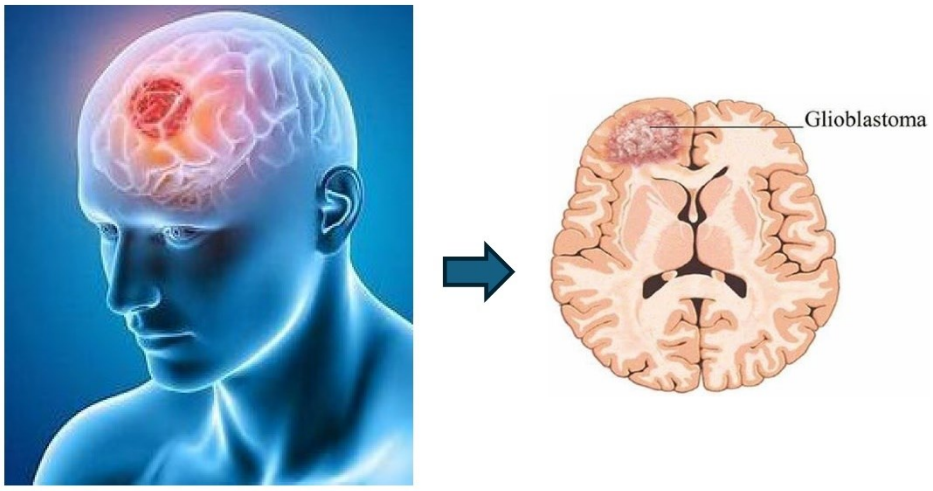

Figure 1.1. Glioblastoma – A Type of Brain Tumor[8].

1.2 Computer-Aided Cancer Detection

The process of diagnosing cancer typically starts with examining tissue samples through biopsies. Histopathology plays a role in getting these samples ready for inspection involving steps like preserving them with chemicals removing moisture and applying stains to highlight cellular structures when viewed under a microscope [9]. Traditionally pathologists have used stains like Hematoxylin and Eosin (H&E) to distinguish between tissue elements an approach that continues to be the benchmark for cancer diagnosis. Nevertheless, the introduction of technologies has transformed this procedure. Digital Pathology has revolutionized the way we approach diagnostics by converting histopathological slides into comprehensive WSIs. This advancement has laid the groundwork for Computer Aided Detection (CAD) systems that support pathologists in automating the process of analyzing, segmenting and diagnosing of

digitized images [10]. These systems leverage machine learning and AI to enhance the accuracy and efficiency of cancer detection tackling various obstacles linked to traditional manual examinations like human errors and inconsistencies in interpretation [11,12].

Early CAD systems relied on algorithms like Support Vector Machines (SVMs) and decision trees (DT) that utilized manually crafted features to classify images [13]. While these approaches had their merits, they fell short in their reliance on feature engineering and struggled to grasp the intricate patterns found in histopathological images. The emergence of deep learning, especially CNNs have brought advancements in assisting computer aided diagnosis by automating the extraction and classification of features from images with impressive precision [14]. This feature makes CNNs highly effective for examining histopathological images that require the recognition of intricate patterns, on different levels. CNNs have showcased performance across various fields such as image classification, object detection and segmentation establishing themselves as a tool in computer assisted cancer detection.

Recent developments in the field of AI for pathology have seen the incorporation of Graph Neural Networks (GNNs) [15] and attention mechanisms. GNNs prove valuable in capturing the connections between various regions in tissue samples offering a deeper understanding of tumor diversity. On the hand attention mechanisms enhance the interpretability and precision of models by directing computational efforts towards the areas of an image that hold diagnostic significance. This capability is essential for diagnosing challenging cases such, as brain tumors. The use of AI in areas like neuropathology proves to be highly beneficial due to the challenges posed by the brains intricate structure during diagnosis. Through the utilization of whole slide imaging and AI driven assessments even smaller medical facilities can now conduct thorough and precise diagnoses improving their clinical and research functions.

1.3 Challenges

Despite the progress made in studying the analysis of histopathology images [16], there are still considerable obstacles to overcome when it comes to seamlessly integrating and classifying data at the cellular, tissue and WSI levels for medical applications. These difficulties impede the

advancement of AI-powered systems that could improve cancer detection, prediction and treatment decision processes.

- **Challenges in Cellular-Level Analysis:** Detecting and segmenting cells in histopathological images is inherently challenging due to the variations in staining techniques, scanner settings and tissue preparation methods. These differences can result in inconsistencies in image quality and appearance across different datasets making it challenging to create algorithms that perform well on diverse clinical samples. Additionally, the complexity and density of structures, especially in tumor areas present further obstacles. Overlapping cells, blurred boundaries, and artifacts add difficulties to accurate segmentation often leading to errors in cell identification and classification.

  The limited availability of annotated datasets worsens these challenges, since training supervised learning models relies on extensive labeled data, that is both time consuming and expensive to acquire. Inadequate training data struggle models from learning the intricate patterns needed for reliable cell detection and segmentation, hindering their capacity to assist in analyses like advanced image classification and predictive modeling.

- **Challenges in Tissue-Level Analysis and Survival Prediction:** Examining histopathological images at the tissue level involves studying how cells are organized and interact within their surroundings. This can offer valuable insights into disease progression and patient outcomes. However, accurately capturing these complex relationships necessitates advanced models that can faithfully depict the intricate structure of tissues. Traditional methods often struggle to capture the subtle distinctions between different cancer types and subtypes limiting their capacity to provide clinically meaningful insights.

  Moreover, constructing and processing graphs from gigapixel-sized WSIs, containing numerous cells, can be computationally intensive, leading to resource heavy processes that impede real time analysis. The challenge is heightened by the requirement for models to adaptively adjust to heterogeneous tissue structures by modifying weights and

connections based on variations in cell arrangements and feature distributions. Achieving this level of adaptability and scalability in graph-based models is crucial for reliable survival classification.

- **Challenges in Multi-Graph Fusion for Cancer Grading:** Integrating data from multiple graphs for a thorough analysis of tissue presents another significant hurdle. Each graph may capture different aspects of the tissue microenvironment, such as cellular morphology, spatial organization, or molecular characteristics. Combining these graphs into a single representation without losing key information is a task that demands advanced fusion methods to capture both intra-graph and inter-graph relationships.

  Accurate cancer grading relies on models that not only consider various data sources but also navigate the intricate interdependencies among them. This necessitates the creation of adaptive multi-graph fusion models that can identify the features and connections most relevant to the classification task, thereby improving the accuracy and robustness of predictions. Moreover, ensuring that these models produce outcomes is vital since healthcare professionals require clear insights into how predictions are made to inform treatment decisions and patient care.

This thesis proposes innovative approaches to tackle these challenges in histopathological image analysis by utilizing a framework that combines cellular and tissue level examinations with state of the art convolutional and graph neural network architectures. The proposed strategies involve using a bifocal CNN (BCNN) for identifying cells, a graph-based attention model (PathoGraph) for survival classification, and an adaptive multi-graph fusion approach (AMGF-GNN) for cancer grading. The goal of these methods is to enhance the robustness, interpretability, and clinical applicability of histopathological image analysis. These contributions collectively address the need for scalable, adaptable, and clinically meaningful AI-driven models in digital pathology.

1.4 Motivation

With the evolution of the digital era and progress in healthcare technology there's a growing influx of medical information being produced rapidly. This data is not vast in volume but also heterogeneous and complex in its data form posing challenges for traditional systems or algorithms to handle. Therefore, there is a pressing demand for precise, automated algorithms capable of interpreting such data efficiently making it a crucial area of research. While AI shows promise in digital pathology with its potential for efficiency, convenience and ability to facilitate remote diagnoses there are still challenges that need to be addressed.

The field of histopathology image processing from intra to the inter level is evolving to enable decision making through the analysis of data. Multi-view/multi-scale data fusion, which refers to combining different perspectives and scales of information, can provide comprehensive information for more precise diagnosis and treatment planning. Our goal is to develop end-to-end machine learning approaches that can extract meaningful features from multi-view/multi-scale diverse data sources to identify higher-level features and enhance diagnostic accuracy and predict patient outcomes. Additionally, we aim to uncover key characteristics within medical data that shed light on the nature and progression of infectious and severe diseases.

This thesis explores the potential of using CNN, GNN and attention mechanisms to automate the processing of WSIs from an intra level to an inter level and achieve outputs at both the cell and slide levels. We employed approaches to analyze WSIs in pathology with the aim of developing more accurate, efficient and comprehensive diagnostic tools. The initial step involves detecting and segmenting individual cells within WSIs as it offers insights into the cellular landscape, which is crucial for making accurate diagnoses. This process enables the identification and localization of cell types along with their morphological characteristics that serve as indicators of disease states. While cell level predictions play a role in tasks like identifying tumour cells, slide-level predictions are crucial for making broader assessments about the entire tissue sample, such as determining cancer survival rates or grading.

Therefore, after cell detection and segmentation, constructing a graph with nodes represent cells

and edges represent their interactions can be used for survival analysis. This graph-based approach goes beyond the capabilities of CNNs by leveraging the spatial and relational information between cells. It allows for capturing complex tissue structures and microenvironment interactions that are often linked to patient outcomes. This method provides a nuanced and comprehensive understanding of the disease progression compared to traditional image analysis techniques potentially enhancing the prediction of survival rates. Moreover, employing multiple graphs to depict different tissue regions or cell populations for grading WSIs enables a more thorough examination of the tumor microenvironment. This multi graph strategy can capture the diversity within the tissue reflecting variations in tumor grading and progression. This multi-graph approach can capture the heterogeneity within the tissue, reflecting variations in tumor grading and progression. By integrating information from various tissue compartments, the method can provide a comprehensive grading of the cancer, which is essential for determining the appropriate treatment strategy.

To summarize, the motivation behind using multi-scale approaches for analysis of WSIs in pathology lies in the need for diagnostic tools that are more precise, efficient and thorough. This includes 1) urgent demand for a detailed understanding of the cellular landscape within vast datasets, 2) making predictions at the slide level to derive insights about the entire tissue sample like cancer survival or grading, 3) improving the accuracy of cancer diagnosis and prognosis by systematically breaking down the intricate structure of WSIs into smaller, manageable parts thus providing pathologists with a resource to enhance clinical outcomes.

1.5 Contributions

In this thesis, we have developed and implemented end-to-end machine learning models to enhance the accuracy, efficiency and thoroughness of diagnosing and predicting cancer diseases. Our key contributions can be outlined as follows.

- At the cellular level, we present a framework that utilizes PathoFusion's BCNN [17] model to detect, identify and quantify cells labeled with immunohistochemistry in WSIs of diagnostic tissue samples. Our method combines information from image tiles

capturing important morphological, texture and spatial features for accurate cell identification. Using CD276 immunoreactive cells in glioblastoma as an example, the BCNN model automatically profiles and counts cells, applying data filtering and thresholding techniques to focus on distinguishable cells for in depth analysis. The detection results are converted into heatmaps that show the distribution of identified cells throughout the WSIs offering a comprehensive view of cellular composition. To enable analyses, such as exploring the relationship between immunohistochemical signals and histological features we utilize DWT to merge heatmaps from adjacent tissues stained with different markers enhancing our understanding of complex tissue interactions.

- Building upon the cellular analysis, we expand our approach to the tissue level through PathoGraph, this is a graph neural network (GNN) framework created for survival prediction and in-depth tissue analysis. PathoGraph transforms gigapixel-sized WSIs into memory-efficient subgraphs, maintaining important spatial connections while simplifying data complexity. This graph representation showcases the relationships among identified cells, reflecting their arrangement in ways that might be linked to prognosis. By employing an attention-enhanced graph convolutional neural network (abGCN), PathoGraph modifies attention weights adaptively based on feature similarity and spatial closeness. This capability enables it to manage feature distributions effectively and capture the diversity of histological structures. The model's transparency is enriched through the visualization of attention scores, highlighting areas within WSIs that influence its predictions, uncovering concealed patterns within the tissue. Moreover, our suggested loss function encourages an attention distribution, enhancing classification performance and offering insights into tissue architecture.

- Expanding further at the tissue level, our final stage addresses the complexity of cancer grading by introducing a multi-graph framework, extending PathoGraph's approach to analyze interactions within the tumor microenvironment. To encompass both local and global tissue structures, we construct multiple graphs, each representing distinct aspects of tissue organization. Acknowledging the difficulty in identifying nodes and contributions within the graphs we present the Adaptive Multi Graph Fusion based

Graph Neural Network (AMGF GNN). This model extracts unique embeddings from node attributes, graph topologies, and their higher-level interactions employing cutting edge attention mechanisms to seamlessly fuse these graphs into a unified representation. By independently learning and updating the embeddings of graphs, the model bolsters resilience and reduces reliance, on the quality of individual graphs. By combining similarity measures within and between graphs in our two-tiered loss optimization approach, we ensure that the representations we learn are consistent and resilient. This allows us to capture intricate relationships and greatly enhance grading accuracy. Thorough experiments conducted on both publicly available datasets and a private cohort confirm the efficacy of our method, showcasing its superiority over existing state of the art techniques. Additionally, it improves diagnostic and predictive processes, in cancer care.

1.6 Thesis Outline

The remainder of this thesis is organized as follows. In Chapter 2 we explore the relevant studies of deep learning and its application in medical diagnosis and prognosis. Chapters 3 to 5 outline the key contributions of the thesis. Chapter 3 presents a multi-scale deep learning framework for detecting, segmenting and counting individual immunochemically labelled cells in whole slide histopathology images. Chapter 4 introduces a graph-based model for detecting and interpreting immunohistochemical labelling results at both cellular and WSIs levels to classify survival outcomes and identify specific areas of diagnostic interest. Chapter 5 describes a multi-graph with an attention mechanism framework for cancer grading on gigapixel WSIs by fusing multiple graphs to enhance feature robustness and graph quality while emphasizing the importance of nodes and graphs. Lastly the final chapter summarizes research findings and discusses future perspectives.

CHAPTER 2

Background and Literature Review

In this chapter, we begin by examining the significance of pathology images that are crucial for disease diagnosis and prognosis in relation to this thesis. Next, we explore various machine learning and deep learning methods including multi-layer perceptron (MLP) and CNNs along with advanced deep learning models that have emerged recently. These techniques are evaluated for their applicability, in analyzing medical data with a specific emphasis on GNN and attention mechanisms. Moreover, we discuss the concept of learning at different levels focusing on both convolutional based and graph-based approaches.

2.1 Histopathology Imaging Nature

The history of histopathology dates back to the 19th century, and it has transformed medical diagnosis by allowing the microscopic examination of tissue samples for accurate disease identification [9]. Histopathology images are very detailed and essential for diagnosing cancer, infection, and other diseases. The pathologist looks at a sample of human body tissue under a microscope. This helps them identify changes in tissue and cells caused by diseases to discover why they came about [18]. Their ability to make conclusive diagnoses and prognosticate patients' outcomes based on changes in both macroscopic and microscopic structures is vital.

Histopathological samples are evaluated based on their imaging quality. This highlights the significance of histological procedures, and the microscope employed in the process. Staining

in histopathology is an important procedure that improves the contrast of microscopic tissue samples. This allows for precise observation and distinction of biological components, tissues, and structures. Several stains and staining techniques are used to fulfil various diagnostic purposes. Histopathology staining takes multiple processes, including rehydration with xylene and alcohol to replace paraffin with water, followed by mounting for long-lasting specimens. Furthermore, cleaning techniques with water, alcohol, and xylene are critical post-staining to prevent the specimen from taking water or other solutions.

Hematoxylin and Eosin (H&E) staining, and Immunohistochemistry (IHC) are among the most common and widely used techniques in histopathology (Figure 2.1). Each of them serves distinct but complementary purposes, providing crucial role for the diagnosis and study of diseases. H&E staining is widely used in histology as it offers a glimpse into the structure of tissues and the appearance of cells. This makes it crucial for preliminary assessments and routine examinations [19] . H&E consists of two stains, hematoxylin and eosin. Hematoxylin imparts a bluish-purple hue to cell nuclei while eosin colors the extracellular matrix and cytoplasm pink. Other structures showing various tints, hues, and combinations of these colors [19]. Consequently, pathologists can easily differentiate between the nuclear and cytoplasmic regions of a cell. The overall coloration patterns from the staining, provide an insight into the arrangement and distribution of cells offering a comprehensive view of the anatomy of a tissue sample [20]. IHC is another brightfield staining method that allows for targeting specific molecular species, often proteins through the use of highly selective antibody reagents. H&E and IHC techniques together have substantially increased the pathologist's diagnostic power [21]. Figure 1 shows example images obtained from H&E staining as well as IHC.

Images in pathology also known as WSIs can be quite massive sometimes containing gigapixels worth of information. These massive image sizes present significant challenges for manual analysis due to the time, effort, and risk of human error required in evaluating such detailed data. The complexity and variability of histology images necessitate professional knowledge and experience from pathologists, making the manual examination also a highly time-consuming process that requires intense concentration.

As a result, there is an increasing demand for quantitative analysis of histology images. The advancement of imaging technology in pathology has led to the growing popularity of computer assisted diagnosis (CAD) algorithms particularly for handling large volumes and intricate cases of histopathological images [13]. CAD systems play a role in supporting pathologists diagnoses providing swift and consistent solutions while enhancing efficiency. The early CAD systems relied on learning algorithms with a significant focus on feature engineering that leverages extensive domain expertise to extract valuable characteristics. In recent years, deep learning techniques have gained prominence in analyzing high resolution histological images enabling more effective and reliable assessments of complex cases. We will delve into these topics in the sections.

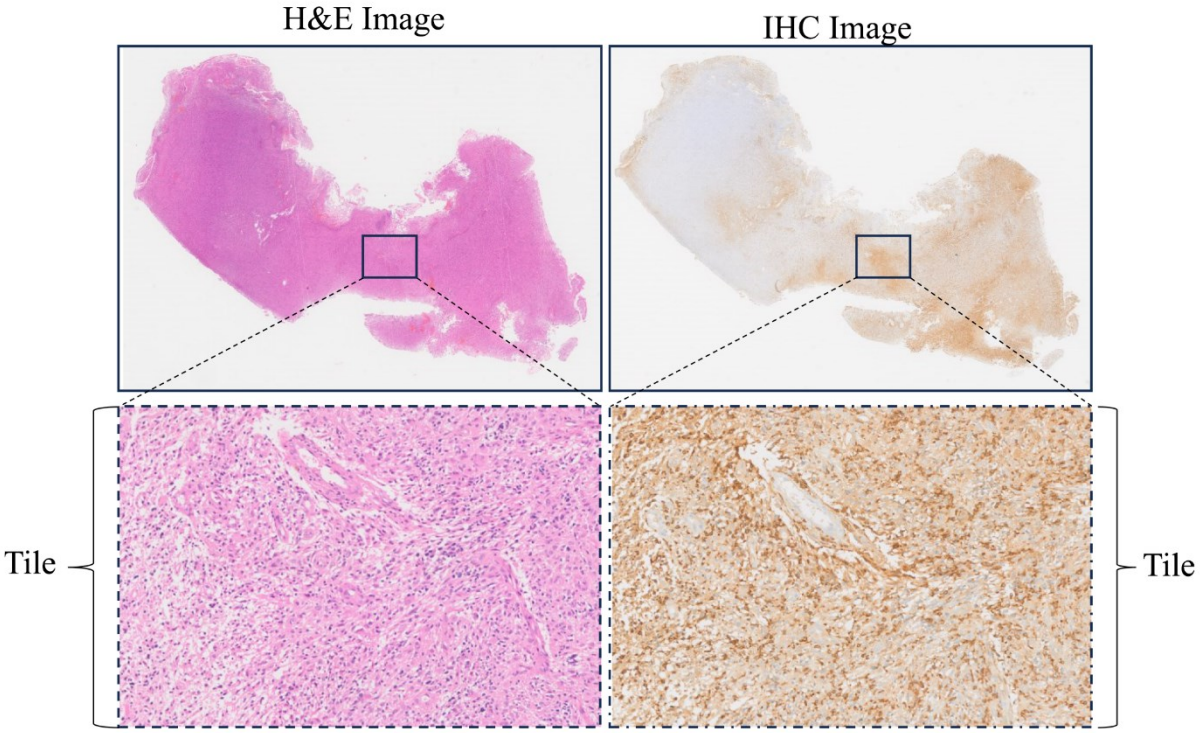

Figure 2.1. Examples of whole slide images [1]. H&E whole slide image in the left panel and a magnified tile of it. IHC whole slide image in the right panel and a magnified tile of it.

2.2 Artificial Intelligence (AI) Research History

AI is an umbrella term that includes machine learning (ML) (Figure 2.2). Deep learning (DL) is

a subtype of ML that relies on the use of artificial neural networks. It is in DL where especially rapid progress is currently being made. The history of ML began in 1943, when McCulloch and Pitts developed a mathematical model of a neuron [22]. ML consists of algorithms that parse data, learn from them, and then apply what they have learned to make informed decisions. McCulloch and Pitts used algorithms known as "threshold logic." A few years later, Turing developed the test named after him to determine whether a computer possesses intelligence [23]. In 1952, Samuel created the first computer learning program [24]. The first neural network for computers, the single-layer perceptron consisting of multiple McCulloch–Pitts neurons, was designed by Frank Rosenblatt in 1958 [25]. Extending the design, this model became a multilayer perceptron (MLP), which essentially represents a "plain vanilla" contemporary neural network while lacking the depth, i.e., many hidden layers, of today's networks.

A step toward allowing computers to use basic pattern recognition was the application of the nearest-neighbor algorithm [26]. In 1979, Fukushima developed the first CNN, which had multiple layers [27]. In 1989, LeCun and colleagues demonstrated the first practical application of backpropagation in combination with CNNs for the recognition of handwritten zip codes [28]. Their original dataset remains in use for validation purposes [29]. Hinton coined the term "deep learning" in 2006 to describe new algorithms that allow computers to discriminate objects and text in images and videos [30]. His group's AlexNet (2012) catapulted CNNs into the spotlight, and their related 2017 paper [31] has earned more than 100 000 citations so far. Rapidly increasing computing power and better software have supported the explosive growth of DL techniques in recent years.

2.3 Traditional Machine Learning

ML-based methods for analyzing WSIs are categorized into two types: traditional ML methods and DL methods, as illustrated in Figure 2.3. Traditional computational methods have been used to analyze images by extracting handcrafted features like morphological and textural characteristics, followed by the design of classifier algorithms like (SVM), Random Forest (RF), and K-Nearest Neighbor (KNN) for downstream analysis tasks.

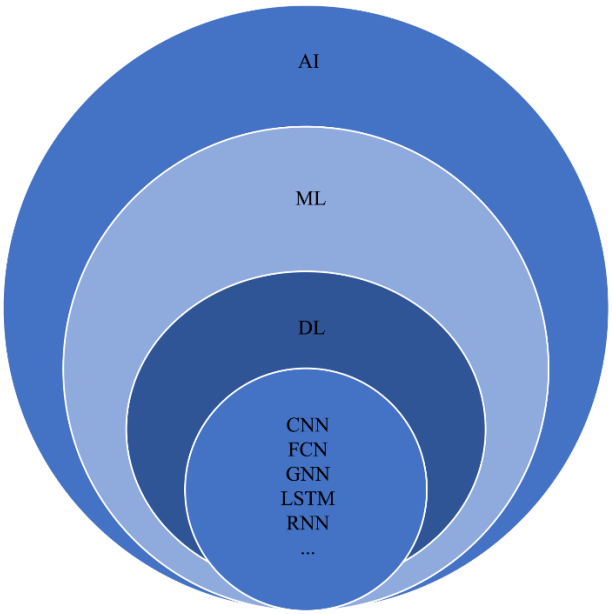

Figure 2.2. Fields of AI research and how they are related [32]. AI, artificial intelligence; CNN, convolutional neural network; DL, deep learning; FCN, fully convolutional network; GNN, graph neural network; ML, machine learning; LSTM, long short-term memory; RNN, recurrent neural network.

The challenge when extracting features from (neuro)pathological WSIs is finding higher-level representations that allow capturing information that is relevant for making a medical diagnosis. In view of the importance of morphometric data like shape and size of individual cells in disease diagnosis, several studies have used them as criteria [33-35]. Several studies have also combined feature categories to extract information found in WSIs [36-43]. Maroof et al. [42] proposed a method of using hybrid space by combining color features with morphological and textural features and then changed the color channel to calculate normalized and cumulative histograms in the wavelet domain for better discrimination of mitotic and nonmitotic nuclei. Barker et al. [44] presented a method for the classification of brain tumor WSIs containing gliomas. They divided WSIs into grids of tiles and localized features of shape, color, and texture from tiled regions, which were then clustered using k-means. A detailed analysis was conducted of a single representative tile from each cluster and a classification decision made for each one using an elastic net model. Their method achieved an accuracy of 93.1%. Recently, the recognition of deep features has gained popularity in a variety of image classification tasks, including of WSIs

[45,46]. Given the rich structures found in histological WSIs, the majority of studies now combine different types of morphometric and textural features. However, traditional ML methods are of limited use in WSI analysis.

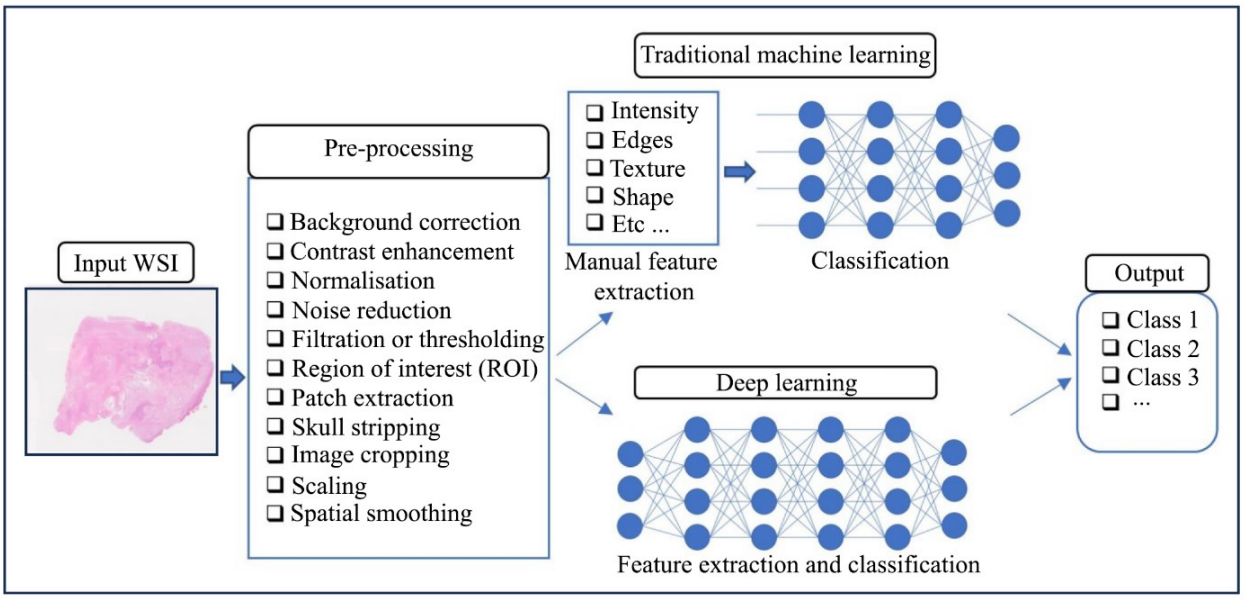

Figure 2.3. Steps and stages of Whole-Slide Image (WSI) analysis using traditional machine learning and deep learning [32].

2.4 Deep Learning Development

Deep learning (DL) is a subset of ML that focuses on using neural networks to learn patterns and insights from data. In contrast to traditional machine learning where individuals need to choose and specify the features for the system, DL algorithms can extract useful features directly from the data input stream, avoiding complex feature extraction steps. As stated in the classical paper by Krizhevsky et al. [31]: "With enough computation and enough data, learning beats programming for complicated tasks that require the integration of many different, noisy cues." Furthermore, DL algorithms appear far more capable of dealing with WSI heterogeneity [32]. DL employ a learning approach where data from one layer serves as input for the subsequent layer. Initially, only edge like areas are recognized in the lower layers of the network [47]. These edge regions are then utilized to delineate corners and contours. By combining corners and

contours it becomes possible to identify object components in the following layer. In DL architectures a sequence of hidden layers is employed to extract characteristics from an image. Ultimately the output layer is responsible for categorizing the image and generating the corresponding class label. However, DL technology has proven extremely successful in the field of computer vision, including WSI analysis [48,49]. We delve into the advancements of DL models in the following sections.

2.4.1 Multi-layer Perceptron

Multilayer perceptron (MLP) is a basic form of neural network and serve as a foundational architecture in deep learning, designed to emulate the cognitive processes of the human brain. It consists of layers of nodes, also known as neurons, connected in a feedforward manner (Figure 2.4). Each layer in an MLP is fully connected to the subsequent layers, where every node in one layer is linked to every node in the subsequent layer. It comprises three primary layers, namely the input layer, one or more hidden layers, and the output layer. Each layer comprises of interconnected neurons that process input data through a series of connections and activation functions. The network is trained on large datasets, using backpropagation, a process which iteratively adjusts the weights based on the difference between predicted and actual outputs [50,51]. This training helps the MLP to recognize complex patterns and relationships within the data, making it a powerful tool for tasks ranging from image recognition to natural language processing.

However, MLPs pose substantial challenges when utilized for imaging data, such as excessive overfitting, high memory consumption, and immense computational intricacy. To address these challenges, a new architecture based on the classic convolution technique was developed. This architecture proposes using "sparse" neuron connections to handle imaging data more efficiently. This approach decreases computing effort while improving the network's ability to generalize, making it better suited for tasks using visual input.

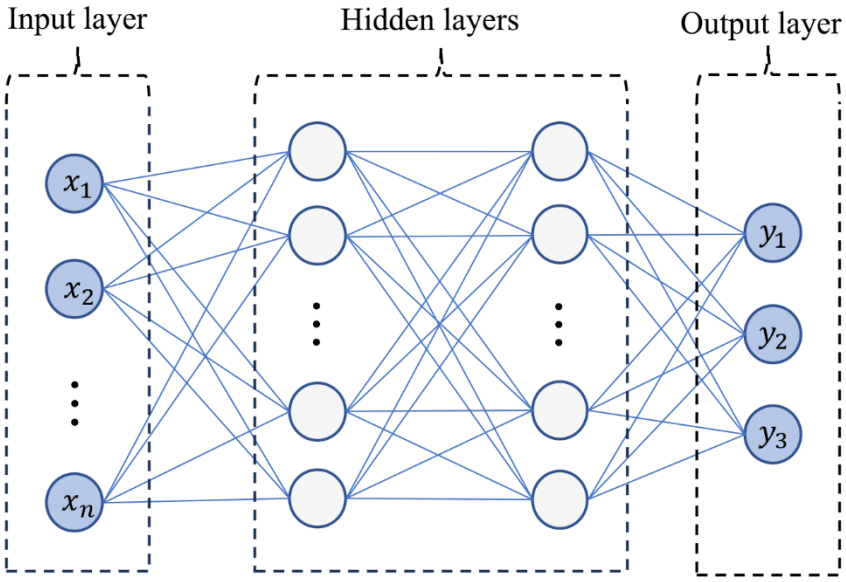

Figure 2.4. A basic neural network demonstrates the message passing mechanism used in deep learning models. Each hidden layer transforms and passes the information forward.

2.4.2 Convolutional Neural Network (CNN)

CNNs currently are the most widely used artificial neural networks (ANNs) [14]. They can be thought of as a feature learning process progressing from a low to a high level. A convolution is a mathematical operation, and convolutional layers are used to learn local features. Convolutional layers are the major building blocks of a CNN. It is here where filters are applied to the original image. The pooling layers are interleaved with convolutional layers to reduce the output from the convolutional layers. The final fully connected layers are used to combine the features learned from the convolutional layers, resulting in a complex and high-level representation. Some ground-breaking CNN architectures include LeNet [52], AlexNet [31], VggNet [53], GoogLeNet [54], and ResNet [55].

To avoid complex feature extraction steps, CNNs use raw images (or large image patches) as input, which are highly resistant to translation, scaling, inclination, and other forms of deformation. Because WSIs consist of highly complex data, DL algorithms are ideally suited for their analysis. Examples include Sharma et al., [56] who designed a CNN for the

classification of gastric carcinoma, and Xu et al., [57] who proposed a framework for the extraction of deep features from an AlexNet-CNN model [31] that had been pretrained on the ImageNet dataset, followed by the use of a linear SVM [58] to identify and segment brain and colon cancer images. Wang et al. [59] adopted the GoogLeNet architecture for the detection of metastatic breast cancer. Jamaluddin et al.[60] used DL on lymph node metastases, and Sirinukunwattana et al. [61] proposed a spatially constrained CNN (SC-CNN) for the classification of colorectal adenocarcinoma images.

Of special interest to neuropathology, Hou et al. [62] proposed a patch-based CNN model and training a decision fusion model at the image level for the classification of glioma and non-small-cell lung carcinoma WSIs. Araújo et al. [63] designed a CNN-based method for the classification of breast cancer histological images. Bejnordi et al. [64] developed a cascaded CNN (CAS-CNN) architecture for the WSI classification of breast histopathology. They first trained a CNN using high-pixel-resolution patches to capture cell-level features. Then the generated features were fed into a second CNN to allow for the incorporation of the interdependence of tissue. Korbar et al. [65] adopted a modified version of the ResNet deep-neural network to classify colorectal polyps in WSIs. Das et al. [66] introduced an ensemble technique of different magnification-specific models for breast histopathology WSI classification. They used four pretrained GoogleNet [54] models and achieved a very high accuracy of 96%.

2.4.3 Fully Convolutional Network

CNNs have gained widespread acceptance for solving medical segmentation problems, with architectures like fully convolutional networks (FCNs) [67] and U-Nets [68] being particularly popular. In an FCN, a classical CNN's fully connected layers are replaced with fully convolutional layers. As a result, instead of classification scores, a FCN model generates a spatial segmentation map. In addition, several models originally developed for medical/biomedical image segmentation were inspired by FCNs and encoder–decoder models, such as the U-Net architecture. A U-Net augments the standard CNN architecture by adding a corresponding expansive path (decoder) with the goal of producing a full-resolution semantic

prediction, whereas the CNN architecture consists of a contracting path (encoder) that iteratively builds feature maps using continuously down-sampled data.

Bándi et al. [69] introduced a FCN and a U-Net network architecture [68] for WSI segmentation. They found that FCN and U-net models outperform older methods. Cui et al. [70] presented a supervised FCN method for nuclear segmentation of histopathology images. First, color normalization is performed on the image, then a FCN is used to detect nuclei and their boundaries in each patch extracted from the WSI. The procedure was tested using different datasets, which showed that the proposed method is accurate, robust, and fast: one 1000 × 1000-pixel image can be segmented in under 5 s. Gecer et al. [71] presented a FCN system to identify ROIs in breast WSIs to classify WSIs into five classes. Wang et al. [72] exploited image-level labels to introduce a weekly supervised method using a FCN to select discriminative patches for the classification of lung cancer WSIs. Sheikhzadeh et al. [73] proposed using a FCN as an extension of a CNN to localize and quantify immunohistochemical WSIs in one integrated step. Lin et al. [74] utilized a fast scanning neural network, referred to as ScanNet, for metastatic breast cancer detection in WSIs. The authors further improved their work by converting the standard FCN layers into an anchor layer that speeds up the network [75].

2.4.4 Advanced Deep Learning Models

In addition to CNNs, other neural network models such as recurrent neural networks (RNNs) [76], long short-term memory (LSTM) [77-79], GANs [80,81], stacked autoencoders (SAEs) [82-84], and genetically optimized neural networks (GONNs) [85,86] have all been applied to WSI analysis. These models offer diverse approaches to capturing spatial and temporal patterns, enhancing the accuracy and robustness of medical image analysis. A novel attention-based multimode RNN architecture called CORAL8n [76] has been proposed for direct renal immunofluorescence (RDIF) detection, capable of dynamically sampling image data. The prior encoder learns to extract context features from clinical notes, whereas VGG16 extracts local and global WSI features. The results of both are then fed into an RNN (sentence generator) to generate the written medical report, and machine-generated text is produced. Ren et al. [77] adopted a LSTM architecture to identify image biomarkers within WSIs by modeling the

spatial relationship of patches as a temporal sequence. The results demonstrate that integration of image features from a LSTM and a genomic pathway score is more closely correlated to a patient's recurrence of disease than standard clinical markers or image texture features.

2.4.5 Graph Neural Network (GNN)

CNNs have demonstrated great success in extracting image-level representations. They are ideal for dealing with linear, grid-like (Euclidean) data. However, when dealing with relation-aware representations, they are inefficient because they limit the model's ability to capture global contextual information and comprehensively model tissue composition. In tissue diagnosis, the phenotypic characteristics and topological distribution of cellular and histological features are critical. Consequently, graphical neural networks (GNNs) have attracted considerable interest because they capture intra- and inter-entity level interactions [15]. In general, graphs describe natural phenomena better. GNNs are generalized versions of CNNs and capable of working with non-Euclidean data.

Traditional CNNs analyze local image areas using fixed connectivity as determined by the convolutional "kernel." Graphs, on the other hand, provide greater flexibility when analyzing unstructured data by encoding the spatial context of individual patches. Thus, graph-based methods in (neuro) pathological analysis can record geometrical and topological properties together, allowing them to model both cell-level information and overall tissue microarchitecture simultaneously. Prior to the advent of DL, many graph-based methods in digital pathology processing relied on traditional ML approaches [87], which are less accurate for graph classification than graph-based DL. However, the capabilities of graph-based DL for disease diagnosis, which bridges the gap between DL methods and traditional cell graphs, have yet to be thoroughly investigated. GNNs have already been used for ROI classification [88], detection of malignancy and invasiveness [89], and survival analysis [90]. We believe that they have great potential for neuropathological analysis.

As illustrated in Figure 2.5, image analysis in (neuro) pathology using graphs involves the following steps: establishment of graph representation (description of entity, embeddings, and

definition of edges), design of GNN-based models for processing graph-structured data, and the use of interpretability methodologies (model-based and post hoc). Pathological images can be translated into entity graphs by means of graph representations where nodes and edges indicate entities and inter entity interactions [91,92], respectively. Individual cells [89,91,93-101], tissue patches [102-110], different types of tissue [111-113], and hierarchical entities [114-117] can all be used depending on the task. Following entity-graph construction, the graph is processed using GNN-based models; thus, GNN output can be either a node- or graph-level-based classification. Explainers for graph-structured data [118] are then applied to the entity graphs, yielding intuitive explanations for (neuro)pathologists. Attention mechanisms are an example of explainable methodologies in which the attention weights highlight the nodes and edges in their relative order of importance; as such, it can direct a network to focus on the most relevant parts of the input, suppressing irrelevant features, reducing computational cost, and improving accuracy.

2.4.5.1 Cell Graphs

In these types of graphs, the graph is presented using cells or nuclei as entities with the locations of identified cells serving as graph nodes, and the spatial distances between them [91,92] are represented as lines called “edges.” Jaume et al. [91] applied cell-graph analysis to breast cancer subtyping. First, they transformed the ROIs in the histology image into a cell graph, and a GNN model was then used to map the entity graph to a corresponding class label. In the next step, they used a post hoc graph explainer. Finally, a set of metrics was used to assess the generated explanations. A cell-graph explainer was used in breast cancer subtyping [95] to eliminate redundant and uninformative graph nodes and edges while retaining tumor cell nuclei relevant for cancer diagnosis.

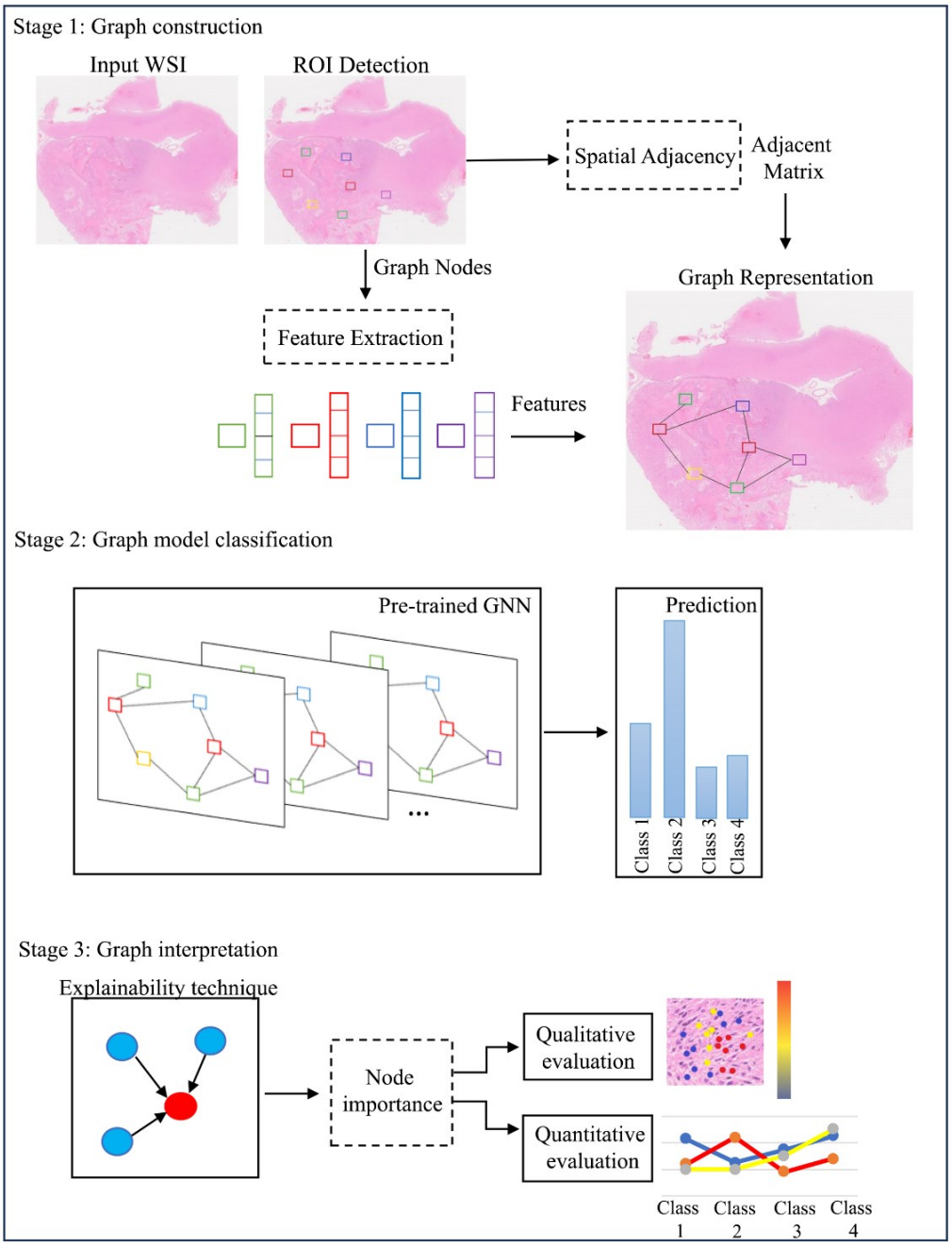

Figure 2.5. Workflow for graph-based WSI analysis (GNN) [32]. Stage 1: Construction of graphs: Morphological entities become feature nodes and distances between nodes become "edges" representing the spatial relationship between the entities. Stage 2: Graph model classification. The graph representation is processed using GNNs for node-level or graph-level prediction. Stage 3: Graph interpretation. GNN model interpretability tools such as graph attention or post hoc graph explainers are used.

Anand et al. [89] proposed a framework named Histograph for classifying breast cancer WSIs. The framework is based on a pretrained model that generates features around the nucleus centroids and models nuclei as graph vertices and internuclear distances as graph edges. Sureka et al. [96] also modeled histopathology images as graphs of nuclei but developed a graph convolutional network (GCN) framework based on an attention mechanism and node occlusion to highlight informative nuclei and internuclear relationships. Nuclei were detected using a U-Net on WSIs, and a graph was constructed by linking pairs of nuclei closer than the distance threshold. The latter method emphasizes the relative contribution of each nucleus and its neighborhood in the WSI and visualizes important nuclei for the final binary classification of the breast cancer biopsy. Zhou et al. [99] developed a GCN for colorectal cancer classification where each node is represented by a nucleus, whereas the edges between these nodes are built according to node similarity. The variety of methods applied again signifies that iterative trial and error and optimization form intrinsic parts of the DL approach.

2.4.5.2 Patch Graphs

In a patch graph, the nodes are the ROIs extracted from the image using color-based, cell density, or attention-based methods. Aygunes et al. [88] proposed a weakly supervised classification method using GCN as a model for ROI classification in breast histopathology. A pretrained ResNet model was used to extract a feature vector for each node represented by fixed-sized patches of the ROI. Then GCN architecture was used to decide on per-patch representations and to propagate information over the neighboring patches in a progressive manner to incorporate the spatial context and, finally, aggregate the resulting patch representations, leading to the assignment of the whole ROI to a diagnostic class. Ye et al. [104] presented a system for segmentation and classification of breast cancer ROI images based on GCNs. The system is composed of a segmentation module and a GCN module. Li et al. [90] developed a GCN with an attention learning–based survival analysis model of WSIs.

2.4.5.3 Tissue Graphs

A cell graph only encodes cellular morphology, and its topology ignores tissue information. Therefore, it is not capable of capturing the tissue microenvironment, i.e., the distribution of tissue areas such as parenchyma or necrosis, for example, which is necessary for the accurate representation of a histopathological context. A tissue graph composed of multiple tissue areas, on the other hand, is incapable of depicting the cell microenvironment. Tissue graphs are built with nodes representing key tissue regions and edges encoding the relationships between these regions.

Anklin et al. [111] proposed a weakly supervised segmentation method for representing a histology image using a "superpixel"-based tissue graph. An input image is first transformed into a tissue graph representation, with graph nodes representing tissue "superpixels." Thus, to learn the intrinsic characteristics of pathological tissue changes, multilevel structural information must be aggregated, with the goal of replicating the tissue diagnostic process used by pathologists when analyzing images at different levels of magnification. A GNN then learns contextualized features of the graph nodes using inexact and incomplete labels. Raju et al. [112] proposed a Graph Attention Multi-instance Learning (Graph Attention MIL) framework for colorectal cancer staging. For each WSI, a texture feature extraction method was used to encode texture from a sample of patches and cluster similar texture features into multiple graphs, with each graph containing nodes that represent different tissues as an instance.

2.4.5.4 Hierarchical Graphs

Histological structures cannot be fully described by abstract cellular or tissue interactions alone, so hierarchical graph representations have been proposed as additional tissue representations. Cell graphs and tissue graphs have been shown to provide valuable complementary information (cellular and tissue interactions) for learning the characteristics of a tissue. Hierarchical graphs aggregate multilevel structural information. Zhang and Li et al. [119] presented a multiscale graph wavelet neural network (MS-GWNN) for the classification of histopathological images of breast cancer. Based on GWNN [120], MS-GWNN was proposed to perform multiscale analyses using spectral graph wavelets after transforming pathological images into graph representations. Multiple GWNNs with different scaling parameters were used in parallel to

integrate the multiscale contextual interactions in graph topology. The final predicted image-level classification was produced by aggregating these features at different scales.

Pati et al. [121] proposed a hierarchical cell-to-tissue-graph (HACT) representation using breast cancer as an example. The almost 50 GB of images used in their study are freely available. This framework, which is also available on GitHub, consists of a low-level cell graph that captures cell morphology and topology, a high-level tissue graph that captures tissue area morphology and the spatial distribution of tissue components, and cell-to-tissue hierarchies that encode the relative spatial distribution of cells in relation to tissue features. Furthermore, a hierarchical graph neural network (HACT-Net) was proposed to map HACT representations.

2.5 Attention Mechanism

In recent years, attention mechanisms have become increasingly important in different areas of deep learning due to their capability to concentrate on the aspects of data when making predictions. In histopathological image analysis, attention mechanisms have shown to be valuable in improving the interpretability and precision of models. By selectively highlighting regions or features within image, attention mechanisms facilitate a more accurate representation of intricate tissue structures and cellular interactions that are vital for tasks like cell detection segmentation and survival prediction.

In the field of computational pathology, attention mechanisms are commonly used alongside techniques to enhance the modeling of WSIs, such as graph-based techniques. These methods leverage structures to capture the intricate connections between cellular and structural elements within a tissue sample. Attention mechanisms are employed to calculate scores at the level aiding in more nuanced predictions, whether for individual instances or entire images. This approach offers a level of interpretability that closely mirrors the analytical processes employed by pathologists.

Attention mechanisms play a role in enhancing predictions at the instance level and aggregating features in multi-instance learning systems, especially in the field of computational pathology.

By utilizing graph structures these approaches capture intricate connections within tissue samples leading to more nuanced and understandable predictions. Ma and Yang [122] proposed a multi-instance learning framework based on GCNs. In this approach graph representations are constructed by treating instances within a package as nodes and establishing similarities between instances as edges. The use of graph attention mechanism helps in scoring features at the instance-level, enhancing the aggregation process and allowing for both instance-level and image-level predictions. By focusing on significant features through attention, this approach improves model interpretability, aligning with the need for explanations in clinical settings.

Wu et al. [123] introduced DeepGCNMIL, that incorporates multi-head attention in a GCN framework. In this method each patch of a histopathology image is treated as a node with edges determined by the Euclidean distances between nodes. The attention mechanism is employed to evaluate prognostic risk by effectively capturing the connections between various regions within the tissue. This enables more accurate predictions. The approach showcases how attention mechanisms can improve the interpretability and accuracy of multi-instance learning models.

In survival outcome prediction and segmentation tasks, attention mechanisms are essential for enhancing graph-based models better understand intricate spatial connections and variations in tissue composition. By incorporating attention these models can flexibly adjust to different features, leading to more informed clinical decision making.

Li et al. [90] proposed a survival prediction model that leverages GCN techniques. It constructs graph representations by using patches as nodes and establishing edges based on Euclidean distances. To improve the graph representation an attention layer is added to concentrate on patches that are crucial for predicting survival outcomes. This method showcases the promise of mechanisms in boosting the ability of models to navigate the intricacies of histopathological images, especially in capturing the subtleties of the tumor microenvironment. Sureka et al. [96] proposed a framework that represents histology tissue in the form of a nucleus graph. They used Robust Spatial Filtering (RSF) combined with a GCN incorporating attention and node occlusion techniques. Their approach emphasizes the contributions of cells similar to how a pathologist thinks. By incorporating an attention layer before pooling operations, they were able

to highlight nuclei making it easier to classify breast cancer cases and determine Gleason grades for prostate cancer. This method highlights the role of attention in improving model interpretability and aligning predictions with clinical practices.

Yoav Alon and Zhou [124] developed a network called the neuroplasticity graph attention network aimed at segmenting cell nuclei in pathology images. This innovative network improves graph structures at different zoom levels using residual graph attention layers to ensure resilience to variations in staining and cell types. By leveraging attention mechanisms, it enhances segmentation precision by prioritizing the most pertinent features showcasing its versatility across various experimental settings.

Integrating multiple data modalities like image features, genomic data, and cellular interactions involves using advanced models to understand their intricate relationships. Attention mechanisms play a role in this process by highlighting important aspects across various data types enhancing both prediction precision and the ability to interpret the models results. Chen et al. [100] introduced multimodal fusion approach to predict survival outcomes. Their model leverages CNNs to extract image features, GNNs for analyzing interactions between cell nuclei, and SNNs to capture genomic information. By integrating these through a Kronecker product and an attention mechanism based on gating the model enhances interpretability compared to conventional network architectures. This integration emphasizes the significance of attention mechanisms in improving generalizability across different tasks. Zheng et al. [109] employed a GNN-based approach for retrieving histopathology images, utilizing hashing to create indexes for quicker retrieval. After breaking down WSIs into smaller patches, a CNN model was used to extracts features. These features serve as nodes to construct a graph structure. By integrating attention mechanisms into their Diagnostic Region Attention Network (DRA Net) they enhance the retrieval process by concentrating on regions relevant to diagnosis. This not only lightens the workload for pathologists but also boosts the efficiency of the recommendation system.

2.6 Neural Network Optimization

Neural network optimization aims to develop models that are not only accurate, but also

efficient, interpretable, and scalable, making them more applicable in a variety of applications. This involves hyperparameter tuning, architecture selection, and the use of various regularization and optimization algorithms. Adjusting hyperparameters such as learning rate, batch size, and architecture of the network to discover the best configuration. Weight initialization and learning rate choosing are critical for guaranteeing smooth and efficient training, while regularization techniques such as L2 regularization and dropout are used to avoid overfitting and improve generalization. The adoption of optimization algorithms such as SGD or Adam has a direct impact on how successfully the network minimizes the loss function, and the network's architecture is carefully constructed to suit the specific goal.

Hyperparameters can be manually assigned or modified using optimizations techniques like Grid Search, Random Search, and Bayesian Optimization [125-127]. Grid Search aims to find the solution by exploring all possible combinations of hyperparameters while Random Search evaluates combinations randomly. Both methods involve dividing hyperparameters into spaces for practical application. On the hand Bayesian Optimization also known as Bayes Search utilizes a model to connect hyperparameter configurations with the neural objective function. The optimization process iteratively refines the model based on the outcomes of the objective function by modifying hyperparameter combinations. According to [128], Bayesian Optimization is more efficient than both Grid and Random Search when it comes to striking a balance between the time taken for evaluations and how closely it gets to the best possible solution.

Parameter reduction is also a crucial characteristic of neural network optimization, especially as deeper and more advanced network architectures have arisen, necessitating significant storage, memory, and computational power. Many strategies have been developed to address these challenges. Network pruning, which was first proposed by LeCun et al. back in 1990 [129], helps eliminate redundant parameters in pretrained models, leveraging weight sparsity in deep CNNs to accelerate inference [130]. Low-rank factorization takes it a step by reducing network weights further by minimizing large weights [131], although this method requires intensive computation due to decomposition procedures. More recent methods involve using compact convolutional filters, such as depth-wise separable convolutions (DSConv), which are integrated during the

training process [132,133].

Data augmentation is also used to improve model resilience through the expansion and diversification of training samples which can enhance the capacity of traditional machine learning techniques to generalize effectively [134]. Early stopping is commonly used to avoid overfitting, and ensemble methods can be used to integrate multiple models, hence increasing performance. These optimization algorithms work together to generate a neural network that is both accurate and efficient, with the ability to generalize well to new, unseen data.

2.7 Deep Learning for Pathology Images Analysis

The examination of WSIs in pathology is critical for increasing diagnostic and prognostic abilities. At the cellular level, accurate cell detection and segmentation help to better comprehend cellular morphology and heterogeneity. At the tissue level, the studies focus on broader issues such as tumor categorization, grading, and survival prediction. Integrated multi-level studies, on the other hand, combine individual cellular properties with the larger tissue environment into a unified framework, providing a comprehensive perspective of tissue architecture. Deep learning networks have emerged as a powerful technique for facilitating and integrating these different levels analyses.

2.7.1 Cellular-Level Analysis

In computational histopathology, cell identification and segmentation are crucial processes that significantly assisting physicians in assessing cellular status [43], performing cell counting [135], identifying pathological changes, understanding cellular behavior, and automating other labor-intensive tasks that traditionally used to be processed manually. It has emerged as an essential application in this field, often used to preprocess histopathology images for advanced tasks, enhancing the model's comprehension of the images. Initially, traditional machine learning methods, like KNN and SVM, were used for these tasks, which relied significantly on manually extracted features. However, with the emergence of deep learning, CNNs have substantially

advanced the field by automatically extracting important features from images. Recently, multi-scale deep learning approaches has been widely implemented, investigating cellular structures at different magnifications and resolutions, significantly enhancing the accuracy and reliability of cell detection and segmentation. Examples of applications include cell counting [136,137] cell type detection (e.g., breast cancer cells [138], liver cells [139,140], blood cells [141,142], cervical cells [143,144], corneal endothelial cells [145,146]) , phase contrast analysis of cells in culture [147], and cell nucleus analysis [148,149].

2.7.1.1 Cell Detection

In the last years, CAD systems have relied on conventional machine learning approaches for cell detection in histopathological images [150,151]. These methods required significant effort on feature extraction based on expert domain knowledge. These features encompass characteristics like density [152], surface texture [153] and cell morphology [154]. Afterwards conventional classification algorithms are employed for analysis, including KNN [155], SVM [156] and random decision forest [157] to obtain the cell probability map [158-163]. Additionally, several image processing approaches have also been employed for nuclei/cell detection including distance transform [164] for highlighting the nuclei by measuring the distance to the nearest background pixel; binary morphological filtering [165] for enhancing the structures of interest; and marker-based watershed [166] algorithms for segment the nuclei/cell by treating them as distinct regions separated by boundaries.

Wang et al. [158] proposed a method locates breast cancer regions of interest (ROIs), via utilizing wavelet transform and multi-scale region growing. They also separate overlapping cells by combining adaptive morphological operation combined with a curvature scale space corner detection algorithm. For classifying cell nuclei, they extracted shape-based features and textural features based on color spaces using a SVM with a chain like agent genetic algorithm (CAGA) for the classification process. Abas et al. [167] suggested the use of SVMs and DTs to classify and segment glioma cells in smear preparations. Manivannan et al. [159] presented a technique for recognizing patterns at both cellular and specimen levels in HEp-2 cell images. Ensembles of SVMs were trained to categorize cells into six classes by sparsely encoding texture features

through cell pyramids that captured spatial and multi scale structures. They applied the same approach to classify specimens into seven categories.

Peikari and Martel et al. [160] introduced a transformation step where colors are transformed by converting the red-green-blue (RGB) color space into a different one by calculation of eigenvectors within the RGB space to aid the detecting and segmenting cells. Kowal et al. [161] employed adaptive thresholding for background-foreground segmentation, followed by testing four clustering algorithms for fast nuclei identification and segmentation, including K-means, fuzzy C-means, competitive learning neural networks, and Gaussian mixture models. Awan et al. [162] presented an method for classifying mitotic cells into their four major phases using eigen-phase representation.

Yuan et al. [163] used the AdaBoost classifier for nuclei segmentation and classification into three types based on intensity, morphology and texture features, providing quantitative information about tumor cellular characteristics in tissue. Tareef et al. [168] proposed an algorithm that uses multiple passes of the watershed technique to segment both nucleus and cytoplasm in densely packed clusters of overlapping cells, using a barrier-based watershed for nuclei location and a watershed transform for segmenting isolated and partially overlapping cells, followed by estimating cell shape in overlapping clusters. Veta et al. [169], and Ali and Madabhushi [170] utilized a set nuclei shape along with the watershed algorithm to distinguish grouped nuclei, whereas Wienert et al. [171] opted for operations excluding the watershed method and Liao et al.[172] employed eclipse fitting techniques to separate clusters.

However, these traditional feature extraction approaches are of limited use in biomedical contexts due to the complex and heterogeneous nature of biological data. Additionally, they struggled with densely clustered cells and showed limited generalizability across different factors like nuclear type, magnification, and imaging modality. The adoption of deep learning methods, in particular CNNs, has shifted interest away from traditional feature extraction methods, and offers end-to-end learning frameworks which enables to capture more complex and nuanced patterns in the data producing more reliable and generalizable outcomes [173].

The common method is to perform pixel-wise binary classification. The work of Malon et al. [174] was the first to use deep learning methods for histological images analyses. Their proposed method utilized deep feature recognition by a CNN to train a SVM classifier to locate mitotic nuclei in histological images. Zerhouni et al. [175] used a wide residual CNN in a pixel-wise classification strategy to detect mitotic figures in histological images. Li et al. [176] proposed a weakly supervised deep segmentation network for mitosis detection. Their method involves expanding the weak label of a mitosis centroid to a novel label consisting of concentric circles, and a concentric loss function is then used to train the network, enabling it to perform the mitosis segmentation. Ciresan et al. [177] performed mitosis detection with deep CNN emphasizing pixel-wise classification, saliency mapping, and integration for higher-level diagnosis by employing sliding windows on breast cancer images.

Kumar et al. [178] proposed a method using three-class CNN to specifically delineate nuclear boundaries for identifying nuclei in H&E-stained images sourced from multiple organs. Albarqouni et al. [179] introduced CNN-based mitosis detection method, using additional crowdsourcing layer to effectively manage data collection and integration throughout the training phase. Xu et al. [84] applied the stacked sparse auto-encoder combined with linear-discriminant analysis and ridge regression for nuclei detection in breast cancer images. Xie et al. [180] proposed a cell counting and detection method based on a fully convolutional regression network involves regressing a gaussian-based density map from a histology image with the density peaks aligning with the centers of nuclei. Bejnordi et al. [64] presented a context-aware stacked CNN for breast cancer classification. their approach involved training a CNN on high-pixel-resolution patches to extract cellular level features, which were then processed by a second CNN for classification. Wang et al. [181] proposed an automated cell type classification pipeline to convert a pathology image into a spatial map of the cells. This map was then used to extract features related to the tumour microenvironment. Based on these features, the authors developed an image feature-based prognostic model.

To enhance the efficiency of algorithms, region proposals [182] were introduced as a substitute for classifying individual pixels. Methods have become standard of object detection, such as RCNN [182], Fast-RCNN [183], Faster-RCNN [184], MaskRCNN [185], Cascade-RCNN [186],

RetinaNet [187], SSD [188], HTC [189] and Detectors [190]. In a study by Shkolyar et al.[191] live microscopic images were processed to identify potential instances of mitosis and CNN was employed to detect mitotic cells from their counterparts. Dong et al. [192] employed a max pooling CNN with supervision to identify cell pixel areas that had been preselected using an SVM classifier. To obtain precise regions of interest, Johnson et al. [193] utilized Mask RCNN to locate the cell nuclei bounding boxes in pathology images. Liu et al. [194] proposed a deep CNN to assigning scores for nuclei candidates then used a programming approach to choose a subset of these candidates as the final detection results. Chandradevan et al. [195] conducted a study for detection and classification bone marrow cells using a fast R-CNN built upon a ResNet101 fully convolutional network.

Yi et al. [196] combined the single shot multi-box detector (SSD) with the standard UNet and ResNet50 to predict both the bounding box and the segmentation mask of individual cells at the same time. The attention mechanisms were used to enhance the precision of detection and segmentation. Cheng et al. [197] proposed the mask regional CNN (Mask-RCNN), which used an output layer mask prediction branch for nuclei instance bounding boxes detection segmentation in microscopic images. Obeid et al. [198] proposed NucDETR model including CNN model for feature extraction, positional encoding for ordinal information, object queries, and Transformer encoder-decoder modules for nucleus detection in histopathology. Huang et al. [199] introduced an Affine-Consistent for nuclei detection using two sub-networks, a global and a local network. The local branch is designed to analyze distorted input images at smaller scales while the global network provides predictions at a larger scale to serve as additional guidance signals.

Although these approaches reduce the computation complexity compared with pixel-wise classification methods, their performance is reliant on the quality of generated boxes, in particular, the overlapping and clustered nuclei which complicate the detection. However, deep learning techniques require large training data and computing resources for nuclei detection. Javed et al. [200] introduced a context aware hierarchical deep correlation filter that takes into account spatial constraints for detecting nuclei in histology images. This approach can be applied using smaller training datasets and run on standard desktop computers.

2.7.1.2 Cell Segmentation

Cell segmentation techniques can be categorized into box-free segmentation and box-based segmentation. Box-free segmentation approaches delineate the cell's boundaries or masks directly without needs to the bounding boxes.

Traditional segmentation methods include edge detection, threshold processing, region growing, and the morphology watershed algorithm [201-203]. Santamaria-Pang et al. [201] presented a method that employs a ranking function for cell shapes. They used shape segmentation as an optimization problem that maximizes shape similarity within specific classes and introduce algorithm to efficiently tackle multi-scale segmentation with minimal steps. Brieu et al. [202] modelled the nuclei detection as an issue of spotting peaks in a probabilistic map, introducing a map illustrating the surface area of nuclei to assist in choosing these peaks without needing any beforehand information about the size or shape of the objects being analyzed. Song et al. [203] developed a sparse convolutional model with multiple layers to segment nuclei by integrating the appearance features at the low-level with contextual features at the high-level into probabilistic binary DTs to accurate final contour inferences.

More recently, DL has become a popular image segmentation method as they achieve higher accuracy than traditional methods.

For instance, Zhou et al. [204] proposed a contour-aware information aggregation network (CIA-Net), which incorporates a module for gathering information at different levels situated between two decoders tailored for specific tasks. By bidirectionally combining features it harnesses the benefits of both spatial and texture relationships, between nuclei and contours. Chen et al. [205] proposed a deep contour-aware network (DCAN) that integrating contour information with object appearance and used two branches of deep learning models for colon gland segmentation and boundary detection, respectively. Koohbanani et al. [206] proposed a spatially-aware network (SpaNet)) that takes spatial awareness into account. This approach involves multiple scales to effectively capture spatial details. The process begins with segmentation at the pixel level and detection the centroids of nuclei. Next it predicts the position information for each

nucleus instance and finally using clustering to identify connected components and assign cluster identifiers.

Oda et al. [207] introduced a new model structure BESNet by adding an extra decoder branch to the U-Net network in order to enhance the cell segmentation achieved in the other decoding branch. Graham et al. [208] proposed Hover-Net deep learning approach for simultaneous segmentation and classification of nuclei within multi-tissue histology images. Their approach includes foreground/ background segmentation and predicting the horizontal and vertical distances of nucleus pixels from their centers of mass. Once an instance is segmented, they identify the type of nucleus through a specialized-up sampling process. A recent extension of this work, Hörst et al. [209], replacing the convolutional encoder with a Vision Transformer (ViT) for simultaneously instance segmentation and detecting of cell nuclei.

Pham et al. [136] evaluated the classic machine learning methods (SVM and RF) alongside deep learning approaches (U-net, and FCN) for segmenting and counting cells in the spinal cord. Manca et al. [138] proposed a CNN segmentation method from urinary bladder urothelial cells. Tai et al. [139] employed a combination of morphological reconstructions, slope detection, and the watershed algorithm to generate a mask and then used CNN to assess the segmentation of liver nucleus. Jha et al. [210] developed a hybrid model based on mutual information and a deep CNN classifier to segment leukocytes, combining active contour model results with the fuzzy C-means algorithm. Features were then extracted from the segmented images, and a CNN classifier was trained using the temporal Sine Cosine Algorithm.

Tran et al. [141] applied deep learning method for identification and segmentation red and white blood cells in images of blood smears. Araújo et al. [143] introduced a CNN-based method to segment abnormal cell clumps even when they are highly overlapping. While they were able to segment clusters of cells regionally using the CNN their approach does not extend to segmenting cells on an individual basis. Joseph et al. [145] developed a method to segment endothelial cells using a combination of image flattening U-Net deep learning and postprocessing steps to achieve distinct cell segmentations. Ayanzadeh et al. [147] introduced a multi-resolution approach using a U-Net architecture to segment cells in two-dimensional phase contrast microscopy images.

This model incorporated a specific form of data augmentation. Long et al. [211] proposed an improved version of U-Net, known as U-Net+, with modified encoder branch designed to operate efficiently on low resource computing systems specifically for the task of segmenting cell nuclei. Mahmood et al. [212] suggested using adversarial training with a CNN model to segment nuclei. They employed a conditional generative adversarial network (cGAN) to create images that would aid in training the network. Cui et al. [70] proposed a contour-aware fully CNN that outputs estimated nuclei and boundary maps.

To further specify, box-free segmentation methods could be classified into two types, including instance-based methods and dots map-based methods. Instance-based methods [208,213-215] segment each cell directly from the input image using pixel-wise classification such as semantic segmentation maps and distance maps, then obtain each nucleus mask with some post-processing methods like watershed algorithm [213].

Doan et al. [213] proposed a neural network based on self-guided ordinal regression to simultaneously segment and classify nuclei. This approach leverages the inherent traits of nuclei and concentrates on regions with high uncertainty during the training process. He et al. [214] proposed a centripetal direction network (CDNet) that takes into account the direction of pixels for nucleus separation. This network consists of two parts namely direction-guided refinement and feature extraction. The attention mechanism used in this approach acts as a feedback loop. Yu et al. [215] proposed a technique using diffusion to enhance the segmentation of nuclei in histopathology images. This approach allows the model to create nuclei structures that closely resemble their actual shapes and spatial arrangements.

Some works [216-218] identify nuclei using standard instance segmentation techniques originally designed for natural images. Zhang et al. [216] proposed Cell R-CNN framework to create panoptic segmentation. Their approach combines semantic and instance segmentation tasks by adding a semantic segmentation branch that can extract features from areas that significantly overlap with other objects. Loh et al. [217] developed an automated deep mask R-CNN to identify and quantify red blood cells malarial parasites from blood smear images. Paing et al. [218] developed Mask R-CNN models to identify and segment multiple myeloma cells in

microscopic images of bone marrow aspiration slides.

However, these instance-based methods computationally intensive due to its need to the pixel-level annotations of nuclei boundaries. For the dots map-based approaches, they predict cell centers (dots) locations and then predict the cell boundaries around these dots without needs to bounding boxes [219-224]. Xie et al. [219] presented CNN based structured regression model for cell detection. Höfener et al. [220] pinpointed various factors and fine-tuned the fundamental PMap method by using different combinations of these parameters to identify nuclei in histopathology images. Feng et al. [221] developed a Mutual Complementing Framework (MCF) to enhance the detection and segmentation of nuclei. Within the MCF, the detection component enriches the pseudo mask generated by the segmentation component. Additionally, the progressively trained segmentation branch fills in the gaps for the detection branch by providing nucleus templates that are missing.

Abousamra et al. [222] proposed an approach that combines detection and classification by taking into account spatial context. Their method employs a statistical function to capture the density of information in a multi class and multi scale fashion. Wang et al. [223] proposed adversarial feature alignment approach to enhance the detection of nuclei in different domains. This approach involves aligning features on a global scale and also focusing on specific areas to effectively transfer information between domains. Wu et al. [224] proposed a 3D nuclei centroid detection method through an analysis of intensity peaks within specific voting areas. Sirinukunwattana et al. [61] introduced a CNN with spatial constraints to identify cell nuclei. They used an adjacent group predictor for classifying these nuclei. Xu et al. [225] employed stacked sparse auto-encoders for nuclei detection.

The box-based segmentation approaches leverage cell detection and segmentation simultaneously following a two-stage process. They first predict bounding boxes around individual cell indicate their locations within the image, then generate a binary mask within each bounding box to segment the cell. For instance, Yousefi et al. [226] explored a transfer learning model builds upon the Faster R-CNN to classify nuclei in histopathology images. They identify nuclei by employing models that are not specific to any particular class and then utilize the CNN

representations of the detected nuclei for clustering and categorizing them. Sun et al. [227] presented a similarity-based region proposal network (SRPN) for identifying nuclei and cells in histology images. The SRPN architecture incorporates an embedding layer that enhances its ability to extract distinguishing features through similarity learning.

Li et al. [228] improved the Faster RCNN-FPN model to detect abnormal cervical cells in Pap smear images. They achieved this by incorporating deformable convolution layers into the feature pyramid network (FPN) and introducing a global contextual aware module in addition to the Region Proposal Network (RPN). These enhancements aimed to enrich the spatial context information, for detection purposes. Wang et al.[229] developed an hierarchical deep learning model effective in accuracy and classification time for detection and classification of blood cells and generating regions of interest for subsequent analysis.

Liang et al. [230] introduced a convolutional network that focuses on regions, integrating a guided anchored into the region proposal network (RPN) for nuclei detection and segmentation. The foundation of this network is built upon Mask RCNN and feature pyramid networks (FPN). Wu et al. [231] proposed a scalable 3D microscopy volumetric nuclei centroid detection. They used RCNN-SliceNet to identify the nuclei centroids in each slice of the volume from various angles and then applied a clustering method, called 3D agglomerative hierarchical clustering (AHC) to determine the 3D positions of the nuclei within the volume.

Mask RCNN [185], a variant of Faster RCNN [184], was frequently employed for cell segmentation, but it struggled to segment overlapping and clustered cells. Some other methods, such as SSD [188] and RetinaNet [187], concentrated on making predictions through a sliding window mechanism, while avoiding the second refinement stage and its showed promising performances. Wang et al. [232] proposed a multipath dilated residual network based on Mask RCNN model [185] for dense nuclei segmentation and detection. Similarly, Cheng et al. [197] designed the mask regional CNN (Mask-RCNN) to detect and segment nuclei bounding boxes in microscopic images. Yi et al. [196] utilized the single shot multi-box detector (SSD) to forecast the boundaries of nuclei by bounding boxes which were then employed to extract patches and assist in the process of instance segmentation.

Chen et al. [233] introduced a network called BRP Net those aids in nuclei instance segmentation by leveraging boundaries. The approach involves utilizing a tailored feature encoding network to generate instance proposals followed by segmenting those proposals through an instance segmentation network. Hatipoglu et al. [234] proposed a method for segmenting cells by examining how they are positioned in relation to each other. This involves gathering both cellular and extracellular samples from histopathological images using a technique called windowing which involves looking at small sections of different sizes.

Transformer-based methods have also been employed for cell identification, using the ability of transformers to handle complex contextual relationships within the data. Huang et al. [235] introduced a transformer framework that utilizes prompts for detecting and classifying nuclei. Their approach involves a classifier that organizes nucleus features into clusters and category embeddings, which are then correlated with nuclei to identify cell types.

2.7.2 Tissue-Level Analysis

In the field of computational histopathology, analyzing tissue samples involves examining WSIs to gather insights that aid in making clinical decisions. This analysis at the level which includes predicting survival outcomes and grading tumors plays a role in offering prognostic information and shaping treatment approaches.

Survival prediction and tumor grading tasks have traditionally relied on statistical models that analyzed patient data in relation to clinical and pathological factors. Early computational approaches involved extracting features manually and using classifiers such as KNN, SVM and RF for classification purposes. With the advent of deep learning, CNNs emerged as tools for detecting intricate patterns in pathology images linked to outcomes. These advancements were significantly enhanced through the automatic learning of features from WSIs. More recently, GNNs have been introduced to capture spatial relationships and interactions among different regions within tissue samples. By integrating multi-level information, they offer a comprehensive perspective on tissue architecture. These graph-based approaches now aim to combine analysis examining tissue at various magnifications and resolutions to further enhance

the accuracy of classification tasks.

2.7.2.1 Survival Prediction

Survival prediction is estimation how long the patient will survive or the likelihood of survival over a certain period [236] based on expert judgment and clinical data. Survival prediction is crucial to assess prognosis, guide options for treatment, and evaluate the efficacy of therapies [237]. Standard survival prediction methods can be categorized into three types: manual prediction, statistical prediction and machine learning prediction.

Manual prediction involves healthcare professionals using their experience of disease progression and patient history as well as some other clinical indicators, biomarkers, imaging results, and qualitative assessments to inform prognosis and treatment decisions. It is widely used some metrics for assessing patients functional statuses and their overall health, such as the Eastern Cooperative Oncology Group score (ECOG) and the Karnofsky functional status score (KPS) [238]. However, the results of the prognoses are too subjective depending on individual expertise and the complexity of the case. A study conducted by Glare et al. [239] revealed that predictions regarding survival made through manual methods were generally more positive about 3 to 5 times compared to the patient's actual survival duration. Therefore, automated prediction methods, such as machine learning, tend to be used as a complement to the manual prediction by creating data-driven conclusions and improving accuracy through systematic examination of massive datasets.

Statistical prediction methods are crucial tools in clinical practice, usually they integrated with the scoring systems to assign weights to different prognostic factors and compute an overall score to classify patients into different risk groups. Statistical models involve Kaplan-Meier (KM) [240] , Nelson-Aalen estimator (NA) [241] , Cox model [242], etc. Based on these models, several scoring models, such as OPS [243], PaP [244], OPPS [245], CPS [246], and PPI [247], have been developed. Although these approaches, provide a systematic and objective prediction, it fails to account individual differences as the prediction results reflect the general trend. Therefore, it is crucial to develop a model that predicts how a person's cancer progresses.

The advancement of pathology has garnered considerable interest in predicting survival outcomes based on pathology images. Traditional methods have played a foundational role in the analysis of pathology images and for survival prediction. For instance, Jiang et al. [248] developed a GC-SVM classifier to predict survival and identify adjuvant chemotherapy benefit of gastric cancer patients using immunomarkers in immunohistochemistry staining slices. Dimitriou et al. [249] employed a K-nearest neighbor approach to predict the survival of stage II colorectal cancer patients based on immunofluorescence images.

Recently, the development of neural networks has led to the introduction of survival models that leverage deep learning techniques to enhance deep representation capabilities. Zhu et al. [250], Kather et al. [251], and Yao et al. [252] all put forward CNN based approaches for predicting survival outcomes by utilizing patch-level images and risk assessment scores. Kather et al. [251] introduced deep residual learning to directly predict microsatellite instability (MSI) tumors based on H&E histology images. Yao et al. [252] present a deep attention multiple instance model that leverages histology images to uncover survival patterns. This approach combines a Siamese MI-FCN with an attention-based MIL pooling mechanism to enhance its learning capabilities from the data. Katzmanet al. [253] proposed a deep fully connected network (DeepSurv) to model the risk function. Similarly, Zhu et al. [103] developed a deep CNN (DeepConvSurv) for survival analysis using pathological images.

Later, Yao et al. [254] combined genomic data with DeepConvSurv to predict survival outcomes based on diverse data sources. It's worth noting that DeepConvSurv requires pathologists to choose specific region of interest patches from WSIs for its convolution processes. In a study by Mobadersany et al. [255] a convolutional neural network (SCNN) was developed to predict the survival of glioma patients. This SCNN merges features from the 19-layer VGG convolutional network with a Cox proportional hazard to predict the survival. In another study, Yu et al. [256] developed an automated system to predict the prognosis of lung cancer patients differentiating between those with shorter survival times and those with longer ones. Wulczyn et al. [257] developed a weakly-supervised deep learning system (DLS) to predict survival rates in various cancer types, and tested three different survival loss functions. Jiang et al. [258]

employed a CNN model combined with a machine classifier to predict the prognosis of stage III colon cancer patients based on H&E-stained images.

Even though these studies showed promising performance, they were usually limited to certain stages of the cancers in question. Additionally, they did not take into account any spatial data from the tumor microenvironment (TME).

Patch-based methods may not completely capture the full details of the patients' tumor morphology, and they often use averaging techniques to obtain predictions at WSI-level which cause limitation in their performance. Therefore, comprehensive methods that consider how different spatial characteristics related to cell distribution are connected to outcomes are important. The unique structure of GNN allows them to effectively capture contextual information, making them a focus of interest in the field of computational pathology. Some graph-based methods have proposed that designed based on the tumor microenvironment.

Li et al. [90] developed DeepGraphSurv framework to integrate spatial information from WSIs for survival prediction. They randomly selected six patches on WSIs and fed them into a pretrained VGG16 [53] and used PCA algorithm [259], the reduced data then used as nodes to build a graph. The connections between nodes were established using Euclidean distances. The resulting graph was then processed by a GCN with an attention layer to predict survival outcomes. Wu et al. [123] proposed DeepGCNMIL framework. They start by segmenting the tissues of the WSIs, followed by extraction the features using pre-trained ResNet50 model and cluster them with K-means. Each patch feature within the cluster treated as a node in a graph with edges based on Euclidean distances. The constructed graph is then fed into a GCN and a multi-headed attention mechanism to evaluate prognostic risk.

Lee et al. [260] introduced graph deep neural network to assess the likelihood of events like mortality or cancer spread and examine the pathological context elements associated with predicting outcomes. Their proposed method for constructing the graph of the tumor microenvironment [261] entails finding comparable images at the patch level using spatial similarity and cosine similarity. This process yields super patches that serve as nodes with spatial

distance functioning as edges. The resulting graph captures contextual features using an enhanced GAT module [262] based on GNN and offers interpretability through attention scores associated with edges and the IG method.

Wang et al. [263] developed a graph network that leverages the tumor microenvironment to digitally stage and predict survival outcomes for gastric cancer images. The process involved using pathology software to segment and extract features from nuclei creating a graph representation. In this graph the nuclei served as nodes while the Euclidean distance between cells meeting a threshold acted as edges. Additionally extracted features were incorporated as attributes. Finally, they employed GINTopK [264] as the GNN model for the survival analysis.

Methods using traditional graph structures may not effectively convey certain concepts. To address this limitation some research has suggested modeling WSIs using hypergraphs instead.

Di et al. [265] developed a neural network that utilizes hypergraph factorization [266] for predicting survival outcomes based on WSIs. This approach incorporates the relationships between vertices and hyperedges into two compact semantic spaces. The model encompasses processes, such as a hypergraph decomposition neural network that generates a holistic representation of WSIs by condensing hypergraph features with numerous vertices obtained through thorough sampling using an innovative factorization module. Additionally, it includes a multi-level ranking survival prediction component that introduces a loss function merging fitting of WSIs with patient level ranking data to reduce the impact of uncertainty.

Clinical professionals face challenges in providing image annotations since there is no direct and straightforward association between pathology images and survival outcomes, making models development and training require large-scale of WSIs. However, using images on a large-scale can lead to redundancy in the data increasing the model's processing requirements. Two common approaches to address this issue involve either downscaling the features extracted from images [90,123,260] and modifying the graph structure [265]. Some datasets include labels for survival outcomes, such as overall survival [267], survival risk scores [268], and risk categories. In many studies, survival prediction is treated as a classification problem rather than

a regression problem [263] by categorizing the survival period into different groups. This significantly simplifies the task complexity and difficulty level.

2.7.2.2 Tumor Grading

Cancer grading is one of the most critical tasks in cancer pathology since it is known to be associated with patient outcomes. The factors causing tumour grade can vary by the type of cancer, however it is often classified into separate grades based on the amount of abnormalities in tumour cells, for example, low, mid, and high-grade tumours [269]. The greater the tumour grade, the more aggressive the tumour cells become. This indicates tumour grades follow a natural ordering, suggesting that cancer grading is fundamentally an ordinal problem.

In clinical practice tumor typing, grading and staging involve assessing cancer based on how differentiated the nuclei are. This assessment includes factors like the size of the nuclei, the ratio of nucleoplasm to nucleus and the division of nuclei [270]. Therefore, it essentially focusing on extract cellular features when developing the model. The typing, grading, and staging of pathology images have been mostly investigated as a multi-class categorical classification problem [63,271,272].

During the initial phase of classifying cancer tissues, the methods put forward primarily relied on handcrafted features to classify tissues. Daskalakis et al. [273] proposed a method that employs classifiers such as k-NN, linear least squares minimum distance (LLSMD), statistical quadratic Bayesian, SVM, and PNN using majority vote, maximum, minimum, average, and product rules to discriminate the H&E stained slides of thyroid nodule either as benign or malign. Wang et al. [158] used SVM to categorize breast cancer based on histopathology images. To identify regions of interest (ROIs), they proposed using a wavelet transform along with a multi-scale region growth method. Additionally, they separate overlapping cells by utilizing an adaptive morphological operation combined with a CSS corner detection algorithm. Reza et al. [274] proposed a method for analyzing the shape characteristics of cell nuclei to differentiate gliomas in pathology images as either GBM or LGG. They employed k-means clustering to extract features based on nuclei morphology such as area, perimeter, eccentricity and major axis

length. These extracted features were then classified into tumor types using a multi-layer perceptron (MLP) model.

Sheikh et al. [275] proposed a multi-scale and multi-feature network that fuses multi-resolution feature maps at different layers for breast cancer classification. This proposed model learns different scale features to account for global and textural cellular features at the same time. Furthermore, Wang et al. [276] employed a combination of morphometric sub-visual metrics and Ki-67 PI data along with the SVM algorithm for brain tumor grading. Mousavi et al. [277] proposed an algorithm to grade glioma tumor into LGGs from GBMs based on image features extracted from regions of interest (ROIs) in WSIs. Brindha et al. [278] introduced a framework for grading brain tumor. The method defines the nuclei boundaries using geometric algorithms, establishes a dictionary of cell nuclei through clustering of nuclei, and uses this dictionary to speed up patch level classification.

Nevertheless, because of the high resolution of WSIs, CNN models often received input in the form of multiple patch-level images extracted from the complete WSIs. For instance, Romo-Bucheli et al. [279] integrated handcrafted features with CNN models to detect and identify tubule nuclei from WSIs. Similarly, Wang et al. [280] introduced a two-step method that integrates manually crafted features like morphology, color and texture with a CNN model to assist in counting mitotic cells and assessing breast cancer grades. Yonekura *et al*. [281] presented a classical CNN model to differentiate between HGG and LGG glioma tumor. Ker et al. [282] used a pre-trained Google Inception v3 model to classify brain histology samples into normal, high or low-grade glioma. Ertosun and Rubin et al. [283] classified LGGs and HGGs by using automated segmentation to examine cell nuclei and morphological features.

Awan et al. [284] presented a context-aware model for classifying breast cancer histology images using contextual information. Their approach involved employing a CNN model to extract features from distinct patches and subsequently utilizing a classifier trained on features from overlapping patches. Similarly, Bejnordi et al. [64] proposed a method that uses a stacked CNN to analyze high resolution image patches. This approach helps to capture detailed cellular features while also enabling the model to learn contextual aspects by stacking larger input

blocks. The results of these two methods better on the cancer classification problem.

Although these methods have made significant improvement in model performance, patch-level images for training causes of losing context across the WSI, limiting the model effectiveness and its generalization abilities. Recently, GNNs have garnered broad interest for their potential to represent unstructured data. Constructing graph structures addresses the problem of losing information that comes with using patch-level images approaches. Specifically, in the classification of pathology images, the graph-based method enhances the understanding of contextual features within WSI, improves classification performance and generalization ability [87].

Zhou et al. [99] developed a network that uses cell graph convolutional techniques to assess the grade of colorectal cancer histology images. For graph construction, each individual cell nucleus is used as a node and the interactions between cells are represented as edges based on how similar the nodes are to one another. Even though this study yielded favorable findings, the model is relatively more complex, and the dataset's performance could be improved further. Therefore Lippoldt [285] proposed a better-performing approach due to its ability to the filtering spatial features and enhancing visuals, which allows the model to better represent its capabilities through spatial modeling.

These two previously discussed methods extract cell nuclei features for nodes manually, which is difficult and time-consuming task. Su et al. [286] proposed Transformer Graph Neural Network (HAT-Net), which leverage both GNN and Transformer for colorectal cancer grading. They introduced a masked nuclei patch (MNP) network to train ResNet50 [55] as a feature encoder to extract more representative nuclei feature. This method enables the extraction of cellular features in a seamless manner and a comparable technique has been employed in many of the later research works.

Lu et al. [287] proposed a GNN to predict HER2 status from breast cancer H&E slides. The method used Hover-Net [208] to segment and classify cell nuclei, and extracted features through ResNet50, which has been pre-trained on ImageNet [288], and clustered regions that are spatially

adjacent and exhibit high similarity. For the graph construction, the geometric centers of the clusters are designated as nodes while the Euclidean distances meeting threshold values serve as edges. Ultimately a GCN model is employed for the purpose of classification.

These studies only targeted one type of cancer. So, a more generalized methods are desirable in which can be used across various cancer classification challenges. Hou et al. [289] proposed a GNN that incorporates a dynamic structure learning (DSL) component to classify colorectal and breast cancers. First, they segmented the nuclei using pre-trained HoverNet [208], merged pixels after cropped the WSIs at various magnifications using the SLIC superpixel algorithm [290]. Feature extraction was done employing a pre-trained ResNet34 network, followed by features aggregation. In the next step, a dynamic structure learning (DSL) module was introduced to adaptively learn the adjacency relationships, among different layers of images for constructing dynamic graph structures. Ultimately a GCN network integrated with an optical transducer (VIT) module [291] was employed to extract dynamic graph features and facilitate subsequent predictions. Through dynamic learning strategies and spatial structure encoding the model aims to simultaneously perform multiple cancer classification tasks.

Hou et al. [292] proposed a heterogeneous GNN to learn representations from a diverse graph at varying levels of detail for tumor typing and staging. A heterogeneous graph is generated to capture the characteristics and spatial relationships of patches with different resolutions. Then, a resolution-aware attention convolution module is employed to enhance the process of updating nodes in the graph and iterative hierarchical pooling module to consolidate the heterogeneous graph based on the scaling relationships among its nodes.

Bai et al. [293] introduced a scalable graph-based framework to classify histology images across multiple organs. After nuclei were segmented and classified, nuclei features extracted using pre-trained ResNet50. An average nuclear sampling strategy is also used, which calculates the average of the channel by considering nuclear features and positions of all nuclei in each grid cell. The updated nuclei coordinates serve as nodes to construct the cell graph and KNN algorithm [294] is utilized to identify the edges. Finally, a multi-stage GNN is employed for predictions incorporating a GIN layer [295] and a Transformer layer [296].

Besides utilizing cell nuclei as nodes in the graph, there are research efforts that employ patch-level nodes. This approach involves extracting features using a pre trained CNN model, which streamlines the process. Gao et al. [297] introduced a quadratic GCN (GQ-GCN) network for classifying breast cancer subtypes. For graph representation, a CNN is used to extract features of the histopathological images and each pixel on the feature map serves as a node with edges formed based on the Euclidean distance between them. Furthermore, the group graph convolutional network (G-GCN) is designed to perform feature selection and compress graph representations. The inclusion of an operation in the graph convolution is aimed at enhancing the capacity of a single neuron to handle intricate data.

In a similar vein, Zhang and Li [114] proposed a multi-scale graph wavelet neural network (MS-GWNN) to classify breast cancer histopathological image. The approach involved considering pixels in the feature maps of image patches as individual nodes and utilizing the similarities among these nodes as edges. A graph wavelet neural network (GWNN) [298] is then utilized to categorize nodes across various scales and combine this information with multilayer embedding for conducting classifications at the graph-level. Baranwal et al. [299] developed Cell Graph Attention Network to classify pancreatic cancer and its precursors based on immunofluorescence histology images. The framework employs a GCN for disease grade classification and incorporates an attention layer to pinpoint the nodes most pertinent to the prediction task.

However, the current GCN approaches tend to create graphs at the cell-level or tissue-level. This process involves preprocessing steps on histopathological images, like detecting and segmenting nuclei. To simplify graph construction, Gao et al. [300] proposed a CNN-GCN framework. This framework integrates graph construction for histopathological images within a network. It builds a graph using feature maps produced by the CNN and employs GCN to capture the spatial feature representation. This offers a different approach for GCN based CAD in histopathological image analysis.

Moreover, some studies used the cell's morphology and topology and improve the learning of contextual information. Pati et al. [121] proposed hierarchical cell-to-tissue-graph (HACT)

representation for breast cancer subtypes classification. It involves a graph to depict cell shape and interactions at a granular level, along with a tissue graph for representing the structure and arrangement of tissue components, and cells-to-tissue hierarchies to illustrate how cells are spatially distributed within the context of tissue distribution. First, the detection of nuclei and tissues is carried out to create a k-NN graph and a region adjacency graph (RAG). Then, a hierarchical map of cells and tissues is developed based on these two graphs. Finally, a graph isomorphic network (GIN) is employed to analyze information and facilitate classification. In their subsequent research, same authors [301] enhanced the HACT-Net by substituting the GIN with the principal neighborhood aggregation (PNA) network [302].

2.7.3 Integrated Multi-Level Analysis

Integrating information from different levels, resolutions, scales, and views provide a holistic view of the data, which is crucial for efficient diagnosis and treatment planning in medical fields. These multi-input strategies capture an extensive range of features and patterns, minimizing the likelihood of missing critical details and resulting to more precisely and reliable results. With these methods, professionals can obtain a deeper and more comprehensive understanding of complex biological phenomena, which ultimately leads to better outcomes in computational pathology field, where precise and context-aware analysis is critical.

Initial work focuses on integrating various types of features, such as texture, shape, and intensity, derived from image patches or entirely images. These fused features were then fed into classic machine learning classifiers such as SVMs or RFs. These studies demonstrated the potential benefits of integrating many feature types, but they frequently encountered difficulties with scalability and adaptability.

Many studies have been conducted to incorporate multi-level information into pathology images for the purpose of predicting cancer, which used a variety of input variants, such as the Multi-View Network [119,303,304], the Multi-Scale Networks [305-316], the Multi-Resolution Networks [275,317], the Multi-Level Networks [318], Multi-Crop Networks [319] Multi-Magnification Networks [320] and the combination of CNN and RNNs [79,321,322].

For instance, Zhang et al. [119] proposed a GAN model called multi-view attention (MVA) that is lightweight and utilizes superpixels to enhance nuclei segmentation and tissue classification tasks. The method utilizes a superpixel segmentation technique that relies on the Bounded Asymmetric Gaussian Mixture Model and a Multi-view Attention Module to segment the nuclei with small areas and fuzzy boundaries. Shen et al. [319] introduced a pooling approach that involves using multi-crops to extract features at different levels. These extracted features are then combined to create the final feature vector. Tokunaga et al. [304] proposed a multi-view CNN capable of utilizing image features from various magnifications to identify distinct cancer subtypes in pathology. The approach involved employing CNNs simultaneously to analyze images at different magnifications and combining the results through a weighted integration.

D'Amato et al. [305] explored two approaches of multiple instance learning framework along with a representation learning algorithm based on clustering to work with multi-scale images, demonstrated the advantages of extracting and merging image tiles across different scales. Mika et al. [307] proposes a multi-scale CNN to estimate the mutational burden of tumors. Also, the authors of these two studies [308,309] incorporate patches of varying sizes for tumor segmentation. Sai et al. [310] introduce a multi-scale multi-branch model for classification purposes, while these methods [311-313] combine MIL and multi-scale inputs.

All of these techniques show that models with multi-scales work much more effectively than those with just one scale. Rijthoven et al. [317] proposed a multi-resolution framework combining context and intricate details using encoder-decoder branches for breast and lung cancer segmentation. The model takes in image patches at various resolutions and views, sending them through distinct branches. The intermediate representations are then merged using a hooking method. Zhou et al. [321] suggested framework for classifying hyperspectral images (HSIs) in terms of their spectrum and spatial characteristics involves the use of CNNs and Gated Recurrent Units (GRUs). CNNs are employed to capture the spatial aspects of the HSI data, and these extracted features are subsequently fed into a fusion network that leverages GRUs to carry out both feature level and decision level fusion processes. Yan et al.[79] suggested a hybrid CNN and RNN network to classify breast cancer histopathological images. CNN is used to extract the multi-level patch wise features and the RNN is used to fuse these features to make the final

image classification.

Sheikh et al. [275] proposed a multi-scale and multi-feature network that fuses multiresolution feature maps at different layers for breast cancer classification. This proposed model learns different scale features to account for global and textural cellular features at the same time. However, the above approaches aim to capture features in a hierarchical manner within Euclidean space where the concept of scale is usually tied to position (Euclidean distance).

Existing graph-based algorithms are mainly generated using the single-path network by stacking graph convolutional layers to derive higher-level features from low-level features in the initial layers [262,323]. Nevertheless, this single-path GNN model is limited in its ability to merge features across various levels of the graph [324]. Indeed, networks with multi-path structure have the potential to capture features enhancing their ability to represent data for cancer prediction tasks.

More recent methods integrate feature fusion within GNN, allowing context-aware learning of node representations in non-Euclidean space. These approaches demonstrated considerable improvements in tasks such as cell classification, tissue segmentation, and survival/grading predictions. Among these methods, Zhang and Li [114] presented multi-scale graph wavelet neural network (MS-GWNN) to classify breast cancer detection. By merging features from various scales MS-GWNN can capture contextual interactions throughout an entire pathological slide. Ding et al. [325] introduced a Fractal GCN (FGCN) to learn multi-level graph representation by combining graph convolution with a fractal structure. In FGCN, the Mixer layer of MLP-mixer and non-local attention mechanism in MMFFU are introduced to fuse multi-level graph features from various paths.

Bazargani et al. [326] introduced multi-scale relational GCN to grade prostate cancer histopathology images by handling different embedding spaces at each magnification. Similarly, Shi et al. [327] introduced a framework for multi-instance learning that takes into account the structure of graphs hierarchically. This approach aims to gather contextual details from various magnifications to aid in the staging of WSIs. Additionally, Bai et al. [293] developed a

hierarchical transformer GNN to handle large-scale cell graphs and learn long-range dependencies. Pati et al. [121,301] proposed a hierarchical cell-to-tissue (HACT) graph to classify breast cancer, integrating both cell and tissue elements via an assignment module based on spatial location. Javed et al. [94] explored cellular communities using cell-level and patch-level graphs in colorectal cancer analysis, facilitating the identification of distinct tissue phenotypes.

Gan et al. [328] introduced a technique that combines multi-graphs to create a well graph and reduce the dimensionality of high dimensional data specifically for the GCN model. Their approach involves extracting both shared and complementary information from several local graphs to construct a cohesive local graph. This unified local graph is then merged with the global graph of the data to form the initial graph used in the GCN model. Ding et al. [329] introduced the Multi-scale Efficient Graph Transformer (MEGT), a dual-branch approach to enhance cancer diagnosis accuracy on WSIs by combining image patches of varying resolutions. They developed multi-scale feature fusion module (MFFM) to combines features at different scales to capture various details and bridge the gap in meaning between patches of different resolutions.

Abbas et al. [330] proposed a network called MMGA-Net that utilizes multiple cell types and levels to grade cancer. This network creates various cell graphs at different levels to capture both intra and inter cell type relationships as well as global and local interactions between cells. Additionally, they employ a CNN to extract contextual information from the tissue. Finally, the tissue and cellular data are combined to predict the cancer grade.

CHAPTER 3

A Multi-Scale Deep Learning Model for Analysis of Single Cells in Whole-Slide Images with a Note on CD276 in Glioblastoma

Routine examination of entire histological slides at cellular resolution poses a significant if not insurmountable challenge to human observers. High-resolution data such as the cellular distribution of proteins in tissues, e.g., those obtained following immunochemical staining, are highly desirable. As the old term glioblastoma multiforme emphasized, glioblastomas are extremely heterogenous morphologically. This applies not only to the macroscopic visual and neuroimaging but also to the histological and even to the cellular and subcellular levels. Not only the tumour cells themselves but also other cells, e.g., microglia and macrophages that are found in these tumours in high numbers, show a great variety of phenotypes. In addition, there is great intratumoral heterogeneity regarding the distribution of the various cell types in glioblastoma, which may include all the cell types present in normal brain tissue that has been infiltrated by the tumour, as well as peripheral immune cells. However, manual microscopic analysis of entire histological slides at cellular resolution is very time-consuming. Depending on the type of tissue and on the labelling results obtained, this task may become prohibitively complex and expensive for human observers [331].

With the development of digital pathology, many researchers expect that AI, especially deep learning (DL), will be able to assist in the workup of diagnostically difficult entities, as well as research into their causes. Thus, AI guidance may play a critical role not only for clinical tumour detection (e.g., neuroimaging, biomarkers) but also for more accurate and timely tissue diagnostics and prognostication. The main challenge when extracting features from histopathological image is the identification of morphological characteristics that are associated

with disease-typical cellular and tissue alterations. The adoption of DL methods has caused a decline in interest in feature extraction methods, as the representation and decision boundaries can be learned in a single optimization process and these methods have therefore become dominant over traditional machine learning approaches. Therefore, an automated system for the immunochemical microscopic analysis of histological WSIs at 40× primary magnification would be very desirable.

Therefore, in this Chapter, we present study extends the applicability of the PathoFusion framework [332] to the cellular and potentially subcellular level. We illustrate our approach using the detection of CD276 immunoreactive cells in glioblastoma as an example. Following automatic identification by means of PathoFusion's bifocal convolutional neural network (BCNN) model, individual cells are automatically profiled and counted. Only discriminable cells selected through data filtering and thresholding were segmented for cell-level analysis. Subsequently, we converted the detection signals into the corresponding heatmaps visualizing the distribution of the detected cells in entire WSIs of adjacent H&E-stained sections using the Discrete Wavelet Transform (DWT).

The data can be used for whole-slide cross-modality analyses, e.g., relationships between immunochemical signals and anaplastic histological features. We have used glioblastoma biopsies and CD276 as an immunochemical marker for testing the extended PathoFusion system. CD276 (B7-H3) is an immune checkpoint molecule of special interest in cancer research [333-335]. On the basis of detailed microscopic analyses of CD276-labelled glioblastoma biopsies, we have identified a subpopulation of cells which, because of their strong perinuclear immunoreactivity, we have tentatively termed "halo cells" (Figure 3.1 and Figure 3.2). Interestingly, the number of "halo cells" varied significantly across biopsies, stimulating questions concerning their role in glioblastoma. According to recently published literature, it seems possible that "halo cells" represent a type of CD276-expressing cancer stem cells [336], i.e., glioblastoma stem cells (GSCs). Our technical contributions include:

- We present an enhanced framework utilizing PathoFusion's BCNN model to detect, segment, and quantify individual cells labelled via immunohistochemistry in WSIs of

diagnostic tissue samples. By combining image tiles of two spatial resolutions (32x32 and 64x64 pixels), the BCNN model captures both fine-grained cellular characteristics and broader tissue context. The framework applies data filtering, thresholding, and edge detection techniques to identify distinguishable cells for accurate analysis, profiling immunoreactive cells such as CD276 in glioblastoma samples.

- Detected cell signals are converted into heatmaps, providing a visual representation of the spatial distribution of immunoreactive cells across entire WSIs. These heatmaps offer a comprehensive overview of cellular composition within a sample, enabling deeper insights into tumor biology.
- To correlate immunohistochemical findings with histological features, we integrate data from adjacent H&E-stained sections. By applying the Discrete Wavelet Transform (DWT), we combine heatmaps from CD276-labelled cells with those from H&E-stained tissues, enabling an enriched analysis of complex tissue interactions and pathological changes characteristic of glioblastoma.

3.1 Materials

3.1.1 Clinical Cases

Whole-slide scans of WHO Grade 4 glioma samples, provided by the Australian Genomics and Clinical Outcomes of Glioma (AGOG) tissue bank, were used for this study (University of Sydney Human Ethics Committee Project number 2016/027). Paraffin sections were stained with H&E and scanned at 40× magnification using an Olympus VS−120 scanner. Adjacent sections were processed for CD276 immunochemistry. Immunochemistry for CD276 was carried out on an automated Ventana system. The SP265 antibody (Abcam, Melbourne, Australia) directed against C-terminal CD276 was used. Whole-slide scans of glioblastoma samples from a second independent cohort provided by the Sydney Brain Tumour Bank, part of the Neuropathology Tumour and Tissue Bank, Royal Prince Alfred Hospital (Royal Prince Alfred Hospital Ethics Committee Project number 2019/ETH07282), were used for validation purposes.

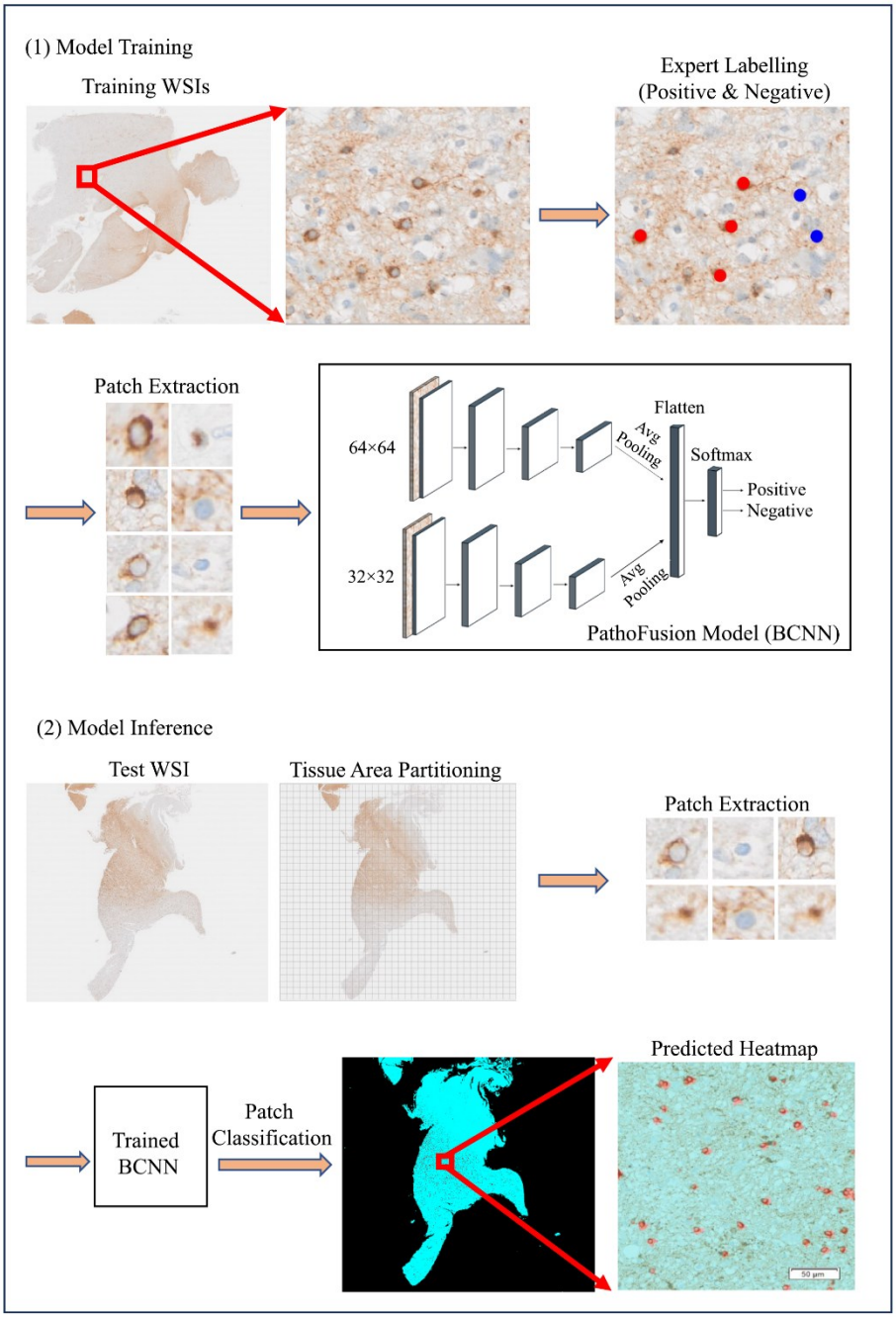

Figure 3.1. Illustration of the extended PathoFusion framework for analysis the cell profiles in whole-slide images (WSI). (**1**) Model training: Patches were extracted from a WSI in line with specialist annotations and passed to BCNN model for training. (**2**) Model inference: Following partitioning of the test image, extracted patches are provided to the trained BCNN model for classification, and the prediction results are converted into the corresponding heatmaps (the pseudo-colour red marks recognized cells). Scale bar: microns.

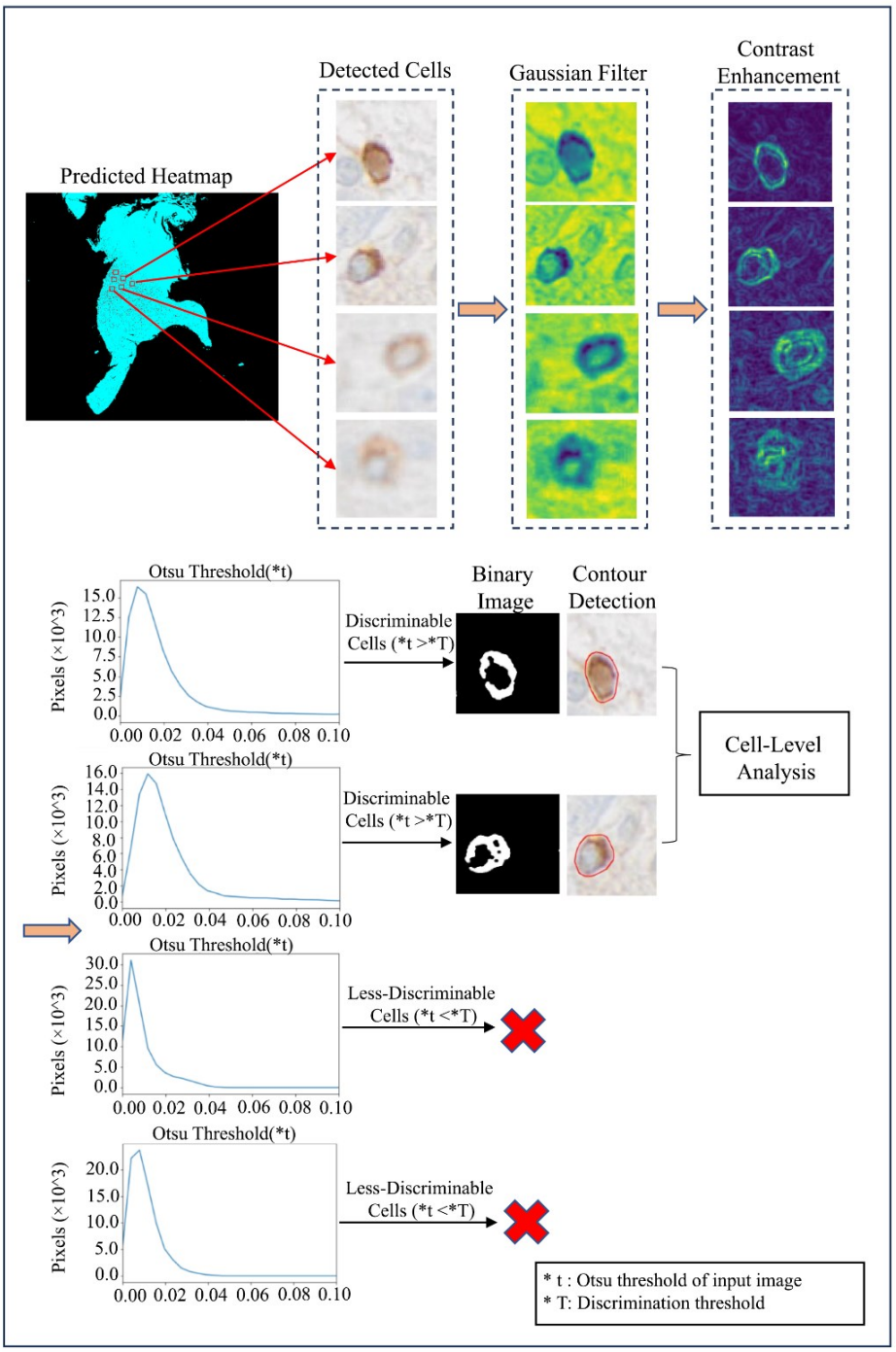

Figure 3.2. Cell-level analysis. Image patches containing cells identified by the BCNN are then processed for cell-level analysis. Filtering and thresholding are followed by edge and contour detection. An image is considered discriminable if the minimum acceptance value for the Otsu threshold is reached. It can then be segmented using edge and contour detection.

3.1.2 Expert Marking and Datasets

Immunoreactive cells were individually marked using the PathoFusion platform, which integrates a labelling website [332]. CD276-positive cells were marked under supervision of a consultant neuropathologist (MBG). Histopathological features (including normal brain tissue and blood vessels, infiltrating tumour, microvascular proliferation, and geographic and palisading necrosis) in adjacent H&E-stained tissue sections were annotated following WHO criteria and correlated by PathoFusion as described earlier [332]. In total, 31,947 cells were manually marked using the website on 8 immunochemically labelled whole-slide scans, which were each split into 4 image tiles to handle the very large file sizes. These cases were selected from the same cohort used previously [332].

The second independent cohort (also described above) was used for validation. In addition, 18,767 regions covering various morphological features in negative areas (no "halo cells" present) were marked resulting in a total of 50,714 marking coordinates. For each one of the 50,714 individually marked coordinates, a pair of image patches was extracted (sizes of 32 × 32 and 64 × 64 pixels, respectively) from the 32 slide image tiles showing immunochemical labelling that had been used for testing (4 image tiles representing 1 WSI per case). Training and testing data were extracted from different cases and there was no overlap between them, i.e., 36,734 paired image patches from the total number of cases were used for training and cross-validation, and 14,005 paired image patches were used for testing.

3.2 A Multi-Scale Classification and Fusion Network

Our methodology for cell profiling in WSIs of immunochemically stained tissue sections consists of two main steps: detection and segmentation of individual cells. The extended PathoFusion framework is illustrated in Figure 3.1: (i) patches of two different sizes with their coordinates based on ground truth labelling are extracted from immunochemically labelled WSIs to allow the BCNN to learn both the cellular information and the surrounding tissue structures; (ii) the BCNN model [17] is trained as a feature extractor and predictor; (iii) at the

prediction stage, the detection results from the whole-slide tissue sections are converted into corresponding heatmaps; (iv) this is followed by filtering and thresholding to segment only the discriminable cells by means of edge and contour detection for further cell-level analysis.

Our aim was to apply PathoFusion at the cellular level for the purpose of profiling individual immunolabelled cells in WSIs. We used CD276 immunoreactive "halo cells" as an example to lay the groundwork for the identification of pathological changes at the cellular level. The performance of the detection process is highly dependent on the features extracted from the images. In our case, analogous to the diagnostic algorithm a microscopist in (neuro)pathology may be using, a specific tissue area of interest (i.e., containing a cell) and its surroundings (the tissue context) were both considered when patches of two different sizes (image tiles of two different sizes) were processed by the CNN. This bifocal design significantly improved the performance of the detection process, as demonstrated earlier [332]. The patches measuring 32 × 32 pixels placed a narrow focus on cell-related characteristics and were extracted as contiguous overlapping patches from the WSIs with a stride of 10 pixels. In addition, we extracted overlapping 64 × 64 patches with a stride of 10 pixels that contained each of the smaller patches and their surroundings, placing a wider focus on the tissue area of interest. All extracted patches were given a label indicating either the presence or absence of "halo cell(s)".

3.2.1 Bifocal Convolutional Neural Network (BCNN)

Our BCNN model is more efficient as a feature extractor and detector than other deep learning models [17]. The BCNN has two input paths and consists of two convolutional sub-nets, one feature concatenation module and a classification layer. Usually, the training of a deep neural network requires many training images to avoid over-fitting. However, our BCNN model requires far less training [17]. To improve model generalization, image augmentation including rotation, contrast and sharpness adjustments was randomly applied to the training dataset. At the evaluation and testing stage, paired image patches were extracted from the upper left to bottom right of the whole-slide test images with a stride of 10 pixels. Each patch pair was classified by the trained BCNN model and assigned to one of two categories, positive or negative for CD276 expression by "halo cells". The prediction results were converted into the

corresponding WSI heatmaps using pseudo-colours illustrating the distribution of the detected cells across entire tissue sections.

3.2.2 Filtering and Thresholding

To distinguish clearly recognizable ("discriminable") CD276 immunoreactive cells from weak ("less discriminable") positivity, a binarization algorithm (Otsu thresholding) was used to post-process the prediction results of the BCNN model. This was preceded by filtering.

Only patches predicted to contain "halo cells" were allowed to pass through the filtering and thresholding process. First, these patches were converted into grayscale images. Next, a Gaussian filter was applied to remove noise caused by various internal and external factors. The Gaussian filter was chosen because of its effectiveness in smoothing images and reducing high-frequency noise while preserving edge details, which is crucial in histopathological image analysis. It works by replacing each pixel value with a weighted average of its neighboring pixels, based on a Gaussian distribution, thereby reinforcing the smoothness of the respective images. After de-noising the images, contrast enhancement was performed by enhancing the variation in the pixel intensity in each neighborhood. Accordingly, each pixel value of the resulting image indicates the contrast intensity in the nearby pixels.

For automatic thresholding, we used Otsu's method [337] to determine the suitability of the detected cells for further analysis. Our thresholding method determines the minimum intensity value of the "halo cell" pixels using Otsu's algorithm. Specially, we utilized images cropped around the cells that had been detected by the BCNN model [17]. Otsu's algorithm [337] assumes that the input image contains two classes of pixels that follow a bimodal distribution (foreground and background pixels, respectively) and sets an optimum threshold such that the intra-class variance is minimized and the inter-class variance is maximized. In our cropped images, "halo cell" pixels have the highest intensity values, whereas there is not much information (immunoreactivity) around the cells. The intensity level returned by the Otsu algorithm refers to the minimum intensity of a cell's pixels. Therefore, if the Otsu algorithm returns a high value, this indicates that the intensity of a cell's pixels in the input image is high (a discriminable cell).

In contrast, if the algorithm returns a small value, this mean that the cell's pixels do not differ much from the background (a less discriminable cell).

After determining the Otsu thresholds (intensity levels) for all patches, we needed to find a sufficiently general minimum value for the pixel intensity (T) of the "halo cells" (discriminable threshold) that could be used as the criterion for distinguishing cells that have been successfully detected. We also needed to minimize inter-class variances among the detected "halo cell" images which could be used in our proposed Equation (3.1) after confirmation of its validity on our cases. If a patch's Otsu threshold surpassed the discriminable threshold, T, as calculated by Equation (3.1), then that patch was deemed to be discriminable. The discriminable threshold can vary from WSI to WSI, as the antibody labelling intensity can differ between WSIs. Figure 3.2 illustrates the filtering and thresholding process in detail.

The formula for finding the minimum acceptance value for the Otsu threshold (discriminable threshold) of "halo cell" images is given by Equation (3.1):

$$T = \frac{\sum \sigma^2}{N} - t_{min} \tag{3.1}$$

where T is the minimum acceptance value for the Otsu threshold of an input image, $\sum \sigma^2$ is the sum of the Otsu threshold values of all input images, N is the number of input images and t_{min} is the lowest threshold found among all input patches.

3.2.3 Edge and Contour Detection

As a part of our post-processing workflow, edge and contour detection were used for further segmentation and more precise measurements of the discriminable "halo cells" in the relevant patches. Edge and contour detection facilitates the location of the cells of interest in an input image, determining their cell area, perimeter, compactness and minimum (min) pixel intensity. Edge detection refers to the process of identifying and locating points in the input image where the intensity changes greatly; it is very high on the edges. A contour is defined as the closed curve joining all points that have the same intensity, corresponding to the shapes of the objects

in the image. Thresholding is used to binarize the discriminable cell images highlighting the cells of interest in white on a black background and allowing the contour detection algorithm to work. Thresholding turns the border of the cells in the image completely white, with all pixels having the same intensity range as shown in Figure 3.2 (binary images). In our "halo cell" images, the color changes typically happen at the boundaries of the cells. This allows the identification of edges simply by observing the change in color. The corresponding intensity class is not constant, but the rate of change in intensity is highest at the edges. The contour detection algorithm can sketch the borders of the cells from the white pixels and find the contour points around the perimeter of the cells. Case-level features were obtained by calculating the mean for cell area, perimeter, compactness and minimum pixel intensity.

3.2.4 Discrete Wavelet Transform for Image Fusion

Image fusion permits precise co-localization of classical diagnostic features defined in H&E-stained sections with immunochemical results, thus increasing the amount of information that can be obtained from biopsy specimens and aiding the diagnostician in their clinical assessment. In other words, the fusion of medical images can generate higher quality content by combining complementary information from multiple sources into a single image.

In order to fuse images of H&E-stained sections with sections immunolabelled for CD276, the Discrete Wavelet Transform (DWT) [338] was implemented by using the average intensities of corresponding pixels from both the input H&E and CD276 ("halo cell") heatmaps. Specifically, single-level discrete wavelet decomposition was applied to the RGB channels of the H&E and CD276 heatmaps using the Haar wavelet kernel, resulting in a four-component image consisting of approximate, horizontal, vertical and diagonal wavelet coefficients for each channel (Figure 3.3). This kernel was chosen for its computational efficiency and ability to balance detail preservation with performance in large-scale data processing. The artifacts introduced during wavelet decomposition and fusion may emphasize subtle differences in intensity and spatial structure that are relevant for classification tasks. These enhanced features provide the classifier with additional discriminative information, potentially improving overall model performance.

The wavelet coefficients were then fused using averaging, where the average value of each corresponding pixel from both input heatmaps was assigned to the corresponding pixel of the output image (Equation (3.2)). Once all bands of each RGB channel had been transformed, an inverse wavelet transform was applied to the fused components to create the fused multi-spectral image for each channel. Finally, concatenation of the RGB channel images was performed.

$$Fused_{Img}(x,y) = (Img1(x,y) + Img2(x,y))/2 \tag{3.2}$$

where $Img1(x,y)$ and $Img2(x,y)$ are the input images (the H&E and CD276/"halo cell" heatmaps, respectively), and $Fused_{Img}(x,y)$ is the resulting fused image.

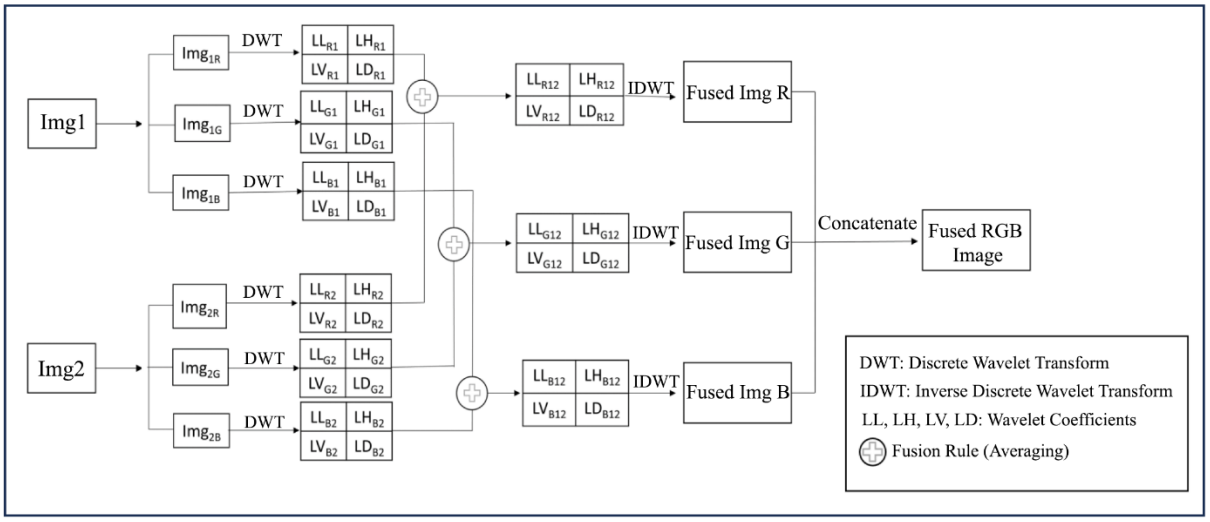

Figure 3.3. Image fusion schematic illustrating our use of the Discrete Wavelet Transform, with pixel averaging as the fusion rule on the RGB channels of input images.

3.3 Results

"Halo cells" were detected using our BCNN model. The performance of the model was evaluated and compared with that of other deep learning models. Heatmaps showing the distribution of the detected cells in entire WSIs were then created. Measurements concerning individual "halo cells" were statistically analyzed. Fusion maps of "halo cells" and H&E heatmaps revealed the exact tissue distribution of "halo cells". Whole-slide cross-modality

analyses of “halo cells” were performed to obtain information on their spatial relationship to anaplastic histological features. The individual analysis steps are detailed in the following sections.

3.3.1 Performance Evaluation

Our BCNN model was used for automated detection of “halo cells” in WSIs [17]. Performance was evaluated using the following criteria: accuracy, precision, specificity and F1-score. In addition, a confusion matrix was created.

The BCNN model, the subnet structure of the BCNN model and ResNet50 [55] were used with identical parameters (learning rate, weight decay, momentum, patch size and number of epochs). The stochastic gradient descent (SGD) algorithm was used to train the models. The learning rate was set to 0.0005, gradually decreasing while training; the batch size was 32 with 70 epochs; the momentum was 0.9 and the weight decay 0.005. The patching strategy described in the Methods section yielded 50,714 pairs of image patches in the training set following extraction from the 32 image tiles, representing eight whole-slide scans with individual patches measuring 64 × 64 and 32 × 32 pixels. To introduce generalization, we used data augmentation techniques while training.

For our experiments (inference on test cases; Figure 3.1, lower panel), between 10 and 35 million patches per case were fed into the BCNN. Specifically, patches of the different sizes (32 × 32 and 64 × 64 pixels) were passed through the two input channels of the BCNN model, providing two categories of patch-level output, including the information as to whether “halo cells” were present. For clarity, the two categories were then assigned two different pseudo-colours. Outcomes that indicated the presence of “halo cells” then passed the filtering and thresholding scheme to select only discriminable cells for further analysis (Figure 3.2). Our BCNN model identified 12,211 “halo cells” in our test WSI, while the filtering and thresholding process discarded 4263 cell profiles following calculation of the minimum acceptance (discrimination) threshold, 0.020; the image threshold values ranged from 0.010 to 0.059. Thus, a total of 7948 “halo cells” were kept in the case of our test WSI.

PathoFusion achieved a high prediction performance for “halo cells” (AUC of 0.992), as shown in Figure 3.4B and as apparent from the confusion matrix in Figure 3.4A. For comparison, two other deep learning methods were used: the subnet structure of our BCNN model and the popular ResNet50 [55]. Figure 3.4C shows the comparison of the ROC curves and AUC scores of the three models. Taken together, our experimental results demonstrate that the PathoFusion model (BCNN) achieved superior results to ResNet50 and the subnet structure in terms of accuracy, precision, specificity and F1-score (97.7%), as shown in Table 3.1.

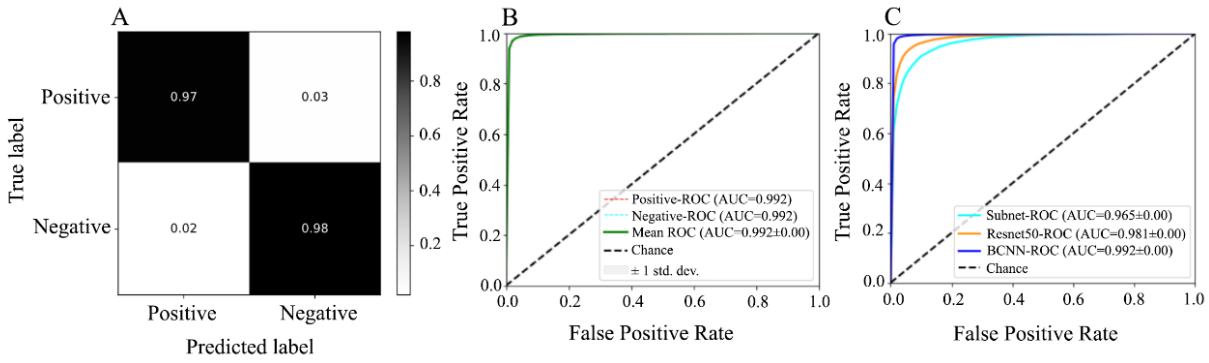

Figure 3.4. (**A**) Confusion matrix for “halo cell” detection. (**B**) Area under the curve (AUC) and receiver operating characteristic (ROC) analysis of “halo cell” detection by our BCNN model. (**C**) Comparison of ROC/AUC performance of the BCNN with that of two other deep learning models.

Table 3.1. Performance evaluation of different deep learning models.

Model	Accuracy	Precision	Recall	F1-Score
BCNN	97.7%	97.7%	97.7%	97.7%
Subnet (BCNN)	90.0%	90.0%	90.0%	89.8%
ResNet-50	94.0%	94.0%	94.0%	94.0%

3.3.2 Visualization

PathoFusion’s proficiency in patch-level recognition [332] formed the basis for the automated analysis of entire histological sections in our study. As demonstrated in Figure 3.5, our methodological improvements allow PathoFusion to be used at the single-cell level. The BCNN

model reliably identified “halo cells” (marked red on the turquoise heatmap; Figure 3.5) after iteratively learning from patches extracted on the basis of ground truth labels. Furthermore, our BCNN model demonstrated the ability to identify “halo cell” profiles efficiently and consistently in WSIs, a task that would be extremely time-consuming and exhausting to achieve for a human observer who also lacks the ability to refer to an accurately defined threshold of detection.

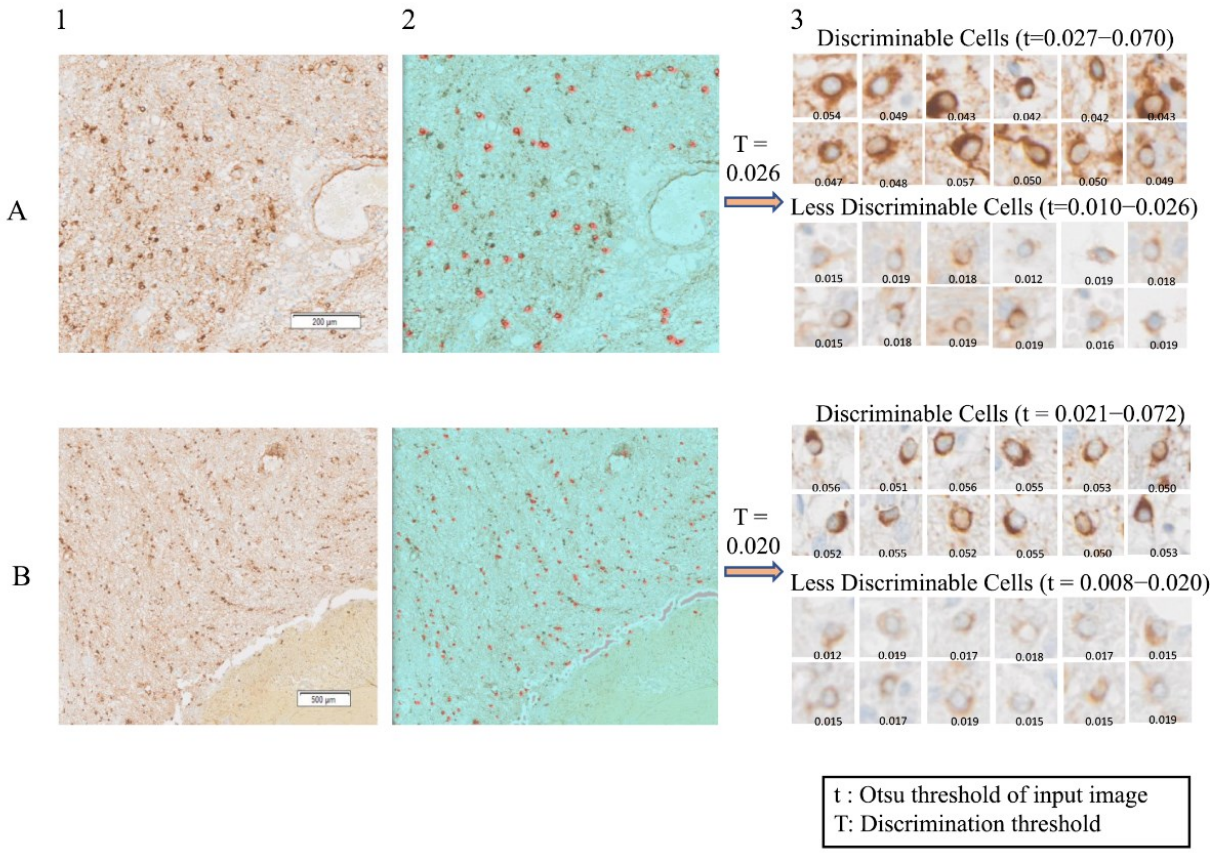

Figure 3.5. (**A**,**B**) The prediction results generated by the BCNN model for two independent cases. (**1**) Low-magnification view of an area showing “halo cells”; (**2**) the corresponding predicted heatmap revealing the detected “halo cells” (in red). (**3**) Visualizations of the results of the filtering and thresholding process: (**3A**,**3B**) upper panel: discriminable cells that returned Otsu values exceeding the discrimination threshold, (**3A**,**3B**) lower panel: Less discriminable cells whose Otsu value fell below the discrimination threshold. Scale bar: microns.

3.3.3 Statistical Analysis

"Halo cells" are amongst the most visible and most discrete objects in our immunochemically labelled images, which is why they were selected for testing. Typically, they are relatively dark brown and have a round to elliptical shape, but some may appear slightly irregular due to variations in their staining density, which can affect their measurement. Therefore, we developed an edge and contour detection algorithm and applied it to all discriminable ("halo cell") patches to analyse them using four of the most visually distinctive properties: area, perimeter, compactness, and minimum pixel intensity, which were then averaged to generate case-level features (Table 3.2).

The area of each cell refers to the total number of pixels occupied by each cell, and the unit is pixels. Perimeter refers to the number of pixels on the circumference of the cell; its unit is also pixels. Area and perimeter were calculated by OpenCV after finding the contour of each cell. The minimum pixel intensity of each cell was calculated by Otsu thresholding. Compactness is the ratio between a cell's area and its perimeter, calculated according to Equation (3.3) [339]:

$$Compactness = \frac{(perimeter)^2}{4\pi \times area} \tag{3.3}$$

Table 3.2. Cell-level features of cases used.

	Feature	Value Range
Cell-Level features	Number of detected "halo cells"	100-15,000 cells
	Density of "halo cells"	0.0003-0.045 cells/mm^2
	Cell Area	7,000-11,000 pixels
	Cell Perimeter	450-800 pixels
	"Halo cell" pixel intensity	0.026-0.070
	Compactness	2.5-4.5

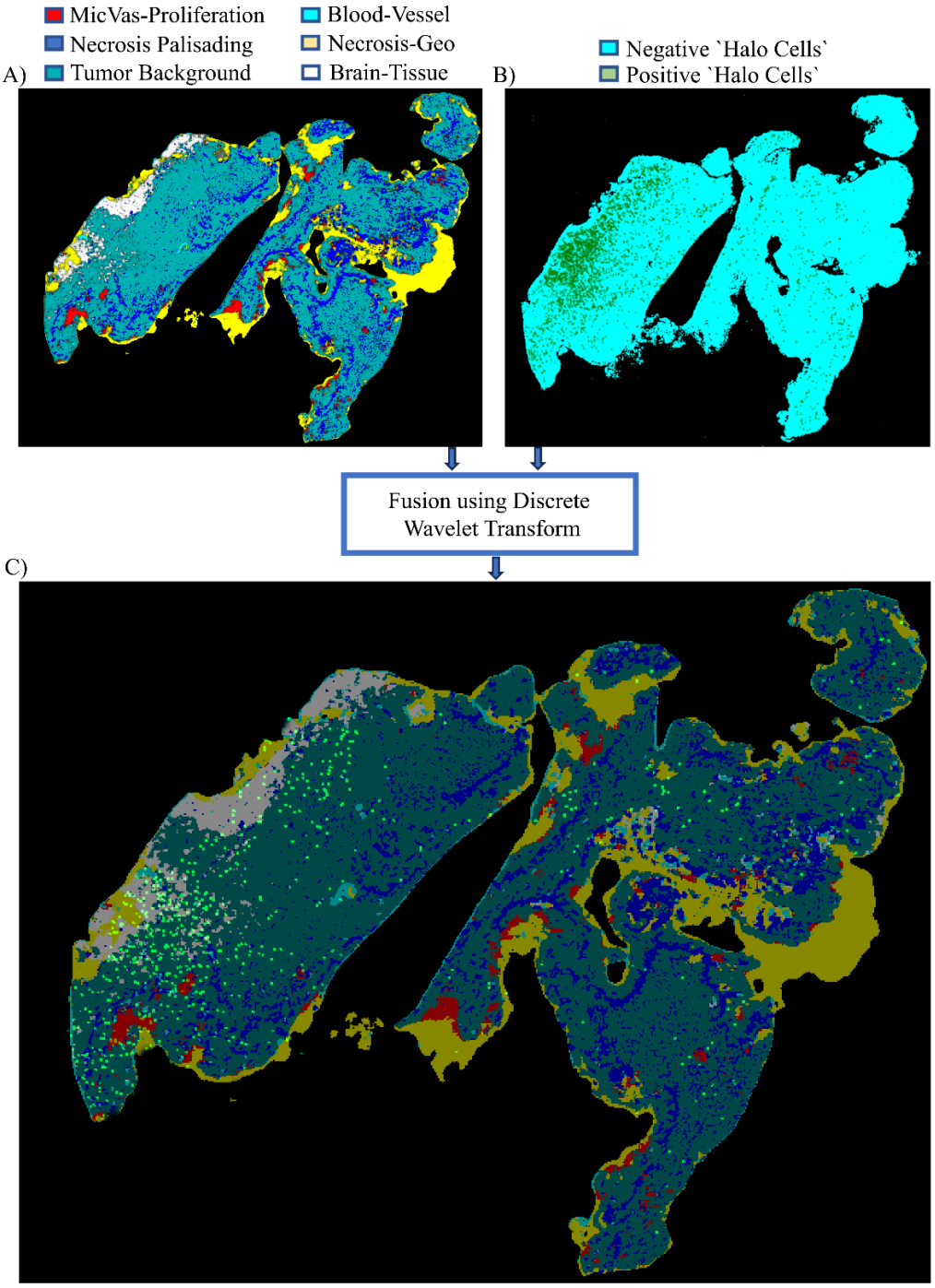

Figure 3.6. Image fusion using discrete wavelet transform. **A)** Original heatmap showing histological features; **B)** heatmap showing cell-level features ("halo cells" and negative areas); and **C)** The discrete wavelet transform is used to fuse heatmaps A and B with greater brightness to emphasize the location of "halo cells" (bright dots).

3.3.4 Fusion of Bimodal Histological Images

The current study extended the versatility of PathoFusion. Pixel averaging of the fused H&E and "halo cell" heatmaps by the DWT [338] allowed detailed comparisons of cases while preserving anatomical references (Figure 3.6). Many different techniques have been proposed for medical image fusion, such as intensity, frequency and Laplacian pyramid-based methods.

Some of these techniques are based on spatial frequency analyses, while others are based on spatial orientation selectivity. DWT allows both frequency and spatial analyses, and is capable of detecting all features contained in a signal [338]. This method effectively preserved both the coarse and fine details of our input images. DWT divided the input image into sub-images, which were then fused (Figure 3.6).

As shown in Figure 3.6B, we applied image registration and alignment with the H&E diagnostic feature map (Figure 3.6A). Different points in the same locations on both heatmaps were then matched by a registration algorithm [340]. Figure 3.6C shows the H&E heatmap (cf. [332]), which was taken as the reference image with which the "halo cell" heatmap was fused.

Table 3.3 Density of "halo cells" in relation to diagnostic morphological features in adjacent H&E-stained sections.

Morphological features	Number of Halo Cells	Density - Number of Halo Cells/mm^2
Normal blood vessels	481	0.051
Normal brain tissue	554	0.036
Geographic necrosis	2247	0.047
Viable tumor tissue	7045	0.034
Palisading necrosis	248	0.031
Microvascular proliferation	1636	0.036

Finally, the correlation of immunolabelled "halo cells" with diagnostic morphological features [332] was visualized in the bimodal WSIs through image fusion of the corresponding predicted heatmaps. There were differences in the number of "halo cells" among cases. The density of

"halo cells" in relation to the diagnostic morphological features is summarized in Table 3.3.

3.4 Discussion

Manual analysis of histological WSIs is very exhausting and time consuming, due to the large diversity and intricate nature of these images, also it requires abundant knowledge and experience of histopathologists as there are substantial intra- and inter-observation variations among experts. With the development of digital imaging in pathology, computer-assisted diagnostic (CAD) algorithms have become popular. CAD systems can now complement diagnoses made by a pathologist, facilitating decision-making and improving the efficiency of the diagnostic process [341]. Early CAD systems were based on conventional learning algorithms, where much of the effort was spent on feature extraction based on expert domain knowledge, followed by the application of traditional classification models, including RFs and SVMs [39,158-162,273].

However, some engineered machine-learning features are of limited use in biomedical applications. Deep learning systems, especially CNNs, are capable of extracting morphological features automatically following suitable training [342]. Since it is impractical to feed a CNN directly with images of enormous size such as WSIs, many workers in the field have adopted a patch-based approach to extract features while preserving essential information for detection and classification tasks [64,181]. In the context of how pathologists make diagnoses, cellular features and global tissue structures taken from two types of patches with different sizes will enhance the effectiveness of detecting abnormalities in WSIs.

In this study, we automatically detected, visualized and segmented cells of interest, i.e., "halo cells", in the given WSIs. We have taken advantage of the capability of PathoFusion [332] to identify pathological changes characteristic of glioblastoma and to correlate the location of CD276 immunoreactive cells. This was achieved by converting the detection signals into corresponding heatmaps visualizing the distribution of the detected cells in entire WSIs. Our new method achieved of 97.7% accuracy, thus eliminating the necessity for extensive pretraining.

The prognostic marker CD276 was chosen as a proof-of-concept example, allowing our model to validate its recognition ability and consequently produce a fusion heatmap that demonstrates the overlap between histological features and CD276 immunoreactivity. CD276 expression by the tumour vasculature is well established [343]. The BCNN could be trained to recognize immunoreactive endothelial cells, and the PathoFusion framework could be used to quantify abnormal blood vessels. In addition, a comprehensive AI-guided correlation between CD276 expression and diagnostic features of malignant glioma, e.g., the different types of necrosis, should be feasible. However, other aspects of CD276 expression are of greater interest in the present context. Wang et al. and Liu et al. [333,336] have previously demonstrated the upregulation of CD276 in cancer stem cells, and CD276 expression has been shown to be associated with GSC self-renewal [344]. Therefore, a detailed analysis of the tissue distribution of "halo cells" and their correlation with other key molecules in glioblastoma would be important and is now feasible.

For instance, Sun et al. [345] demonstrated a positive correlation between CD276 and stemness markers (CD133/PROM1, NGFR, TYH1, SOX2), and Johnston et al. pointed out the correlation with immune modulators (i.e., IFNGR1, IFNGR2, TNFRSF1A and TNFSRF1B) and self-renewal genes, including VAX2, SOX21 and CITED1 [344]. CD276 has also been shown to be involved in the epithelial–mesenchymal transition [333]. Importantly, CD276 expressed by tumour cells may affect the differentiation of tumour-associated macrophages [346], and CD276 is also expressed by at least some macrophages in GBM tissue [347]. Therefore, while the original PathoFusion framework is capable of learning to detect histological features, the improved cellular resolution of the extended framework now allows tumour analyses in fine detail at the cellular and potentially subcellular level, including the mapping of markers for cancer stem cells.

Our present research comes with certain constraints. Firstly, we recognize that the training process for identifying and delineating cells of interest can be enhanced through the establishment of subcategories and a thorough definition of similar looking morphologies (mimics that may lead to diagnostic challenges). Secondly, there are numerous morphological features could be incorporated into the training to achieve fully and accurately segment the

overlapping cells. Thirdly, CD276's role in influencing macrophage differentiation in GBM and the variability of macrophage marker expression warrant further investigation, with image fusion of adjacent tissue sections is needed to clarify this aspect.

3.5 Summary

PathoFusion’s BCNN model makes it now possible to automatically identify and count individual immunochemically labelled cells in routine paraffin sections. Our experimental results indicate that PathoFusion is potentially also suitable for the detection of subcellular structures. It could work as part of a general and clinically applicable (weakly supervised) system for cross-modality analyses in (neuro)pathology.

The known advantages of PathoFusion, such as effective trainability based on a comparatively small number of cases and flexible use of consultant time during training, can thus be applied to the growing number of immunochemical markers which require extensive diagnostic evaluation.

CHAPTER 4

An Attention-Based Graph Neural Network Capable of Prognostication Based on CD276 Labelling of Malignant Glioma Cells

Deep learning approaches are showing promise in achieving high levels of accuracy in image analysis tasks, they allow data reduction and enable feature selection techniques that increase computational efficiency [174-176,275,348]. Deep learning algorithms, particularly CNN, are able to extract morphological features automatically after appropriate training [342], and many researchers in the field have adopted a patch-based approach to extract features while preserving essential information for detection and classification tasks [64,181,195,234,277]. However, while these methods are effective for capturing local patterns and features, it is not well suited for capturing a WSI's global context and structure [349]. Graph-based deep learning methods have been proposed to bridge this gap and to model the geometric structure of the tissue at both local and global levels [15,350].

Cell graph neural networks (CGNNs) are a type of GNN where a graph is constructed using the specific locations of cells or cell nuclei as entities or "nodes" while the spatial distances between them [96] are represented as lines which are called "edges". Attention mechanisms enable the network to focus on specific ROIs in an image resulting in improved accuracy of output. In our previous study, we have explored the capability of BCNN model [17], the core of the PathoFusion framework [332], to identify immunolabeled cells and characterize pathological changes at the cellular level [351]. Hence, GNNs and attention mechanisms can also be used together to further improve the performance in image analysis tasks.

Therefore, in this chapter, we proposed a comprehensive GNN framework (Figure 4.1), termed PathoGraph, which utilizes attention mechanisms that adaptively adjusts weights based on feature similarity and spatial proximity, enhancing its ability to handle diverse feature distributions and incorporate spatial and heterogeneity information in histological features. Specifically, we used a memory-efficient method for encoding gigapixel-sized WSIs enabling detailed analysis of tissue structures and cell connectivity. This approach is based on the assumption that cells in a tissue can organize in a certain way (e.g., clones of cancer cells) which may relate to prognosis and ultimately survival. An optimized graph convolutional neural network (GCN) with attention mechanisms (abGCN) is then used for WSI-level survival prediction and to pinpoint cell populations that deserve the greatest attention. The main contributions of this study are summarized as follows:

- We propose a memory-efficient encoding method for gigapixel-sized WSIs using a set of subgraphs that reduces the data size of a graph while preserving the spatial relationships between image patches, providing a thorough and robust representation of tissue structure for subsequent learning tasks. The constructed graphs are visualized to illustrate the connections between detected cells within each tissue section.
- We introduce a graph convolutional neural network (GCN) with attention mechanisms, referred to as the abGCN model, for WSI-level survival prediction and to identify key cell populations that require the most attention. The model's attention mechanism adaptively adjusts weights based on feature similarity and spatial proximity, enhancing its ability to handle diverse feature distributions and incorporate spatial and heterogeneity information in histological features. By analysing the distribution of attention scores, we provide insights into regions of interest within WSIs, aiding in the understanding of model predictions and uncovering hidden patterns for further pathology research.
- We propose a combined loss function that extends Binary Cross-Entropy Loss with regularization terms to utilize graph structure and attention mechanisms, promoting more balanced attention distribution and better image-level classification performance.

- We conducted experiments to evaluate our proposed approach on histopathology image survival tasks using two datasets, demonstrating superior performance and providing valuable insights into the tissue microenvironment compared to existing methods.

4.1 Datasets

We conducted experiments on two datasets to validate the effectiveness of our model. One imaging dataset was as before in the previous chapter [351]. We have used 32 WSIs of immunohistochemically stained tissue sections with an average size of 100,225 × 150,537 pixels at 40× primary magnification. As described [351], immunohistochemistry for CD276 was carried out on an automated Ventana system. The SP265 antibody (Abcam) directed against C-terminal CD276 was used [351]. Immunoreactive cells were individually marked on digital slide scans using the PathoFusion platform which integrates a labelling website [332] under supervision of a consultant neuropathologist, resulting in the placement of a total of 50,741 marking dots over glioma cells showing strong CD276 expression. They were used for training the BCNN model [17].

Our CD276 dataset [351] consists of high-grade glioma cases from the Australian Genomics and Clinical Outcomes of Glioma (AGOG) repository. Almost all are glioblastomas according to the most recent, revised 2021 WHO classification of CNS tumours but one case needed to be reclassified because it carries an IDH1 mutation (R132H). Importantly, this difference in genotype turned out to be irrelevant for our study because the CD276 expression by GSCs in this case was as strong or stronger as in the other, non-IDH1 mutated tumours suggesting that CD276 expression by malignant glioma cells is a fundamental characteristic of these tumours that dominates the effect of an IDH1 mutation.

A second dataset was obtained from TCGA. This glioblastoma (GBM) dataset contains 255 WSIs [352] and was used to assess our model's effectiveness.

4.2 The Framework

PathoGraph is modelled after a diagnostic workflow, i.e., it follows principles that aid pathologists when they evaluate histopathological slides [332]. Accordingly, high resolution features such as cellular morphology, cellular position, and the expression of certain molecular markers, as well as global features such as anatomical information and landmarks, tissue layers, cellular (community) patterns, and overall tissue organization are taken into account. Figure 4.1 shows an overview of the PathoGraph framework. For any given WSI, we used our pre-trained BCNN model [17] to detect CD276 immunoreactive cells. We then used their positions as vertices as well as their features, which were extracted using the same model. Edges were captured taking into account both spatial relationships and feature similarities between neighbouring cells based on a given Euclidean distance (ED).

The constructed graphs were then visualized illustrating the connection between the detected cells across each tissue section. Each tissue image was analysed by feeding its graph into our abGCN model that had undergone supervised training on our training cohort of cases. The proposed abGCN model enables the learning of a data-driven graph feature hierarchy without the need for hand-crafted features. The attention mechanism in our model allows the network to focus on the most relevant information (important cells) in the graph, making the model more interpretable. The attention scores were finally converted into maps.

4.3 Graph Construction on WSI

A pretrained BCNN model was used to detect the CD276-expressing cell population of cells [17]. Once the cells had been detected, neuropathological features were extracted from 64 × 64 patches, each centered at the detection peak location. These features serve as vertex features derived from the first layers of the BCNN model. Then, an average-pooling layer was utilized to generate a 4096-dimensional feature vector. The spatial distance and cosine similarity of the features extracted by the pre-trained BCNN, which describes the similarity between patches, were taken into account. Since WSIs are gigapixels in size, aiming to represent a raw WSI, using all the information available at the single pixel level in a graph is computationally challenging.

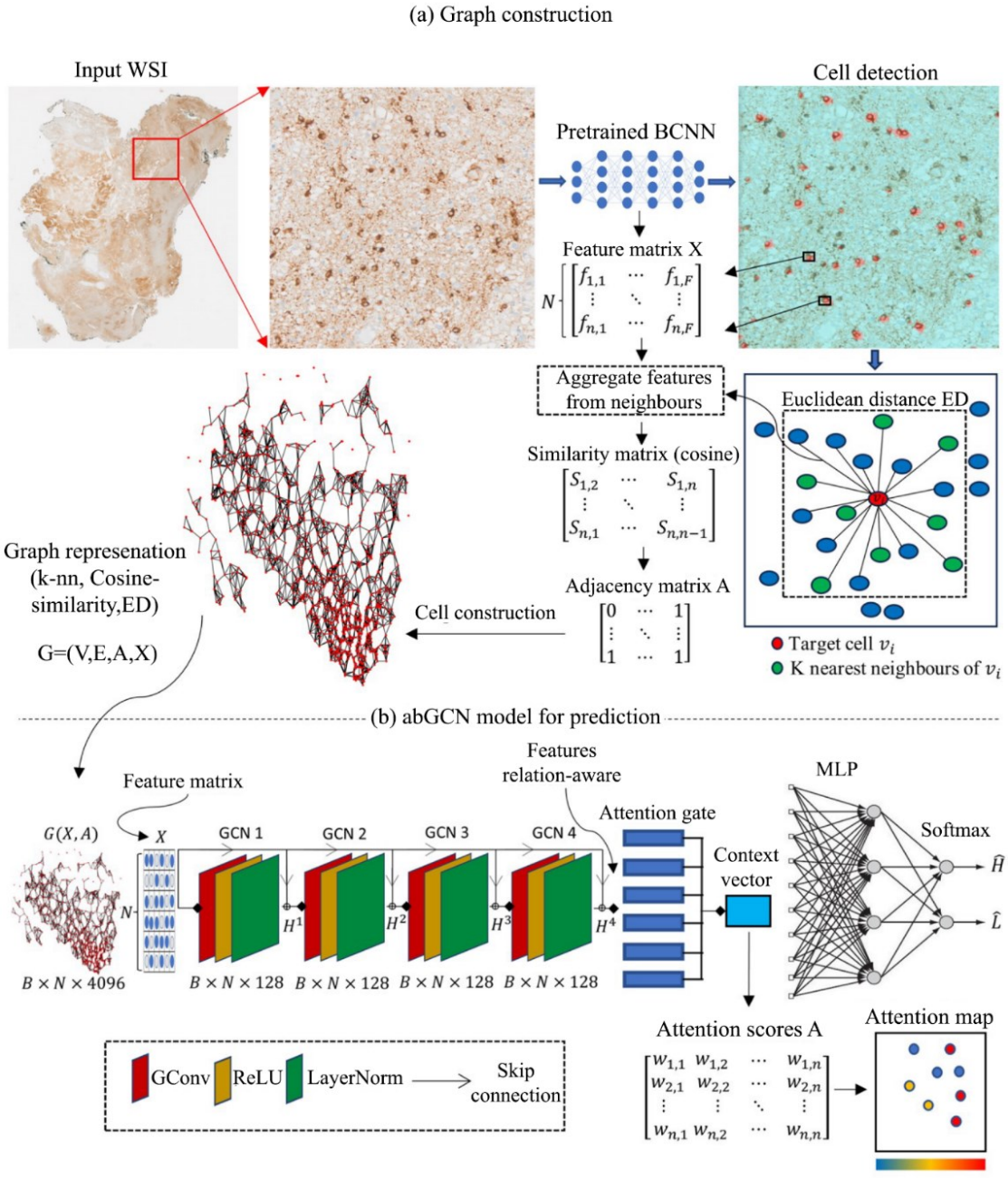

Figure 4.1 PathoGraph framework. (a) Graph construction. BCNN model is used to identify CD276 immunoreactive GSCs, with centroids as graph vertices and features as vertex attributes X. Edge features are derived from spatial distance and feature cosine similarity, mapped via Euclidean distance, ED. The constructed graphs are then visualized. (b) AbGCN model for prediction. The constructed graphs, represented by $B \times N \times 4096$ (B: batch size, N: number of patches, 4096: feature dimensionality), are then fed into our abGCN model to predict survival probabilities, with attention scores visualized as maps (Figure 4.6).

Therefore, we propose a memory-efficient WSI encoding method that uses a set of subgraphs which represent the gigapixel-sized image in a memory-efficient vector space (Algorithm 4.1). We build these sub-graphs by iteratively identifying nearby cells based on their center coordinates and feature characteristics and then draw edges between them based on distance, nearest neighborhood, and cosine similarity for each WSI as described below. This subgraph approach efficiently reduces a graph's data size while preserving spatial relationships between patches.

We employed a maximum ED of 500 pixels as an iteratively determined threshold for the automated construction of edges between the most relevant subset of nodes (CD276 immunoreactive cells [17]). This allows a contextual analysis of the cell's connectivity and its relationship to regions of interest in the WSI defined by pathological features that are recognizable when using other modalities on adjacent sections [332]. The nodes (vertices) that are used have similar features based on cosine similarity. The k −nearest neighbours (k = 8) of each cell that are within the similarity threshold (cosine similarity) of the patch features that have been extracted from the final MLP layer of the pre-trained BCNN model are then determined. Patches are accepted as similar if the value for the cosine similarity is greater than the given threshold. Thus, following the identification of the relevant patches within the WSI, a graph's edges were created based on spatial distance and feature similarity of the patches.

Therefore, for a given WSI, after localizing CD276 immunoreactive cells and extracting their features using our pretrained BCNN model, we construct a graph representation $G(V, E, A, X)$. For this, we treat each patch $n \in N$ as a node to construct a set of vertex patches $V = \{V_n \mid n \epsilon N\}$, such that each node is represented by the patch's geometric center c, and embedding the micro-level features of each cell in a patch. The edge set $E \subseteq V \times V$ represents the connections between nodes expressed by a sparse adjacency matrix A, in which A_{ij} depicts the relationship between node i and node j. WSI graphs were encoded by calculating spatial adjacency and feature similarity between all patches and storing edge level embeddings (features) in the attribute matrix X. The resulting graph was then visualized using pseudo-colors illustrating the connections between the detected cells across the entire tissue section as illustrated in the experiment section.

Algorithm 4.1: Graph construction for WSI

Input:
WSI # whole slide image
Output:
$G = (V, E, A, X)$ # Graph representation of the WSI

Initialize:
$V \Leftarrow \emptyset$ # vertex set
E $\Leftarrow \emptyset$ # edge set
A $\Leftarrow \emptyset$ # Adjacency matrix
X $\Leftarrow \emptyset$ # Feature matrix
K = 8 # the number of nearest neighbors to connect
$ED = 500$ pixels # Euclidean distance threshold
$CT = 0.8$ #cosine similarity threshold
N = patches number detected by BCNN in WSI
for n in N **do**
c_i = geometric center of the patch
f_i = BCNN(n) #extract features for the patch
Add $v_i = (c_i, f_i)$ to V
Add f_i to X
end for

for v_i in V **do**
Neighbors$_i \Leftarrow \emptyset$
Nearest Neighbors$_i \Leftarrow \emptyset$
EDS = {Euclidean Distance $(v_i, v_j) | v_j \epsilon V$ } # Compute Euclidean distances to other vertices
for ed in EDS **do**
if ed $\leq$ ED **do**
CS = Cosine Similarity (v_i, v_j)
Add $\{v_j$, CS$\}$ to Neighbors$_i$
end if
end for

SortedNeighbors$_i$ = Sort (Neighbors$_i$) # Descending by cosine similarity values
Count = 0
for $\{v_j$, CS$\}$ in SortedNeighbors$_i$ **do** # Select top KKK neighbors meeting the threshold
if Count<K and CS $\geq$ CT **do**
Add v_j to Nearest Neighbors$_i$
end if
Count = Count + 1
end for

for nn in Nearest Neighbors$_i$ **do**
Add edge $(v_i$, nn) to E
A_{ij}=1 # edge exists between v_i and v_j
end for
end for

4.4 Network Architecture

PathoGraph was developed to relate cell community characteristics to system outcome. In our specific application, prognosis can be inferred from the distribution of strongly CD276 immunoreactive glioma cells that may represent stem cells which are represented by graphs that are then related to existing survival data of the cases by the abGCN model. From a general application point of view, we suggest that both the input parameters and the system outcome that is being determined can be varied quite freely, e.g., to identify and validate new biological markers that are of a higher system level relevance. Our visualization technique emphasizes tissue areas characterized by high attention scores that appear to match the tissue localization of progressive tumors in neuropathological evaluations but overall, the discriminative power of the abGCN seems much higher when relating the results to survival data.

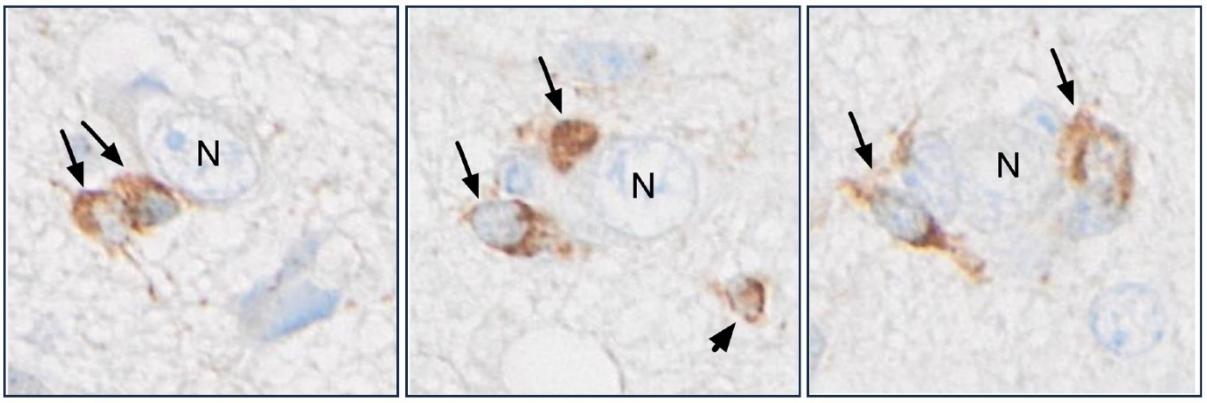

Figure 4.2. CD276 expression by malignant glioma cells/glioma stem cells (GSCs) in a typical perineuronal satellite position (arrows; N, neuron). CD276+ GSCs are also found perivascularly and accumulating in a subpial location (so-called secondary structuring according to Scherer, not shown). The arrowhead points to a free-lying glioma cell.

Thus, within PathoGraph, the task of predicting disease (system) outcome can be considered a graph classification problem at the WSI-level with a WSI representing the disease substrate contained in a biopsy. We anticipate that the approach can be generalized to employ other laboratory data as well as system outcome parameters. The system outcome we are especially interested in for our example is patient survival as the cell population which expresses CD276

strongly in the cytoplasm includes pathognomonic perineuronal satellite glioma cells (Figure 4.2).

We have used the abGCN as our base model (Figure 4.1), which consists of GCN and attention mechanism. The details are presented in the following subsections.

4.4.1 Graph Convolutional Network

A spatial GCN considers the neighbouring nodes up to a certain distance (n-hop neighbours), unlike spectral GCN graph approaches, which take homogeneous graph datasets as inputs with the adjacency matrix fixed across the dataset. Since tissue images can have greatly variable characteristics, PathoGraph must handle different types of nodes and connectivity patterns.
GCNs rely on message passing, which means that vertices exchange information with their neighbours and send messages to each other. GCNs initiate with a node creating a feature vector that represents the message (feature vector values) it passes on to all its neighbours, so each node receives one message per adjacent node. Messages vary across nodes and embeddings of nodes are updated by their neighbours. The total number of messages each node receives is obtained by the GCN layer which is defined as follows [353]:

$$H^{(l+1)} = \sigma\left(D^{-\frac{1}{2}}\acute{A}D^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \quad (4.1)$$

H^l is the node feature matrix at the l^{th} layer, where each row represents the feature vector of a node, $W^{(l)}$ is the weight matrix for the l^{th} layer where the input features are transformed into messages (H^lW^l). The identity matrix I was added to the adjacency matrix A so that each node sends its own message also to itself: $A\acute{} = A + I$. Finally, in order to combine the messages and normalize, we calculated the matrix D which is a diagonal matrix with D_{ii} denoting the number of neighbors node i has to obtain average values. σ represents a non-linear activation function, in particular a ReLU-based activation function.

4.4.2 Graph Attention

In a standard GCN layer, filtering is performed by aggregating the features of neighbouring nodes using a shared weight matrix. In contrast, a graph attention layer extends the traditional GCN by incorporating an attention mechanism. This means the filtering operation is enhanced by considering the importance of neighbouring nodes through assigning specific weights to individual inputs [354]. Analysis of the learned weights after applying the attention mechanism to them can provide insights into how the model is making decisions, enhancing interpretability. Computation in a graph attention layer typically involves computing attention coefficients and weighting features, and aggregating them, which is performed for each node in the graph. The input to this layer is a set of node embeddings of the generated graph obtained from the GCN.

These embeddings are represented as $H = \{h_1, h_2, \ldots., h_N\}$. And the output is a set of a new set of node embeddings $\dddot{H} = \{\dddot{h}_1, \dddot{h}_2, \ldots, \dddot{h}_N\}$, where N is the number of nodes in the graph. First, each input feature vector h_i of node i is transformed into a higher-level representation using a weight matrix W as follows, Equation (4.2):

$$\grave{h}_i = W\, h_i \tag{4.2}$$

Then we calculate the attention scores w_{ij} between a node i and its neighbors j using a shared attention mechanism q as follows:

$$w_{ij} = LeakyReLU\left(q^T\left[\grave{h}_i \parallel \grave{h}_j\right] + d_{ij}\right) \tag{4.3}$$

where $\left[\grave{h}_i \parallel \grave{h}_j\right]$ denotes the concatenation of feature vectors of nodes i and j, and d_{ij} represents the distance between nodes i and j, considering the spatial positional relationship.

The features of neighbouring nodes are aggregated by computing a weighted sum based on the normalized attention coefficients a_{ij}, resulting in the new updated feature vector $\dddot{h}_i$ of node i, Equations (4.4,4.5):

$$a_{ij} = \frac{\exp(w_{ij})}{\sum_{k \in N_i} \exp(w_{ij})} \tag{4.4}$$

$$\hat{\hat{h}}_i = \sum_{j \in N_{(i)}} a_{ij} \; \hat{h}_j \tag{4.5}$$

where N_i is the set of neighbours for node i.

Finally, we applied a non-linear activation function such as ReLU to this aggregated feature vector to introduce non-linearity and enhance the model's capacity to capture complex patterns:

$$\hat{\hat{\hat{h}}}_i = ReLU(\hat{\hat{h}}_i) \tag{4.6}$$

And we get the F-dimension feature representation of N nodes of the input graph, as:

$$\hat{\hat{\hat{H}}} \in \mathbb{R}^{N \times F} = \{\hat{\hat{\hat{h}}}_1, \hat{\hat{\hat{h}}}_2, \ldots, \hat{\hat{\hat{h}}}_N\} \tag{4.7}$$

where F is the original feature dimension. This process of computing the self-attention score ensures that the updated feature representation takes into account the significance of adjacent nodes as well as the spatial connections, between them.

Taken together, during processing by the PathoGraph framework, each WSI is characterized by feeding its corresponding graph into the abGCN model which has undergone supervised training on a cohort of cases.

4.4.3 Optimization Function

In order to enhance the model's learning from both node features and graph structure and to improve binary classification accuracy, we propose a combined loss function to optimize our graph-based model with an attention layer. This combined loss function integrates multiple components: binary cross-entropy loss $\mathcal{L}_{BCE}$, the graph Laplacian regularization $\mathcal{L}_{graph}$, and the attention regularization $\mathcal{L}_{attn}$. By combining these components, the loss function ensures precise classification and enhances the feature extraction and attention mechanisms making them more impactful and efficient. This ultimately results in better performance when it comes to recognizing the crucial features of an input image.

First, to measure the difference between the predicted graph-level class (output of a sigmoid function) $\hat{y}_i$ and the ground truth class y_i for the entire graph. The classification outcome loss is computed as follows:

$$\mathcal{L}_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)) \tag{4.8}$$

where N is the number of WSIs.

By adding graph Laplacian regularization $\mathcal{L}_{graph}$, the model will preserve the structural properties of the graph by encourages neighboring nodes in the graph (connected by an edge) to have similar feature representations. It can be calculated as:

$$\mathcal{L}_{graph} = \lambda \sum_{(i,j)\in E} \| \hat{y}_i - \hat{y}_j \|^2 \tag{4.9}$$

where λ is a regularization hyperparameter, E represents the edges of the graph, and $\hat{y}_i$ and $\hat{y}_j$ are the predicted outputs for nodes i and j.

Moreover, we can add a regularization to the attention mechanism $\mathcal{L}_{attn}$ to ensure that the attention scores exhibit certain preferred characteristics, such as sparsity by concentrating on a limited number of relevant nodes or areas, or smoothness by promoting similar attention scores for comparable nodes or regions we employ regularization techniques. Attention regularization can be computed as:

$$\mathcal{L}_{attn} = -\sum_i \sum_{j\in N(i)} a_{ij} \log(a_{ij}) \tag{4.10}$$

where a_{ij} are the normalized attention coefficients.

Finally, we combine these losses into a single objective function $\mathcal{L}_{total}$ to be optimized during training:

$$\mathcal{L}_{total} = \mathcal{L}_{BCE} + \alpha\mathcal{L}_{graph} + \beta\mathcal{L}_{attn} \tag{4.11}$$

where α and β are hyperparameters of regularization terms to optimize the proposed model via backpropagation.

4.5 Results

To assess the PathoGraph framework's efficacy in predicting system outcome, survival in this case, we have visualized graphs generated from immunohistochemically labelled histopathological image scans, compared the prediction performance of our abGCN model with state-of-the-art architectures, and visualized GSCs bearing high attention scores in tissue. We have used two datasets for our study, our own CD276 dataset [351] and glioblastoma (GBM) dataset from TCGA [352].

It is also worth noting that for the first time according to our knowledge, CD276 expression was shown to occur in the typical perineuronal satellite glioma cells (a classical secondary structure according to Scherer) confirming the identity of the CD276 immunoreactive cells as malignant glioma cells/GSCs and emphasizing the importance of the immune checkpoint-marker, CD276, for the biology of diffuse malignant glioma (Figure 4.2).

4.5.1 Implementation Setting

The proposed abGCN model is implemented using PyTorch and PyTorch Geometric. This network consists of a sequence of GCN layers and non-linearities such as ReLU which enables optimised message exchange between nodes and captures more complex patterns and relationships in the graph. Furthermore, to prevent overfitting, we employ LayerNorm followed by ReLU activation and also skip connections (residual connections) between each layer to mitigate the vanishing gradient problem and improve learning capacities and stability. We used one attention gate (32 dimensions) with three layers, which has a receptive field large enough to capture environmental features with the GCN. We use the MLP layer and the last fully connected layer to make the final classification decision as the WSI output.

4.5.2 Correlation between CD276 Immunoreactive Cells and Survival

One aspect of this work focuses on the evaluation of the relationship between the presence of

CD276 immunoreactive cells and patient survival in glioblastoma as an example of system outcome predicted from WSI graph data. In our previous work, we have demonstrated the ability of our BCNN model to detect and identify cells of interest efficiently and consistently in WSIs based on predefined profiles. We now take advantage of the detection capability of our BCNN model to additionally count the relevant cells.

Interestingly, the number of CD276 GSCs in the material available to us varied significantly in line with the histological heterogeneity of glioblastoma and the inherent variability of the neurosurgical bioptic process. This facilitated illustration of the fact that attention scores do not simply reflect cell numbers but are based on graph features. Since the number of CD276 positive cells varied greatly (from less than 100 to 15,000 per WSI), our analysis also shows that the abGCN method we are proposing is capable of dealing with a wide dynamic range.

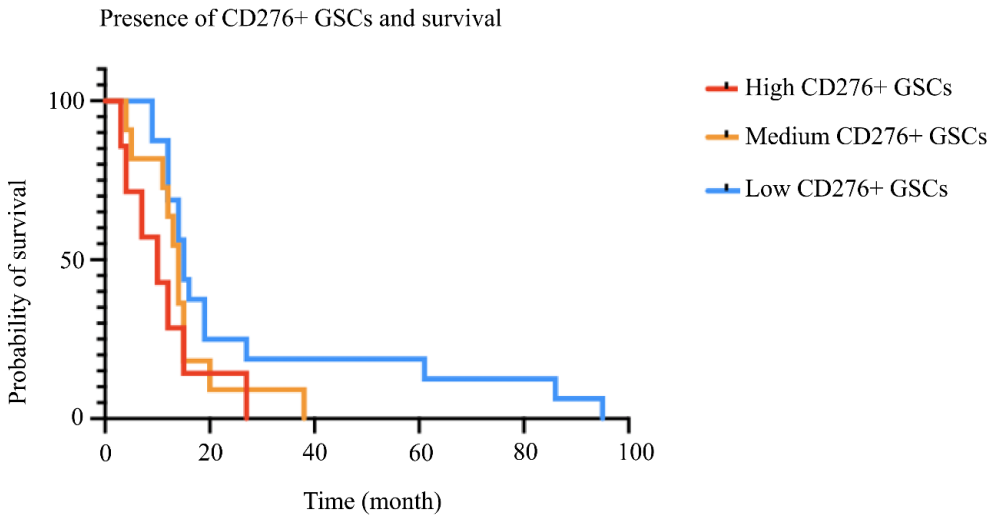

Figure 4.3 Correlation between the number of CD276 immunoreactive GSCs and patient survival.

We have performed a Kaplan–Meier analysis on our dataset to evaluate the correlation between the number of CD276 immunoreactive GSCs and patient survival (Figure 4.3) and found that the number of CD276 positive GSCs is associated with patient survival, with p=0.05 (Figure 4.4).

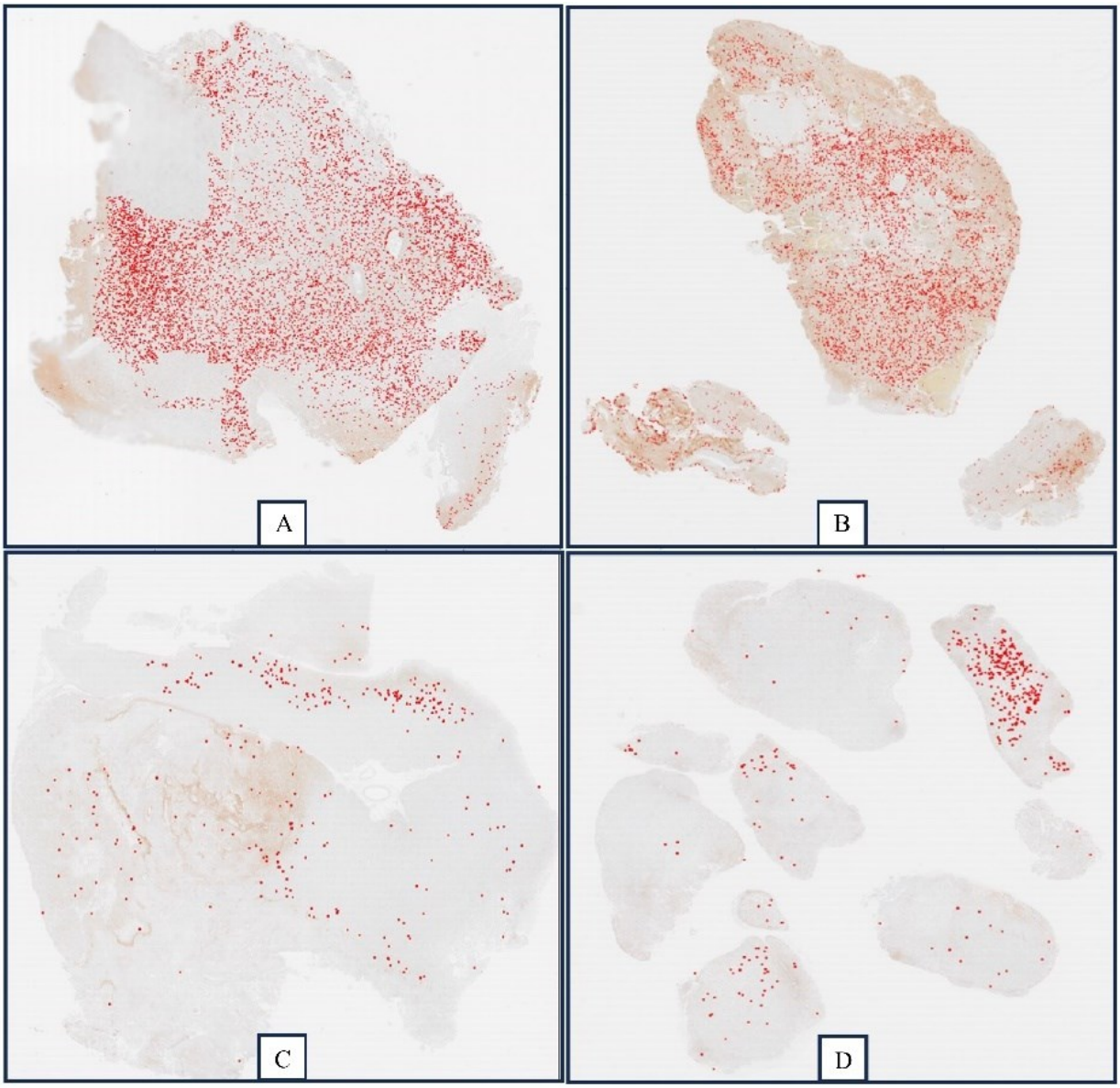

Figure 4.4 (**A**,**B**) Two cases with short survival, and (**C**,**D**) two cases with long survival. The red colour shows the detected CD276 GSCs.

4.5.3 Visualization of the Graphs Constructed by PathoGraph

Each WSI yields a different number of nodes which when connected through edges result in one or more graphs of varying complexity. The number of nodes in our graphs that are based on the CD276 dataset of each WSI ranged from less than 100 to 15,000 per WSI. Sometimes biopsies

can consist of multiple small tissue fragments. It should be noted therefore, that our proposed method requires the size of the fragments that are selected for analysis to be larger than the maximum ED employed. The nodes in the graph network represent the digital patches that contain CD276 immunoreactive cells, and the edges reflect the connection between the patches. Figure 4.5 depicts examples of graphs (right column) and the corresponding immunohistochemically labelled (CD276) WSI (left column) they were derived from. We found that survival correlated not only with the number of CD276 positive GSCs but also with greater graph complexity representing each GSC community contained in the WSI. Graphs were visualized with the help of networkx, a python network module.

4.5.4 Performance Evaluation of the abGCN Model

For our experiment to validate the graph construction method independent of PathoFusion, we used the publicly available TCGA glioblastoma (GBM) dataset [352]. This dataset does not contain CD276 immunohistochemical stains but instead contains H&E histology. Accordingly, we have used the HoVer-Net model [208] for the simple extraction (segmentation) of nuclei from H&E-stained histology images. Subsequently, the corresponding node features were extracted using a ResNet34 pre-trained torch model. Finally, the graph was built using our graph construction method. The resulting graphs were used to train and validate our abGCN model to assess its performance.

For the CD276 dataset, at the graph construction stage, image augmentation including rotation, contrast, and sharpness adjustments was randomly applied to cell patches before they were used to generate the graphs. The actual generated graphs were therefore diversified and expanded greatly (n^2 times) compared to conventional methods (n times), where n is the number of augmentation types.

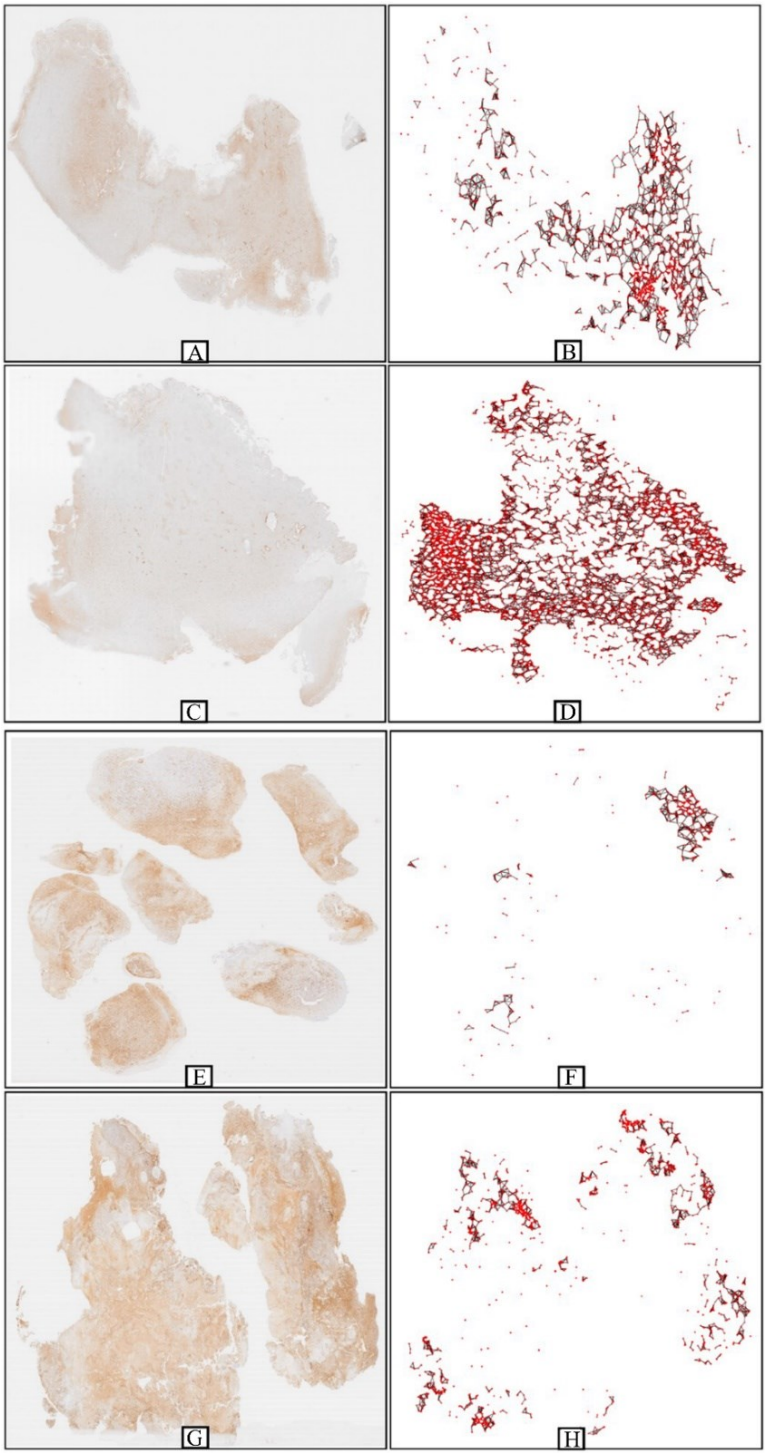

Figure 4.5 Graphs constructed from four sample WSIs on the basis of CD276+ GSCs detected by the BCNN (PathoFusion). (**A**–**D**) two cases with short survival and their corresponding graphs, (**E**–**H**) two cases with long survival and their corresponding much less complex graphs.

We used five-fold cross-validation to assess the performance of our abGCN model for predicting system outcome (patient survival in our example). We split each of the WSI datasets into a training set (80%) (used 25% of them as the validation set) and a testing set (20%). The model was then trained to perform dichotomous categorization (short and long survival, respectively) and to calculate attention scores (risk values) using the training set. Our abGCN model was trained like other state-of-the-art methods for comparison using the Adam optimizer with a weight decay factor of 0.0001. Training was conducted for 100 epochs at a learning rate of 0.001. The predicted risk associated with the input graph is computed by the last fully connected layer. These predicted risks of the input graphs are then fed into the model's Cox proportional hazards regression loss function to minimize those risks and improve the model's accuracy and thus ability to predict system outcome (survival). The model uses back-propagation to update the parameters in order to minimize the loss.

Table 4.1. Performance comparison of different deep learning models.

Model	**C-index**		**Accuracy**	
	TCGA	**CD276**	**TCGA**	**CD276**
MIL (Deep Sets) [355]	0.64	0.63	0.77	0.79
Attention MIL [356]	0.68	0.59	0.80	0.78
DeepGraphConv [90]	0.67	0.56	0.80	0.73
CGAT [299]	0.69	0.73	0.80	0.79
Patch-GNN [103]	0.70	0.79	0.79	0.81
abGCN (this model)	**0.70**	**0.83**	**0.83**	**0.85**

Table 4.1 shows the results of the evaluation of our abGCN model compared to those of existing multiple-instance or contextual feature learning models. Our model outperformed other models with a C-index of approximately 83% and 70% and an accuracy of 85% and 83%, when classifying patient survival (short or long survival) using CD276 dataset and the TCGA GBM dataset, respectively.

4.5.5 Visualization of Attention Scores

Our abGCN model has one attention gate which consists of three layers. It differentially assigns weights to the cells of interest in each of the WSI depending on their feature content and thus highlights their relative importance for the final evaluation. The proposed method produces a rich and interpretable map showing the contribution of each cell and its surroundings to the outcome parameter. The outcome parameter provides an explanation for the patterns found by the deep learning algorithm.

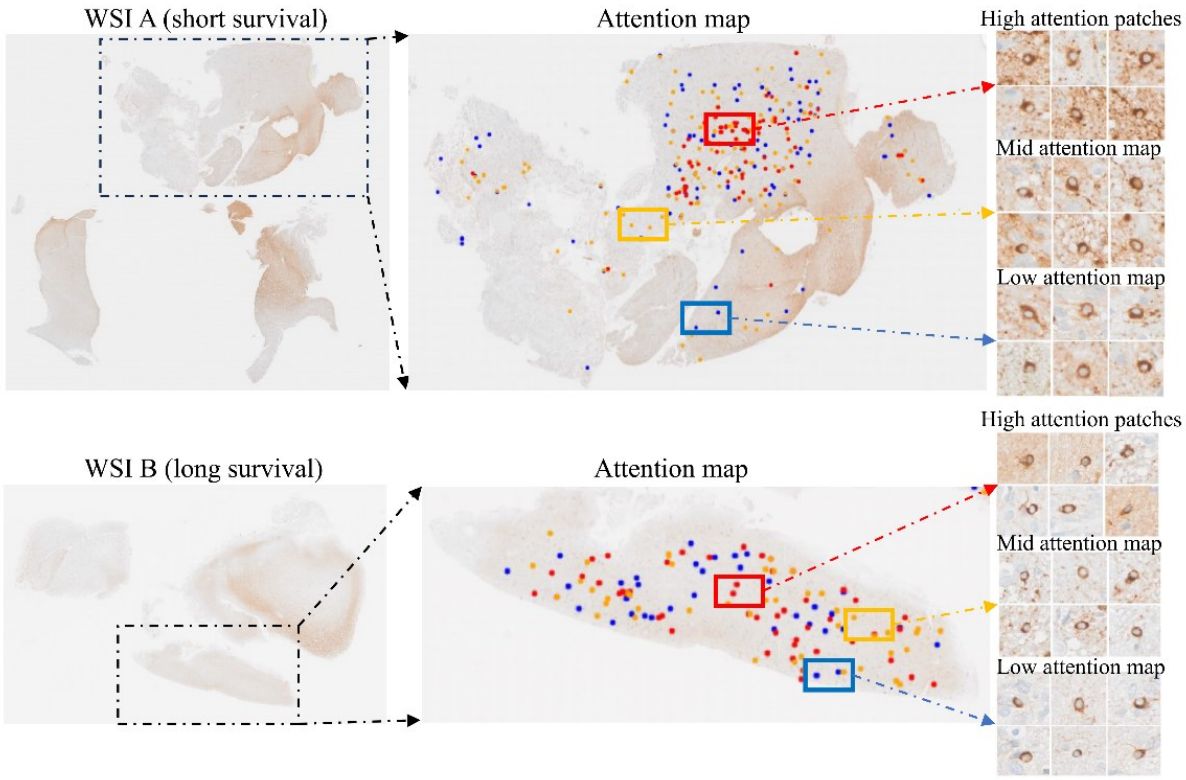

Figure 4.6 Attention scores allow the localization of tissue regions of special interest. This is illustrated using two sample areas from WSIs of cases with different survival outcomes (short and long survival, respectively). The colour coding pinpoints clusters of GSCs showing the highest CD276 expression (red).

In Figure 4.6, we show a visualization of the attention scores of the cells that were detected in the selected regions. We utilize color labels with, e.g., red representing a high weight (score;

please see below). This aids our understanding of which elements of the graph influenced the model's predictions most. Figure 4.6 shows two WSIs with different survival outcomes (short and longer term, respectively). The attention score can be used as a guide on where to focus on the tissue section to find meaningful differences between regions. For instance, they can be used to find areas with especially aggressive tumor growth.

After calculating the attention scores over the entire dataset of each WSI, we group the attention scores into high (>90%, red), medium (60–90%, orange) and low (<60%, blue). The thresholds were strictly defined to ensure that each patch was uniquely assigned to a single group, avoiding any overlap between groups. This is followed by selecting a set of patches from each group based on their attention score. Knowledge of the attention scores allows a contextual analysis of these patches, i.e., how the different groups relate to regions of interest in the WSI that are defined by pathological features which are recognizable when using other modalities on adjacent sections.

4.5.6 Analysis of the K Parameter

To explore how different neighbourhood sizes in the KNN algorithm affect the graph construction, we carried out experiments using K values ranging from 2 to 9 as illustrated in Figure 4.7 The results show that the accuracy of abGCN model initially increases before starting to decline. In our experiments we chose K = 8 to contruct graph in the CD276 dataset K = 6 for the graph in the glioma dataset.

4.5.7 Analysis of the d Parameter

To investigate the impact of the Euclidean distance parameter ED on the clustering process essential for constructing the graph we carried out a range of experiments using ED values of 250, 500, 750 and 1000 as shown in Figure 4.8. Our results show that the best performance of abGCN model at $ED = 500$ for constructing graphs in both CD276 and glioma datasets suggests that this ED value optimally captures the most relevant relationships between nodes. It

likely strikes a balance between overly sparse and overly dense graphs, enabling the model to effectively extract meaningful features for learning.

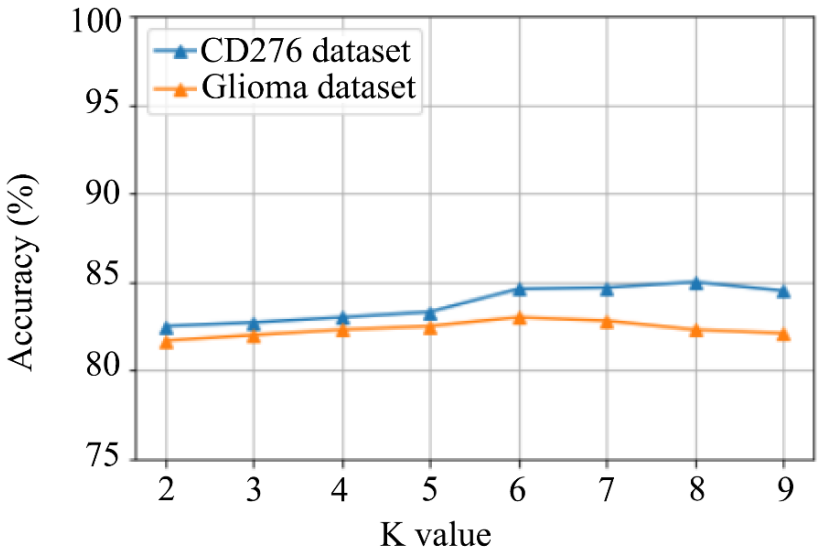

Figure 4.7 Analysis of K values of KNN algorithm for construction the graphs.

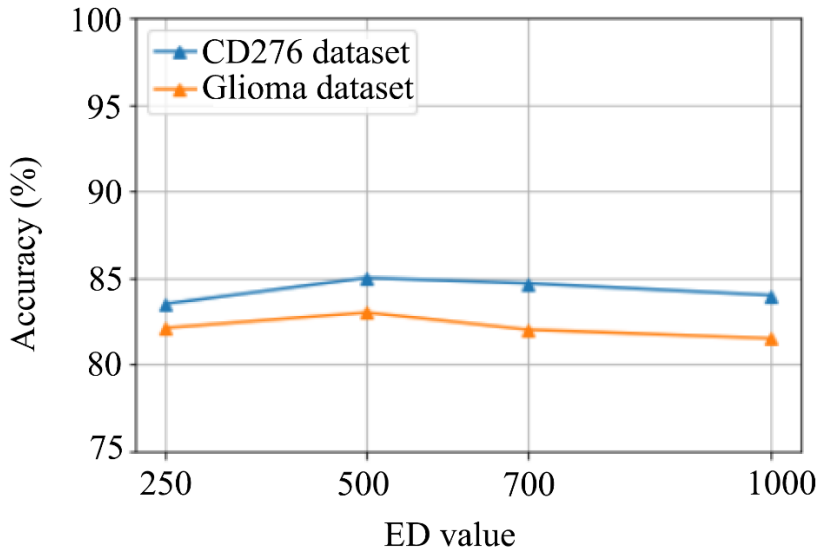

Figure 4.8 Analysis of ED values for construction the graphs.

4.5.8 Analysis of the Hyper-Parameters

In order to optimize our abGCN model performance we carefully adjust the values of α and β within a typical range. Specifically, we explore α values of α = {1, 10^{-1} , 10^{-2} , 10^{-3}, 10^{-4}} and β = {1, 10^{-1} , 10^{-2} , 10^{-3}, 10^{-4}}. The results indicates that abGCN model with parameters α = 10^{-2} and β = 10^{-1} for CD276 data and α = 10^{-3} and β = 10^{-1} for glioma dataset achieves the best performance.

4.5.9 Analysis the Impact of the Constructed Graph

Furthermore, to demonstrate that the proposed method is appropriate for representing the WSI by means of a compressed graph structure, we compared our method to two other graph generation methods, Fully Connected (FC) and Multiple Instance Selection (MIS), respectively. The MIS method randomly samples small patches within a distance threshold independent of the presence of cells of interest whereas the FC graph method includes the extracted cells but constructs one graph without being constrained by limitations on the distances between the connected patches. As shown in Figure 4.9, we found that the proposed graph construction method improves prediction accuracy of the abGCN model significantly compared to the other two methods tested.

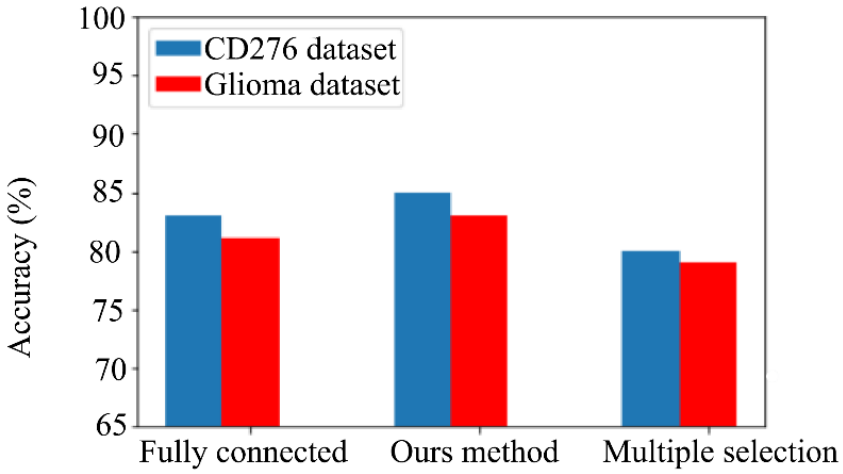

Figure 4.9 Impact of the constructed graph.

4.6 Discussion

Recent studies have demonstrated that elevated expression of CD276 may have a role in inducing the epithelial–mesenchymal transition (EMT) [333]. Jiang and colleagues [357] have shown that upregulated CD276 in primary colon adenocarcinoma can enhance the expression of CD133 and CD44 associated with the EMT in colorectal cancer. CD133 and CD44 are cell membrane markers frequently used to identify cancer stem cells and colorectal tumour-initiating cells. Using digital spatial profiling, Petterson et al. [358] have observed elevated expression of B7-H3 (CD276) and CD44 in areas of hypoxia in GBM tissue. In GBM, B7-H3 is also prominently expressed by pericytes and macrophages in GBM tumour tissue [347]. It is worth noting that elevated B7-H3 expression and increased infiltration of macrophages in colorectal carcinoma have been negatively correlated with patient survival [346]. It has also been suggested that upregulated CD276 expression promotes macrophage differentiation towards an anti-inflammatory phenotype [346]. The elevated expression of B7-H3 by microglia and macrophages in glioblastoma may facilitate immune evasion [358].

Certain existing deep learning methods, for instance CNNs, have been proposed for survival analyses using WSIs and have shown promise when trying to understand disease progression and predicting patient outcome [255,359]. However, when dealing with relation-aware representations, CNN models are inefficient because they suffer from a limited ability to capture global contextual information and thus model tissue composition comprehensively. In tissue diagnosis, the phenotypic characteristics and topological distribution of cellular and histological features are of critical importance.

Consequently, GNNs have attracted considerable interest as they capture intra- and inter-entity level interactions [15] and are able to record geometrical alongside topological properties, thus allowing cell-level information and overall tissue micro-architecture to be modelled simultaneously. GNNs have already been used for ROI classification [88], the detection of malignant growth and tumour invasiveness [89], as well as survival-analysis on whole slide pathological images [90]. While these techniques consider combining the cell-level information and tissue micro-architecture through using the graph networks, a significant challenge lies in

increasing the robustness of the node features, producing a more robust WSI-level representation, and significantly enhancing the model's performance.

Therefore, in this study, we proposed PathoGraph framework for converting selected cellular immunohistochemical labelling (CD276) into graphs in a memory-efficient way with vertices representing cells and their features as the connecting edges which contain information on spatial relationships based on ED mapping and similarities between neighbouring cells. The proposed graph construction method ensures computational efficiency by limiting connections using spatial thresholds or feature similarity constraints, resulting in a sparsely connected graph with an average degree significantly smaller than the total number of vertices. This reduces the computational complexity from $O(V^2)$ in a fully connected graph to $O(V.d)$, where $d \ll V$.

We also demonstrated that employing the abGCN model results in improved performance, reduced computational demand, and better generalization properties. The model leverages a lightweight attention mechanism that adapts its weights based on both feature similarity and spatial proximity, enabling it to effectively capture intricate patterns and relationships in the data. Although the attention mechanism theoretically operates with a complexity of $O(V^2)$, the sparsity of the constructed graph substantially reduces the number of computations in practice. This targeted approach improves the model's ability to manage diverse feature distributions while integrating spatial and heterogeneity aspects within histological characteristics, all contributing to its robust predictive performance.

Our abGCN captures a distinct property of the CD276 high-expressing ($CD276^{high}$) malignant glioma cell population (GSCs) in each slide in the form of graphs that are representative of the respective biopsy/case and based on global information from the WSI. Importantly, more complex graphs are found to correlate with shorter survival. One biological interpretation could be that the more complex graphs represent more aggressively growing clones of GSCs. The underlying information is not readily accessible to a human observer in most cases illustrating the power of the abGCN. In addition, the attention scores allow the localization of such cell communities within the tissue sections. These findings are intriguing and may have implications for histopathological prognostication and treatment decisions.

4.7 Summary

We have developed a graph-based AI model for the detection and interpretation of immunohistochemical labelling results at the cellular and WSI level. In the case of CD276 when used as a marker of GSCs in malignant glioma, the method allowed us to predict system outcome on the basis of the graphs and attention scores that were generated. Specifically, our PathoGraph framework containing an abGCN was able to classify survival outcome and the attention scores allowed pinpointing of WSI areas of particular diagnostic interest. Therefore, our graph visualization method is able to uncover interpretable information that is largely hidden to a human observer. Employing a GCN with attention heads resulted in improved performance.

While our current model shows promising results in classifying scanned slide images based on survival outcomes, we acknowledge that a larger dataset would be advantageous. A more extensive dataset would not only enhance the robustness of our model but also enable more sophisticated analyses and potentially improve its predictive capabilities. We are actively working towards building a larger dataset to further refine and validate the performance of our model. Finally, we would like to point out the potential generalizability of the new method. It might be used for screening for novel biomarkers for example when different system outcomes and input data sets are combined.

CHAPTER 5

An Attentive Multi-Graph Fusion for Tumor Grading in Pathology Images

---

Cancer grading in pathology is crucial for patient care and treatment decisions. Despite advancements in deep learning for cancer diagnosis, integrating diverse knowledge of cancer pathogenesis remains challenging, especially regarding tissue and cellular composition. Traditional GNN techniques typically use a single graph or merge multiple graphs into cohesive structure without considering balancing the contributions of different graph representations. However, histopathological analysis requires considering various features such as cellular morphology, tissue architecture, and spatial relationships, which create complex connections that provide valuable information for classification tasks. Determining which nodes or graphs significantly contribute to classification remains a challenge.

However, despite CNNs have been shown promise in understanding of disease progression and predicting patient outcomes, when dealing with relation-aware representations, they are inefficient because they limit the model's ability to capture global contextual information and comprehensively model tissue composition. As a result, GNN have been proposed addressing the shortcomings of current CNN approaches [350]. GNN have attracted considerable interest as they capture intra- and inter-entity level interactions and are able to record geometrical alongside topological properties, thus allowing cell-level information and overall tissue micro-architecture to be modelled simultaneously. WSI could be represented as a graph, where nodes and edges represent distinct histological entities and their connection/relationship, respectively [15]. GNN have already been used for cancer grading [99], and tissue sub-typing [91,301].

Most of the GNN approaches construct and used a single graph. Given the quantity of cells, the interactions explored and acquired through graph learning are typically limited to a small structure [91]. Recently, various methods have been proposed combining multi-structural information, for a more understanding of tissue structures. Thus, to provide a more effective graph representation of tissues, a GNN technique that can efficiently and effectively leverage both the broader and more detailed aspects of tissue structures is required.

Despite our previously introduced PathoGraph framework [360], that employs a single graph with attention mechanisms, modelling cellular architecture for classifying the survival outcomes efficiently, selecting the optimal graph from multiple options can pose a challenge for achieving the best model prediction performance. Furthermore, combining multiple graphs and determine which nodes or graphs significantly contribute to classification tasks is challenging. Therefore, we propose an Adaptive Multi-Graph Fusion-based Attentive Graph Neural Network, termed AMGF-GNN, in which multiple networks are trained concurrently with different gradient flow routes, which allows for faster convergence while maintaining the same amount of trainable parameters (Figure 5.1). This framework leverages sets of data features and various types of graph connections in parallel to get the best feature representation. Specifically, we construct three graphs: one based on community information, one based on feature similarity, and other based on both, so distinct embeddings were extracted across different spaces and levels by using three parallel GNN with single attention modules. The single-graph attention layer assigns weights to each node in each single graph, enabling the model to distinguish their influence. We further utilize an Adaptive Attentive Multi-Graph Fusion layer to adaptively fuse structural information from multiple graphs, enhancing the model's understanding of complex relationships across different graph instances.

The main contributions of this chapter are summarized as follows:

- We introduce end-to-end multi-graph neural network framework that constructs and processes three distinct graphs based on community information, feature similarity, and a combination of both, providing a comprehensive and robust representation of the tissue structure for subsequent learning tasks.

- While existing multi-graph fusion methods often fail to adaptively balance the contributions of different graph representations, leading to suboptimal integration of information. We develop an Adaptive Attentive Multi-Graph Fusion module that innovatively fuses multiple graphs into a unified dense vector space after learning and updating embeddings of each graph separately. By applying a nonlinear transformation and a shared attention vector, it adaptively adjusts and integrates the feature embeddings from each graph, ensuring that the most relevant information is effectively combined. It learns to balance the contributions of community and feature-based information, addressing the uncertainty of their relative importance for the final grading task.
- We design a dual-level loss optimization strategy incorporating both intra-graph and inter-graph similarity measures. This ensures consistency and robustness in the learned embeddings, capturing intricate relationships within and across graphs, and significantly enhancing the model's performance.
- We conducted experiments to evaluate our proposed approach on histopathology image grading tasks using two public datasets, demonstrating superior performance and providing valuable insights into the tissue microenvironment compared to existing methods.

5.1 Datasets

To evaluate the efficacy of our proposed method, we utilized two different cancer datasets: (1) the glioma cancer dataset and (2) the breast cancer dataset.

The glioma dataset comprises a set of 937 WSIs from 490 patients were collected from The Cancer Genome Atlas (TCGA) via the Genomic Data Commons (https://gdc.cancer.gov/) [361], 654 WSIs were retained for the experiments after excluding those that were corrupted, had unremovable markings, poor of resolution, or lacked clinical information. Glioma images are classified into four grades, with grades 3 and 4 being the most severe. TCGA data includes two types of glioma images: Brain Lower Grade Glioma (LGG, grades 1 and 2) and Glioblastoma Multiforme (GBM, grades 3 and 4). According to the World Health Organization (WHO)

classification, diffuse gliomas are categorized into astrocytic, oligodendroglial, and mixed types, corresponding to grade-II (low-grade), grade-III (anaplastic), and grade-IV (glioblastoma GBM).

Using this dataset, we conducted two experiments in our study, binary grading experiment in which gliomas were classified into LGG (encompassing WHO grade-II and grade-III) and GBM (WHO grade-IV), and three-grade classification experiment in which gliomas were differentiated into three grades, maintaining the distinction between grade-II (low-grade) and grade-III (anaplastic) within LGG, and grade-IV (GBM).

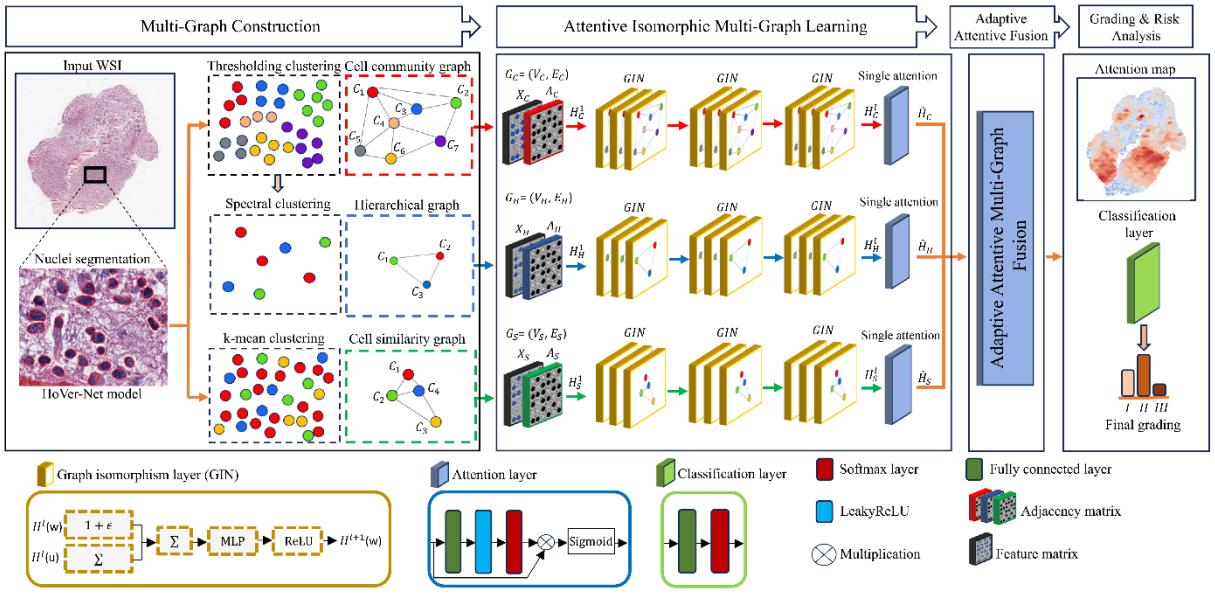

Figure 5.1 The framework of AMGF-GNN model. Nuclei feature matrix X and adjacency matrix A are to construct multiple graphs. AMGF-GNN consists of three dedicated branches, each including graph isomorphism network (GIN) module of three GIN layers and single attention module, and one adaptive attentive multi-graph layer to fuse the outputs of different single graphs modules.

The breast cancer dataset is a histopathological microscopy image dataset for invasive ductal carcinoma (IDC), the most widespread subtype of all breast cancers. This dataset contains 922 images in total, related to 124 patients with IDC [362]. Each specimen is available in four

magnification levels: 4x, 10x, 20x, and 40x. The images are categorized according to their diagnosed grade, which includes three grades of IDC. In our experiment, we only used those of 40x magnification patients with IDC [362].

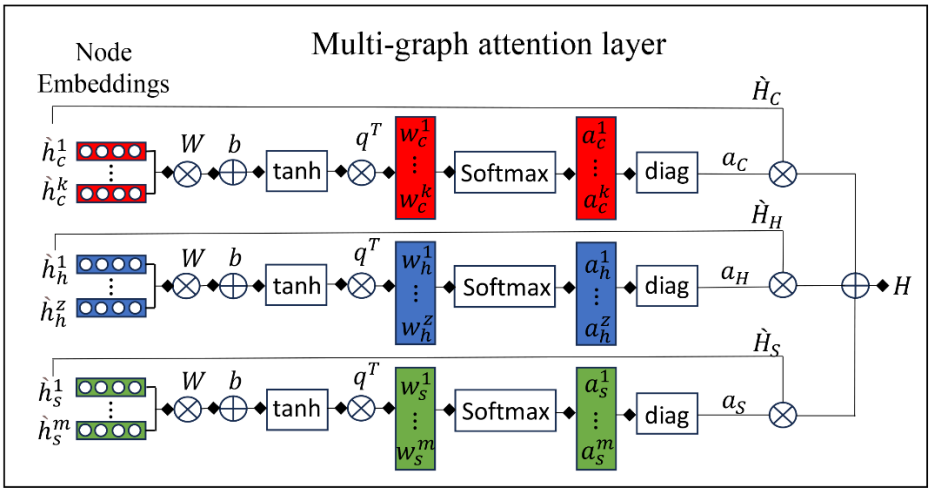

Figure 5.2 Adaptive Attentive Multi-Graph Fusion Module.

5.2 Multi-Graph Construction

5.2.1 Nuclear Segmentation and Features Extraction

A pretrained HoVer-Net [208] was used for nuclear segmentation. HoVer-Net utilizes the information in the vertical and horizontal distances between nuclear pixels and their centers of mass enabling precise segmentation even in regions where instances overlap. The nucleus centroids act as the spatial points for nodes in the graph representation of the entire WSI. For a given WSI, this results in a set of N nuclei $P = \{p_j | \, j = 1, \ldots, N\}$ with nuclear centroid at p_j.

Once the nuclei had been detected, we extract multi-attributed features of 566-dimensional vector corresponding to each nucleus comprised a set of sub-features including:

1) CNN-based features: we utilize ResNet [55] trained on ImageNet dataset [288] to extract features from 72×72 patches, centered around the nuclear centroids. These features are a 512-dimensional feature vector.

2) Gray Level Co-occurrence Matrix (GLCM) features: GLCMs of dimensions 5×5 are computed in both horizontal and vertical directions. When these features are flattened, they result in a feature set comprising 50 dimensions capturing morphological characteristics such as clearing, alignment, and insights into the overlapping or packing density of nuclei.

3) Average intensity (average RGB Value): a patch of size 72×72 undergoes global average pooling, producing a feature vector with 3 dimensions representing the average staining intensity in the nucleus and its surrounding area.

4) Number of neighbors of a nucleus: This feature captures the density of connections of each nucleus. It offers valuable information about the spatial distribution and relationships of nuclei within the WSI.

Concatenating these features of each nucleus from all nuclei in WSI as a feature matrix X of size $N \times F$, where N is the number of nuclei in the WSI and F is the number of features (566 in this case), is crucial for graphs generation purposes.

5.2.2 Cell Community Graph

A cell community graph characterizes cellular spatial patterns and the community information. It can be valuable for understanding the heterogeneity of cell populations and uncovering patterns that might be associated with specific biological properties or conditions. We use thresholding clustering to group spatially neighboring nuclei into clusters C_C, with the threshold distance acting as a key parameter in determining cluster membership.

Therefore, for the detected nuclei N, the clustering process continues until clusters form at a minimum distance threshold, for instance $d_{min} = 500$ pixels for the glioma dataset. The outcome of this clustering process is a set of K clusters, represented by $C_C = \{c_{c,k} | k = 1, \dots, K\}$, where each nucleus p_j is assigned to exactly one cluster of its neighbours in N. This results in the spatial clustering of all nuclei within a given WSI.

Once the clusters are formed, for each cluster $c_{c,k}$ we calculate the centroid $\bar{p}_k$ by taking the

average of the centroids of all nuclei within that cluster. Additionally, we compute the cluster features $\bar{f}_K$ by the average of all the nuclei features within each cluster. So, the centroid $\bar{p}_k$ and features $\bar{f}_K$ of cluster $c_{c,k}$ can be expressed as Equation (5.1) and (5.2):

$$\bar{p}_k = \frac{1}{|c_{c,k}|} \sum_{p_j \in c_{c,k}} p_j \tag{5.1}$$

$$\bar{f}_K = \frac{1}{|c_{c,k}|} \sum_{p_j \in c_{c,k}} f_j \tag{5.2}$$

This clustered representation serves as the basis for subsequent cell community graph generation $G_C = \{V_C, E_C, X_C, A_C\}$, where the nodes of the graph correspond to the centroids of these clusters V_C, along with their corresponding average features X_C, we employ the k-Nearest Neighbours (KNN) algorithm to build edges E_C that represent cellular community interactions and encode cell community topology in the adjacency matrix $A_C \in R^{|V_C| \times |V_C|}$. This approach allows us to represent the spatial organization and average characteristics of nuclei clusters in the form of a graph, facilitating our further analysis and interpretation of the data.

5.2.3 Hierarchical Graph

The hierarchical graph interprets spatial patterns among cellular levels in the tumor microenvironment (TME), offering insights into TME complexity and intra-tumoral heterogeneity through cellular and clonal phenotype analysis. Following the thresholding clustering we did previously to create the cell community graph, we pool the cluster's centroids K together and apply spectral clustering based on the similarity of features (Equation 5.3). This results with Z clusters represented as $C_H = \{C_{H,z} | z = 1, \dots, Z\}$, assist in uncovering the cellular community within a population level (higher-level structure).

For each cluster $C_{H,z}$, we calculate the cluster centroid $\bar{p}_z$ and average features $\bar{f}_z$ that summarize the characteristics of the global generated clusters, using Equation 5.1 and 5.2. Cluster centroids and their average features are used to construct the hierarchical graph $G_H = \{V_H, E_H, X_H, A_H\}$, where the nodes of the graph correspond to the centroids of these clusters V_H, along with their corresponding average features X_H, and we use the KNN algorithm to build

edges E_H, which signify the similarity or closeness between the corresponding clusters in terms of their feature characteristics, and the adjacency matrix $A_H \in R^{|V_H| \times |V_H|}$ encodes these relationships.

5.2.4 Cell Similarity Graph

Cell similarity graph plays a critical role in uncovering the intrinsic relationships and organization present within the WSI. It facilitates the exploration and analysis of complex biological systems, offering valuable insights into the underlying structure and function of cellular communities.

Based on the features extracted for each nucleus, cosine similarity S is employed as a metric to quantify the similarity between the feature profiles of nuclei, as defined in Equation (5.3).

$$S_{ij} = \frac{x_i . x_j}{|x_i||x_j|} \tag{5.3}$$

where S_{ij} measures the similarity between feature vector x_i and feature vector x_j of each pair of nuclei in feature space. This step provides valuable insight into the structural similarities between different nuclei, enabling us to quantify the degree of resemblance between them. Utilizing k-means clustering, results in a set of M clusters of nuclei based on their features similarities, represented as $C_S = \{c_{s,m} | m = 1, \dots, M\}$.

For each cluster $c_{s,m}$ we calculate the cluster centroid and average features that summarize the characteristics of the nuclei within the cluster, same as in Equation 5.1 and 5.2. The result is M of centroids along with the average features of the nuclei within each cluster, allowing us in the construction of cell similarity graph $G_S = \{V_S, E_S, X_S, A_S\}$, where the nodes of the graph correspond to the centroids of these clusters V_S, along with their corresponding average features X_S, we use the KNN algorithm to build edges E_S between the clusters based on their feature similarity, and the adjacency matrix $A_S \in R^{|V_S| \times |V_S|}$ encodes these relationships.

5.3 Network architecture

We focus on the cancer grading task using multi-graph structure denoted as $G = (X, A)$, where the data feature $X \in R^{V \times D}$ consists of V nodes each with D dimensional features, and $A \in R^{|V| \times |V|}$ is the adjacency matrix of the graph that define the relationships between the nodes.

In this section, we present the proposed AMGF-GNN model which includes two main modules that are corresponding to attentive isomorphic multi-graph learning including GIN layers and single attention mechanism, and adaptive attentive multi-graph fusion. The first module is utilized within the dedicated branches (G_C, G_H, G_S) to learn how cells interact with each other in the pathology image from three different views, including cell community graph G_C, hierarchical graph G_H, and cell similarity graph G_S. The adaptive attentive multi-graph fusion component adaptively fuses the outputs of $\grave{H}_C$, $\grave{H}_H$ *and* $\grave{H}_S$ to provide grading label for the pathology image. Our framework is designed to allow information to flow and interact among nodes not just based on their connectivity in the graph structure but also on the characteristics of their features. Therefore, AMGF-GNN considering and integrating information from the topological space (community space), which captures the structural relationships between nodes, the feature space, which encompasses characteristic information associated with each node, and the combination between both, which enables the model to uncover complex patterns and relationships at a higher level, ultimately improving its ability to classify and predict outcomes. Details in the following sections.

5.3.1 Attentive Isomorphic Multi-Graph Learning

5.3.1.1 Graph Isomorphism Network (GIN)

GINs utilize learnable aggregation functions for direct node embedding updates, effectively capturing complex interactions and dependencies among nodes. However, our AMGF-GNN model incorporates three parallel GINs, receiving inputs from multiple graphs including cell community graph, hierarchical graph, and cell similarity graph ($\mathrm{G_C}, \mathrm{G_H}, \mathrm{G_S}$). As shown in

Figure 5.1, each GIN layer in the model includes three "slices": one for aggregating neighboring node features, one for incorporating the node's own features with aggregated neighbors, and one for applying a MLP with ReLU activation to transform the combined features. By using the following Eq. (4), we can learn and generate parallelized feature representations:

$$H_V^{(l)}(w) = MLP\left(\left(1 + \epsilon^{(l)}\right).H_V^{(l-1)}(w) + \sum_{u \in N_V(w)} H_V^{(l-1)}(u)\right) \quad (5.4)$$

where $V = \{C, H, S\}$ indicate the graph type, $H_V^{(l)}(w)$ represents the embedding of node w at layer l, MLP is a multi-layer perceptron (feedforward neural network) applied to each node individually, $\epsilon^{(l)}$ is a learnable parameter at layer l controlling the contribution of the node's own embedding, and $N_V(w)$ denotes the neighbours of node w. The output embeddings of three different paths using three different graphs H_C, H_H, and H_S could be calculated using the above formula.

5.3.1.2 Single Graph Attention

It is crucial to discern the influences of various nodes within a single graph embedding, as each node may play different roles. The self-attention mechanism is trained to identify the nuclei relevant for the grading task. Adaptive weights are learned for every node in the graph, with higher weights assigned to nodes that are more closely related to each other.

The input to this layer is a set of node embeddings of single graph obtained from the GIN. These embeddings include the cell community graph embedding H_C^L, hierarchical graph embedding H_H^L and cell similarity graph embedding H_S^L, represented as $\{h_{v,1}, h_{v,2}, \dots, h_{v,i}\}$. The output is a new set of node embeddings, $\{\acute{h}_{v,1}, \acute{h}_{v,2}, \dots, \acute{h}_{v,i}\}$, where i is the number of nodes in each single graph v.

To transform the input features into higher-level representations, attention weight w_v and bias b are applied to every node in the graph v to compute attention coefficient a. The normalized attention coefficients are utilized to get the better embeddings. Mathematically, the self-attention mechanism can be expressed as follows, Equations (5.5, 5.6, 5.7):

$$p(i) = LeakyReLU(\, w_v . h_i + b_v) \tag{5.5}$$

$$a(i) = \frac{e^{p_i}}{\sum_{j=1}^{N} e^{p_j}} \text{, for all } i \in \{1,2, \dots, N\} \tag{5.6}$$

$$\grave{h}(i) = Sigmoid\ (a(i)\ .\ h_i) \tag{5.7}$$

where h_i is the feature vector of node i and $\grave{h}(i)$ is the resulting better embedding vector. For instance, given a node vector $h_{S,i}^{L}$ from H_S^L, the weight coefficient $a_s\,(i)$ can be expressed as follows:

$$a_s\,(i) = \frac{\exp\left(LeakyReLU\left(w_s^T . h_{S,i}^{L} + b_s\right)\right)}{\sum_m \exp\left(LeakyReLU\left(w_s^T . h_{S,j}^{L} + b_s\right)\right)} \tag{5.8}$$

And utilizing the adaptive selection, we obtain the better embedding $\grave{h}_s^i$ by:

$$\grave{h}_s^i = Sigmoid\left(a_s\,(i)\,.\,h_{S,i}^{L}\right) \tag{5.9}$$

Then we get the F-dimension feature representation of N nodes of v graph, which is the cell similarity graph in our example, as:

$$\grave{H}_S \in\ R^{N \times F} = \left\{\grave{h}_s^1\,, \grave{h}_s^2, \dots, \grave{h}_s^N\right\} \tag{5.10}$$

and same operation for each single graph of $H_1^{(L)}, \dots., H_V^{(L)}$, i.e. $H_C^{(L)}$, $H_H^{(L)}$, and $H_S^{(L)}$ in our application.

5.3.2 Adaptive Attentive Multi-Graph Fusion

To enhance the stability of the self-attention learning process, and to integrate the feature information from different single graphs effectively, we have found extending our mechanism to employ adaptive attentive multi-graph to be beneficial (Algorithm 5.1). We initially apply a nonlinear transformation to the embeddings. Subsequently, we leverage a multi-graph attention method $att(\grave{H}_C\,, \grave{H}_H, \grave{H}_S)$ to further learn more extensive feature embeddings (Figure 5.2). For example, for each ith node in cell similarity graph $\grave{H}_S$ with feature vector $\grave{h}_s^i\ \in R^F$, we first applying a nonlinear transformation to the embedding $\left(\grave{h}_s^i\right)^T$, then we employ a shared attention

vector $q \in R^h$ to get the attention value w_s^i using the following:

$$w_s^i = q^T . tanh\left(W.\left(\grave{h}_s^i\right)^T + b\right) \tag{5.11}$$

where $W \in R^{h \times d}$ represents the weight matrix and $b \in R^h$ denotes the bias. Then, we apply the softmax function to normalize the attention values w_s^i obtaining the final attention weight:

$$a_s^i = \frac{exp(w_s^i)}{\sum_{j=1}^{N} exp\left(w_s^j\right) + \sum_{j=1}^{N} exp\left(w_f^j\right)} \tag{5.12}$$

Finally, we create diagonal matrices with the weights along the diagonal, for example $a_S = diag(a_s^1, a_s^2, \ldots, a_s^N)$ and similar for a_C and a_H, then combined the graph embeddings to obtain the final embedding which complement the information between three different graphs:

$$H = a_C . \grave{H}_C + a_H . \grave{H}_H + a_S . \grave{H}_S \tag{5.13}$$

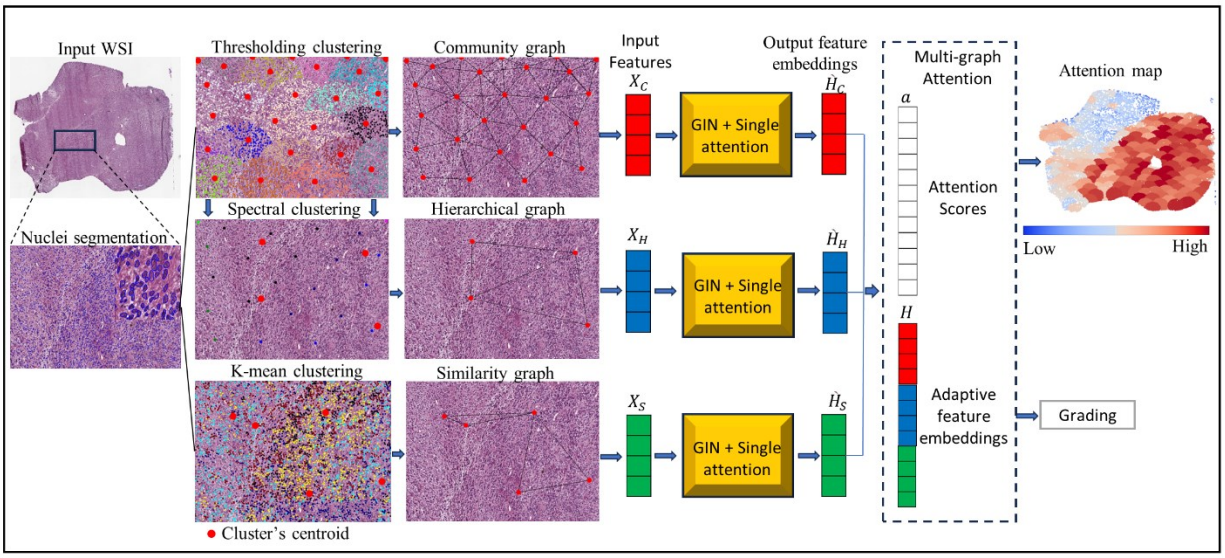

Figure 5.3 Overview of the proposed methodology for analyzing tissue using a fusion of multiple graphs. Following nuclei segmentation, different clustering methods were applied so multiple graphs are constructed, and they are processed via multi-graph based attentive graph neural network to learn the mapping from tissue compositions to corresponding tissue grading and risk analysis.

Algorithm 5.1: Adaptive multi graph fusion GNN model (AMGF-GNN)

Input:

Cell community graph G_C

Hierarchical graph G_H

Cell similarity graph G_S

Output:

{H for grading the input WSI, attention weights (a_C, a_H, a_S) for risk analysis}

begin

$G \in \{G_C, G_H, G_S\}$

for G_V in G **do**

for $l \in 1, \ldots\ldots, L$ **do**

/* calculate the output embedding of each node w in each graph using GIN layers */

$$H_V^{(l)}(w) = MLP\left(\left(1 + \epsilon^{(l)}\right).H_V^{(l-1)}(w) + \sum_{u \in N_V(w)} H_V^{(l-1)}(u)\right)$$

end

/* perform the single graph attention on each node embedding */

$$\grave{\boldsymbol{h}}_{\boldsymbol{v}}^{\boldsymbol{i}} = \boldsymbol{Sigmoid}\left(\frac{e^{LeakyReLU(h_i.w_v + b_v)}}{\sum_{j=1}^{N} e^{LeakyReLU(h_i.w_v + b_v)}} . \boldsymbol{h}_{\boldsymbol{v,i}}^{\boldsymbol{L}}\right)$$

end

/* Using multi-graph attention to unified embeddings $att(\grave{H}_C, \grave{H}_H, \grave{H}_S)$ */

for $\grave{h}_v^i$ in $\grave{H}_V$ **do**

$$w_v^i = q^T . tanh\left(W.\left(\grave{h}_v^i\right)^T + b\right)$$

$$a_v^i = \frac{exp(w_v^i)}{\sum_{j=1}^{N} exp\left(w_v^j\right) + \sum_{j=1}^{N} exp\left(w_f^j\right)}$$

end

/* Fusing the three graph embeddings $\grave{H}_C$, $\grave{H}_H$ and $\grave{H}_S$ */

$$H = a_C . \grave{H}_C + a_H . \grave{H}_H + a_S . \grave{H}_S$$

end

5.3.3 Optimization Function

In order to learn consistent representations across different graphs and improve the overall effectiveness of our grading task, we propose a combined loss function that integrates multiple components: the adapted multi-graph consistency loss function $\mathcal{L}_{cons}$ and the graph-outcome loss L_o. This combined loss ensures that the similarity relationships between clusters within and across different graphs are consistent enabling effective integration of information from multiple graph representations while also accurately predicting the graph-level grades.

First, we normalize the feature matrices $\grave{H}_V$ using L2 normalization where $V = \{C, H, S\}$ indicate the graph type, resulting in normalized matrices $\grave{H}_{nor,V}$. Then, compute a similarity matrix $\grave{S}_V$ for each normalized feature matrix $\grave{H}_{nor,V}$ for nodes within the same graph as follows:

$$\grave{S}_V = \grave{H}_{nor,V} \cdot \grave{H}_{nor,V}^T \tag{5.14}$$

Next, compute the average of the similarity matrices S_{ref} to calculate the intra-graph loss of each graph $\mathcal{L}_V$ as:

$$\mathcal{L}_V = \left\| S_{ref} - \grave{S}_V \right\|_F^2 \tag{5.15}$$

To calculate the inter-graph loss, for each pair of graphs (i, j), calculate the similarity $S_{inter,i,j}$:

$$S_{inter,i,j} = \grave{H}_{nor,i} \cdot \grave{H}_{nor,j}^T \tag{5.16}$$

and compute the loss for each pair of graphs:

$$\mathcal{L}_{inter,i,j} = \left\| S_{inter,i,j} - S_{inter,j,i} \right\|_F^2 \tag{5.17}$$

The overall consistency loss $\mathcal{L}_{cons}$ combines the intra-graph and inter-graph losses:

$$\mathcal{L}_{cons} = \sum_{v=1}^{V} \mathcal{L}_V + \frac{1}{2} \sum_{i=1}^{V} \sum_{j=1}^{V} \left\| S_{inter,i,j} - S_{inter,j,i} \right\|_F^2 \tag{5.18}$$

For the grading task, we use the output embedding H with a linear transformation and a softmax function $\hat{Y}_{graph} = softmax(W.H + B)$, $\hat{Y}_{graph}$ denote the predicted graph-level grade, while Y_{graph} is the ground truth grade for the entire graph. The grading-outcome loss is computed as:

$$\mathcal{L}_o = -\frac{1}{|D|}\sum_{d=1}^{|D|}\sum_{i=1}^{C}\left(Y_{graph,d,i}\ln\hat{Y}_{graph,d,i}\right) \quad (5.19)$$

where $|D|$ represents the total number of graphs in our dataset, and C the number of classes or grades.

Finally, we combine these losses into a single objective function $\mathcal{L}_{Final}$ to be optimized during training:

$$\mathcal{L}_{Final} = \alpha\mathcal{L}_{cons} + \beta\mathcal{L}_o \quad (5.20)$$

α and β are the parameters of the consistency and grading-outcome to optimize the proposed model via back propagation.

5.4 Experiment Results

To assess our AMGF-GNN model for pathology images grading, we visualize graphs generated from histopathology images (Figure 5.3), compare the classification performance of the proposed AMGF-GNN model with state-of-the-art architectures, and carried out thorough ablation experiments to evaluate the graphs used for grading and to deepen our understanding of the model.

5.4.1 Implementation and Hyperparameter Setting

The proposed AMGF-GNN model is implemented using Pytorch alongside Pytorch Geometric. In order to evaluate the performance of our proposed model and compare with state-of-the-art methods (SOTA), all SOTA baselines are initialized using the same parameters as recommended in their respective papers, and we also further fine tune these parameters for optimal results. For our model, we simultaneously train three paths using 3-layer GINs, with each GIN layer uses MLP layer with ReLU activation and a hidden layer dimensions of 64, and a single attention gate, while the outputs aggregated using a multi graph attention and classifier of MLP with 128 hidden neurons and output neurons depend on the grading classes. The model is trained to reduce the difference between the predicted outputs and the actual labels by minimizing cross entropy loss. We employ the Adam optimizer with 5e − 4 learning rate. Moreover, we set a dropout rate

of 0.25 and weight decay 1e − 4. To assess the effectiveness of our models we utilize two metrics, accuracy (ACC) and F1 score (F1).

5.4.2 Evaluating the Performance

In order to explore if our model gains an advantage from using multiple graphs, we set up a comparative experiment with standalone each single graph used in our proposed model. The whole experimental results are summarized in Table 5.1, where cell graph G_{Cell} GNN (the graph generated from cellular interactions among all the cells), cell community graph G_C GNN, hierarchical graph G_H GNN, and cell similarity graph G_S GNN. Each of them indicates that the model using only single path with the corresponding graph. As shown in Table 5.1, illustrates how the quality of a graph impacts classification outcomes with various graphs offering distinct contributions to the results. AMGF-GNN consistently outperforms each of G_{Cell}, G_C, G_H, and G_S, which indicating the limited information provided by single graph and effectiveness of the adaptive fusion mechanism in AMGF-GNN.

On the other hand, for example for the glioma binary grading, the performance of cell graph and cell community graph are significantly better than other single-graphs, it offers the higher quality and delivers information more efficiently, but also AMGF- GNN performances effected by the cell similarity and hierarchical graphs which also means that both are also important. Different single graphs generated from our dataset have varying influences on the final classification results and choosing the optimal graph for classification is challenging.

To address this, our model extracts and combines details from multiple graphs, allowing them to support and enrich one another. The model learns the weights of these graphs through a multi-graph attention layer to prioritize the most relevant one. Experimental results show the benefits of multi-graph information for classification tasks.

Also, to optimize performance, we carefully adjust α and β within a common range, α = {1, 10^{-1} , 10^{-2} , 10^{-3}, 10^{-4}} and β = {1, 10^{-1} , 10^{-2} , 10^{-3}, 10^{-4}}. The proposed model achieves the best performance with parameters α = 10^{-3} and β = 1 on glioma and α = 10^{-4} and β = 1 on

IDC datasets. We compare AMGF-GNN with some of state-of-the-art methods, covering CNN models and different GNN models. Table 5.2 shows the classification results of all methods. It's evident that our proposed model acquires the best classification performance on the glioma TCGA dataset with ACC of 89.68, which demonstrate the effectiveness of AMGF-GNN model.

5.4.3 Analysis of Attention Mechanism

To effectively leverage structural information from multiple graphs, we employ the attention mechanism, which allows the model to weigh the significance of different nodes within a graph (single-graph attention) and different graphs within a set (multi-graph attention). We conduct an ablation experiment to evaluate the impact of the single-graph and multi-graph attention mechanisms on our AMGF-GNN model's ability to learn and potentially gain insights into the underlying data. Thus, we compare the results of the model with and without each of the attention mechanisms to understand their individual contributions.

Table 5.3 illustrates how the single-graph attention layer enhances the proposed model's ability to learn more meaningful and discriminative representations of the graph. The multi graph attention layer aids the model in concentrating on important graphs among a set of graphs. The combination of single graph attention layer and multi-graph attention layer allows the model to integrate insights from multiple graphs, recognize subtle patterns and trends and provides a comprehensive and nuanced understanding of the graphs, which can significantly improve the model's performance in grading tasks. Moreover, we compared these three different configurations of the model in terms of sensitivity and specificity, Figure 5.4. Our model demonstrated the best performance.

Table 5.1 Comparative results for pathology images grading using different graphs.

Model	Glioma (LGG vs GBM)		Glioma (LGG II vs LGG III vs GBM)		BC-IDC (I vs II vs III)	
	Accuracy	**F1-score**	**Accuracy**	**F1-score**	**Accuracy**	**F1-score**
G_{Cell} **GNN**	84.91 ± 0.977	80.92 ± 0.026	82.09 ± 0.918	78.44 ± 0.032	79.82 ± 1.441	77.64 ± 0.039
G_C **GNN**	85.39 ± 0.992	82.38 ± 0.022	83.66 ± 0.985	79.18 ± 0.029	75.59 ± 1.671	74.17 ± 0.036
G_S **GNN**	78.52 ± 1.281	74.91 ± 0.027	75.90 ± 1.663	73.33 ± 0.029	78.62 ± 1.980	76.36 ± 0.018
$G_C + G_S$ **GNN**	88.85 ± 0.871	85.74 ± 0.019	86.61 ± 0.670	82.64 ± 0.013	79.02 ± 1.127	76.59 ± 0.018
G_H **GNN**	68.83 ± 1.652	63.01 ± 0.031	64.83 ± 1.911	61.01 ± 0.039	67.83 ± 1.975	66.01 ± 0.051
$G_C + G_H$ **GNN**	85.92 ± 0.987	82.81 ± 0.023	84.02 ± 1.317	79.81 ± 0.028	76.71 ± 1.574	74.62 ± 0.036
$G_S +$ G_H **GNN**	80.20 ± 1.145	75.63 ± 0.027	76.20 ± 1.455	73.83 ± 0.047	79.82 ± 1.971	75.43 ± 0.068
$AMGF - GNN$ **(ours)**	**89.68 ± 0.391**	**86.51 ± 0.012**	**87.52 ± 0.347**	**83.48 ± 0.014**	**81.82 ± 1.029**	**78.19 ± 0.013**

Table 5.2 Comparative results for pathology images grading using different state-of-the art models.

Model	Glioma (LGG vs GBM)		Glioma (LGG II vs LGG III vs GBM)		BC-IDC (I vs II vs III)	
	Accuracy	**F1-score**	**Accuracy**	**F1-score**	**Accuracy**	**F1-score**
ResNet-50 [55]	84.26 ± 1.362	81.9 ± 0.010	83.85 ± 0.922	80.7 ± 0.021	73.65 ± 1.790	69.9 ± 0.024
ResNet-34 [55]	83.61 ± 0.738	81.4 ± 0.036	82.70 ± 0.981	80.2 ± 0.014	72.82 ± 0.974	69.4 ± 0.029
CGN-Net [99]	84.40 ± 1.980	80.1 ± 0.026	82.28 ± 1.459	77.0 ± 0.033	75.55 ± 1.961	70.2 ± 0.012
Patch-GNN [88]	84.93 ± 1.804	81.3 ± 0.017	83.31 ± 1.225	79.9 ± 0.027	79.17 ± 0.872	74.9 ± 0.015
CGAT [299]	85.33 ± 1.580	82.1 ± 0.022	84.70 ± 1.499	79.6 ± 0.016	75.37 ± 1.887	70.6 ± 0.019
Slide Graph [363]	84.59 ± 1.273	82.4 ± 0.029	82.56 ± 0.871	79.8 ± 0.021	75.75 ± 1.987	70.5 ± 0.016
PathoGraph [360]	85.78 ± 1.182	82.6 ± 0.021	83.45 ± 1.282	80.7 ± 0.018	77.31 ± 1.491	72.0 ± 0.020
Hact-net [121]	87.42 ±0.663	84.2 ± 0.010	86.78 ± 1.201	83.5 ± 0.018	78.60 ± 1.560	74.8 ± 0.019
AMGF-GNN (ours)	**89.68 ± 0.391**	**86.51 ± 0.012**	**87.52 ± 0.347**	**83.48 ± 0.014**	**81.82 ± 1.029**	**78.19 ± 0.013**

Table 5.3 Grading performance on different attention mechanism.

Method		Glioma (LGG vs GBM)		Glioma (LGG II vs LGG III vs GBM)	
Single graph attention	**Multi graph attention**	**Accuracy**	**F1-score**	**Accuracy**	**F1-score**
×	×	84.19 ± 1.095	82.15 ± 0.026	82.92 ± 1.221	81.95 ± 0.026
✓	×	85.45 ± 0.981	82.71 ± 0.018	84.20 ± 0.981	81.53 ± 0.023
✓	✓	**89.68 ± 0.391**	**86.51 ± 0.012**	**87.52 ± 0.347**	**83.48 ± 0.014**

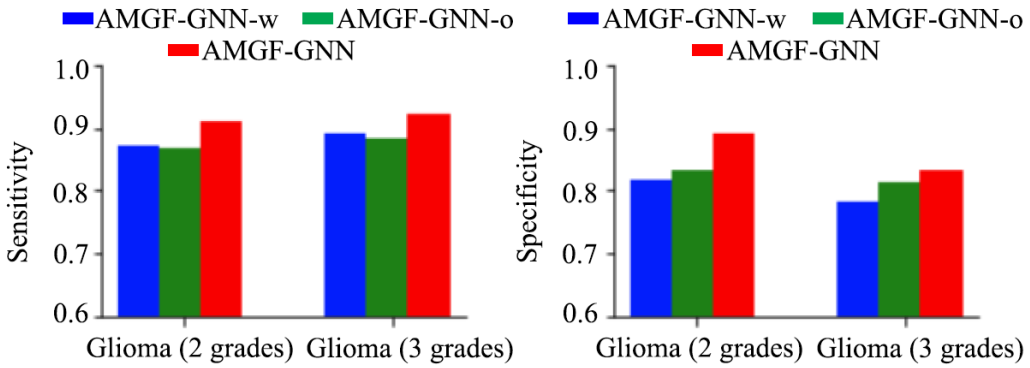

Figure 5.4 Sensitivity and specificity of three different configurations of AMGF-GNN model; without any of the attention mechanisms AMGF-GNN-w, with only single-attention AMGF-GNN-o and our AMGF-GNN with single and multi-graph attention mechanisms.

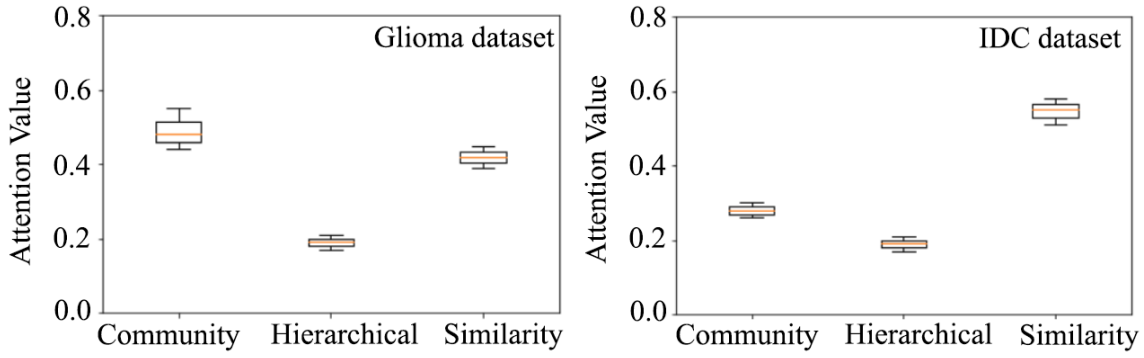

Figure 5.5 Analysis of attention distribution.

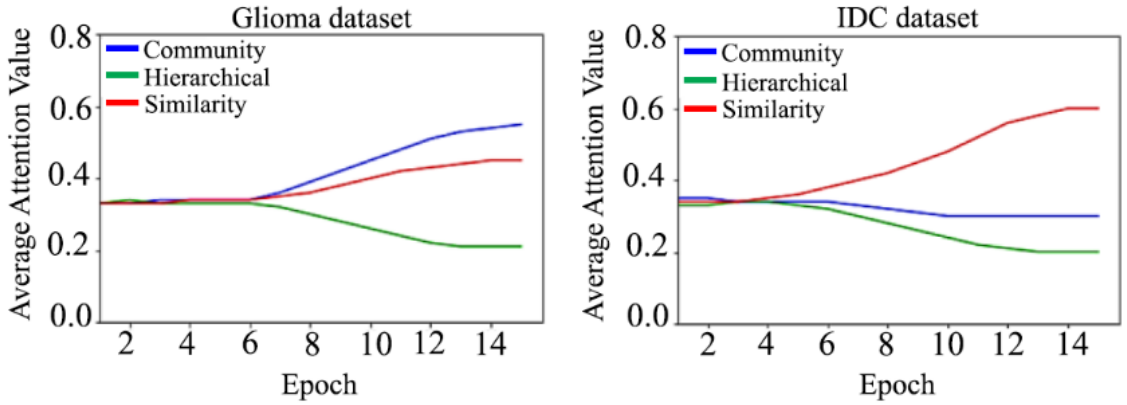

Figure 5.6. Average attention value over epochs.

In order to explore the significance of the attention values acquired by our proposed model, we carry out analyses on attention distribution and the patterns of attention learning over time. Taking the glioma dataset as example, as shown in Figure 5.5 the attention values assigned to certain embeddings in community space are higher than those in similarity space, indicating a greater focus on spatial based information. To confirm this, we can see that the performance of $G_C\ GNN$ is better than $G_S\ GNN$ in Table 5.1.

However, for the IDC dataset, the pattern is different. The attention values assigned to embeddings in the similarity space are higher than those in the community space, suggesting that relational information between nodes plays a more significant role for this dataset. This is further supported by the performance in Table 5.1, where the $G_S\ GNN$ achieves better results than the $G_C\ GNN$. These observations demonstrate the flexibility of our proposed AMGF-GNN in adaptively adjusting to the specific characteristics of different datasets by assigning higher attention values to the most relevant information.

We also examine how attention values evolve throughout training, which can provide insights into how AMGF- GNN learns to focus on different types of information, as shown in Figure 5.6, with the x-axis representing epochs and y-axis showing average attention values. At the beginning, the average attention values for cell community, cell similarity, and hierarchical graphs are almost the same, indicating that the model initially considers all types of information equally important. As training progresses, the attention values diverge, with attention to similarity space gradually decreasing while attention to community space keeps increasing. This observation aligns with the findings in Table 5.1 and Figure 5.5, which indicate that the cell community graph outperforms the cell similarity graph and that information in community space holds greater significance than in similarity space. AMGF-GNN demonstrates the ability to adapt its focus during training, suggests that it can effectively shift its attention based on the performance and feedback during training.

5.4.4 Visualization of Attention Weights

We further used some Glioma TCGA WSIs test set to assess how well our AMGF-GNN can

pinpoint regions of high attention. In the attention heatmaps shown in the second column of Figure 5.7, the attention weights are scaled between 0 and 1 (ranging from blue to red) within each attention map. Therefore, the red areas in the heatmaps indicate the nuclei that contribute most significantly to grading the input WSI. Notably the regions identified by our model tend to encompass broader areas in higher grade WSIs.

5.4.5 Analysis of the K Parameter

To investigate the impact of the top k neighbourhoods of KNN algorithm on the construction the corresponding raw graphs, we conducted a series of experiments with K values ranging from 2 to 9 (shown in Figure 5.8). As shown, the accuracy of AMGF-GNN improves initially before declining. It's possible that as the graph gets denser, the feature becomes easier to smooth, and greater k potentially adds more noisy edges. In our trials, we set K = 2 for cell similarity graph, and K = 7 for cell community graph in the glioma dataset and K = 2 for cell similarity graph, and K = 5 for cell community graph in the IDC dataset.

5.4.6 Analysis of the d Parameter

To explore how the minimum distance parameter d_{min} influences the clustering process which we need for constructing the cell community graph, we conducted a series of experiments with d_{min} values spanning {250, 500, 750,1000}, as illustrated in Figure 5.9. Our findings reveal that the accuracy of AMGF-GNN initially rises and then begins to decline. Consequently, in our trials, we opted to set d_{min} at 500 for constructing cell community and hierarchical graphs in glioma dataset and d_{min} at 250 in IDC dataset.

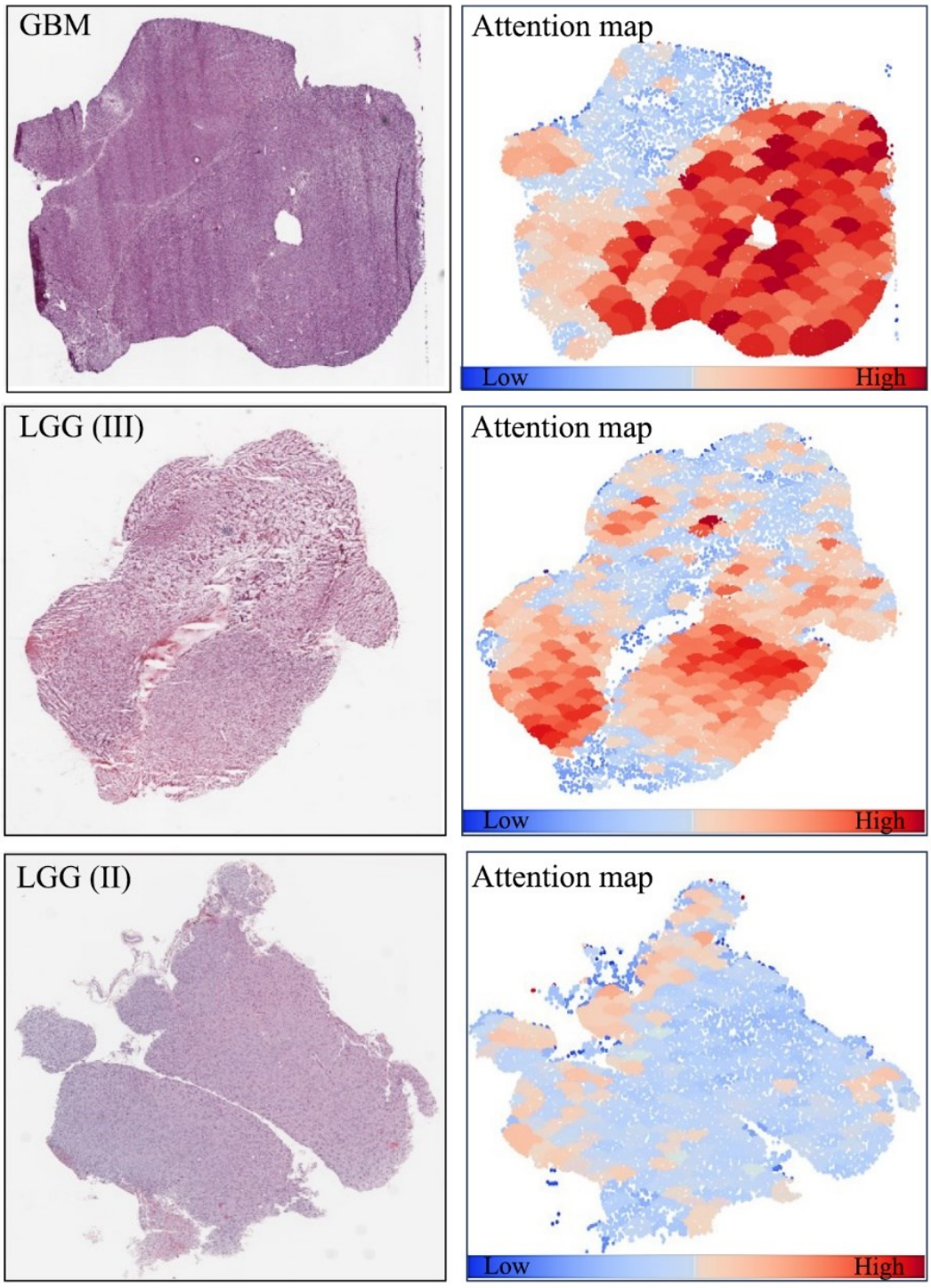

Figure 5.7 Visualization of attention maps using TCGA glioma testing set. Each input WSI in the left, and its corresponding attention map in the right showing high attention patches.

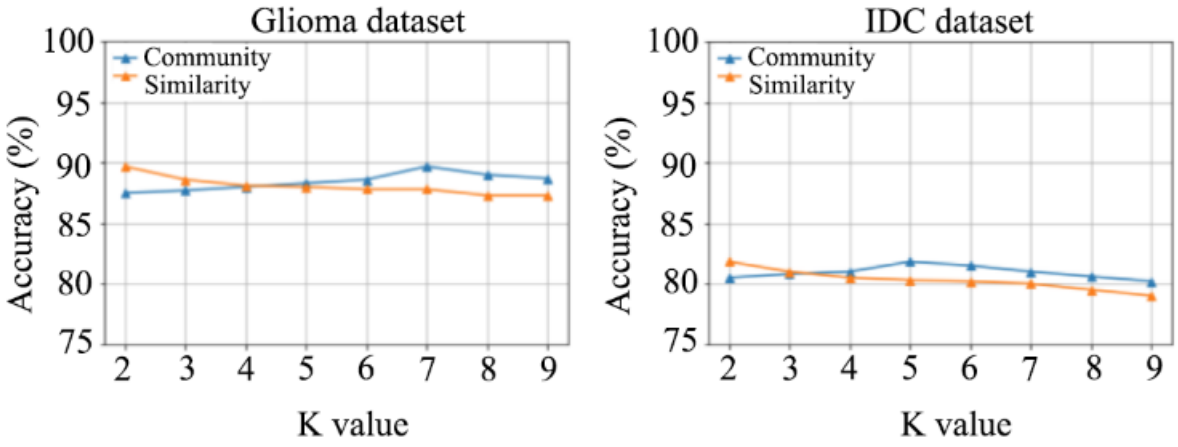

Figure 5.8 Analysis of K values of KNN algorithm for construction the graphs.

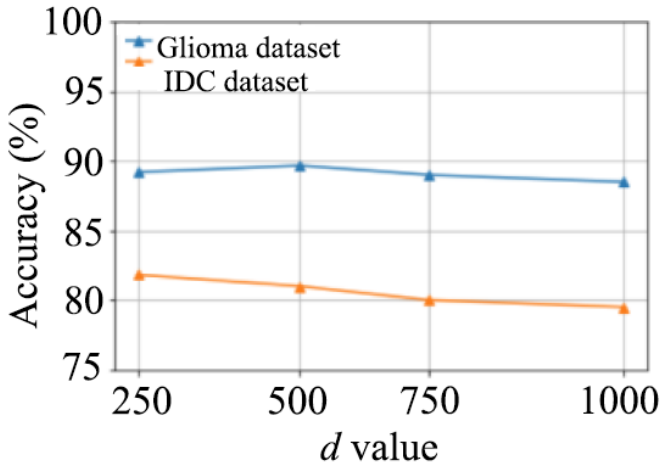

Figure 5.9 Analysis of d_{min} values for constructing the cell community graph.

5.4.7 Analysis the Impact of Node Feature Initialization

The performance of AMGF-GNN heavily depends on the node features. Therefore, we explore the impact of initializing node features to understand how different types of node features impact the performance of AMGF-GNN in our classification tasks. We divide the features into two distinct schemes for testing purposes:

1) CNN-based Features: This scheme will involve initializing node features using the CNN-based features extracted from ResNet34 pre-trained on the ImageNet dataset.

2) Morphological Features: This scheme will include initializing node features using morphological features extracted from GLCMs and average intensity (average RGB value).

The experimental results showed in Figure 5.10 demonstrate the impact of incorporating morphological features with CNN features which significantly improves classification performance. Specifically, graphs with CNN-based morphological features exhibit better performance compared to graphs with only morphological features, suggests that CNN-based features capture richer and more informative information compared to other features. However, incorporating CNN-based features with other morphological features can enhance the discriminative ability of the model and improve classification performance.

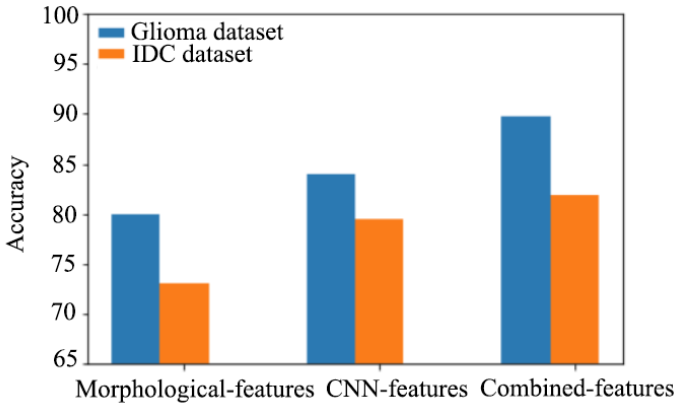

Figure 5.10 Impact of node feature initialization

5.5 Discussion

GNNs have shown great promise in analyzing pathology images. Their key strength lies in their capacity to utilize the features of distinct histological elements like glands and cells represented as nodes as well as their connections represented as edges. This approach aligns with the existing methodology in pathology where pathologists assess different histological objects along with their placements and distributions to arrive at conclusive diagnostic conclusions.

For cancer diagnosis, some approaches rely on hand-crafted features, while others utilize CNN-based methods. Some have suggested using both traditional techniques and CNN-based approaches to extract hand-crafted and CNN-driven features. AMGF-GNN represents a hybrid approach for cancer grading, utilizing both hand-crafted and CNN-based features to characterize tissue images.

In the field of analyzing pathology images, tissue slides are often examined and scrutinized at different scales, resolutions and levels of organization. However, previous GNNs have predominantly consider information from one graph and ignore the crucial topological insights from other graphs, in which the graph's quality has a significant impact on the model performance [88-91,96,99,100,104,111,112]. Various methods have been proposed to tackle the challenges associated with analyzing complex tissue microenvironments combining multilevel structural details, to mimic the diagnostic procedure used by pathologists [94,114,121,301,328-330]. While these techniques take into account combining multiple graphs of the original data to enrich the model's representation; a significant challenge lies in determining the specific contribution of each graph and assigning distinguishable weights to neighboring nodes. AMGF-GNN network is introduced for cancer grading using gigapixel WSIs. AMGF-GNN employs an effective multi-graph with an adaptive attention mechanism framework to aggregate different information, increasing the robustness of the node features and the overall graph quality, which minimizes reliance on the quality of single graph, and to produce stronger WSI-level representation.

AMGF-GNN used multiple graphs as input including one based on community information, one based on feature similarity, and other based on both at the higher level, so distinct embeddings were extracted across different spaces and levels by using three parallel GNN. Additionally, the single attention mechanism employed in our method allows for the acquisition of adaptive importance weights for the embeddings enabling the model to distinguish the node's influence from each single graph. We further utilize an Adaptive Attentive Multi-Graph Fusion module that innovatively fuses multiple graphs into a unified dense vector space. By applying a nonlinear transformation and a shared attention vector, it adaptively adjusts the relevance of

each graph's information. Specifically, the attention vector is shared across all graph embeddings, allowing the model to weigh the contributions of each graph equally while adapting to their individual relevance. It learns to balance the contributions of community and feature-based information, addressing the uncertainty of their relative importance for the final grading task. Additionally, our innovative dual-level loss optimization strategy incorporates both intra-graph and inter-graph similarity measures, promoting consistency and robustness in the learned embeddings, and significantly enhancing the model's performance. Experimental results on both TCGA-glioma and the invasive ductal carcinoma (IDC) datasets indicate the effectiveness of our proposed AMGF-GNN.

5.6 Summary

We propose an Adaptive Multi-Graph Fusion-based Attentive Graph Neural Network for histopathology image analysis. Different from comparative deep graph techniques mainly focused on a single graph, our proposed approach reduces reliance on the quality of a single graph by constructing and processing three distinct graphs based on cell's community distribution, feature similarity, and a combination of both, ensuring a comprehensive and robust representation of the tissue microenvironment. Our Adaptive Attentive Multi-Graph Fusion module adaptively fuses the embeddings from all three paths into a unified dense vector space. This adaptive fusion learns to balance the contributions of community and feature-based information, addressing the uncertainty of their relative importance, ensuring that the most relevant features from each graph are prioritized, and enhancing the overall classification performance. Furthermore, a dual-level loss optimization strategy is proposed which incorporates both intra-graph and inter-graph similarity measures, ensures consistency and robustness in the learned embeddings, and significantly enhancing the model's performance. Based on experimental results on two datasets, our proposed framework outperforms other methods and confirm that fusion of multiple graphs is a viable option for classification tasks.

While our approach integrates multiple graphs for improved histopathology analysis, challenges remain in graph quality, scalability with larger datasets, and generalization to other domains. Further work is needed to enhance graph construction, scalability, and interpretability in clinical

settings. Future directions include extending the model for multi-modal data and exploring integration with other architectures for potential clinical deployment.

CHAPTER 6

Conclusion

6.1 Conclusions

This thesis presents a comprehensive multi-scale framework for histopathological image analysis, addressing key challenges in cellular and tissue-level examination. By leveraging advanced deep learning techniques, GNNs, and multi-graph fusion models, we have developed innovative methods that significantly enhance the accuracy, scalability, and interpretability of AI-driven solutions in pathology. Our framework tackles the following core challenges: (1) robust cell detection and segmentation in complex tissue environments, (2) capturing spatial and functional relationships between cells for survival prediction, and (3) integrating multi-graph representations for accurate cancer grading.

At the cellular level, we addressed the challenges of detecting and segmenting cells in histopathological images, particularly those complicated by variations in staining, tissue preparation, and dense cell structures. To overcome these obstacles, we extended the applicability of PathoFusion's BCNN, which utilizes both fine and coarse image tiles to extract essential morphological and spatial features. This dual-scale approach improves the accuracy of cell detection, enabling the identification, segmentation, and counting of immunohistochemically labeled cells, as demonstrated in our case study of CD276 immunoreactive cells in glioblastoma. Furthermore, we introduced a novel fusion technique using Discrete Wavelet Transform (DWT) for cross-modality analyses, allowing for the correlation between immunohistochemical signals and histological features. This approach

provides a robust tool for cellular-level analysis and visualization across diverse tissue samples.

At the tissue level, our framework expands upon the BCNN by incorporating a graph-based attention model, termed PathoGraph, to capture the intricate spatial relationships between cells within the tissue microenvironment. This study tackled the challenge of constructing memory-efficient graphs from gigapixel WSIs while maintaining spatial fidelity. PathoGraph employs attention mechanisms to adaptively adjust weights based on feature similarity and spatial proximity, enhancing its ability to capture heterogeneity in tissue structures. The proposed model was validated for survival prediction using two datasets: our own CD276 dataset [351] and the glioblastoma (GBM) dataset from TCGA [352]. The validation demonstrated that attention-based graph convolutional networks (abGCN) can effectively interpret and classify WSIs, providing insights into patient prognosis and uncovering regions of interest for further clinical examination.

The third study in this thesis addressed the challenge of integrating multiple graph representations to enhance cancer grading accuracy. Recognizing the complexity of the tumor microenvironment, we extended our graph-based approach by proposing an adaptive multi-graph fusion-based graph neural network (AMGF-GNN). This model simultaneously learns unique embeddings from multiple graphs, one based on the cell community structure, another on feature similarity, and a hybrid graph that combines both perspectives, capturing both local and global tissue structures. The innovative fusion mechanism in AMGF-GNN combines these graphs into a unified dense representation, improving the robustness and reducing reliance on the quality of individual graphs. Additionally, this fusion process ensures that critical information from each graph is preserved, balancing the contribution of structural and feature-based information. Furthermore, our dual-level loss optimization strategy incorporates both intra-graph and inter-graph similarity measures, promoting consistency and robustness in the learned embeddings, leading to improved performance in cancer grading. This framework was rigorously tested on two public datasets, including TCGA glioblastoma dataset [352] and breast cancer dataset [362], outperforming state-of-the-art methods in both diagnostic and prognostic workflows.

6.2 Future outlook

Future research on histopathological image analysis should aim to refine existing methods and explore new strategies to enhance model accuracy and interpretability, which are key for clinical adoption. By focusing on these areas, future studies can create more robust and transparent models that support clinicians in making accurate and reliable decisions, ultimately improving treatment and patient care. The following areas present crucial opportunities for further exploration:

Scaling for Real-Time Clinical Applications: Streamlining computational processes is crucial for immediate clinical applications with an emphasis on minimizing memory consumption and processing speed without compromising model effectiveness.

Integrating Interpretability Mechanisms: Future research should focus on incorporating interpretability methods to enhance the transparency and clinical relevance of model predictions. Techniques such as saliency maps, Class Activation Mapping (CAM), SHAP (Shapley Additive Explanations), and LIME (Local Interpretable Model-agnostic Explanations) can help pinpoint key features or regions in the decision-making process, providing clinicians with insights into how the model arrived at its predictions.

Generalization Across Diverse Pathologies: Exploring the proposed techniques across different types of cancer and other illnesses could help assess how well the models hold up in diverse medical situations, enhancing their usefulness in various healthcare settings.

Automated Annotation and Active Learning: Leveraging semi-supervised learning and automated annotation could address the challenge of limited annotated data, particularly for rare diseases, improving model performance.

Improved Modeling of Tumor Microenvironments: Exploring advanced graph models that incorporate immune cell interactions, vascularization, and other biological factors could further enhance the accuracy of tumor classification.

Bibliography

1. Alzoubi, I.; Bao, G.; Zhang, R.; Loh, C.; Zheng, Y.; Cherepanoff, S.; Gracie, G.; Lee, M.; Kuligowski, M.; Alexander, K.L.; et al. An Open-Source AI Framework for the Analysis of Single Cells in Whole-Slide Images with a Note on CD276 in Glioblastoma. *Cancers (Basel)* **2022**, *14*, doi:10.3390/cancers14143441.
2. World Health Organization (WHO). Global Cancer Burden Growing, Amidst Mounting Need for Services. *WHO*, 2024. Available online: https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services (accessed on January 2, 2024).
3. GLOBOCAN. Brain and Central Nervous System Cancer Statistics. *IARC*, 2024. Available online: https://gco.iarc.fr/en (accessed on January 2, 2024).
4. American Cancer Society (ACS). Brain and Spinal Cord Tumors in Adults. *ACS*, 2024. Available online: https://www.cancer.org/ (accessed on January 2, 2024).
5. Thakkar, J.P.; Dolecek, T.A.; Horbinski, C.; Ostrom, Q.T.; Lightner, D.D.; Barnholtz-Sloan, J.S.; Villano, J.L. Epidemiologic and Molecular Prognostic Review of Glioblastoma. *Cancer Epidemiology, Biomarkers & Prevention* **2014**, *23*, 1985-1996.
6. Ostrom, Q.T.; Gittleman, H.; Farah, P.; Ondracek, A.; Chen, Y.; Wolinsky, Y.; Stroup, N.E.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2006-2010. *Neuro-Oncology* **2013**, *15*, ii1-ii56.
7. Davis, M.E. Glioblastoma: Overview of Disease and Treatment. *Clinical Journal of Oncology Nursing* **2016**, *20*, S2.
8. Glioblastoma Multiforme Explained. *Boston Neurosurgery*, 2024. Available online: https://www.boston-neurosurgery.com/glioblastoma-multiforme-explained/ (accessed on January 2, 2024).
9. Neuber, A.; Nuttall, T. Introduction to Histopathology. *Springer*, **2017**.
10. Litjens, G.; Sánchez, C.I.; Timofeeva, N.; Hermsen, M.; Nagtegaal, I.; Kovacs, I.; Hulsbergen-Van De Kaa, C.; Bult, P.; Van Ginneken, B.; Van Der Laak, J. Deep Learning

as a Tool for Increased Accuracy and Efficiency of Histopathological Diagnosis. *Scientific Reports* **2016**, *6*, 26286.

11. Niazi, M.K.K.; Parwani, A.V.; Gurcan, M.N. Digital Pathology and Artificial Intelligence. *The Lancet Oncology* **2019**, *20*, e253-e261.
12. Bera, K.; Schalper, K.A.; Rimm, D.L.; Velcheti, V.; Madabhushi, A. Artificial Intelligence in Digital Pathology—New Tools for Diagnosis and Precision Oncology. *Nature Reviews Clinical Oncology* **2019**, *16*, 703-715.
13. Jensen, M.P.; Qiang, Z.; Khan, D.Z.; Stoyanov, D.; Baldeweg, S.E.; Jaunmuktane, Z.; Brandner, S.; Marcus, H.J. Artificial Intelligence in Histopathological Image Analysis of Central Nervous System Tumours: A Systematic Review. *Neuropathology and Applied Neurobiology* **2024**, *50*, e12981.
14. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems* **2021**, *33*, 6999-7019.
15. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Philip, S.Y. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* **2020**, *32*, 4-24.
16. De Matos, J.; Ataky, S.T.M.; de Souza Britto Jr, A.; Soares de Oliveira, L.E.; Lameiras Koerich, A. Machine Learning Methods for Histopathological Image Analysis: A Review. *Electronics* **2021**, *10*, 562.
17. Bao, G.; Graeber, M.B.; Wang, X. A Bifocal Classification and Fusion Network for Multimodal Image Analysis in Histopathology. *In Proceedings of the 2020 16th International Conference on Control, Automation, Robotics and Vision* (ICARCV), 2020; pp. 466-471.
18. Saukko, P.; Pollak, S. Brief Historical Introduction to Histopathology. *Forensic Pathology* **2016**, 89.
19. Chan, J.K. The Wonderful Colors of the Hematoxylin–Eosin Stain in Diagnostic Surgical Pathology. *International Journal of Surgical Pathology* **2014**, *22*, 12-32.
20. Wittekind, D. Traditional Staining for Routine Diagnostic Pathology Including the Role of Tannic Acid. 1. Value and Limitations of the Hematoxylin-Eosin Stain. *Biotech Histochem* **2003**, *78*, 261-270, doi:10.1080/10520290310001633725.

21. Jaffer, S.; Bleiweiss, I.J. Beyond Hematoxylin and Eosin—The Role of Immunohistochemistry in Surgical Pathology. *Cancer Investigation* **2004**, *22*, 445-465.
22. McCulloch, W.S.; Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics* **1943**, *5*, 115-133.
23. Turing, A.; Epstein, R.; Roberts, G.; Beber, G. Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer. *Netherlands: Springer* **2009**, 23-65.
24. Wiederhold, G.; McCarthy, J. Arthur Samuel: Pioneer in Machine Learning. *IBM Journal of Research and Development* **1992**, *36*, 329-331.
25. Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review* **1958**, *65*, 386.
26. Cover, T.; Hart, P. Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory* **1967**, *13*, 21-27.
27. Fukushima, K. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biological Cybernetics* **1980**, *36*, 193-202.
28. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* **1989**, *1*, 541-551.
29. ARD, V.; GR, J. Five Ways Deep Learning Has Transformed Image Analysis. *Nature* **2022**, *609*.
30. Hinton, G.E.; Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313*, 504-507.
31. Krizhevsky, A.; Sutskever, I. Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM* 2017, 60. *DOI: https://doi.org/10.1145/3065386*, 84-90.
32. Alzoubi, I.; Bao, G.; Zheng, Y.; Wang, X.; Graeber, M. Artificial Intelligence Techniques for Neuropathological Diagnostics and Research. *Neuropathology* **2022**, *43*, doi:10.1111/neup.12880.

33. Song, J.-W.; Lee, J.-H.; Choi, J.-H.; Chun, S.-J. Automatic Differential Diagnosis of Pancreatic Serous and Mucinous Cystadenomas Based on Morphological Features. *Computers in Biology and Medicine* **2013**, *43*, 1-15.
34. Gorelick, L.; Veksler, O.; Gaed, M.; Gómez, J.A.; Moussa, M.; Bauman, G.; Fenster, A.; Ward, A.D. Prostate Histopathology: Learning Tissue Component Histograms for Cancer Detection and Classification. *IEEE Transactions on Medical Imaging* **2013**, *32*, 1804-1818.
35. Kwak, J.T.; Hewitt, S.M. Multiview Boosting Digital Pathology Analysis of Prostate Cancer. *Computer Methods and Programs in Biomedicine* **2017**, *142*, 91-99.
36. De, S.; Stanley, R.J.; Lu, C.; Long, R.; Antani, S.; Thoma, G.; Zuna, R. A Fusion-Based Approach for Uterine Cervical Cancer Histology Image Classification. *Computerized Medical Imaging and Graphics* **2013**, *37*, 475-487.
37. Kandemir, M.; Feuchtinger, A.; Walch, A.; Hamprecht, F.A. Digital Pathology: Multiple Instance Learning Can Detect Barrett's Cancer. In *Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, 2014; pp. 1348-1351.
38. Pang, W.; Jiang, H.; Li, S. Sparse Contribution Feature Selection and Classifiers Optimized by Concave-Convex Variation for HCC Image Recognition. *BioMed Research International* **2017**, *2017*, 9718386.
39. Kruk, M.; Kurek, J.; Osowski, S.; Koktysz, R.; Swiderski, B.; Markiewicz, T. Ensemble of Classifiers and Wavelet Transformation for Improved Recognition of Fuhrman Grading in Clear-Cell Renal Carcinoma. *Biocybernetics and Biomedical Engineering* **2017**, *37*, 357-364.
40. Basavanhally, A.; Ganesan, S.; Feldman, M.; Shih, N.; Mies, C.; Tomaszewski, J.; Madabhushi, A. Multi-Field-of-View Framework for Distinguishing Tumor Grade in ER+ Breast Cancer from Entire Histopathology Slides. *IEEE Transactions on Biomedical Engineering* **2013**, *60*, 2089-2099.
41. Tashk, A.; Helfroush, M.S.; Danyali, H.; Akbarzadeh-Jahromi, M. Automatic Detection of Breast Cancer Mitotic Cells Based on the Combination of Textural, Statistical and Innovative Mathematical Features. *Applied Mathematical Modelling* **2015**, *39*, 6165-6182.

42. Maroof, N.; Khan, A.; Qureshi, S.A.; ul Rehman, A.; Khalil, R.K.; Shim, S.-O. Mitosis Detection in Breast Cancer Histopathology Images Using Hybrid Feature Space. *Photodiagnosis and Photodynamic Therapy* **2020**, *31*, 101885.
43. Mahmood, T.; Arsalan, M.; Owais, M.; Lee, M.B.; Park, K.R. Artificial Intelligence-Based Mitosis Detection in Breast Cancer Histopathology Images Using Faster R-CNN and Deep CNNs. *Journal of Clinical Medicine* **2020**, *9*, 749.
44. Barker, J.; Hoogi, A.; Depeursinge, A.; Rubin, D.L. Automated Classification of Brain Tumor Type in Whole-Slide Digital Pathology Images Using Local Representative Tiles. *Medical Image Analysis* **2016**, *30*, 60-71.
45. Wang, P.; Wang, J.; Li, Y.; Li, P.; Li, L.; Jiang, M. Automatic Classification of Breast Cancer Histopathological Images Based on Deep Feature Fusion and Enhanced Routing. *Biomedical Signal Processing and Control* **2021**, *65*, 102341.
46. Atwan, J.; Almansour, N.; Hashem Ryalat, M.; Sahran, S.; Aldabbas, H.; Albashish, D. Ensemble of Deep Features for Breast Cancer Histopathological Image Classification. *The Computer Journal* **2024**, bxad127.
47. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience* **2018**, *2018*, 7068349.
48. Depeursinge, A.; Al-Kadi, O.S.; Mitchell, J.R. *Biomedical texture analysis: fundamentals, tools and challenges*; Academic Press: 2017.
49. Van der Laak, J.; Litjens, G.; Ciompi, F. Deep Learning in Histopathology: The Path to the Clinic. *Nature Medicine* **2021**, *27*, 775-784.
50. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436-444.
51. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-Propagating Errors. *Nature* **1986**, *323*, 533-536.
52. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE* **1998**, *86*, 2278-2324.
53. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint* arXiv:*1409.1556* **2014**.

54. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015; pp. 1-9.
55. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016; pp. 770-778.
56. Sharma, H.; Zerbe, N.; Klempert, I.; Hellwich, O.; Hufnagl, P. Deep Convolutional Neural Networks for Automatic Classification of Gastric Carcinoma Using Whole Slide Images in Digital Histopathology. *Computerized Medical Imaging and Graphics* **2017**, *61*, 2-13.
57. Xu, Y.; Jia, Z.; Wang, L.-B.; Ai, Y.; Zhang, F.; Lai, M.; Chang, E.I.-C. Large Scale Tissue Histopathology Image Classification, Segmentation, and Visualization via Deep Convolutional Activation Features. *BMC Bioinformatics* **2017**, *18*, 1-17.
58. Cortes, C.; Vapnik, V. Support-Vector Networks. *Machine Learning* **1995**, *20*, 273-297.
59. Wang, D.; Khosla, A.; Gargeya, R.; Irshad, H.; Beck, A.H. Deep Learning for Identifying Metastatic Breast Cancer. *arXiv preprint* arXiv*:1606.05718* **2016**.
60. Jamaluddin, M.F.; Fauzi, M.F.A.; Abas, F.S. Tumor Detection and Whole Slide Classification of H&E Lymph Node Images Using Convolutional Neural Network. In *Proceedings of the 2017 IEEE International Conference on Signal and Image Processing Applications* (ICSIPA), 2017; pp. 90-95.
61. Sirinukunwattana, K.; Raza, S.E.A.; Tsang, Y.W.; Snead, D.R.J.; Cree, I.A.; Rajpoot, N.M. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. *IEEE Transactions on Medical Imaging* **2016**, *35*, 1196-1206, doi:10.1109/TMI.2016.2525803.
62. Hou, L.; Samaras, D.; Kurc, T.M.; Gao, Y.; Davis, J.E.; Saltz, J.H. Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016; pp. 2424-2433.
63. Araújo, T.; Aresta, G.; Castro, E.; Rouco, J.; Aguiar, P.; Eloy, C.; Polónia, A.; Campilho, A. Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. *PloS One* **2017**, *12*, e0177544.

64. Bejnordi, B.E.; Zuidhof, G.; Balkenhol, M.; Hermsen, M.; Bult, P.; van Ginneken, B.; Karssemeijer, N.; Litjens, G.; van der Laak, J. Context-Aware Stacked Convolutional Neural Networks for Classification of Breast Carcinomas in Whole-Slide Histopathology Images. *Journal of Medical Imaging* **2017**, *4*, 044504-044504.
65. Korbar, B.; Olofson, A.M.; Miraflor, A.P.; Nicka, C.M.; Suriawinata, M.A.; Torresani, L.; Suriawinata, A.A.; Hassanpour, S. Deep Learning for Classification of Colorectal Polyps on Whole-Slide Images. *Journal of Pathology Informatics* **2017**, *8*, 30.
66. Das, K.; Karri, S.P.K.; Roy, A.G.; Chatterjee, J.; Sheet, D. Classifying Histopathology Whole-Slides Using Fusion of Decisions from Deep Convolutional Network on a Collection of Random Multi-Views at Multi-Magnification. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 2017; pp. 1024-1027.
67. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015; pp. 3431-3440.
68. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015; pp. 234-241.
69. Bándi, P.; Van De Loo, R.; Intezar, M.; Geijs, D.; Ciompi, F.; Van Ginneken, B.; Van Der Laak, J.; Litjens, G. Comparison of Different Methods for Tissue Segmentation in Histopathological Whole-Slide Images. In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 2017; pp. 591-595.
70. Cui, Y.; Zhang, G.; Liu, Z.; Xiong, Z.; Hu, J. A Deep Learning Algorithm for One-Step Contour Aware Nuclei Segmentation of Histopathology Images. *Medical & Biological Engineering & Computing* **2019**, *57*, 2027-2043, doi:10.1007/s11517-019-02008-8.
71. Gecer, B.; Aksoy, S.; Mercan, E.; Shapiro, L.G.; Weaver, D.L.; Elmore, J.G. Detection and Classification of Cancer in Whole Slide Breast Histopathology Images Using Deep Convolutional Networks. *Pattern Recognition* **2018**, *84*, 345-356.
72. Wang, X.; Chen, H.; Gan, C.; Lin, H.; Dou, Q.; Tsougenis, E.; Huang, Q.; Cai, M.; Heng, P.A. Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. *IEEE Transactions on Cybernetics* **2020**, *50*, 3950-3962, doi:10.1109/TCYB.2019.2935141.

73. Sheikhzadeh, F.; Ward, R.K.; van Niekerk, D.; Guillaud, M. Automatic Labeling of Molecular Biomarkers of Immunohistochemistry Images Using Fully Convolutional Networks. *PloS One* **2018**, *13*, e0190783.
74. Lin, H.; Chen, H.; Dou, Q.; Wang, L.; Qin, J.; Heng, P.-A. Scannet: A Fast and Dense Scanning Framework for Metastatic Breast Cancer Detection from Whole-Slide Image. In *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018; pp. 539-546.
75. Lin, H.; Chen, H.; Graham, S.; Dou, Q.; Rajpoot, N.; Heng, P.-A. Fast Scannet: Fast and Dense Analysis of Multi-Gigapixel Whole-Slide Images for Cancer Metastasis Detection. *IEEE Transactions on Medical Imaging* **2019**, *38*, 1948-1958.
76. Maksoud, S.; Wiliem, A.; Zhao, K.; Zhang, T.; Wu, L.; Lovell, B. Coral8: Concurrent Object Regression for Area Localization in Medical Image Panels. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22*, 2019; pp. 432-441.
77. Ren, J.; Karagoz, K.; Gatza, M.; Foran, D.J.; Qi, X. Differentiation Among Prostate Cancer Patients with Gleason Score of 7 Using Histopathology Whole-Slide Image and Genomic Data. In *Proceedings of the Spie--The International Society for Optical Engineering*, 2018.
78. Budak, Ü.; Cömert, Z.; Rashid, Z.N.; Şengür, A.; Çıbuk, M. Computer-Aided Diagnosis System Combining FCN and Bi-LSTM Model for Efficient Breast Cancer Detection from Histopathological Images. *Applied Soft Computing* **2019**, *85*, 105765.
79. Yan, R.; Ren, F.; Wang, Z.; Wang, L.; Zhang, T.; Liu, Y.; Rao, X.; Zheng, C.; Zhang, F. Breast Cancer Histopathological Image Classification Using a Hybrid Deep Neural Network. *Methods* **2020**, *173*, 52-60, doi:https://doi.org/10.1016/j.ymeth.2019.06.014.
80. BenTaieb, A.; Hamarneh, G. Adversarial Stain Transfer for Histopathology Image Analysis. *IEEE Transactions on Medical Imaging* **2017**, *37*, 792-802.
81. de Bel, T.; Hermsen, M.; Kers, J.; van der Laak, J.; Litjens, G. Stain-Transforming Cycle-Consistent Generative Adversarial Networks for Improved Segmentation of Renal Histopathology. **2019**.

82. Fakoor, R.; Ladhak, F.; Nazi, A.; Huber, M. Using Deep Learning to Enhance Cancer Diagnosis and Classification. In *Proceedings of the International Conference on Machine Learning*, 2013; pp. 3937-3949.
83. Su, H.; Xing, F.; Kong, X.; Xie, Y.; Zhang, S.; Yang, L. Robust Cell Detection and Segmentation in Histopathological Images Using Sparse Reconstruction and Stacked Denoising Autoencoders. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, Munich, Germany, October 5-9, 2015; Proceedings, Part III 18, 2015; pp. 383-390.
84. Xu, J.; Xiang, L.; Liu, Q.; Gilmore, H.; Wu, J.; Tang, J.; Madabhushi, A. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. *IEEE Transactions on Medical Imaging* **2015**, *35*, 119-130.
85. Bhardwaj, A.; Tiwari, A. Breast Cancer Diagnosis Using Genetically Optimized Neural Network Model. *Expert Systems with Applications* **2015**, *42*, 4611-4620.
86. Bhardwaj, A.; Tiwari, A.; Bhardwaj, H.; Bhardwaj, A. A Genetically Optimized Neural Network Model for Multi-Class Classification. *Expert Systems with Applications* **2016**, *60*, 211-221.
87. Sharma, H.; Zerbe, N.; Lohmann, S.; Kayser, K.; Hellwich, O.; Hufnagl, P. A Review of Graph-Based Methods for Image Analysis in Digital Histopathology. *Diagnostic Pathology* **2015**, *1*.
88. Aygüneş, B.; Aksoy, S.; Cinbiş, R.G.; Kösemehmetoğlu, K.; Önder, S.; Üner, A. Graph Convolutional Networks for Region of Interest Classification in Breast Histopathology. In *Proceedings of the Medical Imaging 2020: Digital Pathology*, 2020; pp. 134-141.
89. Anand, D.; Gadiya, S.; Sethi, A. Histographs: Graphs in Histopathology. In *Proceedings of the Medical Imaging 2020: Digital Pathology*, 2020; pp. 150-155.
90. Li, R.; Yao, J.; Zhu, X.; Li, Y.; Huang, J. Graph CNN for Survival Analysis on Whole Slide Pathological Images. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018; pp. 174-182.
91. Jaume, G.; Pati, P.; Bozorgtabar, B.; Foncubierta, A.; Anniciello, A.M.; Feroce, F.; Rau, T.; Thiran, J.-P.; Gabrani, M.; Goksel, O. Quantifying Explainers of Graph Neural Networks in Computational Pathology. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021; pp. 8106-8116.

92. Francis, K.; Palsson, B.O. Effective Intercellular Communication Distances Are Determined by the Relative Time Constants for Cyto/Chemokine Secretion and Diffusion. *Proceedings of the National Academy of Sciences* **1997**, *94*, 12258-12262.
93. Sharma, H.; Zerbe, N.; Heim, D.; Wienert, S.; Lohmann, S.; Hellwich, O.; Hufnagl, P. Cell Nuclei Attributed Relational Graphs for Efficient Representation and Classification of Gastric Cancer in Digital Histopathology. In *Proceedings of the Medical Imaging 2016: Digital Pathology*, 2016; pp. 238-256.
94. Javed, S.; Mahmood, A.; Fraz, M.M.; Koohbanani, N.A.; Benes, K.; Tsang, Y.-W.; Hewitt, K.; Epstein, D.; Snead, D.; Rajpoot, N. Cellular Community Detection for Tissue Phenotyping in Colorectal Cancer Histology Images. *Medical Image Analysis* **2020**, *63*, 101696.
95. Jaume, G.; Pati, P.; Foncubierta-Rodriguez, A.; Feroce, F.; Scognamiglio, G.; Anniciello, A.M.; Thiran, J.-P.; Goksel, O.; Gabrani, M. Towards Explainable Graph Representations in Digital Pathology. *arXiv Preprint arXiv:2007.00311* **2020**.
96. Sureka, M.; Patil, A.; Anand, D.; Sethi, A. Visualization for Histopathology Images Using Graph Convolutional Neural Networks. In *Proceedings of the 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2020; pp. 331-335.
97. Wang, J.; Chen, R.J.; Lu, M.Y.; Baras, A.; Mahmood, F. Weakly Supervised Prostate TMA Classification via Graph Convolutional Networks. In *Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020; pp. 239-243.
98. Studer, L.; Wallau, J.; Dawson, H.; Zlobec, I.; Fischer, A. Classification of Intestinal Gland Cell-Graphs Using Graph Neural Networks. In *Proceedings of the 2020 25th International Conference on Pattern Recognition* (ICPR), 2021; pp. 3636-3643.
99. Zhou, Y.; Graham, S.; Alemi Koohbanani, N.; Shaban, M.; Heng, P.-A.; Rajpoot, N. CGC-Net: Cell Graph Convolutional Network for Grading of Colorectal Cancer Histology Images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019; pp. 0-0.
100. Chen, R.J.; Lu, M.Y.; Wang, J.; Williamson, D.F.; Rodig, S.J.; Lindeman, N.I.; Mahmood, F. Pathomic Fusion: An Integrated Framework for Fusing Histopathology

and Genomic Features for Cancer Diagnosis and Prognosis. *IEEE Transactions on Medical Imaging* **2020**, *41*, 757-770.

101. Studer, L.; Toneyan, S.; Zlobec, I.; Dawson, H.; Fischer, A. Graph-Based Classification of Intestinal Glands in Colorectal Cancer Tissue Images. In *Proceedings of MICCAI 2019*, *Shenzhen, China* **2019**.
102. Li, R.; Yao, J.; Zhu, X.; Li, Y.; Huang, J. Graph CNN for Survival Analysis on Whole Slide Pathological Images. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, Cham, 2018; pp. 174-182.
103. Aygünes, B.; Aksoy, S.; Cinbis, R.G.; Kösemehmetoğlu, K.; Önder, S.; Üner, A. Graph Convolutional Networks for Region of Interest Classification in Breast Histopathology. In *Proceedings of Medical Imaging: Digital Pathology*, 2020.
104. Ye, H.; Wang, D.-H.; Li, J.; Zhu, S.; Zhu, C. Improving Histopathological Image Segmentation and Classification using Graph Convolution Network. *In Proceedings of the Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition,* Beijing, China, 2020; pp. 192–198.
105. Ozen, Y.; Aksoy, S.; Kösemehmetoğlu, K.; Önder, S.; Üner, A. Self-Supervised Learning with Graph Neural Networks for Region of Interest Retrieval in Histopathology. *In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR)*, 2021; pp. 6329-6334.
106. Zhao, Y.; Yang, F.; Fang, Y.; Liu, H.; Zhou, N.; Zhang, J.; Sun, J.; Yang, S.; Menze, B.; Fan, X.; et al. Predicting Lymph Node Metastasis Using Histopathological Images Based on Multiple Instance Learning With Deep Graph Convolution. *In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020; pp. 4836-4845.
107. Ding, K.; Liu, Q.; Lee, E.; Zhou, M.; Lu, A.; Zhang, S. Feature-Enhanced Graph Networks for Genetic Mutational Prediction Using Histopathological Images in Colon Cancer. *In Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Cham, 2020; pp. 294-304.
108. Adnan, M.; Kalra, S.; Tizhoosh, H.R. Representation Learning of Histopathology Images using Graph Neural Networks. *In Proceedings of the 2020 IEEE/CVF*

Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020; pp. 4254-4261.

109. Zheng, Y.; Jiang, B.; Shi, J.; Zhang, H.; Xie, F. Encoding Histopathological WSIs Using GNN for Scalable Diagnostically Relevant Regions Retrieval. *In Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, Cham, 2019; pp. 550-558.
110. Wu, J.; Zhong, J.-X.; Chen, E.Z.; Zhang, J.; Ye, J.J.; Yu, L. Weakly- and Semi-supervised Graph CNN for Identifying Basal Cell Carcinoma on Pathological Images. *In Proceedings of the Graph Learning in Medical Imaging*, Cham, 2019; pp. 112-119.
111. Anklin, V.; Pati, P.; Jaume, G.; Bozorgtabar, B.; Foncubierta-Rodriguez, A.; Thiran, J.-P.; Sibony, M.; Gabrani, M.; Goksel, O. Learning Whole-Slide Segmentation from Inexact and Incomplete Labels Using Tissue Graphs. *In Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Cham, 2021; pp. 636-646.
112. Raju, A.; Yao, J.; Haq, M.M.; Jonnagaddala, J.; Huang, J. Graph Attention Multi-instance Learning for Accurate Colorectal Cancer Staging. *In Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Cham, 2020; pp. 529-539.
113. Lu, W.; Graham, S.; Bilal, M.; Rajpoot, N.; Minhas, F. Capturing Cellular Topology in Multi-Gigapixel Pathology Images. *In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020; pp. 1049-1058.
114. Zhang, M.; Dong, B.; Li, Q. Ms-Gwnn: Multi-Scale Graph Wavelet Neural Network for Breast Cancer Diagnosis. In Proceedings of the 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), 2022; pp. 1-5.
115. Pati, P.; Jaume, G.; Fernandes, L.A.; Foncubierta-Rodríguez, A.; Feroce, F.; Anniciello, A.M.; Scognamiglio, G.; Brancati, N.; Riccio, D.; Di Bonito, M.; et al. HACT-Net: A Hierarchical Cell-to-Tissue Graph Neural Network for Histopathological Image Classification. In *Proceedings of the Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, Cham, 2020; pp. 208-219.

116. Pati, P.; Jaume, G.; Foncubierta-Rodríguez, A.; Feroce, F.; Anniciello, A.M.; Scognamiglio, G.; Brancati, N.; Fiche, M.; Dubruc, E.; Riccio, D.; et al. Hierarchical Graph Representations in Digital Pathology. *Medical Image Analysis* 2022, *75*, 102264, doi:https://doi.org/10.1016/j.media.2021.102264.
117. Levy, J.; Haudenschild, C.; Barwick, C.; Christensen, B.; Vaickus, L. Topological Feature Extraction and Visualization of Whole Slide Images Using Graph Neural Networks. In *Biocomputing* 2021; pp. 285-296.
118. Pope, P.E.; Kolouri, S.; Rostami, M.; Martin, C.E.; Hoffmann, H. Explainability Methods for Graph Convolutional Neural Networks. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019; pp. 10764-10773.
119. Zhang, H.; Liu, J.; Yu, Z.; Wang, P. MASG-GAN: A Multi-View Attention Superpixel-Guided Generative Adversarial Network for Efficient and Simultaneous Histopathology Image Segmentation and Classification. *Neurocomputing* 2021, *463*, 275-291, doi:https://doi.org/10.1016/j.neucom.2021.08.039.
120. Ertosun, M.G.; Rubin, D.L. Automated Grading of Gliomas using Deep Learning in Digital Pathology Images: A Modular Approach with Ensemble of Convolutional Neural Networks. *AMIA Annu Symp Proc* 2015, *2015*, 1899-1908.
121. Pati, P.; Jaume, G.; Fernandes, L.A.; Foncubierta-Rodríguez, A.; Feroce, F.; Anniciello, A.M.; Scognamiglio, G.; Brancati, N.; Riccio, D.; Di Bonito, M. Hact-Net: A Hierarchical Cell-to-Tissue Graph Neural Network for Histopathological Image Classification. *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging and Graphs in Biomedical Image Analysis: Second International Workshop, UNSURE 2020, and Third International Workshop, GRAIL 2020, Lima, Peru, 2020, pp. 208–219*.
122. Ma, Y.; Yang, Z. Multi-Instance Learning by Utilizing Structural Relationships Among Instances. *arXiv preprint*, arXiv*:2102.01889* 2021.
123. Wu, F.; Liu, P.; Fu, B.; Ye, F. DeepGCNMIL: Multi-Head Attention Guided Multi-Instance Learning Approach for Whole-Slide Image Survival Analysis Using Graph Convolutional Networks. *Proceedings of the 14th International Conference on Machine Learning and Computing*, 2022; pp. 67-73.

124. Alon, Y.; Zhou, H. Neuroplastic Graph Attention Networks for Nuclei Segmentation in Histopathology Images. *arXiv preprint*, arXiv*:2201.03669* **2022**.
125. Pelikan, M.; Goldberg, D.E.; Cantú-Paz, E. BOA: The Bayesian Optimization Algorithm. *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation,* Orlando, Florida, 1999; pp. 525–532.
126. Pelikan, M.; Goldberg, D.E. Hierarchical Bayesian Optimization Algorithm. In Pelikan, M., Sastry, K., & Cantú-Paz, E. (Eds.), *Scalable Optimization via Probabilistic Modeling.* Springer Berlin Heidelberg, 2006; pp. 63-90.
127. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. *Proceedings of the 25th International Conference on Neural Information Processing Systems,* Lake Tahoe, Nevada, 2012; pp. 2951–2959.
128. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for Hyper-Parameter Optimization. *Proceedings of the 24th International Conference on Neural Information Processing Systems,* Granada, Spain, 2011; pp. 2546–2554.
129. Cun, Y.L.; Denker, J.S.; Solla, S.A. Optimal Brain Damage. In *Advances in Neural Information Processing Systems 2,* Morgan Kaufmann Publishers Inc., 1990; pp. 598–605.
130. Chen, W.; Wilson, J.T.; Tyree, S.; Weinberger, K.Q.; Chen, Y. Compressing Neural Networks with the Hashing Trick. *Proceedings of the 32nd International Conference on Machine Learning,* Lille, France, 2015; pp. 2285–2294.
131. Choudhary, T.; Mishra, V.; Goswami, A.; Sarangapani, J. A Comprehensive Survey on Model Compression and Acceleration. *Artificial Intelligence Review* **2020**, *53*, 5113-5155, doi:10.1007/s10462-020-09816-7.
132. Junayed, M.S.; Islam, M.B.; Imani, H.; Aydin, T. PDS-Net: A Novel Point and Depth-Wise Separable Convolution for Real-Time Object Detection. *International Journal of Multimedia Information Retrieval* 2022, *11*, 171-188.
133. Wang, T.; Ray, N. Compact Depth-Wise Separable Precise Network for Depth Completion. *IEEE Access* 2023.
134. Shorten, C.; Khoshgoftaar, T.M. A Survey on Image Data Augmentation for Deep Learning. *Journal of Big Data* 2019, *6*, 60, doi:10.1186/s40537-019-0197-0.

135. Hagos, Y.B.; Narayanan, P.L.; Akarca, A.U.; Marafioti, T.; Yuan, Y. ConCORDe-Net: Cell Count Regularized Convolutional Neural Network for Cell Detection in Multiplex Immunohistochemistry Images. *Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019*, Shenzhen, China, 2019; pp. 667-675.
136. Pham, B.; Gaonkar, B.; Whitehead, W.; Moran, S.; Dai, Q.; Macyszyn, L.; Edgerton, V.R. Cell Counting and Segmentation of Immunohistochemical Images in the Spinal Cord: Comparing Deep Learning and Traditional Approaches. *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018; pp. 842-845.
137. Singh, H.; Kaur, H. A Systematic Survey on Biological Cell Image Segmentation and Cell Counting Techniques in Microscopic Images Using Machine Learning. *Wireless Personal Communications* 2024, *137*, 813-851.
138. Žerovnik Mekuč, M.; Bohak, C.; Hudoklin, S.; Kim, B.H.; Romih, R.; Kim, M.Y.; Marolt, M. Automatic Segmentation of Mitochondria and Endolysosomes in Volumetric Electron Microscopy Data. *Computers in Biology and Medicine* 2020, *119*, 103693, doi:https://doi.org/10.1016/j.compbiomed.2020.103693.
139. Tai, S.-K.; Lo, Y.-S. Using Deep Learning to Evaluate the Segmentation of Liver Cell from Biopsy Image. *2018 9th International Conference on Awareness Science and Technology (iCAST)* **2018**, 232-235.
140. Wu, C.; Chen, Q.; Wang, H.; Guan, Y.; Mian, Z.; Huang, C.; Ruan, C.; Song, Q.; Jiang, H.; Pan, J. A Review of Deep Learning Approaches for Multimodal Image Segmentation of Liver Cancer. *Journal of Applied Clinical Medical Physics* 2024, e14540.
141. Tran, T.; Kwon, O.H.; Kwon, K.R.; Lee, S.H.; Kang, K.W. Blood Cell Images Segmentation using Deep Learning Semantic Segmentation. *Proceedings of the 2018 IEEE International Conference on Electronics and Communication Engineering (ICECE)*, 2018; pp. 13-16.
142. Asha, S.; Gopakumar, G. Deep Learning-Based Semantic Segmentation of Blood Cells from Microscopic Images. Proceedings of the International Conference on Big Data, Machine Learning, and Applications, 2021; pp. 381-394.
143. Araújo, F.H.D.; Silva, R.R.V.; Ushizima, D.M.; Rezende, M.T.; Carneiro, C.M.; Campos Bianchi, A.G.; Medeiros, F.N.S. Deep Learning for Cell Image Segmentation and

Ranking. *Computerized Medical Imaging and Graphics* 2019, *72*, 13-21, doi:https://doi.org/10.1016/j.compmedimag.2019.01.003.

144. Huang, Q.; Zhang, W.; Chen, Y.; Chen, J.; Yang, Z. Review of Cervical Cell Segmentation. *Multimedia Tools and Applications* 2024, 1-40.

145. Joseph, N.; Kolluru, C.; Benetz, B.A.M.; Menegay, H.J.; Lass, J.H.; Wilson, D.L. Quantitative and Qualitative Evaluation of Deep Learning Automatic Segmentations of Corneal Endothelial Cell Images of Reduced Image Quality Obtained Following Cornea Transplant. *J Med Imaging (Bellingham)* 2020, *7*, 014503, doi:10.1117/1.Jmi.7.1.014503.

146. Karmakar, R.; Nooshabadi, S.V.; Eghrari, A.O. Mobile-CellNet: Automatic Segmentation of Corneal Endothelium Using an Efficient Hybrid Deep Learning Model. *Cornea* 2023, *42*, 456-463.

147. Ayanzadeh, A.; Yağar, H.O.; Özuysal, Ö.Y.; Okvur, D.P.; Töreyin, B.U.; Ünay, D.; Önal, S. Cell Segmentation of 2D Phase-Contrast Microscopy Images with Deep Learning Method. In Proceedings of the 2019 Medical Technologies Congress (TIPTEKNO), 2019; pp. 1-4.

148. Cao, R.; Meng, Q.; Tan, D.; Wei, P.; Ding, Y.; Zheng, C. AER-Net: Attention-Enhanced Residual Refinement Network for Nuclei Segmentation and Classification in Histology Images. *Sensors (Basel, Switzerland)* 2024, *24*, 7208.

149. Traoré, M.; Hancer, E.; Samet, R.; Yıldırım, Z.; Nemati, N. CompSegNet: An Enhanced U-Shaped Architecture for Nuclei Segmentation in H&E Histopathology Images. *Biomedical Signal Processing and Control* 2024, *97*, 106699.

150. Hu, Z.; Tang, J.; Wang, Z.; Zhang, K.; Zhang, L.; Sun, Q. Deep Learning for Image-Based Cancer Detection and Diagnosis: A Survey. *Pattern Recognition* 2018, *83*, 134-149.

151. Xing, F.; Yang, L. Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review. *IEEE Reviews in Biomedical Engineering* 2016, *9*, 234-263.

152. Mattie, M.E.; Staib, L.; Stratmann, E.; Tagare, H.D.; Duncan, J.; Miller, P.L. PathMaster: Content-Based Cell Image Retrieval Using Automated Feature Extraction. *Journal of the American Medical Informatics Association* 2000, *7*, 404-415.

153. Fatichah, C.; Tangel, M.L.; Widyanto, M.R.; Dong, F.; Hirota, K. Parameter Optimization of Local Fuzzy Patterns Based on Fuzzy Contrast Measure for White Blood Cell Texture Feature Extraction. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 2012, *16*, 412-419.
154. Chen, S.; Zhao, M.; Wu, G.; Yao, C.; Zhang, J. Recent Advances in Morphological Cell Image Analysis. *Computational and Mathematical Methods in Medicine* 2012, 2012, 101536.
155. Paul, P.R.; Bhowmik, M.K.; Bhattacharjee, D. Automated Cervical Cancer Detection Using Pap Smear Images. In *Proceedings of the Fourth International Conference on Soft Computing for Problem Solving: SocProS 2014*, Volume 1, 2015; pp. 267-278.
156. Bora, K.; Chowdhury, M.; Mahanta, L.B.; Kundu, M.K.; Das, A.K. Automated Classification of Pap Smear Images to Detect Cervical Dysplasia. *Computer Methods and Programs in Biomedicine* 2017, *138*, 31-47.
157. Mao, K.Z.; Zhao, P.; Tan, P.-H. Supervised Learning-Based Cell Image Segmentation for p53 Immunohistochemistry. *IEEE Transactions on Biomedical Engineering* 2006, *53*, 1153-1163.
158. Wang, P.; Hu, X.; Li, Y.; Liu, Q.; Zhu, X. Automatic Cell Nuclei Segmentation and Classification of Breast Cancer Histopathology Images. *Signal Processing* 2016, *122*, 1-13.
159. Manivannan, S.; Li, W.; Akbar, S.; Wang, R.; Zhang, J.; McKenna, S.J. An Automated Pattern Recognition System for Classifying Indirect Immunofluorescence Images of HEp-2 Cells and Specimens. *Pattern Recognition* 2016, *51*, 12-26.
160. Peikari, M.; Martel, A.L. Automatic Cell Detection and Segmentation from H and E Stained Pathology Slides Using Colorspace Decorrelation Stretching. In *Proceedings of the Medical Imaging 2016: Digital Pathology*, 2016; pp. 292-297.
161. Kowal, M.; Filipczuk, P.; Obuchowicz, A.; Korbicz, J.; Monczak, R. Computer-Aided Diagnosis of Breast Cancer Based on Fine Needle Biopsy Microscopic Images. *Computers in Biology and Medicine* 2013, *43*, 1563-1572.
162. Awan, R.; Aloraidi, N.; Qidwai, U.; Rajpoot, N. How Divided Is a Cell? Eigenphase Nuclei for Classification of Mitotic Phase in Cancer Histology Images. In *Proceedings*

of the 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), 2016; pp. 70-73.

163. Yuan, Y.; Failmezger, H.; Rueda, O.M.; Ali, H.R.; Gräf, S.; Chin, S.F.; Schwarz, R.F.; Curtis, C.; Dunning, M.J.; Bardwell, H.; et al. Quantitative Image Analysis of Cellular Heterogeneity in Breast Tumors Complements Genomic Profiling. *Sci Transl Med* 2012, *4*, 157ra143, doi:10.1126/scitranslmed.3004330.
164. Li, F.; Zhou, X.; Ma, J.; Wong, S.T. Multiple Nuclei Tracking Using Integer Programming for Quantitative Cancer Cell Cycle Analysis. *IEEE Transactions on Medical Imaging* 2009, *29*, 96-105.
165. Park, C.; Huang, J.Z.; Ji, J.X.; Ding, Y. Segmentation, Inference, and Classification of Partially Overlapping Nanoparticles. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2012, *35*, 1-1.
166. Mohammed, E.A.; Mohamed, M.M.; Naugler, C.; Far, B.H. Chronic Lymphocytic Leukemia Cell Segmentation from Microscopic Blood Images Using Watershed Algorithm and Optimal Thresholding. In *Proceedings of the 2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2013; pp. 1-5.
167. Abas, F.; Gokozan, H.N.; Goksel, B.; Otero, J.; Gurcan, M. *Intraoperative Neuropathology of Glioma Recurrence: Cell Detection and Classification. Proceedings of SPIE Medical Imaging, 2016*; p. 979109.
168. Tareef, A.; Song, Y.; Huang, H.; Feng, D.; Chen, M.; Wang, Y.; Cai, W. Multi-Pass Fast Watershed for Accurate Segmentation of Overlapping Cervical Cells. *IEEE Transactions on Medical Imaging* 2018, *37*, 2044-2059, doi:10.1109/TMI.2018.2815013.
169. Veta, M.; van Diest, P.J.; Kornegoor, R.; Huisman, A.; Viergever, M.A.; Pluim, J.P. Automatic Nuclei Segmentation in H&E Stained Breast Cancer Histopathology Images. *PLOS One* 2013, *8*, e70221, doi:10.1371/journal.pone.0070221.
170. Ali, S.; Madabhushi, A. An Integrated Region-, Boundary-, Shape-Based Active Contour for Multiple Object Overlap Resolution in Histological Imagery. *IEEE Transactions on Medical Imaging* 2012, *31*, 1448-1460, doi:10.1109/TMI.2012.2190089.
171. Wienert, S.; Heim, D.; Saeger, K.; Stenzinger, A.; Beil, M.; Hufnagl, P.; Dietel, M.; Denkert, C.; Klauschen, F. Detection and Segmentation of Cell Nuclei in Virtual

Microscopy Images: A Minimum-Model Approach. *Scientific Reports* 2012, *2*, 503, doi:10.1038/srep00503.

172. Liao, M.; Zhao, Y.-q.; Li, X.-h.; Dai, P.-s.; Xu, X.-w.; Zhang, J.-k.; Zou, B.-j. Automatic Segmentation for Cell Images Based on Bottleneck Detection and Ellipse Fitting. *Neurocomputing* 2016, *173*, 615-622, doi:https://doi.org/10.1016/j.neucom.2015.08.006.
173. Xing, F.; Xie, Y.; Su, H.; Liu, F.; Yang, L. Deep Learning in Microscopy Image Analysis: A Survey. *IEEE Transactions on Neural Networks and Learning Systems* 2017, *29*, 4550-4568.
174. Malon, C.; Brachtel, E.; Cosatto, E.; Graf, H.P.; Kurata, A.; Kuroda, M.; Meyer, J.S.; Saito, A.; Wu, S.; Yagi, Y. Mitotic Figure Recognition: Agreement Among Pathologists and Computerized Detector. *Analytical Cellular Pathology* 2012, *35*, 97-100.
175. Zerhouni, E.; Lányi, D.; Viana, M.; Gabrani, M. Wide Residual Networks for Mitosis Detection. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 2017; pp. 924-928.
176. Li, C.; Wang, X.; Liu, W.; Latecki, L.J.; Wang, B.; Huang, J. Weakly Supervised Mitosis Detection in Breast Histopathology Images Using Concentric Loss. *Medical Image Analysis* 2019, *53*, 165-178.
177. Cireşan, D.C.; Giusti, A.; Gambardella, L.M.; Schmidhuber, J. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. In Proceedings of the *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference*, Nagoya, Japan, 2013; pp. 411-418.
178. Kumar, N.; Verma, R.; Sharma, S.; Bhargava, S.; Vahadane, A.; Sethi, A. A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology. *IEEE Transactions on Medical Imaging* 2017, *36*, 1550-1560.
179. Albarqouni, S.; Baur, C.; Achilles, F.; Belagiannis, V.; Demirci, S.; Navab, N. AggNet: Deep Learning from Crowds for Mitosis Detection in Breast Cancer Histology Images. *IEEE Transactions on Medical Imaging* 2016, *35*, 1313-1321.
180. Xie, W.; Noble, J.A.; Zisserman, A. Microscopy Cell Counting and Detection with Fully Convolutional Regression Networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 2018, *6*, 283-292.

181. Wang, S.; Wang, T.; Yang, L.; Yang, D.M.; Fujimoto, J.; Yi, F.; Luo, X.; Yang, Y.; Yao, B.; Lin, S. ConvPath: A Software Tool for Lung Adenocarcinoma Digital Pathological Image Analysis Aided by a Convolutional Neural Network. *EBioMedicine* 2019, *50*, 103-110.

182. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014; pp. 580-587.

183. Girshick, R. Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, 2015; pp. 1440-1448.

184. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Advances in Neural Information Processing Systems* 2015, *28*.

185. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, 2017; pp. 2961-2969.

186. Cai, Z.; Vasconcelos, N. Cascade R-CNN: Delving into High-Quality Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018; pp. 6154-6162.

187. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In *Proceedings of the IEEE International Conference on Computer Vision*, 2017; pp. 2980-2988.

188. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot Multibox Detector. In *Proceedings of the Computer Vision–ECCV 2016: 14th European Conference*, Amsterdam, The Netherlands, 2016; pp. 21-37.

189. Chen, K.; Pang, J.; Wang, J.; Xiong, Y.; Li, X.; Sun, S.; Feng, W.; Liu, Z.; Shi, J.; Ouyang, W. Hybrid Task Cascade for Instance Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019; pp. 4974-4983.

190. Qiao, S.; Chen, L.-C.; Yuille, A. Detectors: Detecting Objects with Recursive Feature Pyramid and Switchable Atrous Convolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021; pp. 10213-10224.

191. Shkolyar, A.; Gefen, A.; Benayahu, D.; Greenspan, H. Automatic Detection of Cell Divisions (Mitosis) in Live-Imaging Microscopy Images Using Convolutional Neural Networks. In *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015; pp. 743-746.
192. Dong, B.; Shao, L.; Da Costa, M.; Bandmann, O.; Frangi, A.F. Deep Learning for Automatic Cell Detection in Wide-Field Microscopy Zebrafish Images. In *Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015; pp. 772-776.
193. Johnson, J.W. Adapting Mask-RCNN for Automatic Nucleus Segmentation. *arXiv Preprint* arXiv*:1805.00500* **2018**.
194. Liu, F.; Yang, L. A Novel Cell Detection Method Using Deep Convolutional Neural Network and Maximum-Weight Independent Set. *Deep Learning and Convolutional Neural Networks for Medical Image Computing: Precision Medicine, High Performance, and Large-Scale Datasets,* 2017, 63-72.
195. Chandradevan, R.; Aljudi, A.A.; Drumheller, B.R.; Kunananthaseelan, N.; Amgad, M.; Gutman, D.A.; Cooper, L.A.; Jaye, D.L. Machine-Based Detection and Classification for Bone Marrow Aspirate Differential Counts: Initial Development Focusing on Nonneoplastic Cells. *Laboratory Investigation,* 2020, *100*, 98-109.
196. Yi, J.; Wu, P.; Jiang, M.; Huang, Q.; Hoeppner, D.J.; Metaxas, D.N. Attentive Neural Cell Instance Segmentation. *Medical Image Analysis,* 2019, *55*, 228-240.
197. Cheng, Z.; Qu, A. A Fast and Accurate Algorithm for Nuclei Instance Segmentation in Microscopy Images. *IEEE Access* 2020, *8*, 158679-158689.
198. Obeid, A.; Mahbub, T.; Javed, S.; Dias, J.; Werghi, N. NucDETR: End-to-End Transformer for Nucleus Detection in Histopathology Images. In *Proceedings of the Computational Mathematics Modeling in Cancer Analysis*, Cham, 2022; pp. 47-57.
199. Huang, J.; Li, H.; Wan, X.; Li, G. Affine-Consistent Transformer for Multi-Class Cell Nuclei Detection. In *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023; pp. 21327-21336.
200. Javed, S.; Mahmood, A.; Dias, J.; Werghi, N.; Rajpoot, N. Spatially Constrained Context-Aware Hierarchical Deep Correlation Filters for Nucleus Detection in Histology Images. *Medical Image Analysis,* 2021, *72*, 102104.

201. Santamaria-Pang, A.; Rittscher, J.; Gerdes, M.; Padfield, D. Cell Segmentation and Classification by Hierarchical Supervised Shape Ranking. In *Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, 2015; pp. 1296-1299.
202. Brieu, N.; Schmidt, G. Learning Size Adaptive Local Maxima Selection for Robust Nuclei Detection in Histopathology Images. In *Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017; pp. 937-941.
203. Song, J.; Xiao, L.; Molaei, M.; Lian, Z. Multi-Layer Boosting Sparse Convolutional Model for Generalized Nuclear Segmentation from Histopathology Images. *Knowledge-Based Systems,* 2019, *176*, 40-53.
204. Zhou, Y.; Onder, O.F.; Dou, Q.; Tsougenis, E.; Chen, H.; Heng, P.-A. Cia-Net: Robust Nuclei Instance Segmentation with Contour-Aware Information Aggregation. In *Proceedings of the Information Processing in Medical Imaging: 26th International Conference, IPMI 2019*, 2019; pp. 682-693.
205. Chen, H.; Qi, X.; Yu, L.; Dou, Q.; Qin, J.; Heng, P.-A. DCAN: Deep Contour-Aware Networks for Object Instance Segmentation from Histology Images. *Medical Image Analysis,* 2017, *36*, 135-146.
206. Alemi Koohbanani, N.; Jahanifar, M.; Gooya, A.; Rajpoot, N. Nuclear Instance Segmentation Using a Proposal-Free Spatially Aware Deep Learning Framework. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019,* 2019; pp. 622-630.
207. Oda, H.; Roth, H.R.; Chiba, K.; Sokolić, J.; Kitasaka, T.; Oda, M.; Hinoki, A.; Uchida, H.; Schnabel, J.A.; Mori, K. BESNet: Boundary-Enhanced Segmentation of Cells in Histopathological Images. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2018*, 2018; pp. 228-236.
208. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images. *Medical Image Analysis,* 2019, *58*, 101563.
209. Hörst, F.; Rempe, M.; Heine, L.; Seibold, C.; Keyl, J.; Baldini, G.; Ugurel, S.; Siveke, J.; Grünwald, B.; Egger, J.; et al. CellViT: Vision Transformers for Precise Cell

Segmentation and Classification. *Medical Image Analysis,* 2024, *94*, 103143, doi:https://doi.org/10.1016/j.media.2024.103143.

210. Jha, K.K.; Dutta, H.S. Mutual Information-Based Hybrid Model and Deep Learning for Acute Lymphocytic Leukemia Detection in Single Cell Blood Smear Images. *Computer Methods and Programs in Biomedicine*, 2019, *179*, 104987.
211. Long, F. Microscopy Cell Nuclei Segmentation with Enhanced U-Net. *BMC Bioinformatics*, 2020, *21*, 8, doi:10.1186/s12859-019-3332-1.
212. Mahmood, F.; Borders, D.; Chen, R.J.; Mckay, G.N.; Salimian, K.J.; Baras, A.; Durr, N.J. Deep Adversarial Training for Multi-Organ Nuclei Segmentation in Histopathology Images. *IEEE Transactions on Medical Imaging,* 2020, *39*, 3257-3267, doi:10.1109/TMI.2019.2927182.
213. Doan, T.N.N.; Song, B.; Vuong, T.T.L.; Kim, K.; Kwak, J.T. SONNET: A Self-Guided Ordinal Regression Neural Network for Segmentation and Classification of Nuclei in Large-Scale Multi-Tissue Histology Images. *IEEE Journal of Biomedical and Health Informatics,* 2022, *26*, 3218-3228, doi:10.1109/JBHI.2022.3149936.
214. He, H.; Huang, Z.; Ding, Y.; Song, G.; Wang, L.; Ren, Q.; Wei, P.; Gao, Z.; Chen, J. CDNet: Centripetal Direction Network for Nuclear Instance Segmentation. In *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021; pp. 4006-4015.
215. Yu, X.; Li, G.; Lou, W.; Liu, S.; Wan, X.; Chen, Y.; Li, H. Diffusion-Based Data Augmentation for Nuclei Image Segmentation. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, 2023; pp. 592-602.
216. Zhang, D.; Song, Y.; Liu, D.; Jia, H.; Liu, S.; Xia, Y.; Huang, H.; Cai, W. Panoptic Segmentation with an End-to-End Cell R-CNN for Pathology Image Analysis. In *Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 2018; pp. 237-244.
217. Loh, D.R.; Yong, W.X.; Yapeter, J.; Subburaj, K.; Chandramohanadas, R. A Deep Learning Approach to the Screening of Malaria Infection: Automated and Rapid Cell Counting, Object Detection and Instance Segmentation Using Mask R-CNN. *Computerized Medical Imaging and Graphics,* 2021, *88*, 101845, doi:https://doi.org/10.1016/j.compmedimag.2020.101845.

218. Paing, M.P.; Sento, A.; Bui, T.H.; Pintavirooj, C. Instance Segmentation of Multiple Myeloma Cells Using Deep-Wise Data Augmentation and Mask R-CNN. *Entropy,* 2022, *24*, 134.
219. Xie, Y.; Xing, F.; Kong, X.; Su, H.; Yang, L. Beyond Classification: Structured Regression for Robust Cell Detection Using Convolutional Neural Network. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 2015; pp. 358-365.
220. Höfener, H.; Homeyer, A.; Weiss, N.; Molin, J.; Lundström, C.F.; Hahn, H.K. Deep Learning Nuclei Detection: A Simple Approach Can Deliver State-of-the-Art Results. *Computerized Medical Imaging and Graphics,* 2018, *70*, 43-52, doi:https://doi.org/10.1016/j.compmedimag.2018.08.010.
221. Feng, Z.; Wang, Z.; Wang, X.; Mao, Y.; Li, T.; Lei, J.; Wang, Y.; Song, M. Mutual-Complementing Framework for Nuclei Detection and Segmentation in Pathology Image. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021; pp. 4016-4025.
222. Abousamra, S.; Belinsky, D.; Arnam, J.V.; Allard, F.; Yee, E.; Gupta, R.; Kurc, T.; Samaras, D.; Saltz, J.; Chen, C. Multi-Class Cell Detection Using Spatial Context Representation. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021; pp. 3985-3994.
223. Wang, Z.; Zhu, X.; Li, A.; Wang, Y.; Meng, G.; Wang, M. Global and Local Attentional Feature Alignment for Domain Adaptive Nuclei Detection in Histopathology Images. *Artificial Intelligence in Medicine,* 2022, *132*, 102341, doi:https://doi.org/10.1016/j.artmed.2022.102341.
224. Wu, L.; Chen, A.; Salama, P.; Dunn, K.W.; Delp, E.J. 3D Centroidnet: Nuclei Centroid Detection with Vector Flow Voting. In Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP), 2022; pp. 651-655.
225. Xu, J.; Xiang, L.; Liu, Q.; Gilmore, H.; Wu, J.; Tang, J.; Madabhushi, A. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. *IEEE Transactions on Medical Imaging,* 2016, *35*, 119-130, doi:10.1109/TMI.2015.2458702.

226. Yousefi, S.; Nie, Y. Transfer Learning from Nucleus Detection to Classification in Histopathology Images. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) , 2019; pp. 957-960.
227. Sun, Y.; Huang, X.; Zhou, H.; Zhang, Q. SRPN: Similarity-Based Region Proposal Networks for Nuclei and Cells Detection in Histology Images. *Medical Image Analysi,* 2021, *72*, 102142, doi:https://doi.org/10.1016/j.media.2021.102142.
228. Li, X.; Xu, Z.; Shen, X.; Zhou, Y.; Xiao, B.; Li, T.-Q. Detection of Cervical Cancer Cells in Whole Slide Images Using Deformable and Global Context Aware Faster RCNN-FPN. *Current Oncology,* 2021, *28*, 3585-3601.
229. Wang, C.-W.; Huang, S.-C.; Lee, Y.-C.; Shen, Y.-J.; Meng, S.-I.; Gaol, J.L. Deep Learning for Bone Marrow Cell Detection and Classification on Whole-Slide Images. *Medical Image Analysis,* 2022, *75*, 102270, doi:https://doi.org/10.1016/j.media.2021.102270.
230. Liang, H.; Cheng, Z.; Zhong, H.; Qu, A.; Chen, L. A Region-Based Convolutional Network for Nuclei Detection and Segmentation in Microscopy Images. *Biomedical Signal Processing and Control,* 2022, *71*, 103276, doi:https://doi.org/10.1016/j.bspc.2021.103276.
231. Wu, L.; Han, S.; Chen, A.; Salama, P.; Dunn, K.W.; Delp, E.J. RCNN-SliceNet: A Slice and Cluster Approach for Nuclei Centroid Detection in Three-Dimensional Fluorescence Microscopy Images. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2021; pp. 3750-3760.
232. Wang, E.K.; Zhang, X.; Pan, L.; Cheng, C.; Dimitrakopoulou-Strauss, A.; Li, Y.; Zhe, N. Multi-Path Dilated Residual Network for Nuclei Segmentation and Detection. *Cells,* 2019, *8*, 499.
233. Chen, S.; Ding, C.; Tao, D. Boundary-assisted region proposal networks for nucleus segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2020, 2020; pp. 279-288.
234. Hatipoglu, N.; Bilgin, G. Cell Segmentation in Histopathological Images with Deep Learning Algorithms by Utilizing Spatial Relationships. *Medical & Biological Engineering & Computing,* 2017, *55*, 1829-1848.

235. Huang, J.; Li, H.; Sun, W.; Wan, X.; Li, G. Prompt-Based Grouping Transformer for Nucleus Detection and Classification. In Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, 2023; pp. 569-579.
236. Matsuo, K.; Purushotham, S.; Jiang, B.; Mandelbaum, R.S.; Takiuchi, T.; Liu, Y.; Roman, L.D. Survival Outcome Prediction in Cervical Cancer: Cox Models vs Deep-Learning Model. *American Journal of Obstetrics and Gynecology,* 2019, *220*, 381. e381-381. e314.
237. Fielding, L.P.; Fenoglio-Preiser, C.M.; Freedman, L.S. The Future of Prognostic Factors in Outcome Prediction for Patients with Cancer. *Cancer,* 1992, *70*, 2367-2377.
238. Laird, B.J.; Kaasa, S.; McMillan, D.C.; Fallon, M.T.; Hjermstad, M.J.; Fayers, P.; Klepstad, P. Prognostic Factors in Patients with Advanced Cancer: A Comparison of Clinicopathological Factors and the Development of an Inflammation-Based Prognostic System. *Clinical Cancer Research,* 2013, *19*, 5456-5464.
239. Glare, P.; Virik, K.; Jones, M.; Hudson, M.; Eychmuller, S.; Simes, J.; Christakis, N. A Systematic Review of Physicians' Survival Predictions in Terminally Ill Cancer Patients. *BMJ,* 2003, *327*, 195.
240. Kaplan, E.L.; Meier, P. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association,* 1958, *53*, 457-481.
241. Andersen, P.K.; Borgan, O.; Gill, R.D.; Keiding, N. Statistical Models Based on Counting Processes. *Springer Science & Business Media*: 2012.
242. Tibshirani, R. The Lasso Method for Variable Selection in the Cox Model. *Statistics in Medicine,* 1997, *16*, 385-395.
243. Suh, S.-Y.; Choi, Y.S.; Shim, J.Y.; Kim, Y.S.; Yeom, C.H.; Kim, D.; Park, S.A.; Kim, S.; Seo, J.Y.; Kim, S.H. Construction of a New, Objective Prognostic Score for Terminally Ill Cancer Patients: A Multicenter Study. *Supportive Care in Cancer,* 2010, *18*, 151-157.
244. Pirovano, M.; Maltoni, M.; Nanni, O.; Marinari, M.; Indelli, M.; Zaninetta, G.; Petrella, V.; Barni, S.; Zecca, E.; Scarpi, E. A New Palliative Prognostic Score: A First Step for the Staging of Terminally Ill Cancer Patients. *Journal of Pain and Symptom Management,* 1999, *17*, 231-239.

245. Chen, Y.-T.; Ho, C.-T.; Hsu, H.-S.; Huang, P.-T.; Lin, C.-Y.; Liu, C.-S.; Li, T.-C.; Lin, C.-C.; Lin, W.-Y. Objective Palliative Prognostic Score Among Patients with Advanced Cancer. *Journal of Pain and Symptom Management,* 2015, *49*, 690-696.

246. Chuang, R.-B.; Hu, W.-Y.; Chiu, T.-Y.; Chen, C.-Y. Prediction of Survival in Terminal Cancer Patients in Taiwan: Constructing a Prognostic Scale. *Journal of Pain and Symptom Management,* 2004, *28*, 115-122.

247. Morita, T.; Tsunoda, J.; Inoue, S.; Chihara, S. The Palliative Prognostic Index: A Scoring System for Survival Prediction of Terminally Ill Cancer Patients. *Supportive Care in Cancer,* 1999, *7*, 128-133.

248. Jiang, Y.; Xie, J.; Han, Z.; Liu, W.; Xi, S.; Huang, L.; Huang, W.; Lin, T.; Zhao, L.; Hu, Y. Immunomarker Support Vector Machine Classifier for Prediction of Gastric Cancer Survival and Adjuvant Chemotherapeutic Benefit. *Clinical Cancer Research,* 2018, *24*, 5574-5584.

249. Dimitriou, N.; Arandjelović, O.; Harrison, D.J.; Caie, P.D. A Principled Machine Learning Framework Improves Accuracy of Stage II Colorectal Cancer Prognosis. *NPJ Digital Medicine,* 2018, *1*, 52.

250. Zhu, X.; Yao, J.; Huang, J. Deep Convolutional Neural Network for Survival Analysis with Pathological Images. In *Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2016; pp. 544-547.

251. Kather, J.N.; Pearson, A.T.; Halama, N.; Jäger, D.; Krause, J.; Loosen, S.H.; Marx, A.; Boor, P.; Tacke, F.; Neumann, U.P. Deep Learning Can Predict Microsatellite Instability Directly from Histology in Gastrointestinal Cancer. *Nature Medicine,* 2019, *25*, 1054-1056.

252. Yao, J.; Zhu, X.; Jonnagaddala, J.; Hawkins, N.; Huang, J. Whole Slide Images-Based Cancer Survival Prediction Using Attention-Guided Deep Multiple Instance Learning Networks. *Medical Image Analysis,* 2020, *65*, 101789.

253. Katzman, J.L.; Shaham, U.; Cloninger, A.; Bates, J.; Jiang, T.; Kluger, Y. DeepSurv: Personalized Treatment Recommender System Using a Cox Proportional Hazards Deep Neural Network. *BMC Medical Research Methodology,* 2018, *18*, 1-12.

254. Yao, J.; Zhu, X.; Zhu, F.; Huang, J. Deep Correlational Learning for Survival Prediction from Multi-Modality Data. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2017; pp. 406-414.
255. Mobadersany, P.; Yousefi, S.; Amgad, M.; Gutman, D.A.; Barnholtz-Sloan, J.S.; Velázquez Vega, J.E.; Brat, D.J.; Cooper, L.A. Predicting Cancer Outcomes from Histology and Genomics Using Convolutional Networks. *Proceedings of the National Academy of Sciences,* 2018, *115*, E2970-E2979.
256. Yu, K.-H.; Zhang, C.; Berry, G.J.; Altman, R.B.; Ré, C.; Rubin, D.L.; Snyder, M. Predicting Non-Small Cell Lung Cancer Prognosis by Fully Automated Microscopic Pathology Image Features. *Nature Communications,* 2016, *7*, 12474.
257. Wulczyn, E.; Steiner, D.F.; Xu, Z.; Sadhwani, A.; Wang, H.; Flament-Auvigne, I.; Mermel, C.H.; Chen, P.-H.C.; Liu, Y.; Stumpe, M.C. Deep Learning-Based Survival Prediction for Multiple Cancer Types Using Histopathology Images. *PloS One,* 2020, *15*, e0233678.
258. Jiang, D.; Liao, J.; Duan, H.; Wu, Q.; Owen, G.; Shu, C.; Chen, L.; He, Y.; Wu, Z.; He, D. A Machine Learning-Based Prognostic Predictor for Stage III Colon Cancer. *Scientific Reports,* 2020, *10*, 10333.
259. Jolliffe, I.T.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 2016, *374*, 20150202.
260. Lee, Y.; Park, J.H.; Oh, S.; Shin, K.; Sun, J.; Jung, M.; Lee, C.; Kim, H.; Chung, J.-H.; Moon, K.C. Derivation of Prognostic Contextual Histopathological Features from Whole-Slide Images of Tumours via Graph Deep Learning. *Nature Biomedical Engineering,* 2022, 1-15.
261. Baghban, R.; Roshangar, L.; Jahanban-Esfahlan, R.; Seidi, K.; Ebrahimi-Kalan, A.; Jaymand, M.; Kolahian, S.; Javaheri, T.; Zare, P. Tumor Microenvironment Complexity and Therapeutic Implications at a Glance. *Cell Communication and Signaling,* 2020, *18*, 1-19.
262. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. *arXiv preprint arXiv:1710.10903* 2017.

263. Wang, Y.; Wang, Y.G.; Hu, C.; Li, M.; Fan, Y.; Otter, N.; Sam, I.; Gou, H.; Hu, Y.; Kwok, T. Cell Graph Neural Networks Enable the Precise Prediction of Patient Survival in Gastric Cancer. *NPJ Precision Oncology,* 2022, *6*, 45.
264. Gao, H.; Ji, S. Graph U-Nets. In *Proceedings of the International Conference on Machine Learning*, 2019; pp. 2083-2092.
265. Di, D.; Zhang, J.; Lei, F.; Tian, Q.; Gao, Y. Big-Hypergraph Factorization Neural Network for Survival Prediction from Whole Slide Image. *IEEE Transactions on Image Processing,* 2022, *31*, 1149-1160.
266. Feng, Y.; You, H.; Zhang, Z.; Ji, R.; Gao, Y. Hypergraph Neural Networks. In *Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence*, 2019; pp. 3558-3565.
267. Korn, E.L.; Freidlin, B.; Abrams, J.S. Overall Survival as the Outcome for Randomized Clinical Trials with Effective Subsequent Therapies. *Journal of Clinical Oncology,* 2011, *29*, 2439-2442.
268. Grønnesby, J.K.; Borgan, Ø. A Method for Checking Regression Models in Survival Analysis Based on the Risk Score. *Lifetime Data Analysis,* 1996, *2*, 315-328.
269. Klimstra, D.S.; Modlin, I.R.; Coppola, D.; Lloyd, R.V.; Suster, S. The Pathologic Classification of Neuroendocrine Tumors: A Review of Nomenclature, Grading, and Staging Systems. *Pancreas,* 2010, *39*, 707-712.
270. Blomjous, C.; Vos, W.; Schipper, N.; Uyterlinde, A.; Baak, J.; De Voogt, H.; Meijer, C. The Prognostic Significance of Selective Nuclear Morphometry in Urinary Bladder Carcinoma. *Human Pathology,* 1990, *21*, 409-413.
271. Nagpal, K.; Foote, D.; Liu, Y.; Chen, P.-H.C.; Wulczyn, E.; Tan, F.; Olson, N.; Smith, J.L.; Mohtashamian, A.; Wren, J.H. Development and Validation of a Deep Learning Algorithm for Improving Gleason Scoring of Prostate Cancer. *NPJ Digital Medicine,* 2019, *2*, 48.
272. Mercan, C.; Aksoy, S.; Mercan, E.; Shapiro, L.G.; Weaver, D.L.; Elmore, J.G. From Patch-Level to ROI-Level Deep Feature Representations for Breast Histopathology Classification. In *Proceedings of the Medical Imaging 2019: Digital Pathology*, 2019; pp. 86-93.

273. Daskalakis, A.; Kostopoulos, S.; Spyridonos, P.; Glotsos, D.; Ravazoula, P.; Kardari, M.; Kalatzis, I.; Cavouras, D.; Nikiforidis, G. Design of a Multi-Classifier System for Discriminating Benign from Malignant Thyroid Nodules Using Routinely H&E-Stained Cytological Images. *Computers in Biology and Medicine,* 2008, *38*, 196-203.

274. Reza, S.M.; Iftekharuddin, K.M. Glioma Grading Using Cell Nuclei Morphologic Features in Digital Pathology Images. In *Proceedings of the Medical Imaging 2016: Computer-Aided Diagnosis*, 2016; pp. 735-740.

275. Sheikh, T.S.; Lee, Y.; Cho, M. Histopathological Classification of Breast Cancer Images Using a Multi-Scale Input and Multi-Feature Network. *Cancers,* 2020, *12*, 2031.

276. Wang, X.; Wang, D.; Yao, Z.; Xin, B.; Wang, B.; Lan, C.; Qin, Y.; Xu, S.; He, D.; Liu, Y. Machine Learning Models for Multiparametric Glioma Grading with Quantitative Result Interpretations. *Frontiers in Neuroscience,* 2019, *12*, 1046.

277. Mousavi, H.S.; Monga, V.; Rao, G.; Rao, A.U. Automated Discrimination of Lower and Higher Grade Gliomas Based on Histopathological Image Analysis. *Journal of Pathology Informatics,* 2015, *6*, 15.

278. Brindha, V.; Jayashree, P.; Karthik, P.; Manikandan, P. Tumor Grading Model Employing Geometric Analysis of Histopathological Images with Characteristic Nuclei Dictionary. *Computers in Biology and Medicine,* 2022, *149*, 106008.

279. Romo-Bucheli, D.; Janowczyk, A.; Gilmore, H.; Romero, E.; Madabhushi, A. Automated Tubule Nuclei Quantification and Correlation with Oncotype DX Risk Categories in ER+ Breast Cancer Whole Slide Images. *Scientific Reports,* 2016, *6*, 32706.

280. Wang, H.; Cruz-Roa, A.; Basavanhally, A.; Gilmore, H.; Shih, N.; Feldman, M.; Tomaszewski, J.; Gonzalez, F.; Madabhushi, A. Mitosis Detection in Breast Cancer Pathology Images by Combining Handcrafted and Convolutional Neural Network Features. *Journal of Medical Imaging,* 2014, *1*, 034003-034003.

281. Yonekura, A.; Kawanaka, H.; Prasath, V.S.; Aronow, B.J.; Takase, H. Automatic Disease Stage Classification of Glioblastoma Multiforme Histopathological Images Using Deep Convolutional Neural Network. *Biomedical Engineering Letters,* 2018, *8*, 321-327.

282. Ker, J.; Bai, Y.; Lee, H.Y.; Rao, J.; Wang, L. Automated Brain Histology Classification Using Machine Learning. *Journal of Clinical Neuroscience,* 2019, *66*, 239-245.

283. Ertosun, M.G.; Rubin, D.L. Automated Grading of Gliomas Using Deep Learning in Digital Pathology Images: A Modular Approach with Ensemble of Convolutional Neural Networks. In *Proceedings of the AMIA Annual Symposium Proceedings*, 2015; p. 1899.
284. Awan, R.; Koohbanani, N.A.; Shaban, M.; Lisowska, A.; Rajpoot, N. Context-Aware Learning Using Transferable Features for Classification of Breast Cancer Histology Images. In *Proceedings of the Image Analysis and Recognition: 15th International Conference, ICIAR 2018*, 2018; pp. 788-795.
285. Lippoldt, F. Efficient Colon Cancer Grading with Graph Neural Networks. *arXiv Preprint* arXiv*:2010.01091* **2020**.
286. Su, Y.; Bai, Y.; Zhang, B.; Zhang, Z.; Wang, W. HAT-Net: A Hierarchical Transformer Graph Neural Network for Grading of Colorectal Cancer Histology Images. In *Proceedings of the BMVC*, 2021; p. 412.
287. Lu, W.; Toss, M.; Dawood, M.; Rakha, E.; Rajpoot, N.; Minhas, F. SlideGraph+: Whole Slide Image Level Graphs to Predict HER2 Status in Breast Cancer. *Medical Image Analysis,* 2022, *80*, 102486.
288. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009; pp. 248-255.
289. Hou, W.; Huang, H.; Peng, Q.; Yu, R.; Yu, L.; Wang, L. Spatial-Hierarchical Graph Neural Network with Dynamic Structure Learning for Histological Image Classification. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2022; pp. 181-191.
290. Wu, C.; Zheng, J.; Feng, Z.; Zhang, H.; Zhang, L.; Cao, J.; Yan, H. Fuzzy SLIC: Fuzzy Simple Linear Iterative Clustering. *IEEE Transactions on Circuits and Systems for Video Technology,* 2020, *31*, 2114-2124.
291. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv Preprint* arXiv*:2010.11929* 2020.
292. Hou, W.; Yu, L.; Lin, C.; Huang, H.; Yu, R.; Qin, J.; Wang, L. H^2-MIL: Exploring Hierarchical Representation with Heterogeneous Multiple Instance Learning for Whole

Slide Image Analysis. In *Proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence*, 2022; pp. 933-941.

293. Bai, Y.; Mi, Y.; Su, Y.; Zhang, B.; Zhang, Z.; Wu, J.; Huang, H.; Xiong, Y.; Gong, X.; Wang, W. A Scalable Graph-Based Framework for Multi-Organ Histology Image Classification. *IEEE Journal of Biomedical and Health Informatics,* 2022, *26*, 5506-5517.
294. Steinbach, M.; Tan, P.-N. kNN: k-Nearest Neighbors. In *The Top Ten Algorithms in Data Mining*; Chapman and Hall/CRC: 2009; pp. 165-176.
295. Liu, J.; Wang, C.; Gao, J.; Gu, Q.; Aggarwal, C.; Kaplan, L.; Han, J. GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data. In *Proceedings of the Proceedings of the 2015 SIAM International Conference on Data Mining*, 2015; pp. 388-396.
296. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is All You Need. *Advances in Neural Information Processing Systems,* 2017, *30*.
297. Gao, Z.; Shi, J.; Wang, J. GQ-GCN: Group Quadratic Graph Convolutional Network for Classification of Histopathological Images. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2021*, 2021; pp. 121-131.
298. Xu, B.; Shen, H.; Cao, Q.; Qiu, Y.; Cheng, X. Graph Wavelet Neural Network. *arXiv Preprint* arXiv*:1904.07785* 2019.
299. Baranwal, M.; Krishnan, S.; Oneka, M.; Frankel, T.; Rao, A. CGAT: Cell Graph Attention Network for Grading of Pancreatic Disease Histology Images. *Frontiers in Immunology,* 2021, *12*, 727610.
300. Gao, Z.; Lu, Z.; Wang, J.; Ying, S.; Shi, J. A Convolutional Neural Network and Graph Convolutional Network-Based Framework for Classification of Breast Histopathological Images. *IEEE Journal of Biomedical and Health Informatics,* 2022, *26*, 3163-3173, doi:10.1109/JBHI.2022.3153671.
301. Pati, P.; Jaume, G.; Foncubierta-Rodriguez, A.; Feroce, F.; Anniciello, A.M.; Scognamiglio, G.; Brancati, N.; Fiche, M.; Dubruc, E.; Riccio, D. Hierarchical Graph Representations in Digital Pathology. *Medical Image Analysis,* 2022, *75*, 102264.

302. Corso, G.; Cavalleri, L.; Beaini, D.; Liò, P.; Veličković, P. Principal Neighbourhood Aggregation for Graph Nets. *Advances in Neural Information Processing Systems,* 2020, *33*, 13260-13271.
303. Si, J.; Jia, W.; Jiang, H. Breast Cancer Histopathology Images Classification Through Multi-View Augmented Contrastive Learning and Pre-Learning Knowledge Distillation. *IEEE,* 2024, *12*, 25359-25371, doi:10.1109/ACCESS.2024.3366185.
304. Tokunaga, H.; Teramoto, Y.; Yoshizawa, A.; Bise, R. Adaptive Weighting Multi-Field-of-View CNN for Semantic Segmentation in Pathology. In *Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019; pp. 12597-12606.
305. D'Amato, M.; Szostak, P.; Torben-Nielsen, B. A Comparison Between Single- and Multi-Scale Approaches for Classification of Histopathology Images. *Front Public Health,* 2022, *10*, 892658, doi:10.3389/fpubh.2022.892658.
306. Campanella, G.; Hanna, M.G.; Geneslaw, L.; Miraflor, A.; Werneck Krauss Silva, V.; Busam, K.J.; Brogi, E.; Reuter, V.E.; Klimstra, D.S.; Fuchs, T.J. Clinical-Grade Computational Pathology Using Weakly Supervised Deep Learning on Whole Slide Images. *Nature Medicine,* 2019, *25*, 1301-1309, doi:10.1038/s41591-019-0508-1.
307. Jain, M.S.; Massoud, T.F. Predicting Tumour Mutational Burden from Histopathological Images Using Multiscale Deep Learning. *Nature Machine Intelligence,* 2020, *2*, 356-362, doi:10.1038/s42256-020-0190-5.
308. Schmitz, R.; Madesta, F.; Nielsen, M.; Krause, J.; Steurer, S.; Werner, R.; Rösch, T. Multi-Scale Fully Convolutional Neural Networks for Histopathology Image Segmentation: From Nuclear Aberrations to the Global Tissue Architecture. *Medical Image Analysis,* 2021, *70*, 101996, doi:https://doi.org/10.1016/j.media.2021.101996.
309. Song, Y.; Zhang, L.; Chen, S.; Ni, D.; Lei, B.; Wang, T. Accurate Segmentation of Cervical Cytoplasm and Nuclei Based on Multiscale Convolutional Network and Graph Partitioning. *IEEE Transactions on Biomedical Engineering,* 2015, *62*, 2421-2433, doi:10.1109/TBME.2015.2430895.
310. Kosaraju, S.C.; Hao, J.; Koh, H.M.; Kang, M. Deep-Hipo: Multi-Scale Receptive Field Deep Learning for Histopathological Image Analysis. *Methods,* 2020, *179*, 3-13, doi:https://doi.org/10.1016/j.ymeth.2020.05.012.

311. Hashimoto, N.; Fukushima, D.; Koga, R.; Takagi, Y.; Ko, K.; Kohno, K.; Nakaguro, M.; Nakamura, S.; Hontani, H.; Takeuchi, I. Multi-Scale Domain-Adversarial Multiple-Instance CNN for Cancer Subtype Classification with Unannotated Histopathological Images. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA*, 2020; pp. 3851-3860.

312. Marini, N.; Otálora, S.; Ciompi, F.; Silvello, G.; Marchesin, S.; Vatrano, S.; Caltagirone, G.; Atzori, M.; Müller, H. Multi-Scale Task Multiple-Instance Learning for the Classification of Digital Pathology Images with Global Annotations. *In Proceedings of COMPAY@MICCAI, Virtual Event*, 2021.

313. Li, B.; Li, Y.; Eliceiri, K.W. Dual-Stream Multiple-Instance Learning Network for Whole-Slide Image Classification with Self-Supervised Contrastive Learning. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021; pp. 14313-14323.

314. Bardou, D.; Zhang, K.; Ahmad, S.M. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks. *IEEE Access* 2018, *6*, 24680-24693.

315. Xie, S.; Hu, H.; Wu, Y. Deep Multi-Path Convolutional Neural Network Joint with Salient Region Attention for Facial Expression Recognition. *Pattern Recognition,* 2019, *92*, 177-191, doi:https://doi.org/10.1016/j.patcog.2019.03.019.

316. Tissera, D.; Vithanage, K.; Wijesinghe, R.; Kahatapitiya, K.; Fernando, S.; Rodrigo, R. Feature-Dependent Cross-Connections in Multi-Path Neural Networks. *In Proceedings of the 25th International Conference on Pattern Recognition*, 2021; pp. 4032-4039.

317. van Rijthoven, M.; Balkenhol, M.; Siliņa, K.; van der Laak, J.; Ciompi, F. HookNet: Multi-Resolution Convolutional Neural Networks for Semantic Segmentation in Histopathology Whole-Slide Images. *Medical Image Analysis,* 2021, *68*, 101890, doi:https://doi.org/10.1016/j.media.2020.101890.

318. Chen, Y.; Li, X.; Hu, K.; Chen, Z.; Gao, X. Nuclei Segmentation in Histopathology Images Using Rotation Equivariant and Multi-level Feature Aggregation Neural Network. *In Proceedings of the 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM),* 2020; pp. 549-554.

319. Shen, W.; Zhou, M.; Yang, F.; Yu, D.; Dong, D.; Yang, C.; Zang, Y.; Tian, J. Multi-Crop Convolutional Neural Networks for Lung Nodule Malignancy Suspiciousness Classification. *Pattern Recognition,* 2017, *61*, 663-673.

320. Nakach, F.Z.; Zerouaoui, H.; Idri, A. Binary Classification of Multi-Magnification Histopathological Breast Cancer Images Using Late Fusion and Transfer Learning. *Data Technologies and Applications,* 2023, *57*, 668-695.

321. Zhou, F.; Hang, R.; Liu, Q.; Yuan, X. Integrating Convolutional Neural Network and Gated Recurrent Unit for Hyperspectral Image Spectral-Spatial Classification. In *Proceedings of the Pattern Recognition and Computer Vision*, Cham, 2018; pp. 409-420.

322. Guo, Y.; Liu, Y.; Bakker, E.M.; Guo, Y.; Lew, M.S. CNN-RNN: A Large-Scale Hierarchical Image Classification Framework. *Multimedia Tools and Applications,* 2018, *77*, 10251-10271, doi:10.1007/s11042-017-5443-x.

323. Chen, J.; Ma, T.; Xiao, C. FastGCN: Fast Learning with Graph Convolutional Networks via Importance Sampling. *ArXiv* 2018, *abs/1801.10247*.

324. Quan, P.; Shi, Y.; Lei, M.; Leng, J.; Zhang, T.; Niu, L. A Brief Review of Receptive Fields in Graph Convolutional Networks. *In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence - Companion Volume, Thessaloniki, Greece*, 2019; pp. 106–110.

325. Ding, S.; Gao, Z.; Wang, J.; Lu, M.; Shi, J. Fractal Graph Convolutional Network with MLP-Mixer Based Multi-Path Feature Fusion for Classification of Histopathological Images. *Expert Systems with Applications,* 2023, *212*, 118793, doi:https://doi.org/10.1016/j.eswa.2022.118793.

326. Bazargani, R.; Fazli, L.; Gleave, M.; Goldenberg, L.; Bashashati, A.; Salcudean, S. Multi-Scale Relational Graph Convolutional Network for Multiple Instance Learning in Histopathology Images. *Medical Image Analysis,* 2024, *96*, 103197.

327. Shi, J.; Tang, L.; Li, Y.; Zhang, X.; Gao, Z.; Zheng, Y.; Wang, C.; Gong, T.; Li, C. A Structure-Aware Hierarchical Graph-Based Multiple Instance Learning Framework for PT Staging in Histopathological Image. *IEEE Transactions on Medical Imaging,* 2023, *42*, 3000-3011.

328. Gan, J.; Hu, R.; Mo, Y.; Kang, Z.; Peng, L.; Zhu, Y.; Zhu, X. Multigraph Fusion for Dynamic Graph Convolutional Network. *IEEE Transactions on Neural Networks and Learning Systems,* 2022, *35*, 196-207.
329. Ding, S.; Li, J.; Wang, J.; Ying, S.; Shi, J. Multi-Scale Efficient Graph-Transformer for Whole Slide Image Classification. *IEEE Journal of Biomedical and Health Informatics* 2023.
330. Abbas, S.F.; Le Vuong, T.T.; Kim, K.; Song, B.; Kwak, J.T. Multi-Cell Type and Multi-Level Graph Aggregation Network for Cancer Grading in Pathology Images. *Medical Image Analysis,* 2023, *90*, 102936.
331. Angel Arul Jothi, J.; Mary Anita Rajam, V. A Survey on Automated Cancer Diagnosis from Histopathology Images. *Artificial Intelligence Review,* 2017, *48*, 31-81.
332. Bao, G.; Wang, X.; Xu, R.; Loh, C.; Adeyinka, O.D.; Pieris, D.A.; Cherepanoff, S.; Gracie, G.; Lee, M.; McDonald, K.L. PathoFusion: An Open-Source AI Framework for Recognition of Pathomorphological Features and Mapping of Immunohistochemical Data. *Cancers,* 2021, *13*, 617.
333. Liu, S.; Liang, J.; Liu, Z.; Zhang, C.; Wang, Y.; Watson, A.H.; Zhou, C.; Zhang, F.; Wu, K.; Zhang, F. The Role of CD276 in Cancers. *Frontiers in Oncology,* 2021, *11*, 654684.
334. Zhang, J.; Wang, J.; Marzese, D.M.; Wang, X.; Yang, Z.; Li, C.; Zhang, H.; Zhang, J.; Chen, C.C.; Kelly, D.F. B7H3 Regulates Differentiation and Serves as a Potential Biomarker and Theranostic Target for Human Glioblastoma. *Laboratory Investigation,* 2019, *99*, 1117-1129.
335. Lemke, D.; Pfenning, P.-N.; Sahm, F.; Klein, A.-C.; Kempf, T.; Warnken, U.; Schnölzer, M.; Tudoran, R.; Weller, M.; Platten, M. Costimulatory Protein 4IgB7H3 Drives the Malignant Phenotype of Glioblastoma by Mediating Immune Escape and Invasiveness. *Clinical Cancer Research,* 2012, *18*, 105-117.
336. Wang, C.; Li, Y.; Jia, L.; koo Kim, J.; Li, J.; Deng, P.; Zhang, W.; Krebsbach, P.H.; Wang, C.-Y. CD276 Expression Enables Squamous Cell Carcinoma Stem Cells to Evade Immune Surveillance. *Cell Stem Cell,* 2021, *28*, 1597-1613. e1597.
337. Sezgin, M.; Sankur, B.l. Survey over Image Thresholding Techniques and Quantitative Performance Evaluation. *Journal of Electronic Imaging,* 2004, *13*, 146-168.

338. Pajares, G.; De La Cruz, J.M. A Wavelet-Based Image Fusion Tutorial. *Pattern Recognition,* 2004, *37*, 1855-1872.

339. Brady, M.; Yuille, A. An Extremum Principle for Shape from Contour. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 1984, 288-301.

340. Dubrofsky, E. Homography Estimation. *Diplomová Práce*, Vancouver: *University of British Columbia,* 2009, *5*.

341. Veta, M.; Pluim, J.P.; Van Diest, P.J.; Viergever, M.A. Breast Cancer Histopathology Image Analysis: A Review. *IEEE Transactions on Biomedical Engineering,* 2014, *61*, 1400-1411.

342. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J. Recent Advances in Convolutional Neural Networks. *Pattern Recognition,* 2018, *77*, 354-377.

343. Seaman, S.; Zhu, Z.; Saha, S.; Zhang, X.M.; Yang, M.Y.; Hilton, M.B.; Morris, K.; Szot, C.; Morris, H.; Swing, D.A. Eradication of Tumors through Simultaneous Ablation of CD276/B7-H3-Positive Tumor Cells and Tumor Vasculature. *Cancer Cell,* 2017, *31*, 501-515. e508.

344. Johnston, M.J.; Nikolic, A.; Ninkovic, N.; Guilhamon, P.; Cavalli, F.M.; Seaman, S.; Zemp, F.J.; Lee, J.; Abdelkareem, A.; Ellestad, K. High-Resolution Structural Genomics Reveals New Therapeutic Vulnerabilities in Glioblastoma. *Genome Research,* 2019, *29*, 1211-1222.

345. Sun, F.; Yu, X.; Ju, R.; Wang, Z.; Wang, Y. Antitumor Responses in Gastric Cancer by Targeting B7H3 via Chimeric Antigen Receptor T Cells. *Cancer Cell International,* 2022, *22*, 50.

346. Mao, Y.; Chen, L.; Wang, F.; Zhu, D.; Ge, X.; Hua, D.; Sun, J. Cancer Cell-Expressed B7-H3 Regulates the Differentiation of Tumor-Associated Macrophages in Human Colorectal Carcinoma. *Oncology Letters,* 2017, *14*, 6177-6183.

347. Digregorio, M.; Coppieters, N.; Lombard, A.; Lumapat, P.N.; Scholtes, F.; Rogister, B. The Expression of B7-H3 Isoforms in Newly Diagnosed Glioblastoma and Recurrence and Their Functional Role. *Acta Neuropathologica Communications,* 2021, *9*, 1-13.

348. Li, X.; Li, C.; Rahaman, M.M.; Sun, H.; Li, X.; Wu, J.; Yao, Y.; Grzegorzek, M. A Comprehensive Review of Computer-Aided Whole-Slide Image Analysis: From

Datasets to Feature Extraction, Segmentation, Classification, and Detection Approaches. *Artificial Intelligence Review,* 2022, *55*, 4809-4878.

349. Bandi, P.; Geessink, O.; Manson, Q.; Van Dijk, M.; Balkenhol, M.; Hermsen, M.; Bejnordi, B.E.; Lee, B.; Paeng, K.; Zhong, A. From Detection of Individual Metastases to Classification of Lymph Node Status at the Patient Level: The Camelyon17 Challenge. *IEEE Transactions on Medical Imaging,* 2018, *38*, 550-560.
350. Ahmedt-Aristizabal, D.; Armin, M.A.; Denman, S.; Fookes, C.; Petersson, L. A Survey on Graph-Based Deep Learning for Computational Histopathology. *Computerized Medical Imaging and Graphics,* 2022, *95*, 102027.
351. Alzoubi, I.; Bao, G.; Zhang, R.; Loh, C.; Zheng, Y.; Cherepanoff, S.; Gracie, G.; Lee, M.; Kuligowski, M.; Alexander, K.L.; et al. An Open-Source AI Framework for the Analysis of Single Cells in Whole-Slide Images with a Note on CD276 in Glioblastoma. *Cancers,* 2022, *14*, 3441.
352. Kandoth, C.; McLellan, M.D.; Vandin, F.; Ye, K.; Niu, B.; Lu, C.; Xie, M.; Zhang, Q.; McMichael, J.F.; Wyczalkowski, M.A. Mutational Landscape and Significance across 12 Major Cancer Types. *Nature,* 2013, *502*, 333-339.
353. Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv Preprint arXiv:1609.02907* **2016**.
354. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. *stat,* 2017, *1050*, 10-48550.
355. Zaheer, M.; Kottur, S.; Ravanbakhsh, S.; Poczos, B.; Salakhutdinov, R.R.; Smola, A.J. Deep Sets. *Advances in Neural Information Processing Systems,* 2017, *30*.
356. Ilse, M.; Tomczak, J.; Welling, M. Attention-Based Deep Multiple Instance Learning. In *Proceedings of the International Conference on Machine Learning*, 2018; pp. 2127-2136.
357. Jiang, B.; Zhang, T.; Liu, F.; Sun, Z.; Shi, H.; Hua, D.; Yang, C. The Co-Stimulatory Molecule B7-H3 Promotes the Epithelial-Mesenchymal Transition in Colorectal Cancer. *Oncotarget,* 2016, *7*, 31755.
358. Petterson, S.A.; Sørensen, M.D.; Burton, M.; Thomassen, M.; Kruse, T.A.; Michaelsen, S.R.; Kristensen, B.W. Differential Expression of Checkpoint Markers in the Normoxic and Hypoxic Microenvironment of Glioblastomas. *Brain Pathology,* 2023, *33*, e13111.

359. Zhu, X.; Yao, J.; Zhu, F.; Huang, J. WSISA: Making Survival Prediction from Whole Slide Histopathological Images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017; pp. 7234-7242.

360. Alzoubi, I.; Zhang, L.; Zheng, Y.; Loh, C.; Wang, X.; Graeber, M.B. PathoGraph: An Attention-Based Graph Neural Network Capable of Prognostication Based on CD276 Labelling of Malignant Glioma Cells. *Cancers* 2024, *16*, 750.

361. Weinstein, J.N.; Collisson, E.A.; Mills, G.B.; Shaw, K.R.; Ozenberger, B.A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J.M. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nature Genetics,* 2013, *45*, 1113-1120.

362. Bolhasani, H.; Amjadi, E.; Tabatabaeian, M.; Jassbi, S.J. A Histopathological Image Dataset for Grading Breast Invasive Ductal Carcinomas. *Informatics in Medicine Unlocked,* **2020**, *19*, 100341.

363. Lu, W.; Graham, S.; Bilal, M.; Rajpoot, N.; Minhas, F. Capturing Cellular Topology in Multi-Gigapixel Pathology Images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020; pp. 260-261.